Public Economics: Tools and Topics

Ben J. Heijdra

January 2017

Contents

List of Figures

List of Tables

List of Intermezzos

Preface

Chapter 1

What is public economics?

The purpose of this chapter is to discuss the following topics:

- What do we mean with public economics? What do public economists study?
- What have previous generations of economists written about this important field? A brief overview of the relevant history of economic thought.
- What are the topics we treat in detail? Which topics are left out and why?

1.1 What is public economics?

1.1.1 Definition

??? to be added. Based on Atkinson and Stiglitz (1980, Lecture 1).

1.1.2 Potential reasons for government intervention

- Basic role to establish and enforce the "rules of the economic game".
- Recall the basic theorems of welfare economics: if the economy is perfectly competitive and there is a full set of markets then the equilibrium (if it exists) is Pareto-efficient (no one can be made better off without someone else being made worse off).
- Further intervention may nevertheless be called for because:
  - Pareto efficiency does not ensure that the resulting distribution of resources is in accordance with the prevailing concepts of equity (redistribution).
  - Economy may not be perfectly competitive so that market allocation is not Pareto-efficient.
  - Full set of future and insurance markets may not exist (missing markets).

- Full equilibrium may not always be reached (e.g. unemployment).
- There may be externalities like pollution or congestion (corrective taxation/subsidization).
- For public goods, the market-based supply will not generally be correct (e.g. defence, basic research).
- "Merit wants": state discourages "bad" things (e.g. alcohol, tobacco) and encourages "good" things (e.g. education, social behaviour).

1.2 Who is who in public economics

??? to be added. Based on Musgrave (1959, 1985).

1.3 Topics included

These are the topics that will be covered in this course:

- "Positive" analysis of taxation policy.
  - Taxation and labour supply.
  - Taxation and saving.
  - Taxation and risk taking.
  - Taxation and the firm.
  - Tax incidence.
  - Taxation and economic growth.
- "Positive economics": concerns with **what is**.
- "Normative" analysis of policy.
  - Social welfare function.
  - Structure of indirect taxation.
  - Structure of income taxation.
  - Provision of public goods.
  - Dealing with externalities.
  - Pensions and pension reform.
- "Normative": concerns with **what should be**.

1.4 Topics deliberately left out

It is impossible to cover the entire field of public economics in a one-semester masters course. I have left out several topics that are very interesting:

- The political economy approach to public economics.
  - There are some recent text books by Persson and Tabellini (2000), and Drazen (2000).
  - Rationale for leaving out this topic: it warrants entire course in itself. Few robust results.
- Environmental taxation. Some examples can be used throughout the course. This is not a specialist environmental economics book.
- Tax avoidance and evasion.
- Law and economics.
- Public sector pricing [Belongs in a micro-regulation economics course].
- Local public sector [Belongs in a regional economics course].
- Cost benefit analysis [Interesting when very applied].
- International taxation issues [Belongs in an open-economy macro course].
- Public provision of education.
- Public health care system.
- Fiscal federalism: tax competition, debt and stability pact [Belongs in an open-economy macro course].
- Social choice theory [very very technical].

1.5 Related literature

- There does not at this stage exist a single good textbook for this course. All students are nevertheless advised to purchase the following book which covers the taxation part quite well: Salanié (2003).
- Since this book does not cover all the topics, additional teaching materials may be taken from the following sources: Myles (1995), Atkinson and Stiglitz (1980), Atkinson (1995), Tresch (2002), Jha (1998), Auerbach and Feldstein (1985, 1987, 2002a, 2002b), Heijdra and van der Ploeg (2002), Ihori (1996), and de la Croix and Michel (2002).
- There is a website under construction for this book containing slides, problem sets, and model solutions: `http://www.heijdra.org/pubecon.htm`

Further reading

??? What should interested students read?

Key literature

- Atkinson & Stiglitz (1980, lecture 1), Myles (1995, ch. 1).
- Musgrave (1959, 1985).
- History of thought: Schumpeter (1954), Robbins (1998), and Ekelund and Hebert (1990).
- Tools: Hausman (1981)

Part I

Positive economics

Chapter 2

Taxation and the supply of labour

The purpose of this chapter is to discuss the following topics:

- How can we model the labour supply decision of a representative agent?
- Can we move beyond the representative-agent model and study the supply of labour by the family?
- How do linear taxes, non-linear taxes, and unemployment benefits affect the labour supply decision by households?
- What does the empirical evidence say about the crucial elasticities appearing in the various models?

In this chapter we will focus on *static* models of labour supply (dynamic models are treated in Chapter 3). We study the effects of both *linear* and *non-linear tax systems* on labour supply. In addition to the supply of *hours* decision we also briefly touch on the so-called labour market *participation* decision. All models discussed in this chapter abstract from risk and uncertainty (decision making under uncertainty is studied in Chapter 4).

2.1 A basic model of labour supply

In this section we present a basic model of labour supply. This model is then used to study the effects of various taxes on goods consumption and labour supply. The key assumptions we make are the following. First, we postulate a so-called *representative agent* who (among other things) chooses the optimal number of hours to be supplied to the labour market. This representative agent can be interpreted either (literally) as comprising a single-person household, or (more realistically) as the head of a multi-person household making the family labour supply decisions. We abstract from heterogeneity by assuming that all households are exactly the same.[1] The price of this assumption is that we cannot study the distributional aspects of taxation with this basic labour supply model. There are many identical households but to cut down on the notation we normalize their number to unity.

Second, we assume that the representative household possesses all relevant information, i.e. he knows all final goods and factor prices and tax rates with complete certainty. Third, we abstract from dynamic considerations, such as the household's decision to save, by employing a static model. Fourth, in this section we restrict attention to the study of so-called linear (or "flat rate") taxes. In such a tax system, the marginal tax rate is constant, i.e. it does not depend on the size of the object or activity that is being taxed.

2.1.1 Setting up the model

The utility function of the representative household is given by:

$$U = U(C, \bar{L} - L), \tag{2.1}$$

where U is utility, C is goods consumption, L is the number of hours supplied to the labour market, and $\bar{L}$ is the exogenously given time endowment ($\bar{L} - L$ is thus the amount of leisure enjoyed by the household). We make the usual assumptions regarding the utility function, i.e. marginal utility of both consumption and leisure is positive, though each at a diminishing rate, and indifference curves bulge toward the origin (see, for example, Figure 2.1). In technical terms, the assumptions are represented as follows:[2]

$$\begin{array}{ll} U_C \equiv \frac{\partial U}{\partial C} > 0, & U_{\bar{L}-L} \equiv \frac{\partial U}{\partial(\bar{L}-L)} > 0, \\ U_{CC} \equiv \frac{\partial^2 U}{\partial C^2} < 0, & U_{\bar{L}-L,\bar{L}-L} \equiv \frac{\partial^2 U}{\partial(\bar{L}-L)^2} < 0, \\ U_{C,\bar{L}-L} \equiv \frac{\partial^2 U}{\partial C \partial(\bar{L}-L)} \gtreqless 0, & U_{CC}U_{\bar{L}-L,\bar{L}-L} - \left(U_{C,\bar{L}-L}\right)^2 > 0. \end{array} \tag{2.2}$$

The household budget restriction is given by:

$$P(1+t_C)C = M + WL - T\,(WL)\,, \tag{2.3}$$

where P is the price of the consumption good, t_C is the consumption tax, M is exogenous non-labour income, W is the pre-tax wage rate, and $T\,(WL)$ is the labour income tax function. This tax function is linear in wage income and takes the following form:

$$T\,(WL) \equiv T_0 + t_L WL, \tag{2.4}$$

[1] In Chapter 11 we relax this assumption and allow households to differ in various aspects.

[2] The assumptions ensure that the utility function is *strictly quasi-concave* in consumption and leisure. See Silberberg and Suen (2001, pp. 140, 260) for the proof that indifference curves bulge toward the origin for such a function.

where T_0 is the lump-sum part of the labour income tax and t_L is the marginal tax rate on labour. It is assumed that the marginal tax rate is positive but less than unity, i.e. $0 < t_L < 1$. Of course, if T_0 is negative then it constitutes a lump-sum transfer from the tax authority to the household.

The household chooses C and L in order to maximize utility (2.1) subject to the budget constraint (2.3) and taking into account the tax schedule (2.4). The Lagrangian associated with this standard optimization problem is:

$$\mathcal{L} \equiv U(C, \bar{L} - L) + \lambda \left[M - T_0 + (1 - t_L)WL - P(1 + t_C)C \right], \tag{2.5}$$

where λ is the Lagrange multiplier. The first-order necessary conditions are $\partial\mathcal{L}/\partial\lambda = 0$ (yielding the budget constraint) as well as:

$$\frac{\partial \mathcal{L}}{\partial C} = U_C - \lambda P(1 + t_C) = 0, \tag{2.6}$$

$$\frac{\partial \mathcal{L}}{\partial L} = -U_{\bar{L}-L} + \lambda W(1 - t_L) = 0. \tag{2.7}$$

By eliminating the Lagrange multiplier from these first-order conditions we obtain an expression characterizing the optimum choices:

$$\lambda = \frac{U_C}{P(1+t_C)} = \frac{U_{\bar{L}-L}}{W(1-t_L)} \qquad \Rightarrow$$

$$\frac{U_{\bar{L}-L}}{U_C} = w^*, \tag{2.8}$$

where w^* is the *after-tax* real wage rate:

$$w^* \equiv \frac{W}{P} \frac{1 - t_L}{1 + t_C}. \tag{2.9}$$

The key thing to note about (2.9) is that, since the choice of hours is a *marginal* decision,[3] it is also the marginal labour income tax rate which features in the relevant real wage rate expression. Furthermore, this wage rate depends not only on the gross real wage, $w \equiv W/P$, but also on both t_L and t_C. Although t_C is not directly applied to labour income, it nevertheless affects labour supply because it affects the tax-inclusive price of goods and thus influences what the household can buy with his labour- and non-labour income.

Equation (2.8) is an important expression which we will see time and again in various contexts. According to (2.8), the household chooses goods consumption and labour supply in such a way that the marginal rate of substitution between leisure and consumption (left-hand side of (2.8)) is equal to the

[3] With a marginal decision we mean that the household can vary L (and C) by infinitesimal amounts and can choose exactly the number of seconds per day he wants to supply to the labour market. In reality, of course, the length of the working day may well be fixed (say at 8 hours) so that the household faces the choice between working 8 hours or not working at all. We return to this important issue of *indivisible labour* in Section 2.2 below where we study the labour market participation issue.

after-tax real wage rate. We can illustrate the optimum choice of the household with the aid of Figure 2.1. In this figure, consumption is measured on the vertical axis and leisure on the horizontal axis. Given the assumptions made about utility in (2.2) above, indifference curves bulge toward the origin—see the line labeled $U = U_0$. Using (2.4), the budget line (2.3) can be written as:

$$C + w^* \left(\bar{L} - L\right) = \frac{M - T_0 + (1 - t_L) W\bar{L}}{P(1 + t_C)} \; [\equiv Y_0] \, , \tag{2.10}$$

which is downward sloping in $(C, \bar{L} - L)$ space—see the line AB in Figure 2.1. The left-hand side of (2.10) is real spending on consumption goods and on leisure (so-called *full consumption*) whereas the right-hand side is *full income*, i.e. total resources available to the household were it to supply its entire labour endowment to the labour market. Of course, by definition the household is unable to supply more labour than its time endowment, i.e. the total amount of leisure cannot exceed $\bar{L}$. The feasible region is thus the area 0ABC in Figure 2.1.

In the interior optimum, the slope of the indifference curve is equal to the slope of the budget line. As drawn in the figure, this optimum occurs at point E_0, where consumption is C^* and leisure is $(\bar{L} - L)^*$. We call an optimum such as at point E_0 an *interior solution* to the optimization problem because the first-order conditions hold with equality. If the household has a very strong preference for leisure, however, it may very well be the case that a *corner solution* will be optimal, i.e. maximum utility will be attained at point B. By using the Kuhn-Tucker conditions (see the Mathematical Appendix) we find that $U_{\bar{L}-L}/U_C > w^*$ in point B. Given w^*, the household would like to consume more leisure but the feasibility constraint ($L \geq 0$) prevents him from doing so. In the remainder we will largely ignore corner solutions and instead focus on interior solutions.

2.1.2 Tax effects

Despite its simplicity, the model can already be used for tax policy analysis. There are three different tax rates in the model, namely the lump-sum tax (T_0), the marginal labour income tax rate (t_L), and the consumption tax rate (t_C). In order to illustrate the effects on labour supply of these tax rates, we first conduct a graphical analysis for the *very special* (but often used) case of *homothetic preferences.* The household has homothetic preferences if its utility function, $U(C, \bar{L} - L)$, can be written as $G\left(\bar{U}(C, \bar{L} - L)\right)$ with $G(\cdot)$ strictly increasing and $\bar{U}(C, \bar{L} - L)$ homogeneous of degree one (*linear homogeneous*) in its arguments.[4]

It is easy to see why homothetic preferences are so convenient to work with. Indeed, if $U(C, \bar{L} - L)$

[4] An example of a utility function representing homothetic preferences is:

$$U = \ln\left(C^\alpha \left[\bar{L} - L\right]^\beta\right),$$

with $\alpha > 0$ and $\beta > 0$. By writing $G \equiv (\alpha + \beta) \ln \bar{U}$ and $\bar{U} \equiv C^\eta \left[\bar{L} - L\right]^{1-\eta}$ (with $\eta \equiv \alpha / (\alpha + \beta)$) we easily find that U is homothetic.

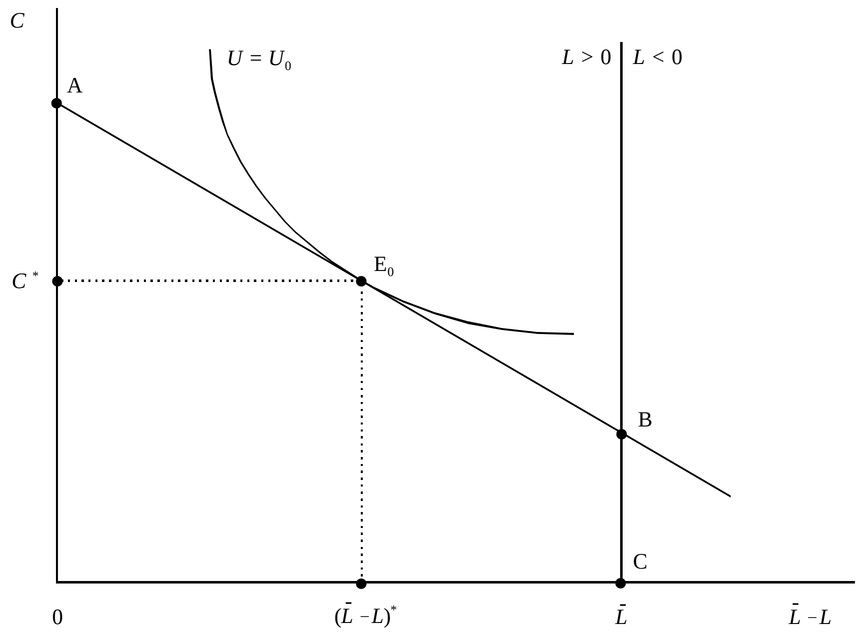

Figure 2.1: Optimal consumption leisure choice

is homothetic, the marginal rate of substitution between leisure and consumption can be written as:

$$\frac{U_{\bar{L}-L}}{U_C} = \frac{G'(\cdot)\,\bar{U}_{\bar{L}-L}}{G'(\cdot)\,\bar{U}_C} = \frac{\bar{U}_{\bar{L}-L}}{\bar{U}_C}, \tag{2.11}$$

where $\bar{U}_{\bar{L}-L}$ and $\bar{U}_C$ are the partial derivatives of the $\bar{U}(\cdot)$ function. But since $\bar{U}(\cdot)$ is homogeneous of degree one in its arguments, it follows that both $\bar{U}_{\bar{L}-L}$ and $\bar{U}_C$ are homogeneous of degree zero, i.e. they depend only on the consumption-leisure ratio, $C/(\bar{L}-L)$. By (2.11) we then obtain the result that $U_{\bar{L}-L}/U_C$ is a unique function of $C/(\bar{L}-L)$ also, i.e. the so-called *Income Expansion Path* is linear and passes through the origin. The advantage of working with a linear income expansion path lies in the fact that once we know *one* tangency between an indifference curve and a budget line, we know the position of *all* tangencies between budget lines of equal slope (but varying incomes) and corresponding indifference curves.

To see this principle at work, consider the effects of a rise in the lump-sum tax. As is illustrated in Figure 2.2, an increase in T_0 leads to a parallel shift downward of the budget line (see also equation (2.10) above). The feasible region shrinks from 0ABC to 0A′B′C. If the initial optimum is at E_0 then it must be the case that the new optimum is at point E_1. The slope of the budget line is unchanged so the household stays on the original income expansion path and just scales down consumption and leisure in equal proportions (both consumption and leisure are thus so-called *normal goods*). A change in the lump-sum tax thus only has an *income effect*. Utility falls, from U_0 to U_1, so even lump-sum taxes are painful from the perspective of the household!

A rise in the labour income tax or consumption tax leads to both a shift and a rotation of the budget

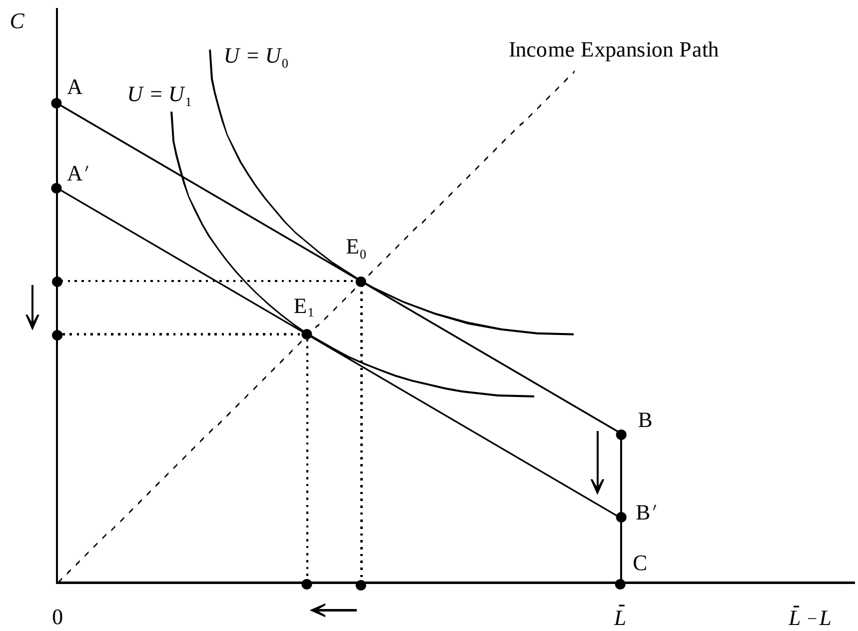

Figure 2.2: Increasing the lump-sum tax (homothetic case)

line (see (2.10) above). As a result there are both income effects (as before) and *substitution effects*. In Figure 2.3 we illustrate the effects of an increase in the labour income tax, t_L. The initial feasible region is given by 0ABC and the initial optimum is at E_0. As a result of the shock, the budget line rotates counter-clockwise around point B (at that point, $L = 0$ and t_L drops out of the budget line altogether). The new feasible region is given by 0A′BC. We assume that the optimum shifts from E_0 to E_1, so that C falls, $\bar{L} - L$ rises, and thus L falls. Utility also falls as a result of the tax increase, i.e. the indifference curve tangent to E_1 (not drawn) is associated with a lower utility level than the one for the initial situation ($U_1 < U_0$).

As usual, we can decompose the total effect on the variables into a *pure substitution effect* and an *income effect*. To compute the pure substitution effect, we find out which point on the original indifference curve the household would choose for the new after-tax real wage rate. In Figure 2.3 this is point E′ (the dashed line is parallel to the new budget line A′B). The pure substitution effect thus involves the move from E_0 to E′. It is easy to see that this effect is always positive for leisure and thus negative for labour supply. Of course, the household is not actually able to move to point E′, because it does not lie in its feasible region. The household would have to be compensated by a lump-sum transfer in order to be able to choose E′. For that reason we often refer to the hypothetical choices holding utility constant, *compensated* (or Hicksian) solutions.

The income effect consists of the move from E′ to E_1. It is easy to see that the income effect is negative for both consumption and leisure and is thus positive for labour supply. The case drawn in Figure 2.3 assumes that the substitution effect dominates the income effect in labour supply. It is quite possible, however, for the labour supply curve to be backward bending, i.e. for the income effect to dominate the substitution effect. For example, if there would be no substitution possible between consumption

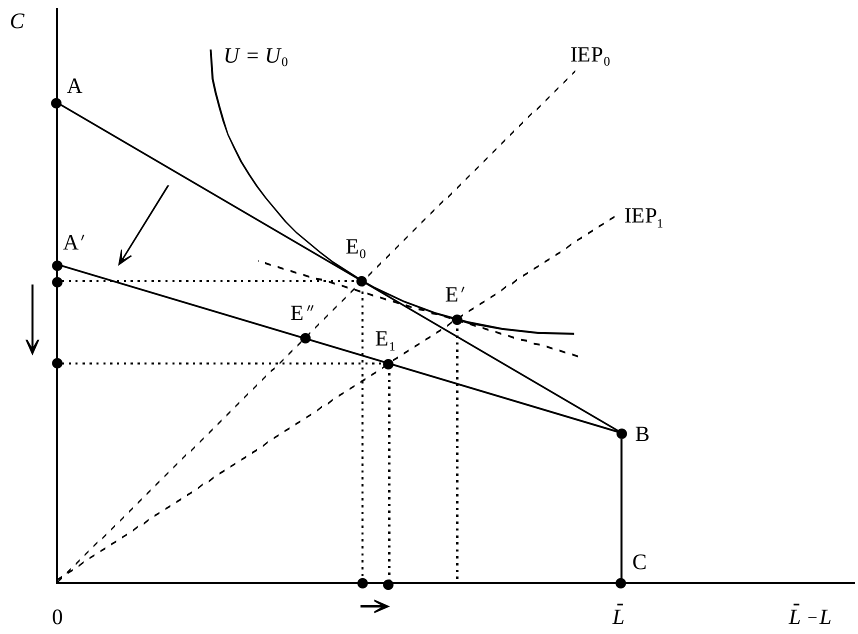

Figure 2.3: Increasing the labour income tax (homothetic case, dominant SE)

and leisure, so that $U(\cdot)$ takes the Leontief form,[5] then there is no pure substitution effect and the new optimum would be at E″!

A more formal method to decompose the total effect into its constituent parts involves the famous *Slutsky equation*. For labour supply this equation takes the following form:

$$\frac{\partial L}{\partial w^*} = \left(\frac{\partial L}{\partial w^*}\right)_{U=U_0} + L\frac{\partial L}{\partial m_0}, \tag{2.12}$$

where $m_0 \equiv (M - T_0) / [P(1 + t_C)]$ is real non-labour income. This expression is derived formally below, but let us first interpret it intuitively. The first term on the right-hand side is the pure substitution effect (non-negative) whereas the second term is the income effect (negative if leisure is a normal good). We reach an immediate insight from (2.12): if leisure is a normal good, then labour supply declines with full income, i.e. the income effect is relatively unimportant for rich households (whose high non-labour income will ensure a high level of full income). This rather subtle effect can be explained with the aid of Figure 2.4 which is based on the assumption that the utility function is of the Leontief type, i.e. there is no substitutability between consumption and leisure, the indifference curves are right angles, and the pure substitution effect is zero. There are two agents. The *poor* agent has no non-labour income ($m_0 = 0$) and initially faces the budget line CD. In contrast, the *rich* agent has a very high level of non-

[5] A Leontief utility function has the following form:

$$U = \min\left(\left[\frac{C}{\alpha}, \frac{\bar{L} - L}{\beta}\right]\right),$$

with $\alpha > 0$ and $\beta > 0$. The slope of the income expansion path is $C/(\bar{L} - L) = \alpha/\beta$ and the household keeps a constant consumption-leisure proportion no matter what the real wage rate is.

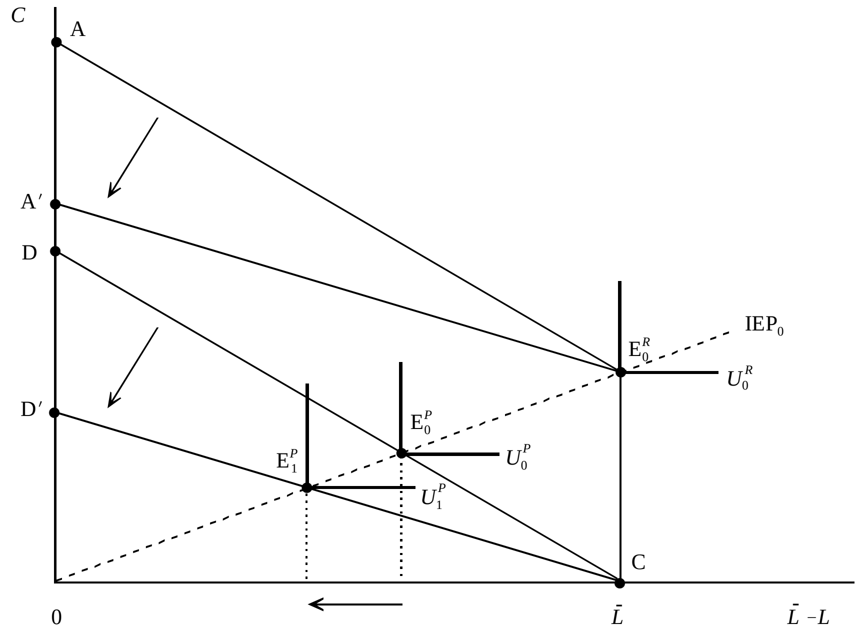

Figure 2.4: Increasing the labour income tax: the rich and the poor

labour income and initially faces the feasible choice set $CE_0^R A$. By assumption, the level of this agent's non-labour income is such that he wants to consume exactly $\bar{L}$ units of leisure, i.e. at point E_0^R labour supply for the rich agent is zero. Since the agents have the same utility function, the income expansion path curve passes through point E_0^R and it can be deduced that the poor agent's initial optimum is at point E_0^P. Now consider what happens if the labour income tax is increased. The budget lines rotate counter-clockwise to $E_0^R A'$ (for the rich agent) and CD' (for the poor agent). The rich agent stays at point E_0^R but the poor agent experiences a large income effect which induces him to move from E_0^P to E_1^P.

To investigate the effect of the labour income tax for the general (non-Leontief) case, we first note that labour supply can be written in general terms as $L = L(w^*, m_0)$, where w^* is defined in (2.9) above and the definition of m_0 is given directly below (2.12). Next we differentiate this expression with respect to t_L (noting that t_L affects w^* but not m_0):

$$\frac{\partial L}{\partial t_L} = \frac{\partial L}{\partial w^*}\frac{\partial w^*}{\partial t_L}. \tag{2.13}$$

By noting the definitions of w^* and using the Slutsky equation (2.12), we can rewrite this expression as follows:

$$\frac{\partial L}{\partial t_L} = -\frac{w}{1+t_C}\left[\left(\frac{\partial L}{\partial w^*}\right)_{U=U_0} + L\frac{\partial L}{\partial m_0}\right]. \tag{2.14}$$

If labour supply is upward sloping, the term in round brackets on the right-hand side of (2.14) is positive so that an increase in the tax rate reduces labour supply. Since the income effect is close to zero for wealthy households (provided leisure is a normal good), the supply of labour will certainly fall for such households because the pure substitution effect is non-negative.

The Slutsky equation (2.12) can be derived by making use of standard duality results (see the Intermezzo for a quick review of the basic duality tools used in this chapter). First, we define leisure as $H \equiv \bar{L} - L$ and note that the budget equation in real terms can be written as $\mathbf{1}\, C + w^* H = Y_0$, where Y_0 is full income (see the left-hand side of equation (2.10) above). Note that consumption is the numeraire commodity so its tax-inclusive price is $\mathbf{1}$ and w^* represents the relative price of leisure. Second, we define the expenditure and indirect utility functions as:

$$\begin{aligned} E\left(\mathbf{1}, w^*, U_0\right) &\equiv \quad \min_{\{C,H\}} \mathbf{1}\, C + w^* H \quad \text{subject to:} \quad U\left(C, H\right) = U_0, \\ V\left(\mathbf{1}, w^*, Y_0\right) &\equiv \quad \max_{\{C,H\}} U\left(C, H\right) \quad \text{subject to:} \quad Y_0 = \mathbf{1}\, C + w^* H. \end{aligned}$$

It follows that the Hicksian and Marshallian demands (denoted by superscripts "H" and "M", respectively) are given by:

$$\begin{aligned} C^H\left(\mathbf{1}, w^*, U_0\right) &= \frac{\partial E\left(\mathbf{1}, w^*, U_0\right)}{\partial \mathbf{1}}, \qquad H^H\left(\mathbf{1}, w^*, U_0\right) = \frac{\partial E\left(\mathbf{1}, w^*, U_0\right)}{\partial w^*}, \\ C^M\left(\mathbf{1}, w^*, Y_0\right) &= -\frac{\frac{\partial V(\mathbf{1}, w^*, Y_0)}{\partial \mathbf{1}}}{\frac{\partial V(\mathbf{1}, w^*, Y_0)}{\partial Y_0}}, \qquad H^M\left(\mathbf{1}, w^*, Y_0\right) = -\frac{\frac{\partial V(\mathbf{1}, w^*, Y_0)}{\partial w^*}}{\frac{\partial V(\mathbf{1}, w^*, Y_0)}{\partial Y_0}}. \end{aligned}$$

From here on in we focus on the demand for leisure, leaving the demand for consumption goods as an exercise.

Differentiating the Marshallian demand for leisure with respect to w^* we get:

$$\begin{aligned} \frac{\partial H^M}{\partial w^*} &= \left(\frac{\partial H^M}{\partial w^*}\right)_{Y_0 \text{ constant}} + \frac{\partial H^M}{\partial Y_0}\frac{\partial Y_0}{\partial w^*} \\ &= \left(\frac{\partial H^M}{\partial w^*}\right)_{Y_0 \text{ constant}} + \bar{L}\frac{\partial H^M}{\partial Y_0}, \end{aligned} \tag{2.15}$$

where we have used the definition of Y_0 in the second step (i.e. $\partial Y_0 / \partial w^* = \bar{L}$). Next, we note that by definition we can derive the Hicksian demand by substituting the expenditure function into the Marshallian demand (because $Y_0 = E\left(\mathbf{1}, w^*, U_0\right)$):

$$H^H\left(\mathbf{1}, w^*, U_0\right) = H^M\left(\mathbf{1}, w^*, E\left(\mathbf{1}, w^*, U_0\right)\right).$$

By differentiating this expression with respect to w^* we obtain an expression for the slope of the Hicksian demand for leisure:

$$\frac{\partial H^H}{\partial w^*} = \left(\frac{\partial H^M}{\partial w^*}\right)_{Y_0 \text{ constant}} + \frac{\partial H^M}{\partial Y_0}\frac{\partial E\left(\mathbf{1}, w^*, U_0\right)}{\partial w^*}$$

$$= \left(\frac{\partial H^M}{\partial w^*}\right)_{Y_0 \text{ constant}} + H^M \frac{\partial H^M}{\partial Y_0}, \tag{2.16}$$

where we have used Shephard's Lemma in the second step.

By combining (2.15) and (2.16), and noting that $H^H = H^M = \bar{L} - L^M$, we obtain the Slutsky equation for leisure demand:

$$\frac{\partial H^M}{\partial w^*} = \frac{\partial H^H}{\partial w^*} + L^M \frac{\partial H^M}{\partial Y_0}. \tag{2.17}$$

Of course, (2.17) can also be written in terms of labour supply (as in (2.12) above) by noting that $\partial H^M/\partial w^* = -\partial L^M/\partial w^*$, $\partial H^H/\partial w^* = -\partial L^H/\partial w^*$, and $\partial H^M/\partial Y_0 = -\partial L^M/\partial Y_0$.

Intermezzo 2.1

The expenditure function. This intermezzo introduces some very useful tools derived from duality theory, which will be used time and again. A good and accessible source for this material is Varian (1992) and Diamond and McFadden (1974). See also the Mathematical Appendix for further details.

We focus on the two-good case for expositional purposes. In particular, X_1 and X_2 are the two goods, P_1 and P_2 are their respective prices, and Y_0 is lump-sum income. We define the *expenditure function*, $E(\cdot)$, as the minimum level of lump-sum income the household needs to spend in order to attain a given level of utility, U_0, when faced with consumer prices P_1 and P_2. In formal terms we have:

$$E(P_1, P_2, U_0) \equiv \min_{\{X_1, X_2\}} P_1 X_1 + P_2 X_2 \quad \text{subject to:} \quad U(X_1, X_2) = U_0.$$

We next define the *indirect utility function*, $V(\cdot)$, as the maximum achievable utility level, given prices P_1 and P_2 and lump-sum income Y_0. Formally, we have:

$$V(P_1, P_2, Y_0) \equiv \max_{\{X_1, X_2\}} U(X_1, X_2) \quad \text{subject to:} \quad Y_0 = P_1 X_1 + P_2 X_2.$$

The following key properties can be derived for these functions.

First, under *local non-satiation* indirect utility, $V(\cdot)$, is strictly increasing in lump-sum income, Y_0, and we can find $E(\cdot)$ by inverting $V(\cdot)$ or vice versa. This property indicates that the two functions are intimately linked to each other.

Second, the expenditure function, $E(P_1, P_2, U_0)$, is homogeneous of degree one in prices. Intuitively, if all prices double then the household's income needs to double also in order to attain the same utility level. Since relative prices do not change in this experiment, the optimal consumption point is unchanged.

Third, the expenditure function, $E(P_1, P_2, U_0)$, is concave in prices.

Fourth, the expenditure function, $E(P_1, P_2, U_0)$, is strictly increasing in U_0 and non-decreasing in prices.

The fifth, very convenient property of the expenditure function concerns the derivation of the *compensated (Hicksian) demand curves*. Indeed, using the superscript 'H' for 'Hicksian' we find that the Hicksian demand for good i is simply the derivative of the expenditure function with respect to the price of good i:

$$X_i^H(P_1, P_2, U_0) = \frac{\partial E(P_1, P_2, U_0)}{\partial P_i}. \tag{I.1}$$

This relationship is *Shephard's Lemma* (also often referred to as the *derivative property*).

The sixth property deals with the derivation of the *uncompensated (Marshallian) demand curves*. They are obtained from the indirect utility function by applying *Roy's Identity*:

$$X_i^M(P_1, P_2, Y_0) = -\frac{\frac{\partial V(P_1, P_2, Y_0)}{\partial P_i}}{\frac{\partial V(P_1, P_2, Y_0)}{\partial Y_0}}, \tag{I.2}$$

where the superscript 'M' stands for 'Marshallian'.

The seventh property deals with the relationship between Hicksian and Marshallian demands. Indeed, it is identically true that substitution of the expenditure function into the Hicksian demand yields the Marshallian demand, i.e.:

$$X_i^M\big(P_1, P_2, \underbrace{E(P_1, P_2, U_0)}_{=Y_0}\big) \equiv X_i^H(P_1, P_2, U_0). \tag{I.3}$$

This expression is quite useful because it allows a short-cut derivation of the Slutsky equation. In the first step, we differentiate (I.3) with respect to P_j and obtain:

$$\frac{\partial X_i^M}{\partial P_j} + \frac{\partial X_i^M}{\partial Y_0}\frac{\partial E(P_1, P_2, U_0)}{\partial P_j} = \frac{\partial X_i^H}{\partial P_j}. \tag{I.4}$$

In the second step we use (I.1) and (I.2) in (I.4) to get:

$$\frac{\partial X_i^M}{\partial P_j} = \frac{\partial X_i^H}{\partial P_j} - X_j^M \frac{\partial X_i^M}{\partial Y_0}. \tag{I.5}$$

Note that (I.5) contains all Slutsky terms, i.e. not only "own" effects (like $\partial X_i^M/\partial P_i$) but also "cross" effects (such as $\partial X_i^M/\partial P_j$ for $i \neq j$).

****

2.1.3 Non-linear taxes

Up to this point we have limited attention to the graphical analysis of linear taxes. We now extend the basic model in the following directions. First, we recognize that most (labour) income tax systems are *progressive*, in the sense that the average tax rate rises with the tax base. Second, we move beyond the qualitative-graphical analysis by pursuing a quantitative-mathematical analysis of tax effects. This approach will be used extensively throughout the book, for example in Chapters 6-7 on the theory of tax incidence.

We augment the basic model by replacing the linear tax function (2.4) by the following general tax function $T(WL)$. We thus continue to assume that non-labour income is untaxed. We define the *marginal tax rate* (t_M) as the derivative of the tax function with respect to the tax base, i.e. $t_M \equiv dT(WL)/d(WL)$, and the *average tax rate* (t_A) is defined as the tax bill divided by the tax base, i.e. $t_A \equiv T(WL)/(WL)$.[6] Of course, both t_M and t_A are not constants, but rather are *functions* of the tax base, WL.

The rest of the model unchanged, i.e. the utility function is still given by (2.1) and the household's budget equation is given by:

$$\begin{aligned} P(1+t_C)C &= M + WL - T(WL) \\ &\equiv M + (1-t_A)WL, \end{aligned} \tag{2.18}$$

where we have used the definition of the average tax rate in the final expression. The household chooses consumption and leisure in order to maximize utility (2.1) subject to the budget constraint (2.18) and recognizes the progressivity of the tax system. The Lagrangian expression is now given by:

$$\mathcal{L} \equiv U(C, \bar{L} - L) + \lambda \left[M + (1-t_A)WL - P(1+t_C)C \right], \tag{2.19}$$

so that the first-order necessary conditions are (2.18) and:

$$\frac{\partial \mathcal{L}}{\partial C} = U_C - \lambda P(1+t_C) = 0, \tag{2.20}$$

$$\frac{\partial \mathcal{L}}{\partial L} = -U_{\bar{L}-L} + \lambda W \left[(1-t_A) - L\frac{dt_A}{dL} \right] = 0. \tag{2.21}$$

Equation (2.20) is the same as before (see (2.6) above) but the first-order condition for labour supply is more complex (compare (2.21) with (2.7) above). In making the marginal decision about hours, the

[6] Note that the linear tax schedule (2.4) also represents a progressive tax system provided T_0 is negative. The non-linear tax schedule used in this subsection is more general because it can also incorporate the notion of *rate progressivity*, according to which both the average and the marginal tax rate increase with the tax base.

household not only takes into account that part of the additional wage income will be taxed away but it also recognizes that the expansion in the tax base will increase the average tax rate due to the progressivity of the tax system.

By differentiating $t_A \equiv T(WL) / (WL)$ with respect to L we can simplify the second term in square brackets on the right-hand side of (2.21) to:

$$\begin{aligned} L\frac{dt_A}{dL} &= L\frac{(WL)\frac{dT(WL)}{dL} - T(WL)\frac{dWL}{dL}}{(WL)^2} \\ &= L\frac{(WL)\frac{dT(WL)}{d(WL)}\frac{dWL}{dL} - T(WL)\frac{dWL}{dL}}{(WL)^2} = t_M - t_A. \end{aligned} \tag{2.22}$$

It thus follows from (2.20)-(2.22) that the income expansion path can be written as:

$$\begin{aligned} \lambda = \frac{U_C}{P(1+t_C)} &= \frac{U_{\bar{L}-L}}{W(1-t_M)} \qquad \Rightarrow \\ \frac{U_{\bar{L}-L}}{U_C} &= w\frac{1-t_M}{1+t_C}, \end{aligned} \tag{2.23}$$

where $w \equiv W/P$ is the gross real wage rate. Just as before, we reach the intuitive conclusion that the marginal rate of substitution between leisure and consumption depends on the *marginal* (and not on the *average*) tax rate facing households!

We continue to assume that the utility function is homothetic and define the *substitution elasticity* between consumption and leisure as follows:

$$\sigma = \frac{\text{percentage change in } C/(\bar{L}-L)}{\text{percentage change in } U_{\bar{L}-L}/U_C} \equiv \frac{d\ln(C/(\bar{L}-L))}{d\ln(U_{\bar{L}-L}/U_C)} \geq 0, \tag{2.24}$$

where σ measures how "easy" it is (in utility terms) for the household to substitute consumption for leisure (or vice versa). Intuitively, if σ is low, substitution is rather difficult and there are sharp kinks in the household's indifference curves. Large changes in the relative price (w^*) are needed to produce a given change in the consumption-leisure mix. Conversely, if σ is high, substitution is very easy and indifference curves are fairly flat. Small changes in the relative price suffice to produce a given change in the consumption-leisure mix.

Our quantitative analysis proceeds by loglinearizing the model around a given initial point and investigating changes in the variables that occur as a result of infinitesimal changes in the tax parameters. Since this technique is used throughout the book, we show the details of the derivation here.[7] First, linearization of (2.23) is straightforward and results in:

$$d\ln\left(\frac{U_{\bar{L}-L}}{U_C}\right) = \tilde{w} - \tilde{t}_M - \tilde{t}_C, \tag{2.25}$$

[7] See also the Mathematical Appendix for further details.

where $\tilde{w} \equiv dw/w$, $\tilde{t}_M \equiv dt_M/(1-t_M)$, and $\tilde{t}_C \equiv dt_C/(1+t_C)$ are proportional changes in, respectively w, t_M, and t_C.[8] The left-hand side of (2.25) represents the proportional change in the marginal rate of substitution between leisure and consumption (recall that $d\ln x = dx/x = \tilde{x}$). By using the definition for σ (given in (2.24) above) we find that this term can be rewritten as:

$$\begin{aligned}\sigma d\ln\left(\frac{U_{\bar{L}-L}}{U_C}\right) &= \tilde{C} - (\widetilde{\bar{L}-L}) \\ &= \tilde{C} + (1/\omega_L)\tilde{L},\end{aligned} \tag{2.26}$$

where $\tilde{C} \equiv dC/C$, $\tilde{L} \equiv dL/L$, and $\omega_L \equiv (\bar{L}-L)/L$ is the ratio of leisure consumption to labour supply. By combining (2.25) and (2.26) we find that the loglinearized version of (2.23) is given by:

$$\tilde{C} + (1/\omega_L)\tilde{L} = \sigma\left[\tilde{w} - \tilde{t}_M - \tilde{t}_C\right]. \tag{2.27}$$

Linearizing the household's budget constraint (2.18) is straightforward and results in:

$$\tilde{C} + \tilde{t}_C = \omega_M\tilde{m} + (1-\omega_M)\left[\tilde{w} + \tilde{L} - \tilde{t}_A\right], \tag{2.28}$$

where $m \equiv M/P$ is real non-labour income, $\omega_M \equiv m/(m+(1-t_A)wL)$ is the initial share of non-labour income in total income, and $\tilde{t}_A \equiv dt_A/(1-t_A)$. By definition, we have that $0 \leq \omega_M \leq 1$. Furthermore, $\omega_M \approx 0$ for poor households who are heavily reliant on labour income, and $\omega_M \approx 1$ for rich households, who rely mostly on non-labour income. Note finally that it is the *average* (and not the marginal) labour income tax rate which directly influences the budget restriction of the household.

We now have two expressions relating $\tilde{C}$ and $\tilde{L}$ to the variables that are exogenous to the household, i.e. $\tilde{w}$, $\tilde{m}$, $\tilde{t}_M$, $\tilde{t}_A$, and $\tilde{t}_C$. By combining (2.27) and (2.28) into one matrix equation we find:

$$\begin{bmatrix} 1/\omega_L & 1 \\ -(1-\omega_M) & 1 \end{bmatrix}\begin{bmatrix} \tilde{L} \\ \tilde{C} \end{bmatrix} = \begin{bmatrix} \sigma\left[\tilde{w} - \tilde{t}_M - \tilde{t}_C\right] \\ \omega_M\tilde{m} + (1-\omega_M)\left[\tilde{w} - \tilde{t}_A\right] - \tilde{t}_C \end{bmatrix}. \tag{2.29}$$

By inverting the matrix on the left-hand side of (2.29) we obtain the solution:

$$\begin{aligned}\begin{bmatrix} \tilde{L} \\ \tilde{C} \end{bmatrix} &= \frac{\omega_L}{1+\omega_L(1-\omega_M)}\begin{bmatrix} 1 & -1 \\ 1-\omega_M & 1/\omega_L \end{bmatrix} \\ &\quad\times\begin{bmatrix} \sigma\left[\tilde{w} - \tilde{t}_M - \tilde{t}_C\right] \\ \omega_M\tilde{m} + (1-\omega_M)\left[\tilde{w} - \tilde{t}_A\right] - \tilde{t}_C \end{bmatrix}.\end{aligned} \tag{2.30}$$

Equation (2.30) presents *all* the comparative static results allowed for by the model. Focusing on the

[8] Note, however, the slightly different treatment of tax variables. For example, instead of working with dt_C/t_C we prefer to express the model in terms of $dt_C/(1+t_C)$ which is well-defined even if the initial tax rate is zero.

expression for labour supply (the first row of (2.30)), for example, we find:

$$\tilde{L} = \varepsilon_w^H \left[\tilde{w} - \tilde{t}_M - \tilde{t}_C\right] - \varepsilon_I \left[\omega_M \tilde{m} + (1-\omega_M)\left[\tilde{w} - \tilde{t}_A\right] - \tilde{t}_C\right], \tag{2.31}$$

where ε_w^H and $-\varepsilon_I$ are, respectively, the *compensated* (Hicksian) wage elasticity and the income elasticity of labour supply:

$$\varepsilon_w^H \equiv \frac{\sigma\omega_L}{1+\omega_L(1-\omega_M)} > 0, \tag{2.32}$$

$$\varepsilon_I \equiv \frac{\omega_L}{1+\omega_L(1-\omega_M)} > 0. \tag{2.33}$$

Equation (2.31) is quite useful because it allows us to disentangle the income and substitution effects associated with changes in the different variables. First, as we illustrated graphically above (in Figure 2.3), a change in the real wage rate has both an income and a substitution effect. Indeed, the *uncompensated* (Marshallian) wage elasticity is defined as $\varepsilon_w^M \equiv \varepsilon_w^H - \varepsilon_I(1-\omega_M)$, which has an ambiguous sign. Of course, for rich households $\omega_M \approx 1$, the income effect is negligible, $\varepsilon_w^M \approx \varepsilon_w^H$, and labour supply is likely to be upward sloping (see also Figure 2.4 above). Second, we observe from (2.31) that a change in the marginal labour income tax rate, ceteris paribus, isolates the pure substitution effect, i.e. an increase in t_M produces a decrease in labour supply ($\tilde{L} = -\varepsilon_w^H \tilde{t}_M < 0$). Third, and in stark contrast to the previous result, a change in the average labour income tax rate isolates the income effect, i.e. an increase in t_A produces an increase in labour supply ($\tilde{L} = \varepsilon_I(1-\omega_M)\tilde{t}_A > 0$). Intuitively, the reduction in income makes the household poorer and, since leisure is a normal good, labour supply is increased. Fourth, leisure is a normal good, which also explains why an increase in non-labour income results in a decrease in labour supply ($\tilde{L} = -\varepsilon_I \omega_M \tilde{m} < 0$). Finally, an increase in the consumption tax has both a negative substitution effect and a positive income effect, so that the net effect depends on sign of $\sigma - 1$, i.e. $\tilde{L} = \left(\varepsilon_I - \varepsilon_w^H\right)\tilde{t}_C \lesseqgtr 0$).

In Figure 2.5 we illustrate the effect of an increase in the marginal tax rate t_M *holding constant* the average tax rate t_A. For convenience we assume in the construction of the diagram that (i) there is no non-labour income ($m = 0$), (ii) the tax on consumption is zero ($t_C = 0$), and (iii) there is a linear progressive tax schedule:

$$\frac{T}{P} = t_M wL - z_0, \tag{2.34}$$

where $z_0 > 0$ is the real lump-sum transfer and t_M is a constant tax rate. The budget line (2.3) then simplifies to:

$$C = z_0 + (1-t_M)\,w\bar{L} - (1-t_M)\,w\left(\bar{L} - L\right). \tag{2.35}$$

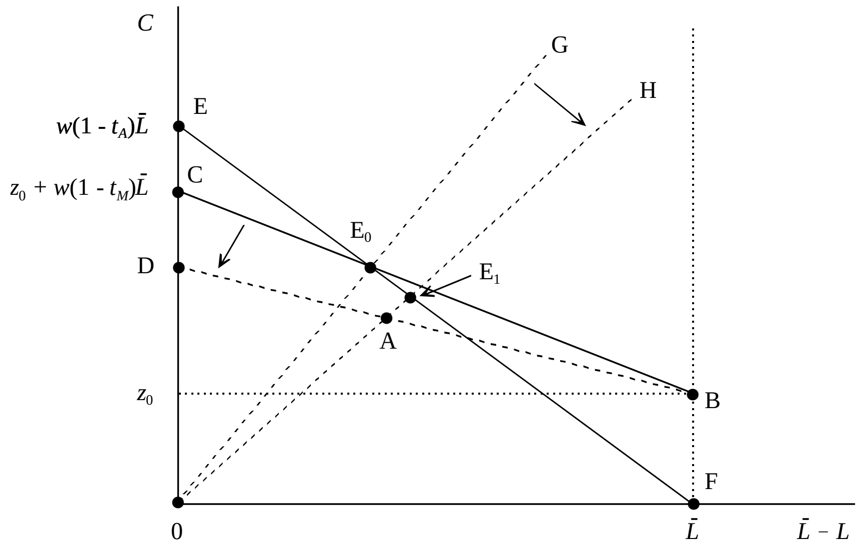

Figure 2.5: Increasing the marginal tax rate (constant average tax rate)

Equation (2.35) is drawn as the line BC in Figure 2.5. Alternatively, using the definition of the average tax rate, the budget line can be written as:

$$
\begin{aligned}
C &= (1 - t_A)\, wL \\
&= (1 - t_A)\, w\bar{L} - (1 - t_A)\, w\, (\bar{L} - L)\,, \qquad (2.36)
\end{aligned}
$$

which is the line EF in Figure 2.5. Anywhere along the line EF the average tax rate is constant. Of course, with a progressive tax system, EF is steeper than BC because $t_M > t_A$.

In the tax experiment, t_M rises but t_A is held constant. The initial equilibrium is at point E_0, where there is a tangency of an indifference curve (which is not drawn) and the line BC. As t_M rises, the budget line BC rotates in a counter-clockwise direction to BD. Since t_A is unchanged, the line EF remains in the same position. In the absence of compensation measures, the new equilibrium would be at point A, where there is a tangency between an indifference curve (not drawn) and the new budget line BD. But at point A, the average tax rate would be too high. In order to keep t_A unchanged z_0 must rise (causing a parallel shift in BD) so that the new equilibrium is at E_1. It follows from the figure that C falls but $\bar{L} - L$ rises (i.e. labour supply falls).

2.1.4 Non-convex choice set

We have demonstrated in the previous subsection that a progressive tax system can be easily handled within the context of the basic household decision model. The reason for this convenient result is that the choice set remains convex so that the household optimum is unique. Many features of actual tax systems, however, may make choice sets (for some households) non-convex.[9] In the non-convex case, there may be multiple tangencies, standard comparative static effects (e.g. Slutsky decomposition) are no longer valid, and econometric testing is much more complicated.

Consider the following example of a *means-tested transfer program*. We assume that any transfers the household receives from the government are means-tested, i.e. they depend on the recipient's income. For simplicity, we assume the following transfer scheme (in real terms):

$$z = \begin{cases} z_0 + t_Z w \left(L_{\text{MIN}} - L \right) & \text{for } 0 < L \leq L_{\text{MIN}} \\ & \\ z_0 & \text{for } L > L_{\text{MIN}} \end{cases}, \tag{2.37}$$

where L_{MIN} is the policy-determined critical number of hours (yielding a subsistence level of income) and t_Z (> 0) is the means-testing parameter. If the household works fewer hours than this critical number ($L \leq L_{\text{MIN}}$) and thus has a below subsistence level of income, it receives additional transfers from the government. The more the household works, however, the lower are the transfers it receives. The means-testing parameter thus operates as an effective tax!

We assume that the tax system is given by $T/P = t_M wL$ so that the household budget constraint (in real terms and setting $t_C = 0$) is given by:

$$C = m + z + wL - t_M wL. \tag{2.38}$$

Substituting (2.37) into (2.38) we find that the budget constraint features two straight segments:

$$C = \begin{cases} \left[m + z_0 + t_Z w L_{\text{MIN}} \right] + \left(1 - t_M - t_Z \right) wL & \text{for } 0 < L \leq L_{\text{MIN}} \\ & \\ \left[m + z_0 \right] + \left(1 - t_M \right) wL & \text{for } L > L_{\text{MIN}} \end{cases}. \tag{2.39}$$

In terms of Figure 2.6, the budget line features a kink at point A where leisure is equal to $\bar{L} - L_{\text{MIN}}$. For low levels of labour supply, the fact that transfers are means-tested implies that the effective tax rate on labour income is higher than for high levels of labour supply, i.e. BA is flatter than CA. As a result, the household choice set is non-convex.

In the particular case drawn in Figure 2.6, there are two tangencies between the kinked budget line and the indifference curve, namely at points E_0 and E_1. Standard comparative statics methods are invalid in this case because infinitesimal changes in tax rates may produce large changes in the endogenous variables in the optimum. Take, for example the effects of an infinitesimal reduction in t_Z. In Figure 2.6 this results in a clockwise rotation of the lower branch of the budget line from AB to AB$'$. If the household was initially at E_1, it would now clearly prefer E_0, i.e. a marginal change has produced

[9] A set S is convex if the straight line connecting any two points in S also lies in the set S. In Figure 2.6 this is clearly not the case: E_0 and E_1 are both in the set, but *none* of the points on the line connecting E_0 and E_1 are in the set.

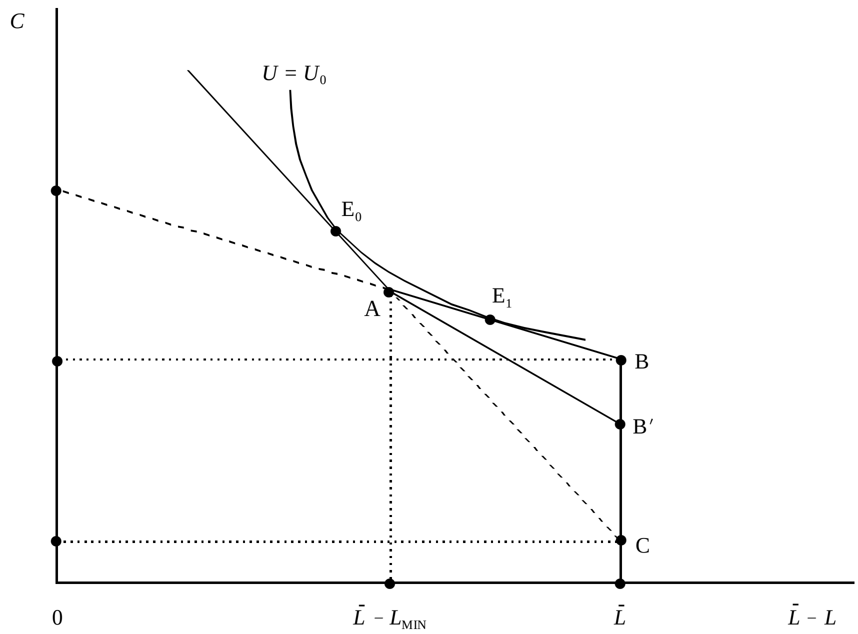

Figure 2.6: Means-tested transfer system

an inframarginal response.

Matters are even more complex if the tax system features increasing marginal tax rates (here shown as two different marginal tax rates rather than one single one). Take, for example, the tax schedule:

$$\frac{T}{P} = \begin{cases} t_M^1 wL & \text{for } 0 < L \leq L_1 \\ & \\ t_M^1 wL_1 + t_M^2 w\,(L - L_1) & \text{for } L > L_1 \end{cases},$$

with $t_M^2 > t_M^1$ and $L_1 > L_{\text{MIN}}$). Figure 2.7 shows that the budget line features two kinks in that case, i.e. one at point A because of the benefit system and the other at point E_0 because of the tax progression. As a result, there may be multiple local optima. In Figure 2.7 we have drawn one tangency optimum (at E_1) and one corner solution (at E_0). In order to determine the global optimum we must thus check all the local optima in order to determine which one features the highest utility level. In order to do this, we must know the form of the individual's utility function. In the figure the corner solution turns out to be better than the interior solution ($U_0 > U_1$).

2.2 Labour force participation

Up to now, we have assumed that the choice of hours can be made freely, i.e. L can take on any value between 0 and 1 (we normalize the time endowment $\bar{L}$ to unity). This is, of course, rather unrealistic since labour hours are typically indivisible. There are not too many bosses who will want to employ you for 6 hours, 24 minutes, and three seconds if that happens to be your optimum labour supply! The

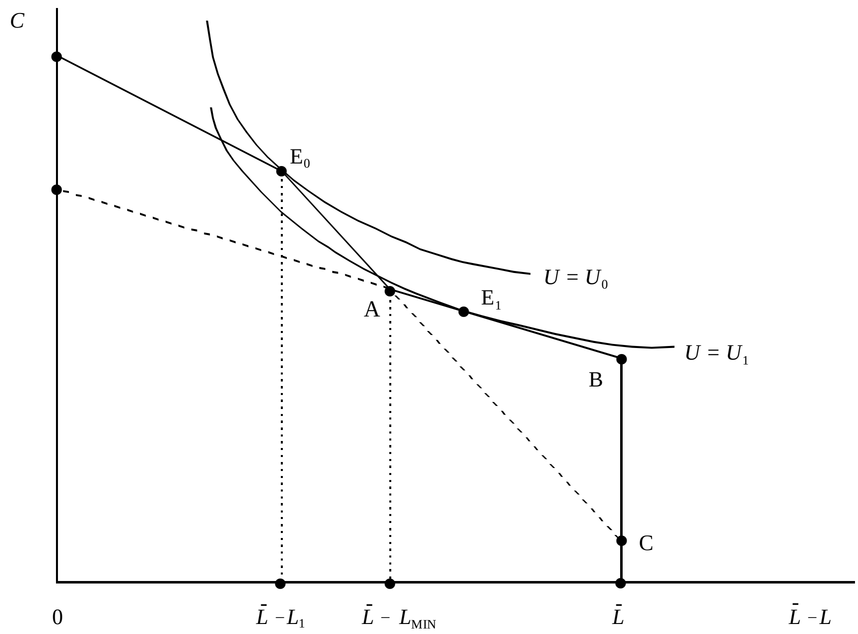

Figure 2.7: Means-tested transfers and marginal tax rate progression

typical option confronting a household is either to work full time ($L = L_F$) or not to work at all ($L = 0$).[10] The objective of this section is to study individual and aggregate labour supply in such a setting. The key ingredient of the model is the notion that different people may have different attitudes toward work.

2.2.1 A simple discrete-choice model

Assume that the utility function of household i takes the following Cobb-Douglas form:

$$U^i(C, 1 - L) \equiv C^\alpha [1 - L]^{\beta_i}, \tag{2.40}$$

where $\alpha > 0$ and $\beta_i \geq 0$. All households have the same α but there exists a frequency distribution for β_i across the population (see below). The budget constraint of a working household is given by:

$$C = wL_F(1 - t_L), \tag{2.41}$$

where $w \equiv W/P$ is the real wage rate, t_L is the labour income tax rate, and we abstract from non-labour income. In contrast, the budget constraint of a non-working household is:

$$C = b, \tag{2.42}$$

where b is the real unemployment transfer received from the government (assumed to be untaxed).

[10]Obviously, the possibility of part-time work can be easily introduced ($L = L_P$ where $0 < L_P < L_F$). The key notion in this section is that L can only take on a finite number of values.

There is a rudimentary unemployment benefit system according to which benefits are linked to after-tax real wage income:

$$b = \gamma w L_F (1 - t_L), \qquad 0 < \gamma < 1, \tag{2.43}$$

where γ is the *replacement rate*.

In this context, the labour supply decision is really a participation decision of the "yes-no" type; either the households works L_F hours or it does not work any hours. For household i, utility while working ($L = L_F$ and $C = wL_F(1 - t_L)$) is given by:

$$U^i_{\text{worker}} \equiv (wL_F(1 - t_L))^{\alpha} (1 - L_F)^{\beta_i}, \tag{2.44}$$

whereas utility when unemployed ($L = 0$ and $C = b$) is:

$$U^i_{\text{unemployed}} \equiv b^{\alpha}. \tag{2.45}$$

It follows that the relative utility of labour force participation for this household is:

$$\frac{U^i_{\text{worker}}}{U^i_{\text{unemployed}}} = \frac{(wL_F(1 - t_L))^{\alpha} (1 - L_F)^{\beta_i}}{(\gamma w L_F (1 - t_L))^{\alpha}} = \frac{(1 - L_F)^{\beta_i}}{\gamma^{\alpha}}. \tag{2.46}$$

The optimal labour supply choice is thus given by:

$$L_i = \begin{cases} 0 & \text{if } \gamma^{-\alpha}(1 - L_F)^{\beta_i} < 1 \\ \\ L_F & \text{if } \gamma^{-\alpha}(1 - L_F)^{\beta_i} > 1 \end{cases}. \tag{2.47}$$

According to this expression, a *workaholic* household—which does not value leisure at all ($\beta_i = 0$)—will certainly work full time (recall that for $0 < \gamma < 1$, it follows that $\gamma^{-\alpha} > 1$). At the other extreme, households with a high valuation of leisure ($\beta_i \gg 0$) will choose not to work at all.

The *marginal household* is indifferent between working and not working, i.e. it has a $\beta_i = \beta_M$ such that $\gamma^{-\alpha}(1 - L_F)^{\beta_M} = 1$. By taking logarithms on both sides of this expression we can solve for β_M:

$$-\alpha \ln \gamma + \beta_M \ln(1 - L_F) = 0 \qquad \Leftrightarrow \qquad \beta_M = \frac{\alpha \ln \gamma}{\ln(1 - L_F)} > 0, \tag{2.48}$$

where the sign follows from the fact that $0 < \gamma < 1$ and $0 < L_F < 1$ (so that $\ln \gamma < 0$ and $\ln(1 - L_F) < 0$). According to (2.47)-(2.48), all households whose β_i exceeds β_M prefer not to work (they like leisure "too much") whereas households with a β_i smaller than β_M choose to work.

We assume that the β_i's are distributed uniformly over the interval $[0, \beta_{\max}]$. The frequency distribution is drawn in Figure 2.8. In addition, we assume that the population size is N. Given that all households with a $\beta_i \leq \beta_M$ are workers and all households with a $\beta_i > \beta_M$ are "loungers", we easily find that there are $\beta_M N/\beta_{\max}$ workers (who each work L_F hours) and $(\beta_{\max} - \beta_M)N/\beta_{\max}$ non-participants. See Figure 2.8. Aggregate labour supply (L^S) is thus given by:

$$L^S = \frac{\beta_M N L_F}{\beta_{\max}}, \tag{2.49}$$

where β_M is defined in (2.48) above.[11] The macroeconomic labour supply curve is drawn in Figure 2.9. The key thing to note is that the aggregate labour supply curve is vertical because β_M does not depend on the wage rate or the labour income tax, reflecting that unemployment benefits are linked to *after-tax* wage income. Of course, the replacement rate exerts a negative influence on aggregate labour supply in this model:

$$\frac{\partial \beta_M}{\partial \gamma} = \frac{\alpha}{\gamma \ln(1 - L_F)} < 0, \qquad \frac{\partial L^S}{\partial \gamma} = \frac{N L_F}{\beta_{\max}} \frac{\partial \beta_M}{\partial \gamma} < 0, \tag{2.50}$$

where the signs follow from the fact that $\ln(1 - L_F) < 0$. The reduction in β_M causes the aggregate labour supply curve to shift to the left, as is indicated in Figure 2.9. Some households, whose β_i was close to β_M in the initial situation, withdraw from the labour market if the benefit system becomes more generous.

2.2.2 A different benefit system

In the model developed in the previous subsection, the labour income tax did not affect aggregate labour supply. It is easy to demonstrate that the results are sensitive to the details of the benefit system. For example, instead of (2.43) we now assume that the unemployment benefit is linked to gross wage income:

$$b = \gamma w L_F, \tag{2.51}$$

where we assume that $\gamma < 1 - t_L < 1$ (otherwise nobody would choose to work!). Reworking the earlier steps, we find that utility when unemployed is equal to:

$$U^i_{\text{unemployed}} = b^\alpha = (\gamma w L_F)^\alpha = \left(\frac{\gamma}{1 - t_L}\right)^\alpha (w L_F (1 - t_L))^\alpha . \tag{2.52}$$

Then, the utility comparison amounts to:

$$\frac{U^i_{\text{worker}}}{U^i_{\text{unemployed}}} = \frac{(w L_F (1 - t_L))^\alpha (1 - L_F)^{\beta_i}}{\left(\frac{\gamma}{1 - t_L}\right)^\alpha (w L_F (1 - t_L))^\alpha} = \left(\frac{1 - t_L}{\gamma}\right)^\alpha (1 - L_F)^{\beta_i}, \tag{2.53}$$

[11] Note that L^S/N is the labour *participation rate* implied by this model, i.e. the proportion of the population that participates in the labour market.

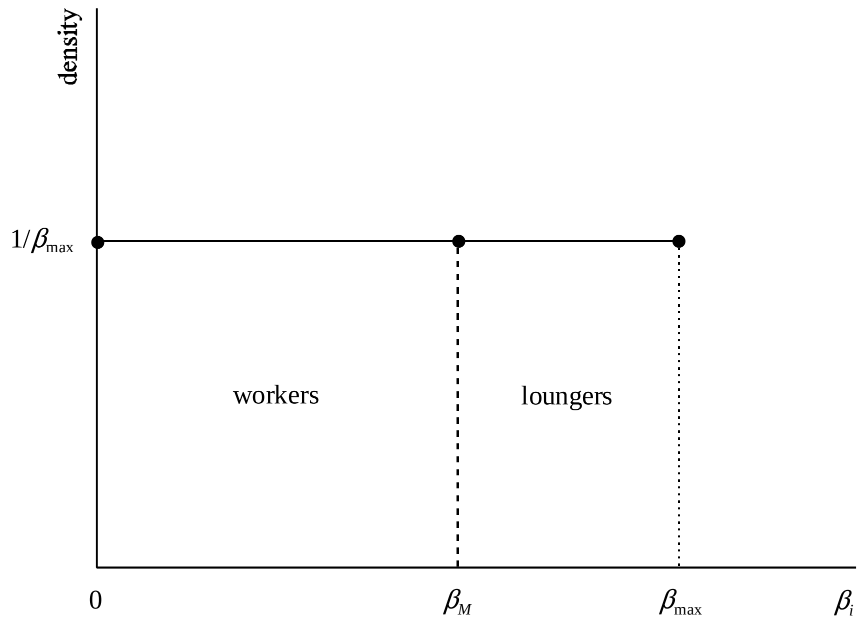

Figure 2.8: Frequency distribution of β_i coefficients

and the critical value—analogously to (2.48) above—is given by:

$$\alpha\left[\ln(1-t_L)-\ln\gamma\right]+\beta_M\ln(1-L_F)=0 \quad \Leftrightarrow$$
$$\beta_M=\frac{\alpha\left[\ln\gamma-\ln(1-t_L)\right]}{\ln(1-L_F)}>0, \tag{2.54}$$

where the sign follows from the fact that $\ln(1-L_F)<0$ (as before) and the assumption that $\gamma<1-t_L$.

In this modified model, an increase in the tax rate leads to an increase in $\gamma^*\equiv\gamma/(1-t_L)$ and thus to an increase in the *effective* replacement rate, γ^*. This implies that β_M falls so that aggregate labour supply falls.

2.3 Other theoretical approaches

In this section we discuss a number of alternative theoretical approaches regarding consumption and labour supply decisions. Because the relevant literature is rather large, the survey is both selective and incomplete. Although we de-emphasize taxation issues, it is safe to conclude that in all these alternative models the effects of taxes can be quite different from those obtained with the basic model.

2.3.1 Household production

In the basic labour supply model used so far, leisure is assumed to yield direct utility to the household, i.e. $\bar{L}-L$ enters as an argument into the utility function (2.1). Though analytically convenient, this assumption has not remained unchallenged in the literature. As was argued forcefully by Becker (1965),

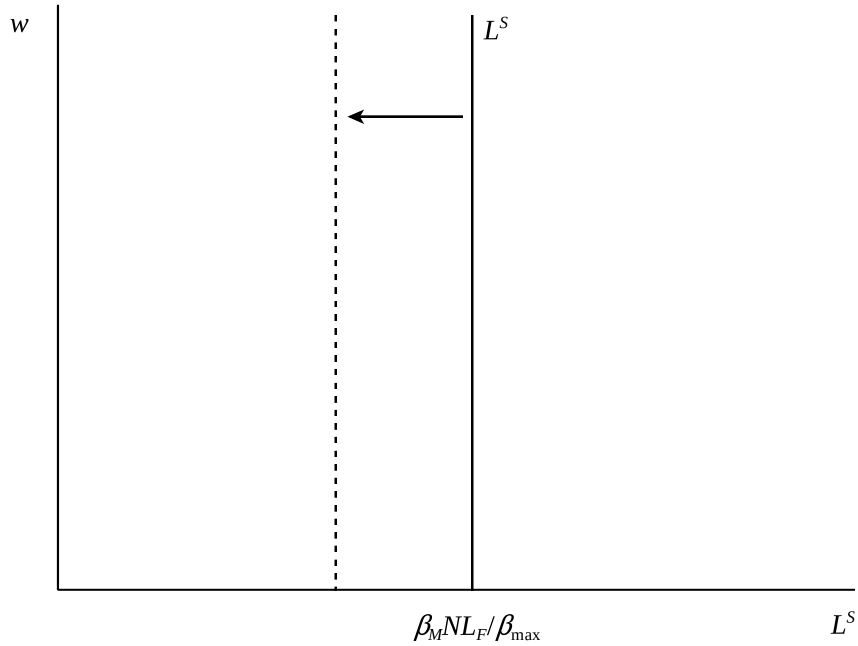

Figure 2.9: Aggregate labour supply

many consumption activities require not only goods themselves but also the input of (valuable) time. For example, if the utility-yielding activity is the consumption of beer, then one must not only purchase the bottle of beer itself but one must also spend some time drinking it. Similarly, to enjoy the activity of playing golf one needs not only golf clubs and balls but also a lot of time. Goods and time together are used to produce the activity which yields utility. But this begs the question what we mean by pure leisure as a utility-yielding activity. Indeed, as is pointed out by Atkinson and Stern, "Pure 'leisure' would require time only and no other inputs, but, apart from sunbathing naked, it is hard to think of an activity which requires no complementary inputs" (1981, p. 268). In this subsection we show a simple example of how the time-consuming nature of consumption can be modelled.[12]

The utility function of the representative household is given by:

$$U = U\left(C_1, C_2\right), \tag{2.55}$$

where C_1 and C_2 are both consumption *activities*, and the utility function possesses the usual properties, i.e. it features positive but diminishing marginal utility for both activities and is strictly quasi-concave in its arguments. Partial derivatives of the utility function are denoted by $U_i \equiv \partial U/\partial C_i > 0$, $U_{ii} = \partial^2 U/\partial C_i^2 < 0$ (for $i = 1, 2$), and $U_{12} \equiv \partial^2 U/\partial C_1 \partial C_2 \gtreqless 0$. The key thing to note is that pure leisure does not exist and that labour does not enter the utility function directly.

Following Becker (1965) and Kleven (2004), it is assumed that the consumption activities are "pro-

[12]The model is a simplified version of the one formulated by Atkinson and Stern (1979, 1981). It is also close in spirit to Kleven (2004).

duced" by the household using market goods and household time as inputs:

$$C_i = \min\left[\frac{X_i}{\alpha_i}, \frac{L_i}{\beta_i}\right], \qquad (\text{for } i = 1,2), \tag{2.56}$$

where α_i and β_i are fixed input coefficients. In the production of activity C_i, the household uses two inputs, namely X_i units of the market good i and L_i units of labour. The household "production function" is of the Leontief type, i.e. there is no substitutability between the two inputs X_i and L_i. Conditional on the level of C_i, the household will thus choose its inputs according to:

$$X_i = \alpha_i C_i, \qquad L_i = \beta_i C_i, \qquad (\text{for } i = 1,2). \tag{2.57}$$

The household budget constraint is given by:

$$P_1 X_1 + P_2 X_2 = M - T_0 + (1 - t_L)\, WL, \tag{2.58}$$

where P_i is the market price of good X_i. Just as in the basic model, M is exogenous non-labour income, T_0 is the lump-sum part of the labour income tax, t_L is the marginal tax rate on labour ($0 < t_L < 1$), and L is labour supply. According to (2.58), total spending on market goods (left-hand side) must equal total after-tax income (right-hand side).

In addition to facing the *monetary* budget constraint (2.58), the household also faces a *time* budget constraint of the form:

$$L_1 + L_2 + L = \bar{L}, \tag{2.59}$$

where $\bar{L}$ is the exogenous time endowment. The available time is allocated to home production (L_1 and L_2) and to labour supply (L). As was pointed out by Becker (1965, p. 496), the constraints (2.58) and (2.59) are not independent and can be combined into a single constraint.[13] Indeed, by using (2.57) in (2.58)-(2.59) and eliminating market labour supply (L, which we assume to be non-zero), we obtain the following overall budget constraint:

$$P_1^* C_1 + P_2^* C_2 = M - T_0 + W^* \bar{L}, \tag{2.60}$$

where $W^* \equiv (1 - t_L)\, W$ is the after-tax wage rate, and P_i^* denotes the total price of activity C_i:

$$P_i^* \equiv \alpha_i P_i + \beta_i W^*. \tag{2.61}$$

The total price of an activity is thus equal to the weighted price of the two inputs needed to produce it,

[13] As is pointed out by Atkinson and Stern (1981, p. 270), this procedure is only valid if labour does not enter the utility function directly (as is assumed here).

with the input coefficients acting as weights.[14]

The household chooses activities C_1 and C_2 in order to maximize utility (2.55) subject to the overall budget constraint (2.60). As is stressed by Atkinson and Stern (1981, p. 269), this maximization program is formally identical to the standard program studied above. Apart from the fact that total prices (P_1^* and P_2^*) rather than market goods prices (P_1 and P_2) must be used, standard demand theory applies. The expenditure function is thus defined as:

$$E\left(P_1^*, P_2^*, U_0\right) \equiv \min_{\{C_1, C_2\}} P_1^* C_1 + P_2^* C_2 \quad \text{subject to: } U\left(C_1, C_2\right) = U_0, \tag{2.62}$$

and it follows from Shephard's Lemma that the compensated demand for C_i is equal to:

$$C_i^H \equiv \frac{\partial E\left(P_1^*, P_2^*, U_0\right)}{\partial P_i^*}, \qquad (\text{for } i = 1, 2), \tag{2.63}$$

where (once again) the superscript "H" stands for Hicksian. In view of (2.59) and the second expression in (2.57), the Hicksian labour supply can be written as:

$$L^H = \bar{L} - \beta_1 C_1^H - \beta_2 C_2^H. \tag{2.64}$$

Consider the effects of an increase in the marginal labour income tax rate, t_L, on the supply of labour. In view of (2.61) the resulting decrease in the after-tax wage leads to a decrease in the total price of all activities; more so the more labour-intensive the activity is ($\partial P_i^* / \partial t_L = -\beta_i W < 0$). The effect on Hicksian labour supply is obtained from (2.64):

$$\begin{aligned}
\frac{\partial L^H}{\partial t_L} &= -\beta_1 \left[\frac{\partial C_1^H}{\partial P_1^*} \frac{\partial P_1^*}{\partial t_L} + \frac{\partial C_1^H}{\partial P_2^*} \frac{\partial P_2^*}{\partial t_L} \right] - \beta_2 \left[\frac{\partial C_2^H}{\partial P_1^*} \frac{\partial P_1^*}{\partial t_L} + \frac{\partial C_2^H}{\partial P_2^*} \frac{\partial P_2^*}{\partial t_L} \right] \\
&= \beta_1 W \left[\beta_1 \frac{\partial C_1^H}{\partial P_1^*} + \beta_2 \frac{\partial C_1^H}{\partial P_2^*} \right] + \beta_2 W \left[\beta_1 \frac{\partial C_2^H}{\partial P_1^*} + \beta_2 \frac{\partial C_2^H}{\partial P_2^*} \right] \\
&= W \left[\begin{array}{cc} \beta_1 & \beta_2 \end{array} \right] \left[\begin{array}{cc} \frac{\partial C_1^H}{\partial P_1^*} & \frac{\partial C_1^H}{\partial P_2^*} \\ \frac{\partial C_2^H}{\partial P_1^*} & \frac{\partial C_2^H}{\partial P_2^*} \end{array} \right] \left[\begin{array}{c} \beta_1 \\ \beta_2 \end{array} \right].
\end{aligned} \tag{2.65}$$

The expression on the right-hand side of (2.65) is a so-called *quadratic form* of the type $\mathbf{x}'\mathbf{S}\mathbf{x}$, where $\mathbf{x}' \equiv \left[\begin{array}{cc} \beta_1 & \beta_2 \end{array} \right]$ is a row vector and $\mathbf{S}$ is a two-by-two matrix.[15] We would like to know the sign of this quadratic form. It turns out that this sign depends on certain properties of the matrix $\mathbf{S}$. The $\mathbf{S}$ matrix contains all the Hicksian substitution terms and is therefore often called the *Substitution Matrix*. It follows from (2.63) that it is equal to the *Hessian* of the expenditure function, i.e. the matrix of its

[14] The basic consumption-leisure model is a special case of the Becker model developed in this section. Indeed, by setting $\alpha_1 = 1$, $\beta_1 = 0$, $\alpha_2 = 0$, and $\beta_1 = 1$ we arrive at the model discussed in Section 2.1 above.

[15] See the Matematical Appendix for further details on the concepts and properties introduced here.

second-order derivatives:

$$\mathbf{S} \equiv \begin{bmatrix} \frac{\partial C_1^H}{\partial P_1^*} & \frac{\partial C_2^H}{\partial P_1^*} \\ \frac{\partial C_1^H}{\partial P_2^*} & \frac{\partial C_2^H}{\partial P_2^*} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 E}{\partial (P_1^*)^2} & \frac{\partial E}{\partial P_1^* \partial P_2^*} \\ \frac{\partial E}{\partial P_2^* \partial P_1^*} & \frac{\partial^2 E}{\partial (P_2^*)^2} \end{bmatrix}. \tag{2.66}$$

Hence, by Young's theorem it follows that $\mathbf{S}$ is a *symmetric* matrix, i.e. the off-diagonal elements are equal to each other ($S_{ij} = S_{ji}$). Furthermore, concavity of the expenditure function, $E\left(P_1^*, P_2^*, U_0\right)$, implies that $\mathbf{S}$ is *negative semidefinite*, i.e. $\mathbf{x}'\mathbf{S}\mathbf{x} \leq 0$ for all $\mathbf{x}$. Provided we allow for some substitutability between C_1 and C_2 (and thus rule out the Leontief case), $S_{ii} < 0$ and we reach the stronger result that $\mathbf{S}$ is *negative definite*, i.e. $\mathbf{x}'\mathbf{S}\mathbf{x} < 0$ for all $\mathbf{x} \neq 0$. Using this result in (2.65) we find that $\partial L^H / \partial t_L < 0$. Just as in the basic labour supply model, the compensated labour supply function in the home production model is upward sloping in the after-tax wage (and downward sloping in the tax rate).

It is clear from (2.65) that the tax increase leads to a reduction in (Hicksian) labour supply and thus an increase in the amount of labour employed at home, i.e. $\partial \left[L_1^H + L_2^H\right] / \partial t_L > 0$. But what happens to the components L_1^H and L_2^H or, equivalently, to C_1^H and C_2^H? As is demonstrated by Atkinson and Stern (1979), the answer depends critically on the relative labour intensity of the two activities. By differentiating (2.63) with respect to t_L we get:

$$\begin{aligned} \frac{\partial C_1^H}{\partial t_L} &= S_{11} \frac{\partial P_1^*}{\partial t_L} + S_{12} \frac{\partial P_2^*}{\partial t_L} \\ &= -W \left[\beta_1 S_{11} + \beta_2 S_{12}\right], \end{aligned} \tag{2.67}$$

where S_{ij} is the element in row i and column j of the substitution matrix $\mathbf{S}$. But $C_1^H\left(P_1^*, P_2^*, U_0\right)$ is homogeneous of degree zero in P_1^* and P_2^* so that $0 = S_{11}P_1^* + S_{12}P_2^*$ and (2.67) can be rewritten as:

$$\frac{\partial C_1^H}{\partial t_L} = -\frac{S_{11}P_1^*}{1 - t_L} \left[\frac{\beta_1 W^*}{P_1^*} - \frac{\beta_2 W^*}{P_2^*}\right]. \tag{2.68}$$

The term in front of the square brackets is positive (because $S_{11} < 0$) so the effect on the Hicksian demand for activity C_1 is determined by the relative labour intensity of the two activities. If C_1 is relatively labour intensive ($\beta_1 W^*/P_1^* > \beta_2 W^*/P_2^*$), then the term in square brackets is positive and $\partial C_1^H / \partial t_L > 0$. The decrease in the after-tax wage rate induces a shift toward the labour-intensive home production activity.[16]

[16]Using similar steps the effect on the Hicksian demand for C_2 is obtained:

$$\frac{\partial C_2^H}{\partial t_L} = \frac{S_{22}P_2^*}{1 - t_L} \left[\frac{\beta_1 W^*}{P_1^*} - \frac{\beta_2 W^*}{P_2^*}\right].$$

Hence, if C_1 is labour intensive then $\partial C_2^H / \partial t_L < 0$. This result is obvious because we are focusing on the pure substitution effect, i.e. along a given indifference curve.

2.3.2 Collective decision making

In the basic labour supply model of Section 2.1, the behaviour of a representative household is modelled. Implicitly, therefore, the household, rather than the individual, is seen as the basic spending unit in the economy. Of course, in reality, most households are made up out of more than one person. The methodology of neoclassical economics requires economic behaviour to be modelled at the level of *individual agents* rather than at the group level (such as a multi-person family). As is stressed by Chiappori (1992, p. 440), adherence to methodological individualism calls for an explicit description of the behaviour of all members of a household, i.e. the black box of the household must be opened up. In this subsection we discuss some of the recent literature on multi-person household behaviour. To keep matters as simple as possible we restrict attention to the case of two-person households.

2.3.2.1 Household welfare function

Almost half a century ago, Paul Samuelson (1956) suggested two possible justifications for treating the representative household as the basic unit of analysis. The first justification is that the tastes of a household's members are such that they can be aggregated into one household utility function. An advanced treatment of this case is found in Varian (1992, pp. 153-154), but here we focus on the simple case of *homothetic preferences*. Suppose that the family consists of two members whose direct utility functions can be written as:

$$U^i = U^i\left(C^i, \bar{L} - L^i\right), \qquad \text{(for } i = 1, 2\text{)}, \tag{2.69}$$

where U^i, C^i, $\bar{L}$ and L^i denote, respectively, utility, consumption, the time endowment, and labour supply of person i. Ignoring the tax system, the budget constraint faced by each person can be written as:

$$PC^i + W\left(\bar{L} - L^i\right) = Y^i, \tag{2.70}$$

where Y^i is (exogenous) full income of person i. For homothetic preferences, the indirect utility function can be written as:

$$V^i = v\left(P, W\right) Y^i, \tag{2.71}$$

where $v(P,W)$ is the same for all i.[17] By using Roy's Identity the Marshallian demands for consumption and leisure are obtained:

$$C^i = -\frac{\partial V^i/\partial P}{\partial V^i/\partial Y^i} = -\frac{\partial v(P,W)/\partial P}{v(P,W)} Y^i, \tag{2.72}$$

$$\bar{L} - L^i = -\frac{\partial V^i/\partial W}{\partial V^i/\partial Y^i} = -\frac{\partial v(P,W)/\partial W}{v(P,W)} Y^i. \tag{2.73}$$

The key thing to note about these expressions is that the marginal propensities to consume goods and leisure out of full income are the same for all household members and independent of full income of each member. But this means that aggregate family demands for goods and leisure can be written as:

$$C = -\frac{\partial v(P,W)/\partial P}{v(P,W)} Y, \tag{2.74}$$

$$\bar{L}' - L = -\frac{\partial v(P,W)/\partial W}{v(P,W)} Y, \tag{2.75}$$

where $C \equiv C^1 + C^2$, $Y \equiv Y^1 + Y^2$, $\bar{L}' \equiv 2\bar{L}$, and $L \equiv L^1 + L^2$ denote, respectively, household consumption, full income, time endowment, and labour supply. The fact that family consumption and labour supply only depends on family full income suggests that an aggregate approach is valid. To demonstrate that this is so, we postulate a representative agent whose indirect utility function is given by:

$$V \equiv V^1 + V^2 = v(P,W)\, Y. \tag{2.76}$$

It is easy to see that the application of Roy's Identity to (2.76) yields exactly the same family demands for consumption and leisure as given in (2.74)-(2.75). Hence, with homothetic preferences the focus on the representative household is justified.[18] The household is just a blown-up version of each household member.

The second justification suggested by Samuelson (1956) does not require individual preferences to be of a particular form. Instead, Samuelson assumes the existence of a so-called *household welfare function*. In the context of our simple 2-person household, the individual utility functions are as given by (2.69)

[17] The direct utility function (2.69) can be written as $U^i\left(C^i, \bar{L} - L^i\right) = G^i\left(\bar{U}\left(C^i, \bar{L} - L^i\right)\right)$, where $\bar{U}(\cdot)$ is homogeneous of degree one and the same for all household members. The function G^i is strictly increasing but may differ between household members. The $v(P,W)$ function is then defined as:

$$v(P,W) \equiv \max_{\{y_1, y_2\}} \bar{U}(y_1, y_2) \quad \text{subject to:} \quad 1 = Py_1 + Wy_2.$$

[18] As is argued by Varian (1992, p. 154), aggregation from the individual to the aggregate household level is possible if and only if the indirect utility function of household members is of the Gorman form and can be written as:

$$V^i = f^i(P,W) + v(P,W)\, Y^i,$$

where $f^i(P,W)$ is allowed to differ across household members. Note that the indirect utility function for the homothetic case (2.71) is a simple example of a Gorman form (with $f^i(P,W) = 0$ for all i). See also Gorman (1953).

(which need not be homothetic), and the household acts *as if* it maximizes the following function:

$$HW \equiv \Psi\left(U^1, U^2\right), \tag{2.77}$$

where HW is an indicator for household welfare and $\Psi\left(U^1, U^2\right)$ is some function featuring positive partial derivatives, i.e. $\Psi_i \equiv \partial\Psi/\partial U^i > 0$ and each member's utility contributes toward household welfare. The only joint decision that has to be made at the family level deals with the division of full income among its members (such that $Y^1 + Y^2 = Y$). For a given level of full income, each household member makes its *own* consumption and labour supply decision in order to maximize its *own* utility U^i subject to its *own* budget constraint (2.70). The key first-order condition characterizing each member's private optimum calls for an equalization of the marginal rate of substitution between leisure and consumption to the real wage rate:

$$\frac{\partial U^i/\partial\left(\bar{L} - L^i\right)}{\partial U^i/\partial C^i} = w. \tag{2.78}$$

Samuelson's basic insight can now be stated as follows. Provided household full income is distributed *optimally* across its members (in a manner to be explained below), the choices of individual household members will be such as to maximize a household welfare function involving only aggregate household quantities. Put differently, the household welfare function can be written directly in terms of aggregate consumption and labour supply:

$$HW \equiv \Phi\left(C, \bar{L}' - L\right), \tag{2.79}$$

and individual choices are such that (2.79) is maximized subject to the household budget constraint:

$$PC + W\left(\bar{L}' - L\right) = Y. \tag{2.80}$$

A heuristic proof of Samuelson's important theorem runs as follows. First, by substituting (2.69) into (2.77) we obtain an expression for household welfare directly in terms of individual quantities:

$$HW \equiv \Psi\left(U^1\left(C^1, \bar{L} - L^1\right), U^2\left(C^2, \bar{L} - L^2\right)\right). \tag{2.81}$$

The household budget constraint (2.80) can similarly be rewritten as:

$$P\left(C^1 + C^2\right) + W\left(\bar{L} - L^1 + \bar{L} - L^2\right) = Y. \tag{2.82}$$

In the household optimum, C^i and L^i are chosen such that (2.81) is maximized subject to (2.81). This gives the key first-order conditions:

$$\frac{\partial \Psi}{\partial U^i}\frac{\partial U^i}{\partial C^i} = \lambda P, \tag{2.83}$$

$$\frac{\partial \Psi}{\partial U^i}\frac{\partial U^i}{\partial \left(\bar{L} - L^i\right)} = \lambda W, \tag{2.84}$$

where λ is the Lagrange multiplier for the household budget constraint (2.82). Obviously, by combining (2.83)-(2.84) for the same person we obtain (2.78), i.e. from the family point of view individual members make the correct marginal decision regarding consumption and leisure. Furthermore, by using (2.83) for the two household members we obtain:

$$[\lambda P =] \;\frac{\partial \Psi}{\partial U^1}\frac{\partial U^1}{\partial C^1} = \frac{\partial \Psi}{\partial U^2}\frac{\partial U^2}{\partial C^2}. \tag{2.85}$$

Since $\partial U^i / \partial C^i$ is the marginal utility of income to person i, and $\partial \Psi / \partial U^i$ is the weight in household welfare of that person, equation (2.85) requires the marginal household utility of income to be the same for all household members (Samuelson, 1956, p. 11). This is the important family income distribution condition mentioned above. Lump-sum redistribution of income across household members must be such that (2.85) holds.

But the satisfaction of conditions (2.78) and (2.85) implies that the maximization of the alternative household welfare function (2.79) subject to the household budget constraint (2.80) yields exactly the same solutions as in the previous paragraph. Indeed, we can relate the two approaches by noting that:

$$\frac{\partial \Phi}{\partial C} = \frac{\partial \Phi}{\partial C^i} = \frac{\partial \Psi}{\partial U^i}\frac{\partial U^i}{\partial C^i} = \lambda P, \tag{2.86}$$

$$\frac{\partial \Phi}{\partial \left(\bar{L}' - L\right)} = \frac{\partial \Phi}{\partial \left(\bar{L} - L^i\right)} = \frac{\partial \Psi}{\partial U^i}\frac{\partial U^i}{\partial \left(\bar{L} - L^i\right)} = \lambda W, \tag{2.87}$$

where we have used (2.83)-(2.84) to establish the equivalency between the two sets of first-order conditions. If one is willing to assume that the household welfare function (2.79) is strictly quasi-concave in its arguments, C and $\bar{L}' - L$, then the analysis can proceed entirely at the household level (as was done in Section 2.1 above) and all the usual duality methods can be employed.[19]

In the discussion so far, aggregate household labour supply is determined as a unique function of full household income. For example, for the homothetic model aggregate labour supply takes the form as given in equation (2.75). Similarly, for the household-welfare-function model, an expression for aggregate labour supply results from the maximization of (2.79) subject to (2.80). But what about the labour supply of different household members? For both models, labour supply of household member i is a unique function of the full income level of that person (Y^i). In the homothetic model, the division of full income across households members is indeterminate so the theory itself does not determine individual

[19]See also Gorman (1959), who argues that the indifference curves associated with household welfare (2.79) are likely to be convex toward the origin.

labour supplies. In contrast, in the household-welfare-function model the division of household full income is determined within the model (by the condition stated in (2.85)). Depending on the properties of the household welfare function (2.81), labour supplies of both household members are determined.

In the empirical literature, the determinateness of individual labour supplies is often imposed explicitly by writing the household utility function as:

$$U = U\left(C, \bar{L} - L^1, \bar{L} - L^2\right), \tag{2.88}$$

where C is total household consumption, $\bar{L}$ is each member's time endowment, and L^i is labour supply by household member i. The household budget equation is then written as:

$$PC + W\left(\bar{L} - L^1\right) + W\left(\bar{L} - L^2\right) = Y, \tag{2.89}$$

where Y is total full income of the household. In this formulation, (a) leisure for the two household members need not be perfectly substitutable in utility (as it is in the household welfare function (2.79)), and (b) the household engages in full income sharing. In this model, the household chooses total consumption (C) and the labour supplies (L^1 and L^2) in order to maximize utility (2.88) subject to the budget constraint (2.89). Provided leisure of the two members are less-than-perfect substitutes in utility, individual labour supplies are determinate and take the form $L^i = L^i(W, P, Y)$.[20]

2.3.2.2 Family bargaining

In recent years, a number of authors have applied the tools of cooperative bargaining theory to the issue of household decision making. In this subsection we present a simplified version of the model suggested by Manser and Brown (1980) to illustrate some of the key points which emerge in such a setting. As before, attention is restricted to a two-person (potential) household. We first study each individual's behaviour if they remain single. Next, we illustrate what happens if they join resources and "get married."

The utility functions of the two persons take the following form:

$$U^1 = U^1\left(C_1^1, C_2^1, \bar{L} - L^1, Z^2\right), \tag{2.90}$$

$$U^2 = U^2\left(C_1^2, C_2^2, \bar{L} - L^2, Z^1\right), \tag{2.91}$$

where C_j^i is the consumption of good j by person i ($j = 1,2$ and $i = 1,2$) and Z^i is an efficiency parameter which depends on the marital state and is meant to capture the notion of love and companionship.[21] If

[20] See, for example, Ashenfelter and Heckman (1974) for a model of this type. They also allow the wage rates for the two household members to differ, so that labour supply functions will take the form $L^i = L^i\left(W^1, W^2, P, Y\right)$, where W^i is the wage rate of member i. Boskin and Sheshinski (1983) use this type of model to study the optimal tax treatment of married couples.

[21] McElroy and Horney (1981) instead assume that one partner's consumption and leisure directly enters the other partner's utility function.

the two persons remain single, it is assumed that $Z^1 = Z^2 = 0$. In contrast, if they decide to get married, then Z^1 and Z^2 take on positive values and, under the assumption that the two partners actually like each other (so that $\partial U^1/\partial Z^2 > 0$ and $\partial U^2/\partial Z^1 > 0$), the utility of each partner rises as a result of the marriage. There are two types of consumption goods in this model. Good 2 is a traditional good, but good 1 is a so-called *household good* (or *shared* good). This good has the property that in the married state, one partner's consumption of the good does not affect the amount available for consumption by the other partner.[22] As an example of a shared good one could think of a (large) house, a king-size bed, or heating.

If the two persons are not married, they each face a budget constraint of the form:

$$P_1 C_1^i + P_2 C_2^i + W\left(\bar{L} - L^i\right) = Y^i, \tag{2.92}$$

where $Y^i \equiv M^i + W\bar{L}$ is exogenous full income of person i (of which M^i is non-labour income), P_j is the price of good j, and W is the (common) wage rate. As before we abstract from consumption and labour income taxes. In the *single state*, person i maximizes utility U^i subject to the own budget constraint (2.92), and taking as given $Z^1 = Z^2 = 0$. The resulting indirect utility function for person i can be written in general form as:

$$V_S^i = V^i\left(P_1, P_2, W, Y^i\right), \tag{2.93}$$

where the subscript "S" refers to the single state. For given prices, wage rate, and full income, person i achieves the utility level V_S^i when single.

If the two persons decide to get married, there are three things which change. First, as was mentioned above, the efficiency parameters Z^1 and Z^2 take on positive values. Second, the two partners can profit from the consumption of the household good. Third, it is typically assumed that the two engage in income sharing once married, i.e. instead of facing the individual budget constraints (2.92), the married couple faces the following *household budget constraint*:

$$P_1 C_1 + P_2\left(C_2^1 + C_2^2\right) + W\left(\bar{L} - L^1 + \bar{L} - L^2\right) = Y, \tag{2.94}$$

where $Y \equiv Y^1 + Y^2$ is the household's full income. The key thing to note in the comparison between (2.92) and (2.94) is the cost saving that is possible for the shared good. In the single state, both partners must buy such goods but in the married state one purchase will affect both partner's utility directly (i.e. when married, $C_1^1 = C_1^2 = C_1$ in (2.90)-(2.91)).

The gain from being married (instead of single) can now be defined as the difference in utility in the

[22]The household good is thus like a pure public good studied in Chapter 13 below. Good 1 is a pure private good.

married and single states for the two (prospective) partners:

$$\Gamma^1 \equiv U^1\left(C_1, C_2^1, \bar{L} - L^1, Z^2\right) - V_S^1 \geq 0, \tag{2.95}$$

$$\Gamma^2 \equiv U^2\left(C_1, C_2^2, \bar{L} - L^2, Z^2\right) - V_S^2 \geq 0, \tag{2.96}$$

where V_S^i is defined in (2.93) above. Following Manser and Brown (1980), it is assumed that $\Gamma^i \geq 0$ (for $i = 1,2$) for existing marriages, i.e. all marriages for which this does not hold are dissolved. In this setting, V_S^1 and V_S^2 are often referred to as the *threat points* of, respectively, person 1 and 2. For feasible marriages it remains to decide how the household makes its decisions. Here we consider two approaches, namely the *dictatorial* model and the *Nash-bargaining* model.

2.3.2.2.1 Dictatorial model In the dictatorial model, it is assumed that one partner (say person 1) has the power to fully determine the family allocation of resources. Person 1 thus chooses C_1, C_2^i, and L^i (for $i = 1,2$) in order to maximize U^1 (given in (2.90) above) subject to the household budget constraint (2.94) and the *participation constraint* of the subservient partner, equation (2.96). This participation constraint limits the dictatorial powers of person 1 because the marriage will be dissolved if the subservient partner is better off in the single state. The Lagrangian for this optimization problem is:

$$\begin{aligned}\mathcal{L} \equiv\, & U^1\left(C_1, C_2^1, \bar{L} - L^1, Z^2\right) + \mu\left[U^2\left(C_1, C_2^2, \bar{L} - L^2, Z^2\right) - V_S^2\right] \\ & + \lambda\left[Y - P_1C_1 - P_2\left(C_2^1 + C_2^2\right) - W\left(\bar{L} - L^1 + \bar{L} - L^2\right)\right],\end{aligned}$$

where μ and λ are the Lagrange multipliers for, respectively, the participation constraint and the household budget constraint. Assuming the participation constraint to hold with equality ($\mu > 0$), the first-order conditions for the dictatorial optimum are given by the two constraints and:

$$\frac{\partial \mathcal{L}}{\partial C_1} = \frac{\partial U^1}{\partial C_1} + \mu\frac{\partial U^2}{\partial C_1} - \lambda P_1 = 0, \tag{2.97}$$

$$\frac{\partial \mathcal{L}}{\partial C_2^1} = \frac{\partial U^1}{\partial C_2^1} - \lambda P_2 = 0, \tag{2.98}$$

$$\frac{\partial \mathcal{L}}{\partial C_2^2} = \mu\frac{\partial U^2}{\partial C_2^2} - \lambda P_2 = 0, \tag{2.99}$$

$$\frac{\partial \mathcal{L}}{\partial\left(\bar{L} - L^1\right)} = \frac{\partial U^1}{\partial\left(\bar{L} - L^1\right)} - \lambda W = 0, \tag{2.100}$$

$$\frac{\partial \mathcal{L}}{\partial\left(\bar{L} - L^1\right)} = \mu\frac{\partial U^2}{\partial\left(\bar{L} - L^2\right)} - \lambda W = 0. \tag{2.101}$$

The two constraints plus the first-order conditions (2.97)-(2.101) jointly determine optimal solutions for C_1, C_2^i, L^i (for $i = 1,2$), λ, and μ as a function of the exogenous variables, P_1, P_2, W, Z^1, Z^2, Y^1, and Y^2.

The labour supply functions for the two partners can thus be written in general functional form as:

$$\bar{L} - L^1 = H^1 \left(P_1, P_2, W, Z^1, Z^2, Y^1, Y^2 \right), \tag{2.102}$$

$$\bar{L} - L^2 = H^2 \left(P_1, P_2, W, Z^1, Z^2, Y^1, Y^2 \right), \tag{2.103}$$

where the partial derivatives can be obtained in the usual fashion by applying the implicit function rule.[23] The key thing to note about these labour supply functions is that Y^1 and Y^2 exert separate effects on L^1 and L^2. Despite the fact that the household engages in income pooling (see (2.94) above), pre-marital income levels of both partners matter because Y^2 affects the minimum utility, V_S^2, required by the subservient partner to remain married (see (2.93) above).

2.3.2.2.2 Nash-bargaining model In the bargaining model, it is assumed that the married couple negotiate over the optimal choices for C_1, C_2^i, and L^i (for $i = 1, 2$). The outcome of this bargaining process is modelled as a so-called generalized Nash bargaining solution (see e.g. Binmore and Dasgupta, 1987). According to this solution concept, the variables are chosen such that the geometrically weighted average of the gains to the two partners is maximized subject to the constraints. In logarithmic terms we have:

$$\ln \Omega \equiv \zeta \ln \Gamma^1 + (1 - \zeta) \ln \Gamma^2, \tag{2.104}$$

where Γ^i is the (non-negative) gain that person i derives from being married and ζ represents the relative bargaining strength of person 1 ($0 \leq \zeta \leq 1$). Obviously, the dictatorial model is obtained as a special case of the Nash bargaining model by setting $\zeta = 1$.

In the (generalized) Nash bargaining model, the married couple choose C_1, C_2^i, and L^i (for $i = 1, 2$) in order to maximize (2.104) subject to the household budget constraint (2.94) and the non-negativity conditions (2.95)-(2.96). The Lagrangian for this optimization problem is:

$$\begin{aligned} \mathcal{L} \equiv\; & \zeta \ln \left[U^1 \left(C_1, C_2^1, \bar{L} - L^1, Z^2 \right) - V_S^1 \right] + (1 - \zeta) \ln \left[U^2 \left(C_1, C_2^2, \bar{L} - L^2, Z^2 \right) - V_S^2 \right] \\ & + \lambda \left[Y - P_1 C_1 - P_2 \left(C_2^1 + C_2^2 \right) - W \left(\bar{L} - L^1 + \bar{L} - L^2 \right) \right]. \end{aligned}$$

The first-order conditions for an interior optimum are given by the household budget constraint and:

$$\frac{\partial \mathcal{L}}{\partial C_1} = \frac{\zeta}{\Gamma^1} \frac{\partial U^1}{\partial C_1} + \frac{1 - \zeta}{\Gamma^2} \frac{\partial U^2}{\partial C_1} - \lambda P_1 = 0, \tag{2.105}$$

$$\frac{\partial \mathcal{L}}{\partial C_2^1} = \frac{\zeta}{\Gamma^1} \frac{\partial U^1}{\partial C_2^1} - \lambda P_2 = 0, \tag{2.106}$$

$$\frac{\partial \mathcal{L}}{\partial C_2^2} = \frac{1 - \zeta}{\Gamma^2} \frac{\partial U^2}{\partial C_2^2} - \lambda P_2 = 0, \tag{2.107}$$

[23] McElroy and Horney (1981) present the details of such a derivation in a much more general model with Nash bargaining.

$$\frac{\partial \mathcal{L}}{\partial (\bar{L}-L^1)} = \frac{\zeta}{\Gamma^1}\frac{\partial U^1}{\partial (\bar{L}-L^1)} - \lambda W = 0, \tag{2.108}$$

$$\frac{\partial \mathcal{L}}{\partial (\bar{L}-L^2)} = \frac{1-\zeta}{\Gamma^2}\frac{\partial U^2}{\partial (\bar{L}-L^2)} - \lambda W = 0. \tag{2.109}$$

In combination with the household budget constraint (2.94), the first-order conditions (2.105)-(2.109) jointly determine the optimal solutions for C_1, C_2^i, L^i (for $i = 1, 2$) and λ as a function of the exogenous variables, P_1, P_2, W, Z^1, Z^2, Y^1, Y^2, and ζ. For the labour supply functions we thus obtain:[24]

$$\bar{L} - L^1 = H^1\left(P_1, P_2, W, Z^1, Z^2, Y^1, Y^2, \zeta\right), \tag{2.110}$$

$$\bar{L} - L^2 = H^2\left(P_1, P_2, W, Z^1, Z^2, Y^1, Y^2, \zeta\right). \tag{2.111}$$

Both the bargaining parameter (ζ) and the pre-marital income levels (Y^1 and Y^2) exert separate effects on L^1 and L^2. The bargaining parameter influences the outcome because it affects the relative weight that each partner gets in the optimization problem. The income levels matter because they affect the threat points, V_S^1 and V_S^2.

2.4 Empirical evidence

There is a huge econometric literature attempting to estimate labour supply equations for households or individuals. While it is clearly impossible to do this literature justice here, in this section we nevertheless present a brief discussion of the empirical evidence regarding the key elasticities appearing in the static labour supply model. The interested reader is referred to the excellent literature surveys on, respectively, male and female labour supply by Pencavel (1986) and Killingsworth and Heckman (1986), and to the more recent general survey by Blundell and MaCurdy (1999).

Before discussing the elasticity estimates in detail, two general remarks are in order. First, real world tax and welfare systems are very complex indeed. This means that the non-convex model (discussed in Subsection 2.1.4 above) is probably relevant for at least some—and probably most—households. Blundell and MaCurdy, for example, suggest that the *effective* tax rate faced by poor California households in the monthly income bracket $750-$1500 is 89%; much higher than the rates faced by wealthier Californians (1999, p. 1566). Second, it is a well-established fact of life in most countries that men and women exhibit different labour supply behaviour (hours decision) and different participation behaviour (decision to be in or out of the labour force).

In his survey on the labour supply behaviour of men, Pencavel (1986, p. 69) presents the results from fourteen key US studies. He reports three summary estimates, namely the uncompensated labour supply elasticity, $\varepsilon_w \equiv \frac{w^*}{L}\frac{\partial L}{\partial w^*}$, the compensated (Hicksian) labour supply elasticity, $\varepsilon_w^H \equiv \frac{w^*}{L}\left(\frac{\partial L}{\partial w^*}\right)_{U=U_0}$,

[24] McElroy and Horney (1981) present the comparative static properties of a slightly different Nash bargaining model in which both partners have equal bargaining power ($\zeta = 1/2$).

and (what he calls) the marginal propensity to earn out of non-wage income, $mpe \equiv w^* \frac{\partial L}{\partial m_0}$. Of course, it follows from the Slutsky equation for labour supply, given in (2.12) above, that ε_w, ε_w^H, and *mpe* are related according to:

$$\varepsilon_w = \varepsilon_w^H + mpe. \tag{2.112}$$

Eleven of the fourteen studies report *negative* estimates ε_w, i.e. the majority of studies suggest that the labour supply function for US males is backward sloping (dominant income effect). Only one study, by Wales and Woodland (1979), yields a positive estimate for ε_w of 0.14. The average ε_w for the thirteen remaining studies is -0.12. Interestingly, twelve out of the fourteen studies report *negative* estimates for *mpe*, suggesting that leisure is a normal good as in the standard labour supply models (only one study yields a small positive estimate). Though there appears to be consensus about the sign of *mpe*, the magnitude differs greatly between studies. Finally, eight of the fourteen studies report strictly positive estimates for the pure substitution elasticity ε_w^H (as the theory implies), but five report *negative* estimates (which is inconsistent with the theory). In summary, an "average and reasonable" guess for the parameters that emerges from the studies surveyed by Pencavel (1986) seems to be $\varepsilon_w = -0.12$, $\varepsilon_w^H = 0.11$, and $mpe = -0.23$. The labour supply curve for men is nearly vertical as income and substitution effects virtually cancel out. More or less the same conclusion emerges from the recent survey of non-linear budget constraint models by Blundell and MaCurdy (1999, Table 1, pp. 1646-1648).

In their survey on labour supply behaviour of women, Killingsworth and Heckman (1986) discuss a large number of different studies (and even results for different estimation procedures or subgroups within studies). The general impression that one gets from these studies is as follows. Estimates for the uncompensated wage elasticity, ε_w, are generally positive and quite large, suggesting that labour supply is upward sloping for women. Relatively few studies yield negative estimates for the pure substitution elasticity ε_w^H, and virtually all studies yield negative estimates for *mpe*, suggesting that leisure is a normal good for female workers. Again, more or less the same conclusions emerge from the recent survey of non-linear budget constraint models for married women by Blundell and MaCurdy (1999, Table 2, pp. 1649-1651).

In summary, the available econometric evidence on static labour supply models seems to suggest that (a) the uncompensated wage elasticity is near-zero for men but positive for (married) women, and (b) income elasticities are negative (and leisure is a normal good) for both men and (married) women.

2.5 Punchlines

In this chapter we study the static labour supply decision of individuals or households. In the basic labour supply model, the representative household maximizes a well-behaved utility function by choosing the optimal levels of goods and leisure consumption. On the one hand, supplying an additional unit of labour is bad for utility because it reduces the amount of leisure consumption. On the other hand, it is good for utility because it allows the household to buy more consumption goods. In the optimum the household equates marginal costs and benefits and sets the marginal rate of substitution between leisure and consumption equal to the after-tax real wage rate. Both consumption taxes and labour income taxes affect this wage rate and thus distort the labour supply decision. A tax or wage change causes both a pure substitution effect and an income effect on labour supply. The former is always non-negative whilst the latter is negative (positive) if leisure is a normal (inferior) good. In the standard case, with leisure as a normal good, the income effect is relatively unimportant for wealthy households so for them the pure substitution effect dominates.

The basic labour supply model is applied to a number of issues. First, by loglinearizing the model around an initial optimum, a quantitative-analytical analysis of small tax changes is made possible. In this analysis the crucial parameters are the substitution elasticity between leisure and consumption in the utility function, the initial share of non-labour income in total household income, and the initial ratio between leisure consumption and labour supply. Depending on the magnitudes of these parameters, the effects of changes in the tax system can be computed.

The second application of the basic labour supply model assumes that the labour income tax system displays rate progressivity, i.e. both the average and marginal labour income tax rates rise with the tax base (wage income). Since the choice set remains convex, the standard tools of comparative static continue to be applicable.

Many features of real world tax and social benefit systems render the household's choice set non-convex. The example that we discuss in detail shows how in a means-tested benefit program the means-testing parameter (relating benefits to wage income over and above subsistence) acts as an implicit tax on labour income for low-income households. With a non-convex choice set, standard calculus methods (based on infinitesimal changes) are no longer applicable and both the economic theorist's task (of deriving comparative static results) and the econometrician's task (of estimating income and substitution elasticities) are much more complicated.

Next we present a brief discussion of the labour force participation decision and the tax effects thereon. If labour supply is indivisible, and can only take on a finite number of discrete values (say zero, part-time, and full-time), then the attitude towards leisure in a household's utility function determines which state the household finds most to its liking. Those with a relatively high valuation of leisure will not work at all (and just rely on the unemployment benefit), whilst those households with a low valuation of leisure will work either part-time or full-time. Details of the unemployment benefit system are shown to exert a major influence on both the participation decision by individuals and on aggregate labour supply.

In the third section of this chapter we briefly discuss a number of alternative theoretical approaches to the labour supply decision. In the household production model, leisure does not enter utility directly. Instead, both household time and market commodities are used as inputs to "home-produce" consumption activities which enter the utility function. Any part of the time endowment not used in this manner (as an intermediate input) is supplied to the labour market. It is demonstrated that the household production model can be analyzed using the very same tools that were used for the basic labour supply model.

The second alternative theoretical approach to labour supply opens the black box of household decision making. We first discuss the classic household welfare function approach of Samuelson (1956). If individual preferences are of a particular type then arguing on the basis of the aggregate family entails no loss of generality at all. On the other hand, if individual preferences are sufficiently different, then maximization of aggregate household welfare is still valid provided the income distribution within the household is set appropriately. In the first case, aggregation from the individual to the aggregate is trivial because preferences are identical (in the relevant sense). In the second case, aggregation is made possible because the income distribution leads to an equalization of marginal household utility of income for each household member.

In recent years the tools of cooperative bargaining theory have been applied to the issue of household decision making. Using a simple model, we demonstrate that labour supply by the two partners depends both on total family income and on the two partners' pre-marital income levels. Despite the fact that the household members pool their incomes in marriage, the premarital income levels affect the fallback positions that are important in the bargaining setting.

The chapter concludes with a brief discussion of the empirical evidence. The available econometric evidence on static labour supply models seems to suggest that for most developed countries the following two features are relevant. First, the uncompensated wage elasticity of labour supply is near-zero for men but positive for (married) women. Second, the income elasticities of labour supply are negative (and leisure is a normal good) for both men and (married) women.

Further reading

Basic labour supply model. Atkinson and Stiglitz (1980, lecture 2) cover much of the same topics as we do. Deaton and Muellbauer (1980) present very thorough discussions of the basic neoclassical labour supply model (in Chapter 4) and the labour force participation decision (in Chapter 11). Their exposition makes extensive use of duality methods. See also Hausman (1981b, 1985) on the theory of labour supply with non-linear budget sets. Stern (1986) presents a very comprehensive study of different functional forms for labour supply.

Home production. The classic papers on home production are by Becker (1965), Lancaster (1966), and Muth (1966). Good theoretical expositions are Atkinson and Stern (1979, 1980) and Kleven (2004). Early empirical studies using the basic notion of home production include Wales and Woodland (1977), Gronau (1977, 1980), Atkinson and Stern (1979, 1980), and Graham and Green (1984). In the home production model used in the text, all activities must be produced within the household itself—they cannot be bought in the market. An alternative formalization of the home production idea is due to Gronau (1977). In his model, the household uses its time endowment for three activities, namely leisure consumption, labour supply, and home production of goods or services (e.g. cooking, child rearing, etc.) which can also be bought in the market. This approach has been used for tax policy analysis by, inter alia, Sandmo (1990) and Kleven, Richter, and Sørenson (2000).

Family labour supply (often combined with home production). The classic source on the household welfare function approach is Samuelson (1956). The key theoretical contributions to the household bargaining model are Manser and Brown (1980) and McElroy and Horney (1981). A critique of the bargaining approach is found in Chiappori (1988a). He develops an alternative approach which is based on a direct assumption of Pareto efficiency in the household. See Chiappori (1988b, 1992) and Browning and Chiappori (1998). Apps and Rees (1997) and Chiappori (1997) combine household production and collective decision making. Nice surveys on household (bargaining) theory are Bourguignon and Chiappori (1994), Bergstrom (1996), and Pollak (2005). Taxation issues in a family context are discussed by *inter alia* Boskin and Sheshinski (1983), Apps and Rees (1988,1999a,1999b), Piggott and Whalley (1996), and Kleven and Kreiner (2004).

Empirical evidence. Good general surveys are Pencavel (1986) (for male labour supply) and Killingsworth and Heckman (1986) (female labour supply). See also Blundell and MaCurdy (1999) for a survey of recent empirical approaches. Atkinson and Stern (1980, 1981) estimate a system of commodity demands in which labour is used for household production. The impact of the 1986 US Tax Reform Act on female labour supply are studied by Eissa (1995) and Eissa and Liebman (1996).

Chapter 3

Taxation and intertemporal choice

The purpose of this chapter is to discuss the following topics:

- How can we model the consumption-savings decision by households and how is it affected by the various taxes?
- Which *tax equivalency* results be established between the various taxes in a dynamic context?
- What are the key implications of endogenizing the labour supply decision in the simple two-period Fisherian model of consumption and saving?
- How can we extend the Fisherian model and incorporate real world phenomena such human capital investment, borrowing restrictions, and intergenerational altruism and bequests?
- To what extent does the empirical evidence support the intertemporal substitution hypothesis for consumption, saving, and labour supply?

In this chapter, we focus on simple dynamic representative-agent models. We abstract from risk and uncertainty (decision making under uncertainty is studied in Chapter 4 below). The models are partial equilibrium in nature, i.e. factor prices and tax rates are taken parametrically by the agents and are exogenous to the models. Only linear taxes are considered in the policy experiments (non-linear taxes were studied in a static framework in Chapter 2).

3.1 A basic intertemporal model

Whereas the previous chapter restricted attention to static models, it is not very difficult to introduce an intertemporal dimension in the model. In doing so we are able to study the household's consumption-saving choice. The models that are constructed and used throughout this chapter build on the pioneering approach by Irving Fisher (1930). In his honour we shall therefore refer to these models as the Fisherian model of consumption, saving, and labour supply.

The basic Fisherian model makes the following simplifying assumptions. First, calender time is split into two main segments, namely period 1, which we call the *present*, and period 2, which stands for the remaining *future*. Obviously, by construction, there is no period 3 and period 0 refers to the past (which cannot be undone). Second, we continue to abstract from risk and uncertainty and assume that the representative household possesses *perfect foresight* about wages, prices, interest rates, and taxes. Third, in the most basic version of the Fisherian model we assume that household labour supply is exogenous. Fourth, there are perfect capital markets (no constraints on borrowing or lending), and there are no bequests (no intergenerational links).[1]

The representative household's *lifetime utility* is given by:

$$\Lambda = U(C_1, C_2), \tag{3.1}$$

where Λ is lifetime utility and C_t is consumption in period t. We assume positive but diminishing marginal utility of consumption in both periods, that is $U_t \equiv \partial U/\partial C_t > 0$ and $U_{tt} \equiv \partial^2 U/\partial C_t^2 < 0$ (for $t = 1, 2$). For the time being, we place no restriction on the sign of $U_{12} \equiv \partial U_1/\partial C_2 \equiv \partial U_2/\partial C_1 \equiv U_{21}$ but we do assume that the indifference curves bulge toward the origin (as in Figure 3.1 below). That is, $U(\cdot)$ is strictly quasi-concave and $U_{11}U_{22} - U_{12}^2 > 0$.

The budget identities (in nominal terms) for the two periods are given by:

$$A_1 = (1 + R_0)A_0 + W_1\bar{L} - P_1C_1, \tag{3.2}$$

$$A_2 = (1 + R_1)A_1 + W_2\bar{L} - P_2C_2, \tag{3.3}$$

where A_t represents financial assets at the end of period t (A_0 was accumulated in period 0, i.e. in the "past"), R_t is the nominal interest rate in period t, $L_t = \bar{L}$ is labour supply in period t ($\bar{L}$ is the exogenous time endowment), P_t is the price of consumption goods in period t, and W_t is the nominal wage rate in period t. We choose the price of the consumption good, P_t, as the numeraire so that the real budget identities can be written as:

$$a_1 = (1 + r_0)a_0 + w_1\bar{L} - C_1, \tag{3.4}$$

$$a_2 = (1 + r_1)a_1 + w_2\bar{L} - C_2, \tag{3.5}$$

where $a_t \equiv A_t/P_t$, $w_t \equiv W_t/P_t$, and r_t are, respectively, real financial assets, the real wage rate, and the real interest rate in period t. The real and nominal interest rates are related according to:

$$(1 + r_t) \equiv (1 + R_t)\frac{P_t}{P_{t+1}}. \tag{3.6}$$

[1]Below we also study endogenous labour supply, imperfect capital markets, and intergenerational bequests. Risk and uncertainty are studied in Chapter 4. The multi-period version of the model is discussed briefly in Section 3.4 and more extensively in Chapter 8.

Finally, in order to turn the budget *identities* into budget *constraints*, we impose a so-called solvency condition:

$$a_2 = 0. \tag{3.7}$$

This expression can be understood as follows. First, in the absence of preference satiation and bequests, the household would certainly not plan to possess positive assets at the end of period 2. Indeed, it would love to be heavily indebted at the end of period 2, i.e. from the household's point of view the relevant constraint is thus that $a_2 \leq 0$. The capital market (the lenders) will, however, not allow the household to be indebted at the end of period 2, i.e. the constraint $a_2 \geq 0$ must also hold. By combining the two constraints we obtain (3.7).

In the basic model, we abstract from any additional lending- or borrowing constraints, i.e. in period 1 the household can freely borrow or lend at the going interest rate r_1, a_1 can have either sign, and the budget identities can be *consolidated* into a single *lifetime-budget constraint*.[2] In step 1 we set $a_2 = 0$ and solve (3.4)-(3.5) for a_1:

$$a_1 = \frac{C_2 - w_2\bar{L}}{1+r_1} = (1+r_0)a_0 + w_1\bar{L} - C_1. \tag{3.8}$$

In step 2 we re-write the final equality in (3.8) as follows:

$$C_1 + \frac{C_2}{1+r_1} = (1+r_0)a_0 + h_0 \equiv \Omega, \tag{3.9}$$

where Ω is total wealth and h_0 is *human wealth* representing the after-tax value of the time endowment:

$$h_0 \equiv w_1\bar{L} + \frac{w_2\bar{L}}{1+r_1}. \tag{3.10}$$

Intuitively, the life-time budget constraint (3.9) says that, for the solvent household, the present value of spending on goods (left-hand side of (3.9)) equals initial total wealth (Ω on the right-hand side). The prices of C_1 and C_2 are, respectively, **1** and $1/(1+r_1)$. We could thus use duality theory just as for the static model (discussed in the previous chapter) to derive Hicksian and Marshallian demand expressions for C_1 and C_2. Obviously, we expect to find income and substitution effects to play a crucial role.

Here we solve the household optimization problem in the usual (primal) manner. The household chooses C_1 and C_2 in order to maximize lifetime utility (3.1) subject to the lifetime budget constraint (3.9). The Lagrangian expression is given by:

$$\mathcal{L} \equiv U(C_1, C_2) + \mu \left[\Omega - C_1 - \frac{C_2}{1+r_1}\right],$$

[2] With a binding borrowing constraint (of the type $a_1 \geq 0$) in the first period, the household is forced to set $a_1 = 0$ and the model is essentially a static one. See below.

where μ is the Lagrange multiplier for the constraint. The first-order necessary conditions are given by the constraint and the so-called *Euler equation*:[3]

$$\frac{U_1(C_1, C_2)}{U_2(C_1, C_2)} = 1 + r_1. \tag{3.11}$$

According to (3.11), the marginal rate of substitution between C_1 and C_2 is equated to the *relative* price of C_1. Note also that U_1 and U_2 in general depend on both C_1 and C_2, because $U_{12} \neq 0$ is not excluded a priori.

In order to study the key properties of the household's optimal consumption-saving plan, we make use of the implicit function theorem. Expressions (3.9) and (3.11) define implicit functions relating C_t to Ω and r_1. We write these functions as $C_t = C_t(\Omega, r_1)$ for $t = 1, 2$ and we wish to determine the various partial derivatives of these implicit functions. By totally differentiating (3.9) and (3.11) we obtain the following matrix expression:

$$\Delta \begin{bmatrix} dC_1 \\ dC_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} d\Omega + \begin{bmatrix} \frac{C_2}{(1+r_1)^2} \\ U_2 \end{bmatrix} dr_1, \tag{3.12}$$

where Δ is:

$$\Delta \equiv \begin{bmatrix} 1 & \frac{1}{1+r_1} \\ U_{11} - (1+r_1)U_{12} & U_{12} - (1+r_1)U_{22} \end{bmatrix}. \tag{3.13}$$

In deriving (3.12) we have made use of Young's theorem (which says that $U_{12} = U_{21}$). Furthermore, the second-order conditions for utility maximization ensure that $|\Delta| > 0$. Indeed, the *bordered Hessian* associated with the constrained maximization problem is given by:

$$\bar{H} \equiv \begin{bmatrix} 0 & 1 & \frac{1}{1+r_1} \\ 1 & U_{11} & U_{12} \\ \frac{1}{1+r_1} & U_{12} & U_{22} \end{bmatrix}, \tag{3.14}$$

and the second-order sufficient conditions for a maximum are that the principal minors of $\bar{H}$ alternate

[3]This Euler equation is obtained as follows. The first-order conditions for C_1 and C_2 are:

$$\frac{\partial \mathcal{L}}{\partial C_1} = U_1(C_1, C_2) - \mu = 0,$$
$$\frac{\partial \mathcal{L}}{\partial C_2} = U_2(C_1, C_2) - \frac{\mu}{1+r_1} = 0.$$

By eliminating μ from these expressions we find (3.11).

in sign, starting negative:

$$|H_2| \equiv \begin{vmatrix} 0 & 1 \\ 1 & U_{11} \end{vmatrix} = -1 < 0, \qquad |\bar{H}_3| \equiv \begin{vmatrix} 0 & 1 & \frac{1}{1+r_1} \\ 1 & U_{11} & U_{12} \\ \frac{1}{1+r_1} & U_{12} & U_{22} \end{vmatrix} = |\bar{H}| > 0. \tag{3.15}$$

Finally, it is easy to derive from (3.13) that $|\Delta| = (1+r_1)\,|\bar{H}|$. The utility maximization hypothesis thus yields very useful information that is needed in the comparative statics exercises.

Let us first consider the effects of a marginal increase in wealth. Holding constant r_1, we obtain from equation (3.12):[4]

$$\frac{\partial C_1}{\partial \Omega} = \frac{U_{12} - (1+r_1)U_{22}}{|\Delta|} \gtreqless 0, \tag{3.16}$$

$$\frac{\partial C_2}{\partial \Omega} = \frac{(1+r_1)U_{12} - U_{11}}{|\Delta|} \gtreqless 0. \tag{3.17}$$

The effect of a wealth change on consumption in both periods is ambiguous in general, i.e. we know that $U_{tt} < 0$ but $U_{12} \gtreqless 0$. If $U_{12} \geq 0$ then $\partial C_t / \partial \Omega > 0$ for $t = 1, 2$, and present and future consumption are both *normal goods*. In contrast, if $U_{12} < 0$ then either present consumption or future consumption may be an inferior good ($\partial C_t / \partial \Omega < 0$). It can be easily shown, however, that *at most* one good can be inferior. Indeed, it follows from (3.9) that:

$$\frac{\partial C_1}{\partial \Omega} + \frac{1}{1+r_1}\frac{\partial C_2}{\partial \Omega} = 1, \tag{3.18}$$

so that $\partial C_1 / \partial \Omega$ and $\partial C_2 / \partial \Omega$ cannot both be negative.

Next, we consider the effects of a marginal increase in the real interest rate r_1. This has two effects. First, the relative price of future consumption decreases. Second, the value of human wealth (and thus total wealth) falls. Indeed, we derive from (3.9)-(3.10) that:

$$\frac{\partial \Omega}{\partial r_1} = \frac{\partial h_0}{\partial r_1} = -\frac{w_2 \bar{L}}{(1+r_1)^2} < 0. \tag{3.19}$$

An increase in the interest rate causes future wage income to be discounted more heavily.

By taking both effects into account we obtain from (3.12):

$$\frac{\partial C_1}{\partial r_1} = \frac{U_{12} - (1+r_1)U_{22}}{|\Delta|}\frac{a_1}{1+r_1} - \frac{1}{|\Delta|}\frac{U_2}{1+r_1} \gtreqless 0, \tag{3.20}$$

$$\frac{\partial C_2}{\partial r_1} = \frac{(1+r_1)U_{12} - U_{11}}{|\Delta|}\frac{a_1}{1+r_1} + \frac{1}{|\Delta|}U_2 \gtreqless 0, \tag{3.21}$$

[4] In deriving the comparative static effects it is useful to note that Δ^{-1} is equal to:

$$\Delta^{-1} \equiv \frac{1}{|\Delta|}\begin{bmatrix} U_{12} - (1+r_1)U_{22} & -\frac{1}{1+r_1} \\ (1+r_1)U_{12} - U_{11} & 1 \end{bmatrix}.$$

where we have used the second period budget identity $(1+r_1)a_1 = C_2 - w_2\bar{L}$ to simplify these expressions. Without further restrictions on U_{12} and a_1 the effects are ambiguous. It is nevertheless possible to deduce a number of properties. First, by differentiating the lifetime budget equation (3.9) with respect to r_1 we find:

$$\frac{\partial C_1}{\partial r_1} + \frac{1}{1+r_1}\frac{\partial C_2}{\partial r_1} = \frac{a_1}{1+r_1}. \tag{3.22}$$

For an agent who chooses to save in the first period ($a_1 > 0$), either present or future consumption (or both) rise if the interest rate rises. Second, if $a_1 > 0$ and $U_{12} \geq 0$ then $\partial C_1/\partial r \gtreqless 0$ and $\partial C_2/\partial r > 0$. Third, for the special case where the agent's utility maximum happens to coincide with its endowment point (so that $a_1 = 0$), it follows that $\partial C_1/\partial r < 0$ and $\partial C_2/\partial r > 0$.

Just as in the previous chapter, it is convenient to focus on the case of homothetic preferences because this imposes more structure on the general specification of the model. Recall that $U(C_1, C_2)$ represents homothetic preferences if it can be written as $U(C_1, C_2) = G\left(\bar{U}(C_1, C_2)\right)$, where $G(\cdot)$ is a strictly increasing function and $\bar{U}(C_1, C_2)$ is homogeneous of degree one in C_1 and C_2. The $\bar{U}(C_1, C_2)$ function has the following properties:

(P1) $\bar{U}_1 C_1 + \bar{U}_2 C_2 = \bar{U}$;

(P2) $\bar{U}_1$ and $\bar{U}_2$ are homogeneous of degree zero in C_1 and C_2;

(P3) $\bar{U}_{12} = -(C_1/C_2)\bar{U}_{11} = -(C_2/C_1)\bar{U}_{22}$ and thus $\bar{U}_{11} = (C_2/C_1)^2\bar{U}_{22}$;

(P4) the substitution elasticity is $\sigma \equiv -d\ln(C_1/C_2)/d\ln(\bar{U}_1/\bar{U}_2) = \bar{U}_1\bar{U}_2/(\bar{U}\bar{U}_{12}) \geq 0$.

Note that homotheticity imposes quite a bit of structure on the model. Indeed, since $\bar{U}_{tt} < 0$ it follows from (P3) that $\bar{U}_{12} > 0$ (and thus $U_{12} > 0$). Hence, it follows from (3.16)-(3.17) that $\partial C_1/\partial\Omega > 0$ and $\partial C_2/\partial\Omega > 0$, i.e. present and future consumption are both normal goods! Returning to the interest rate shock, we find that for a homothetic utility function the Euler equation (3.11) reduces to $\bar{U}_1/\bar{U}_2 = 1 + r_1$. Since the $\bar{U}_t$ functions are homogeneous of degree zero, this Euler equation thus pins down a unique C_1/C_2 ratio as a function of $1 + r_1$.

By loglinearizing the Euler equation and the lifetime budget restriction (3.9), holding constant $(1 + r_0)a_0$, $w_1\bar{L}$, and $w_2\bar{L}$, we obtain the following expression:

$$\begin{bmatrix} \omega_1 & 1-\omega_1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{dC_1}{C_1} \\ \frac{dC_2}{C_2} \end{bmatrix} = \begin{bmatrix} a_1/\Omega \\ \sigma \end{bmatrix} \frac{dr_1}{1+r_1}, \tag{3.23}$$

where $\omega_1 \equiv C_1/\Omega$ and $1 - \omega_1 \equiv C_2/[(1+r_1)\Omega]$ are the budget shares of, respectively, present and future consumption. By inverting the matrix on the left-hand side and interpreting the derivatives in a partial sense (since all other determinants of consumption are held constant), we obtain the comparative

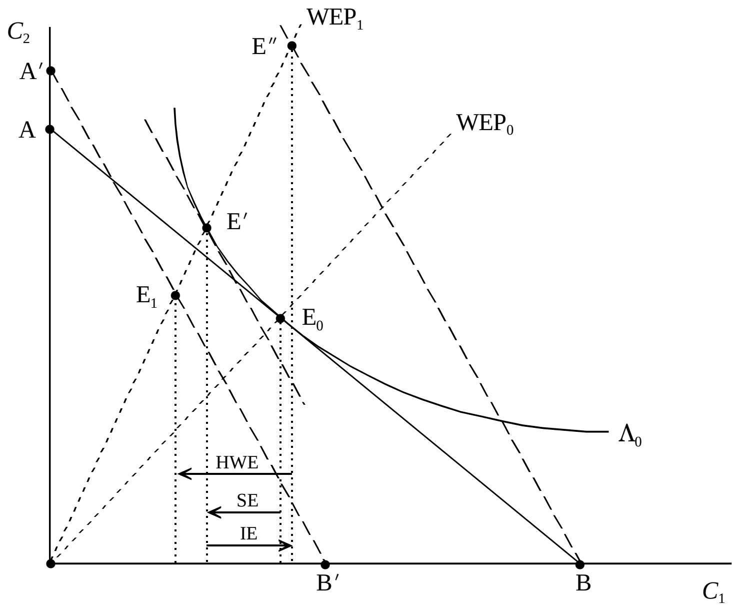

Figure 3.1: Income, substitution, and human wealth effects (homothetic case)

static effects:

$$\frac{\partial C_1}{\partial r_1} = \frac{C_1}{1+r_1}\left[(1-\omega_1) - \frac{w_2\bar{L}}{(1+r_1)\Omega} - (1-\omega_1)\sigma\right], \tag{3.24}$$

$$\frac{\partial C_2}{\partial r_1} = \frac{C_2}{1+r_1}\left[(1-\omega_1) - \frac{w_2\bar{L}}{(1+r_1)\Omega} + \omega_1\sigma\right], \tag{3.25}$$

where we have also used $(1+r_1)a_1 = C_2 - w_2\bar{L}$. The three terms appearing in square brackets on the right-hand sides of (3.24)-(3.25) represent, respectively, the *income effect*, the *human wealth effect*, and the *substitution effect*. We illustrate these effects in Figure 3.1.

The initial lifetime budget restriction (3.9) is given by the solid line AB and the initial equilibrium is at point E_0. The line WEP_0 from the origin is the initial *wealth expansion path* characterizing the optimal C_2/C_1 ratio—consistent with the homothetic version of (3.11)—for different levels of total wealth, Ω. As a result of the increase in r_1, the budget line shifts to A′B′ and the wealth expansion path rotates in a counter-clockwise fashion to WEP_1. The ultimate effect of the shock is given by the move from E_0 to E_1. Graphically, the total effect can be decomposed into constituent partial effects. Here we have shown the Hicksian decomposition. The move from E_0 to E′ is the pure substitution effect (SE) whilst the move from E′ to E″ is the income effect (IE). If the household were to have no non-interest income in the second period ($w_2\bar{L} = 0$) then the so-called human wealth effect would be absent (see (3.19) above). However, since $w_2\bar{L} > 0$ by assumption, the increase in the interest rate reduces the value of human capital and shifts the budget restriction inward. Hence, the human wealth effect (HWE) is represented by the move from E″ to E_1.

3.1.1 Some tax equivalency results

Armed with the basic Fisherian model it is possible to introduce various (linear) taxes and to investigate their effects on the choice set of a representative household. In the process it is possible to demonstrate some well known tax equivalency results (see also Atkinson and Stiglitz, 1980, pp. 70-72). The first equivalency result is that between (broadly defined) proportional labour taxes and proportional consumption taxes. A proportional tax (t_L) on wage income plus *initial* assets (e.g. inheritance) augments the budget constraints as follows:

$$a_1 = (1-t_L)\left[(1+r_0)a_0 + w_1\bar{L}\right] - C_1, \tag{3.26}$$

$$0 = (1+r_1)a_1 + w_2(1-t_L)\bar{L} - C_2, \tag{3.27}$$

so that the consolidated budget constraint becomes:

$$C_1 + \frac{C_2}{1+r_1} = (1-t_L)\left[(1+r_0)a_0 + w_1\bar{L} + \frac{w_2\bar{L}}{1+r_1}\right] \equiv \Omega_1. \tag{3.28}$$

In contrast, a proportional tax (t_C) on consumption augments the budget constraints as follows:

$$a_1 = (1+r_0)a_0 + w_1\bar{L} - (1+t_C)C_1, \tag{3.29}$$

$$0 = (1+r_1)a_1 + w_2\bar{L} - (1+t_C)C_2, \tag{3.30}$$

so that the consolidated budget constraint becomes:

$$(1+t_C)\left[C_1 + \frac{C_2}{1+r_1}\right] = (1+r_0)a_0 + w_1\bar{L} + \frac{w_2\bar{L}}{1+r_1} \equiv \Omega_2. \tag{3.31}$$

The comparison between (3.28) and (3.31) reveals that the respective effects of t_L and t_C on the choice set of the household is the same if (and only if):

$$1-t_L = \frac{1}{1+t_C}. \tag{3.32}$$

Provided the condition in (3.32) holds, it follows that total lifetime wealth is the same under the two tax systems, i.e. $\Omega_1 \equiv (1-t_L)\Omega_2 = \Omega_2/(1+t_C)$. Since relative prices are also the same, the two systems are equivalent, i.e. they yield the same solutions for C_1 and C_2.[5] Of course, as follows readily from the comparison of (3.26) and (3.29), the private saving plans are not identical because the paths of tax bills differ between the two tax systems.[6]

[5]Note that the equivalency result is quite special as it hinges on the time-constancy of the consumption tax, t_C! Indeed, if $t_{C1} \neq t_{C2}$ then the Euler equation is affected by consumption taxation but not by labour taxation. The consumption tax is like an interest income tax in that case. We return to this issue below.

[6]Denoting the optimal choices for a_1 under the two tax systems by a_1^L and a_1^C, respectively, we find that if the condition (3.32) holds, $a_1^L - a_1^C = t_C a_1^C$. It follows that the savings choices are only the same in the trivial case with $a_1^L = a_1^C = 0$.

There is also an equivalency result between an interest income tax and a wealth tax. A proportional tax (t_R) on interest income affects the budget constraints as follows:

$$a_1 = [1 + r_0 (1 - t_R)] a_0 + w_1 \bar{L} - C_1, \tag{3.33}$$
$$0 = [1 + r_1 (1 - t_R)] a_1 + w_2 \bar{L} - C_2, \tag{3.34}$$

so that the consolidated budget constraint becomes:

$$C_1 + \frac{C_2}{1 + r_1 (1 - t_R)} = [1 + r_0 (1 - t_R)] a_0 + w_1 \bar{L} + \frac{w_2 \bar{L}}{1 + r_1 (1 - t_R)} \equiv \Omega_3. \tag{3.35}$$

Time-varying proportional wealth taxes (t_{W1} and t_{W2}) affect the budget constraints as follows:

$$a_1 = (1 + r_0 - t_{W1}) a_0 + w_1 \bar{L} - C_1, \tag{3.36}$$
$$0 = (1 + r_1 - t_{W2}) a_1 + w_2 \bar{L} - C_2, \tag{3.37}$$

so that the consolidated budget constraint becomes:

$$C_1 + \frac{C_2}{1 + r_1 - t_{W2}} = (1 + r_0 - t_{W1}) a_0 + w_1 \bar{L} + \frac{w_2 \bar{L}}{1 + r_1 - t_{W2}} \equiv \Omega_4. \tag{3.38}$$

The comparison between (3.35) and (3.38) reveals that the two tax systems are equivalent if the following two conditions hold:

$$t_{W1} = r_0 t_R \quad \text{and} \quad t_{W2} = r_1 t_R. \tag{3.39}$$

Again several points are worth noting. First, even though t_R is constant over time, the wealth tax must take into account that the interest rate may be time-varying. Second, if the household is a net borrower in the first period ($a_1 < 0$), then according to (3.34), interest paid on loans are deductible for tax purposes (as $t_R r_1 a_1$ is negative in that case). The equivalency condition then implies that the wealth tax actually leads to receipts from the government (as $-t_{W2} a_1$ is positive in that case). Finally, the equivalency results are only possible if there exists only one asset in the economy, a rather unlikely situation (see Chapter 4 for models with more than one type of financial assets).

3.1.2 Application: The effects of consumption taxes

In this subsection we study the effects on consumption and saving of (potentially time-varying) consumption taxes, t_{C1} and t_{C2}. Instead of working with the general utility function (3.1), we employ an often-used specification which assumes *intertemporal additive separability* in preferences. In particular, the life-time utility function is written as:

$$\Lambda = U(C_1) + \frac{1}{1+\rho} U(C_2), \tag{3.40}$$

where $U(\cdot)$ is the *instantaneous utility* function (often called *felicity function*), and $\rho > 0$ is the constant *pure rate of time preference*, representing the effects of "impatience." The higher is ρ, the heavier future felicity is discounted, and the more impatient is the household. In addition, it is often assumed in the literature that the felicity function features a constant *intertemporal substitution elasticity*, σ:

$$U(C_t) \equiv \begin{cases} \frac{C_t^{1-1/\sigma}-1}{1-1/\sigma} & \text{for } \sigma \neq 1 \\ \ln C_t & \text{for } \sigma = 1 \end{cases}. \tag{3.41}$$

In the absence of other taxes and transfers, the consolidated lifetime budget constraint is:

$$(1+t_{C1})\, C_1 + \frac{(1+t_{C2})\, C_2}{1+r_1} = (1+r_0)a_0 + h_0 \equiv \Omega, \tag{3.42}$$

where Ω is again total wealth and h_0 is human wealth:

$$h_0 \equiv w_1 \bar{L} + \frac{w_2 \bar{L}}{1+r_1}. \tag{3.43}$$

The household chooses C_1 and C_2 in order to maximize lifetime utility subject to the consolidated budget constraint. The Lagrangian for this optimization problem is:

$$\mathcal{L} \equiv \frac{C_1^{1-1/\sigma}-1}{1-1/\sigma} + \frac{1}{1+\rho}\frac{C_2^{1-1/\sigma}-1}{1-1/\sigma} + \mu \left[\Omega - (1+t_{C1})\, C_1 - \frac{(1+t_{C2})\, C_2}{1+r_1} \right],$$

where μ is the Lagrange multiplier. The first-order conditions consist of the constraint (3.42) and:

$$C_1^{-1/\sigma} = \mu\,(1+t_{C1}), \tag{3.44}$$

$$\frac{1}{1+\rho} C_2^{-1/\sigma} = \frac{\mu\,(1+t_{C2})}{1+r_1}. \tag{3.45}$$

Since the optimized value of μ represents the marginal utility of wealth, equations (3.44) and (3.46) can be interpreted as instructing the household to equate the marginal utility of consumption in both periods (left-hand sides) to the marginal utility of wealth, corrected for the prices of C_1 and C_2 (of course, the price of C_1 is $1+t_{C1}$ whilst the price of C_2 is $(1+t_{C2})\,/\,(1+r_1)$).

By dividing (3.44) by (3.45) and rearranging we find the Euler equation:

$$\frac{C_2}{C_1} = \left[\frac{1+t_{C1}}{1+t_{C2}} \frac{1+r_1}{1+\rho} \right]^{\sigma}. \tag{3.46}$$

Several aspects of this Euler equation are worth noting. First, if the interest rate exceeds the rate of time preference ($r_1 > \rho$) then, ceteris paribus, the household chooses a high ratio between C_2 and C_1, i.e. present consumption is postponed in favour of future consumption and saving in the first period is high. The effect is more pronounced, the higher is the value of the intertemporal substitution parameter, σ. Second, if the consumption tax falls over time ($t_{C1} > t_{C2}$) then, ceteris paribus, C_2/C_1 is high. Again, current consumption is postponed and saving is high (more so the higher is σ). Third, the relative price of current and future consumption is affected by the consumption tax provided the tax rate is time-varying ($t_{C1} \neq t_{C2}$). Conversely, if t_C is time-invariant then it drops out of the Euler equation altogether (see also above).

Quantitative tax policy analysis can be conducted by loglinearizing the model and considering infinitesimal tax changes. Following the same steps as in Chapter 2 above, we find that (3.42) and (3.46) can be loglinearized to obtain the following matrix expression:

$$\begin{bmatrix} \omega_1 & 1-\omega_1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{C}_1 \\ \tilde{C}_2 \end{bmatrix} = \begin{bmatrix} -\omega_1 \\ \sigma \end{bmatrix} \tilde{t}_{C1} - \begin{bmatrix} 1-\omega_1 \\ \sigma \end{bmatrix} \tilde{t}_{C2}, \tag{3.47}$$

where $\tilde{C}_t \equiv dC_t/C_t$, $\tilde{t}_{Ct} \equiv dt_{Ct}/\left(1+t_{Ct}\right)$, $\omega_1 \equiv C_1\left(1+t_{C1}\right)/\Omega$, and $1-\omega_1 \equiv C_2\left(1+t_{C2}\right)/\left[\left(1+r_1\right)\Omega\right]$. There are three different cases which can be studied, namely: (a) a present tax change only ($\tilde{t}_{C1} > 0$ and $\tilde{t}_{C2} = 0$); (b) a future tax change only ($\tilde{t}_{C2} > 0$ and $\tilde{t}_{C1} = 0$); and (c) an equal tax change in both periods ($\tilde{t}_{C1} = \tilde{t}_{C2} =$) $\tilde{t}_C > 0$. Once the effects on consumption in the two periods have been determined, the effect on *net* saving, that is $a_1 - a_0$, can be deduced from the first-period (or second-period) budget identity (whichever happens to be most convenient):

$$\begin{aligned} S_1 &\equiv a_1 - a_0 \\ &= r_0 a_0 + w_1 \bar{L} - \left(1+t_{C1}\right) C_1 \qquad (3.48) \\ &= \frac{\left(1+t_{C2}\right) C_2 - w_2 \bar{L}}{1+r_1} - a_0. \qquad (3.49) \end{aligned}$$

3.1.2.1 Raising the current consumption tax

In the first case to be studied, only the current consumption tax is increased. The shock can thus be seen as an unanticipated and temporary increase in the consumption tax. The shock is "unanticipated" because it occurs immediately in the current period, and it is "temporary" because there is no tax change in the future.

Setting $\tilde{t}_{C2} = 0$ we find from equation (3.47) that current consumption and future consumption change according to:

$$\begin{bmatrix} \tilde{C}_1 \\ \tilde{C}_2 \end{bmatrix} = \begin{bmatrix} 1 & -\left(1-\omega_1\right) \\ 1 & \omega_1 \end{bmatrix} \begin{bmatrix} -\omega_1 \\ \sigma \end{bmatrix} \tilde{t}_{C1}$$

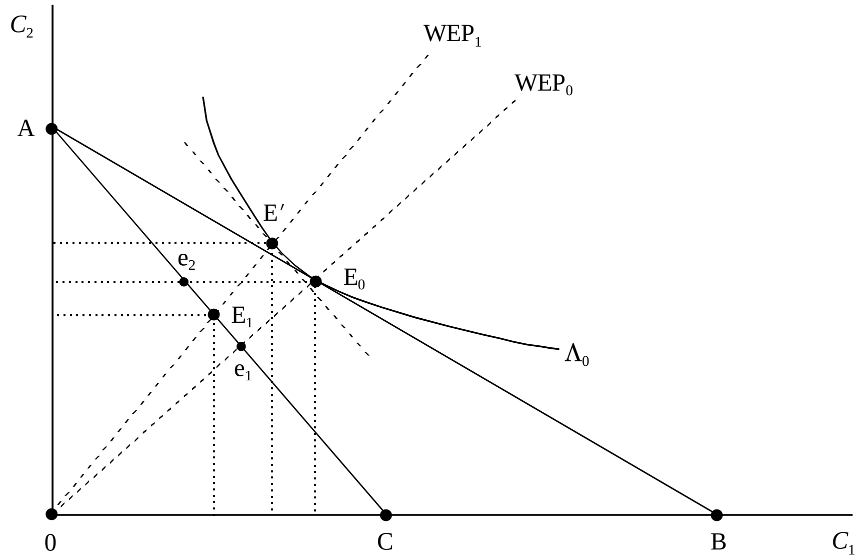

Figure 3.2: Raising the current consumption tax (low σ)

$$= \begin{bmatrix} -\left[\omega_1 + (1-\omega_1)\,\sigma\right] \\ (\sigma - 1)\,\omega_1 \end{bmatrix} \tilde{t}_C. \tag{3.50}$$

Present consumption falls unambiguously ($\tilde{C}_1 < 0$); the current tax increase induces the household to postpone current consumption. The effect on future consumption is ambiguous due to offsetting income and substitution effects. (In the Cobb-Douglas case (where $\sigma = 1$) these two effects exactly match and future consumption is unchanged.) It follows from (3.49) that the effect on net saving is also ambiguous. Indeed, saving rises (falls) if and only if future consumption rises (falls).

In Figure 3.2 we illustrate the effects on present and future consumption for the case of a relatively low substitution elasticity, i.e. for $0 < \sigma < 1$. The tax increase rotates the budget line in a clockwise fashion from AB to AC. The total effect consists of the move from E_0 to E_1. The pure substitution effect is represented by the move from E_0 to E', whilst the income effect is the move from E' to E_1. Note that for Leontief preferences—with no substitution possible at all ($\sigma = 0$)—the new equilibrium would be at e_1, whereas for the Cobb-Douglas case (with $\sigma = 1$) it would be at e_2 (where a wealth expansion path, WEP_2 (not drawn) would pass through e_2). Not surprisingly, the intermediate case with $0 < \sigma < 1$ lies in between the two corner cases of Leontief and Cobb-Douglas preferences, i.e. point E_1 lies in between points e_1 and e_2.

3.1.2.2 Raising the future consumption tax

In the second case to be studied, only the future consumption tax is increased. This shock can be seen as an anticipated and permanent increase in the consumption tax. The shock is "anticipated" because it is known in the current period to occur in the future, and it is "permanent" because the future tax remains at the higher level (recall that there is no period 3).

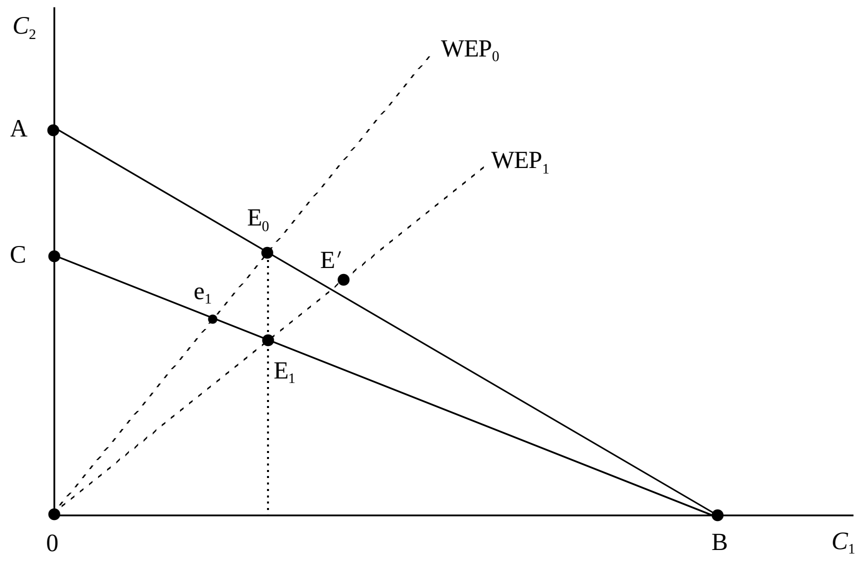

Figure 3.3: Raising the future consumption tax: the Cobb-Douglas case ($\sigma = 1$)

Setting $\tilde{t}_{C1} = 0$ we find from equation (3.47) that current consumption and future consumption change according to:

$$\begin{aligned}\begin{bmatrix} \tilde{C}_1 \\ \tilde{C}_2 \end{bmatrix} &= -\begin{bmatrix} 1 & -(1-\omega_1) \\ 1 & \omega_1 \end{bmatrix} \begin{bmatrix} 1-\omega_1 \\ \sigma \end{bmatrix} \tilde{t}_{C2} \\ &= \begin{bmatrix} (\sigma-1)(1-\omega_1) \\ -[1-\omega_1+\omega_1\sigma] \end{bmatrix} \tilde{t}_{C2}. \end{aligned} \tag{3.51}$$

Now future consumption falls unambiguously (due to a reverse postponement effect), but the effect on present consumption is ambiguous due to offsetting income and substitution effects. (The two effects match for a Cobb-Douglas utility function ($\sigma = 1$) so that $\tilde{C}_1 = 0$ in that case.) It follows from (3.48) that net saving falls (rises) if present consumption rises (falls).

In Figure 3.3 we illustrate the effects for the Cobb-Douglas case ($\sigma = 1$). To avoid cluttering the diagram, the indifference curves have been omitted. The tax increase rotates the budget line in a counter-clockwise fashion from BA to BC. The total effect consists of the move from E_0 to E_1, the substitution effect is the move from E_0 to E', and the income effect is the move from E' to E_1.[7]

3.1.2.3 Intertemporally-neutral increase in the consumption tax

In the third case to be studied, both the current and future consumption tax are increased equiproportionally so that the slope of the wealth expansion path (representing the Euler equation (3.46)) is unaffected. This shock can be seen as an unanticipated and permanent increase in the consumption tax.

[7]The initial indifference curve passes through points E_0 and E'. Note also that for the Leontief case (with $\sigma = 0$) the new equilibrium would be at e_1. The reader can easily verify that for the intermediate case, with $0 < \sigma < 1$, the equilibrium lies somewhere between point e_1 and point E_1.

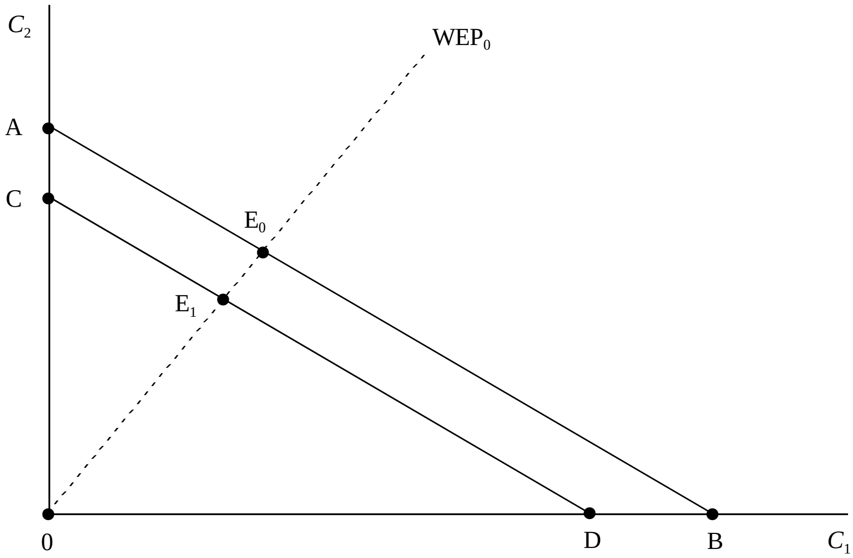

Figure 3.4: Intertemporally neutral increase in the consumption tax

By setting $\tilde{t}_{C1} = \tilde{t}_{C2} = \tilde{t}_C > 0$ in equation (3.47) we find that current and future consumption change according to:

$$\begin{bmatrix} \tilde{C}_1 \\ \tilde{C}_2 \end{bmatrix} = \begin{bmatrix} 1 & -(1-\omega_1) \\ 1 & \omega_1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} \tilde{t}_C = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \tilde{t}_{C1}. \tag{3.52}$$

Not surprisingly, the intertemporal substitution elasticity (σ) does not feature in this expression. The relative price of present consumption vis-a-vis future consumption is unaffected in this case, so the C_2/C_1 ratio is constant regardless of the value of σ. In terms of Figure 3.4, the tax shock shifts the budget line in a parallel fashion from AB to CD and shifts the equilibrium from E_0 to E_1. Regardless of the value of σ, present and future consumption both fall unambiguously (due to the income effect). There is no effect on net saving as can be deduced from the loglinearized versions of (3.48) and (3.49):

$$\frac{dS_1}{\Omega} = -\omega_1 \left[\tilde{C}_1 + \tilde{t}_{C1}\right] = (1-\omega_1) \left[\tilde{C}_2 + \tilde{t}_{C2}\right]. \tag{3.53}$$

Since $\tilde{C}_t + \tilde{t}_{Ct} = \tilde{C}_t + \tilde{t}_C = 0$ (for $t = 1, 2$) it follows from (3.53) that $dS_1 = 0$. Figure 3.4 illustrates the effects for the general case (σ may be anything). (Indifference curves have again been omitted.) The total effect equals the income effect and is given by the move from E_0 to E_1. Note that the substitution effect is absent.

3.2 Endogenous labour supply

Up to this point we have focused attention on the *inter*temporal trade-offs allowed by the Fisherian model, i.e. we have ignored the *intra*temporal substitution possibilities between consumption and leisure by assuming that labour supply is exogenous. As was argued in the previous chapter, however, one of the crucial aspects of many kinds of taxes is that they affect the labour supply decision by households. In this subsection we extend the Fisherian model by endogenizing the labour supply decision of the representative household.[8]

In order to endogenize labour supply, the general lifetime utility function (3.1) is replaced by:

$$\Lambda(\cdot) = U(C_1, \bar{L} - L_1) + \frac{1}{1+\rho} U(C_2, \bar{L} - L_2), \tag{3.54}$$

where $U(\cdot)$ is the instantaneous utility function, $\rho > 0$ is the pure rate of time preference, C_t and L_t are, respectively, consumption and labour supply in period t (for $t = 1, 2$), and $\bar{L}$ is the exogenous time endowment. Leisure consumption in period t is thus given by $\bar{L} - L_t$. Note that the utility function (3.54), like (3.40) above, is quite special in that it is *intertemporally additively separable*.

Choosing the price of the consumption good, P_t, as the numeraire, adding labour income taxes, and imposing the solvency condition (3.7), the periodic budget constraints can be written as:

$$a_1 = (1 + r_0)a_0 + (1 - t_{L1})w_1 L_1 - C_1 \tag{3.55}$$

$$0 = (1 + r_1)a_1 + (1 - t_{L2})w_2 L_2 - C_2, \tag{3.56}$$

where r_t, a_t, and w_t are respectively, the real interest rate, real financial assets, and the real wage in period t. The (linear) labour income tax rates are t_{L1} and t_{L2}. Under the assumption that the household can freely borrow or lend at the going interest rate, the periodic constraints (3.55)-(3.56) can be consolidated into one lifetime budget constraint:

$$C_1 + w_1^* (\bar{L} - L_1) + \frac{C_2 + w_2^* (\bar{L} - L_2)}{1 + r_1} = (1 + r_0)a_0 + h_0 \equiv \Omega, \tag{3.57}$$

where $w_t^* \equiv (1 - t_{Lt}) w_t$ is the after-tax real wage rate in period t, and h_0 is net *human wealth* representing the after-tax value of the time endowment:

$$h_0 \equiv w_1^* \bar{L} + \frac{w_2^* \bar{L}}{1 + r_1}. \tag{3.58}$$

Equations (3.57) and (3.58) generalize, respectively, (3.9) and (3.10) for the case of endogenous labour supply and non-zero labour income taxes.

The life-time budget constraint (3.57) says that for the solvent household the present value of spending on goods *and* leisure (left-hand side) equals initial total wealth (right-hand side). The prices of C_1, $\bar{L} - L_1$, C_2, and $\bar{L} - L_2$ are, respectively, **1**, w_1^*, $\frac{1}{1+r_1}$, and $\frac{w_2^*}{1+r_1}$. Once a particular functional form has been chosen for the felicity function, $U(\cdot)$, the model can be solved by means of *two-stage budgeting* (see the

[8] This approach was pioneered in the late 1960s by Lucas and Rapping (1969). They used it for macroeconomic purposes, however, and in doing so placed the intertemporal substitution mechanism in labour supply at the core of modern real business cycle theory.

Intermezzo).

For our tax policy analysis, we make use of the often-used felicity function:

$$U(F_t) \equiv \frac{F_t^{1-1/\sigma} - 1}{1 - 1/\sigma}, \tag{3.59}$$

$$F_t \equiv \left[\varepsilon (C_t)^{(\eta-1)/\eta} + (1-\varepsilon)(\bar{L} - L_t)^{(\eta-1)/\eta}\right]^{\eta/(\eta-1)}, \tag{3.60}$$

where F_t is sub-felicity, which itself depends on consumption, C_t, and leisure, $\bar{L} - L_t$. This specification is a so-called *nested* utility structure featuring constant elasticity of substitution (CES) functions. In the top-level function (3.59), σ represents the *intertemporal* substitution elasticity, and in the bottom-level function (3.60), η is the *intratemporal* substitution elasticity between consumption and labour.

The household chooses C_t and L_t (and thus F_t) for $t = 1, 2$ in order to maximize lifetime utility (3.54) subject to the lifetime budget constraint (3.57). Using the method of two-stage budgeting we find the following solutions, conditional on the level of *full consumption*, X_t:

$$C_t = \omega_{Ct} X_t, \tag{3.61}$$

$$w_t^* (\bar{L} - L_t) = (1 - \omega_{Ct}) X_t, \tag{3.62}$$

where full consumption, X_t, and the consumption share, ω_{Ct}, are defined as follows:

$$X_t \equiv C_t + w_t^* (\bar{L} - L_t), \tag{3.63}$$

$$\omega_{Ct} \equiv \frac{\varepsilon^\eta}{\varepsilon^\eta + (1-\varepsilon)^\eta (w_t^*)^{1-\eta}}. \tag{3.64}$$

The dynamic part of the solution consists of:

$$\frac{F_2}{F_1} = \left(\frac{P_{F2}}{P_{F1}}\right)^{-\sigma} \left(\frac{1+r_1}{1+\rho}\right)^{\sigma}, \tag{3.65}$$

$$\frac{X_2}{X_1} \equiv \frac{P_{F2} F_2}{P_{F1} F_1} = \left(\frac{P_{F2}}{P_{F1}}\right)^{1-\sigma} \left(\frac{1+r_1}{1+\rho}\right)^{\sigma}, \tag{3.66}$$

where P_{Ft} is the *true price index* linking F_t and X_t:

$$P_{Ft} \equiv \begin{cases} \left[\varepsilon^\eta + (1-\varepsilon)^\eta (w_t^*)^{1-\eta}\right]^{1/(1-\eta)} & \text{for } \eta \neq 1 \\ \left(\frac{1}{\varepsilon}\right)^{\varepsilon} \left(\frac{w_t^*}{1-\varepsilon}\right)^{1-\varepsilon} & \text{for } \eta = 1 \end{cases}. \tag{3.67}$$

Finally, the closed-form solution for current full consumption is:

$$X_1 = \xi_1 \Omega, \tag{3.68}$$

where Ω is defined in (3.57) above and ξ_1 represents the marginal propensity to consume out of total wealth:

$$\xi_1 \equiv \begin{cases} \left[1 + \left(\frac{1}{1+\rho}\right)^{\sigma} \left(\frac{P_{F1}(1+r_1)}{P_{F2}}\right)^{\sigma-1}\right]^{-1} & \text{for } \sigma \neq 1 \\ \frac{1+\rho}{2+\rho} & \text{for } \sigma = 1 \end{cases} . \tag{3.69}$$

The key thing to note about these expressions is the different types of substitution margins that can be affected by taxes. For example, the labour income tax affects both static and dynamic trade-offs:

(a) It affects the static allocation of a given amount of full consumption, X_t, over goods consumption, C_t, and the consumption of leisure, $(\bar{L} - L_t)$—see equations (3.61)-(3.62) and (3.64). The *intra*temporal substitution elasticity η matters here; and

(b) It affects the true cost-of-living index, P_{Ft}, and thus also the dynamic choices regarding F_2/F_1 and X_2/X_1. Both η and the *inter*temporal substitution elasticity σ matter here.

Not surprisingly, the effects of labour taxes are quite complex. In general, a tax change will have income effects, substitution effects, and human wealth effects. In principle, the same linearization technique can be used as in Subsection 3.1.2 above. This is left as an exercise to the interested reader.

Intermezzo 3.1

Two-stage budgeting. The method of two-stage budgeting is a useful technique both for theoretical and empirical work. The basic idea is to split up a dynamic optimization problem into a static part (which is easy to solve) and a dynamic part (which is almost as easy to solve). In the static part, a given amount of *full consumption* is allocated optimally over its components. In the dynamic part, the optimal time path for full consumption itself is chosen.

The two-stage budgeting procedure is valid provided (i) preferences are intertemporally separable and (ii) the felicity function is homothetic. This Intermezzo shows the details of the derivations leading to equations (3.61)-(3.69). The utility function is given in (3.59)-(3.60).

We define full consumption as total spending on goods and leisure in a period:

$$X_t \equiv \mathbf{1}C_t + w_t^* \left(\bar{L} - L_t\right), \tag{I.1}$$

and write the budget constraints (3.55)-(3.56) as follows:

$$a_1 = (1 + r_0)\, a_0 + w_1^* \bar{L} - X_1 \tag{I.2}$$

$$0 = (1 + r_1)\, a_1 + w_2^* \bar{L} - X_2, \tag{I.3}$$

where we recall that $w_t^* \equiv (1 - t_{Lt}) w_t$ is the after-tax real wage rate in period t. (Note that we include the real price of consumption, **1**, in order to indicate at which places it features in the derivations.)

Stage 1: Optimal static choice. In the **first stage**, the household chooses consumption, C_t, and leisure, $\bar{L} - L_t$, in order to maximize sub-felicity, F_t, given the constraint (I.1), and holding constant full consumption, X_t. The Lagrangian expression for this problem is:

$$\mathcal{L}_1 \equiv \left[\varepsilon (C_t)^{(\eta-1)/\eta} + (1-\varepsilon)\left(\bar{L} - L_t\right)^{(\eta-1)/\eta}\right]^{\eta/(\eta-1)} + \mu_t \left[X_t - C_t - w_t^* \left(\bar{L} - L_t\right)\right],$$

where μ_t is the Lagrange multiplier. The first-order necessary conditions (FONCs) are given by the constraint as well as:

$$\frac{\partial \mathcal{L}_1}{\partial C_t} = \varepsilon \left(\frac{F_t}{C_t}\right)^{1/\eta} - \mu_t \mathbf{1} = 0, \tag{I.4}$$

$$\frac{\partial \mathcal{L}_1}{\partial \left(\bar{L} - L_t\right)} = (1-\varepsilon)\left(\frac{F_t}{\bar{L} - L_t}\right)^{1/\eta} - \mu_t w_t^* = 0. \tag{I.5}$$

We can eliminate μ_t by combining the two FONCs and obtain:

$$\frac{(1-\varepsilon) / \left(\bar{L} - L_t\right)^{1/\eta}}{\varepsilon / C_t^{1/\eta}} = \frac{w_t^*}{\mathbf{1}}. \tag{I.6}$$

In the optimal static choice, the marginal rate of substitution (MRS) between leisure and consumption is equated to the relative price of leisure (i.e. the after-tax real wage rate).

By combining (I.6) and the 'budget constraint' (I.1) we find the solutions for consumption and leisure conditional on full consumption:

$$\mathbf{1} C_t = \omega_{Ct} X_t, \tag{I.7}$$

$$w_t^* \left(\bar{L} - L_t\right) = (1 - \omega_{Ct}) X_t, \tag{I.8}$$

where ω_{Ct} is a complicated function of the parameters and the after-tax wage rate:

$$\omega_{Ct} \equiv \frac{\varepsilon^\eta}{\varepsilon^\eta \mathbf{1}^{1-\eta} + (1-\varepsilon)^\eta \left(w_t^*\right)^{1-\eta}}. \tag{I.9}$$

Several features are worth noting in (I.9). First, for the often-used Cobb-Douglas subfelicity function ($\eta = 1$) the spending shares on consumption and leisure are constant (i.e. $\omega_{Ct} = \varepsilon$ in that case). Second, in the general CES case, the following elasticity can be derived from

(I.9):

$$\frac{\partial \omega_{Ct}}{\partial w_t^*} \frac{w_t^*}{\omega_{Ct}} = (\eta - 1)(1 - \omega_{Ct}). \tag{I.10}$$

Since $0 < \omega_{Ct} < 1$ this expression implies that an increase in w_t^* leads to an increase (decrease) in the spending share of consumption if the household finds it easy (difficult) to substitute consumption for leisure, i.e. if $\eta > 1$ ($\eta < 1$).

Armed with the expressions (I.7)-(I.8) we can deduce the true price index linking X_t and F_t. First, we postulate the link by writing:

$$P_{Ft} F_t = X_t, \tag{I.11}$$

where P_{Ft} is the (yet unknown) cost of living index. By substituting (I.7)-(I.8) into the subfelicity function (3.60) we obtain:

$$\begin{aligned} F_t^{(\eta-1)/\eta} &\equiv \varepsilon (C_t)^{(\eta-1)/\eta} + (1-\varepsilon)(\bar{L} - L_t)^{(\eta-1)/\eta} \\ &= \varepsilon \left(\frac{\omega_{Ct} X_t}{\mathbf{1}} \right)^{(\eta-1)/\eta} + (1-\varepsilon) \left(\frac{(1-\omega_{Ct}) X_t}{w_t^*} \right)^{(\eta-1)/\eta} \quad \Leftrightarrow \\ \frac{F_t}{X_t} &= \left[\varepsilon \left(\frac{\omega_{Ct}}{\mathbf{1}} \right)^{(\eta-1)/\eta} + (1-\varepsilon) \left(\frac{1-\omega_{Ct}}{w_t^*} \right)^{(\eta-1)/\eta} \right]^{\eta/(\eta-1)}. \end{aligned} \tag{I.12}$$

By using the definition for ω_{Ct} (from (I.9) above), the term in square brackets on the right-hand side of (I.11) can be simplified to:

$$[\cdot] = \left[\varepsilon^\eta \mathbf{1}^{1-\eta} + (1-\varepsilon)^\eta (w_t^*)^{1-\eta} \right]^{1/\eta}. \tag{I.13}$$

By using (I.13) in (I.12) and comparing the result with (I.11) we find the expression for P_{Ft}:

$$P_{Ft} = \left[\varepsilon^\eta \mathbf{1}^{1-\eta} + (1-\varepsilon)^\eta (w_t^*)^{1-\eta} \right]^{1/(1-\eta)}, \tag{I.14}$$

where $\mathbf{1}$ and w_t^* represent the price of, respectively, consumption and leisure. Of course, expression (I.11) is really an *expenditure function*, that is we can write $E(\mathbf{1}, w_t^*, F_t) = P_{Ft} F_t$ and recover the Hicksian demands for consumption and leisure in the usual fashion by applying Shephard's Lemma. Similarly, the *indirect utility function* can be recovered by writing $V(\mathbf{1}, w_t^*, X_t) = X_t / P_{Ft}$ from which we can recover the Marshallian demands by using Roy's Identity.

Stage 2: Optimal dynamic choice. In the **second stage** the household chooses full consumption in the two periods (X_1 and X_2) in order to maximize lifetime utility $\Lambda(\cdot)$ subject to the budget constraints. Formally, for the general CES case, the household maximizes:

$$\Lambda(\cdot) \equiv \frac{(X_1/P_{F1})^{1-1/\sigma}-1}{1-1/\sigma} + \frac{1}{1+\rho}\frac{(X_2/P_{F2})^{1-1/\sigma}-1}{1-1/\sigma},$$

subject to (I.2)-(I.3). The Lagrangian expression for this problem is:

$$\begin{aligned}\mathcal{L}_2 \equiv\ & \frac{(X_1/P_{F1})^{1-1/\sigma}-1}{1-1/\sigma} + \frac{1}{1+\rho}\frac{(X_2/P_{F2})^{1-1/\sigma}-1}{1-1/\sigma} \\ & + \lambda_1 \left[a_1 - (1+r_0)\,a_0 - \bar{w}_1 \bar{L} + X_1\right] \\ & + \lambda_2 \left[0 - (1+r_1)\,a_1 - \bar{w}_2 \bar{L} + X_2\right],\end{aligned}$$

where λ_1 and λ_2 are the Lagrange multipliers for the budget constraints in the two periods. The first-order necessary conditions again consist of the two constraints as well as:

$$\frac{\partial \mathcal{L}_2}{\partial X_1} = (P_{F1})^{1/\sigma-1}\frac{1}{X_1^{1/\sigma}} - \lambda_1 = 0, \tag{I.15}$$

$$\frac{\partial \mathcal{L}_2}{\partial X_2} = \frac{(P_{F2})^{1/\sigma-1}}{1+\rho}\frac{1}{X_2^{1/\sigma}} - \lambda_2 = 0, \tag{I.16}$$

$$\frac{\partial \mathcal{L}_2}{\partial a_1} = \lambda_1 - \lambda_2\,(1+r_1) = 0. \tag{I.17}$$

By substituting (I.15)-(I.16) into (I.17) we can eliminate λ_1 and λ_2 in order to derive:

$$\frac{(P_{F1})^{1/\sigma-1}/X_1^{1/\sigma}}{(P_{F2})^{1/\sigma-1}/[(1+\rho)\,X_2^{1/\sigma}]} = 1 + r_1. \tag{I.18}$$

In the optimal dynamic choice, the household equates the MRS between present and future full consumption to the relative price of present full consumption (i.e. the gross interest rate). This is the key dynamic condition characterizing the intertemporal optimum. Note that we obtain the Euler equation for full consumption by rewriting (I.18):

$$\frac{X_2}{X_1} = \left(\frac{1+r_1}{1+\rho}\right)^{\sigma}\left(\frac{P_{F2}}{P_{F1}}\right)^{1-\sigma}. \tag{I.19}$$

Of course, (I.19) only says something about the relative amount of full consumption in the two periods. In order to obtain the level solutions for X_1 and X_2, we note that the consolidated lifetime budget constraint is given by:

$$X_1 + \frac{X_2}{1+r_1} = (1+r_0)\,a_0 + h_0 \equiv \Omega, \tag{I.20}$$

where Ω and h_0 are, respectively, total wealth and full human wealth:

$$h_0 \equiv w_1^* \bar{L} + \frac{w_2^* \bar{L}}{1+r_1}. \tag{I.21}$$

By combining (I.18) (or equivalently (I.19)) with (I.20) we find the full consumption levels after some straightforward steps:

$$X_1 = \xi_1 \Omega, \tag{I.22}$$

$$\frac{X_2}{1+r_1} = (1-\xi_1)\,\Omega, \tag{I.23}$$

$$1/\xi_1 \equiv 1 + \left(\frac{1}{1+\rho}\right)^{\sigma} \left(\frac{P_{F1}\,(1+r_1)}{P_{F2}}\right)^{\sigma-1}. \tag{I.24}$$

Note that for the logarithmic felicity function (for which $\sigma = 1$), $\xi_1 = (1+\rho)\,/\,(2+\rho)$, i.e. the household dedicates a constant fraction of total wealth to full consumption in the present period. Neither the interest rate nor the true price index affects this proportion in that case.

We have now completely characterized the optimal solution for the household optimization problem. Indeed, equations (I.22)-(I.24) determine X_1 and X_2, whilst (I.7)-(I.9) determine C_t and $\bar{L} - L_t$ for $t = 1, 2$. We are done!

****

3.3 Extensions to the two-period model

In this section we develop some extensions to the basic two-period household model. In the first subsection we study the human capital investment decision and its reaction to a labour income tax and an interest income tax. In the second subsection we look at imperfections in the capital market and in the third subsection we allow for endogenous bequests. To keep things simple we restrict attention to the case of exogenous labour supply in the last two subsections.

3.3.1 Human capital accumulation

In the basic two-period model, the household can transfer resources across time by lending or borrowing in the first period. The savings instrument consists of a single financial asset, a_1, carrying a given interest rate. In this subsection we extend the model by assuming that the household has an additional instrument by which it can achieve the optimal life-cycle consumption path, namely investment in its own *human capital*. Human capital is embodied in the household itself and thus cannot be borrowed—it must be built up by the household itself by means of time-consuming educational and training efforts.

Intuitively, human capital incorporates things like the level of the household's sophistication and technical skills which are valuable in the market.

We study the human capital investment decision in a simple model which makes use of the insights of Eaton and Rosen (1980b) and Heckman (1976). The household lifetime utility function is given by:

$$\Lambda = U\left(C_1, C_2, \bar{L} - L_1 - I_1\right), \tag{3.70}$$

where I_1 is time spent in the first period on human capital formation. Since L_1 is labour supply in the first period and $\bar{L}$ is the time endowment, leisure is equal to $\bar{L} - L_1 - I_1$. Educational activities are costly because, for a given amount of labour supply, they reduce the amount of leisure that can be consumed. These activities do not yield direct utility themselves, i.e. it is not fun to read difficult textbooks or go to school in this model. To keep the model as simple as possible, it is assumed that future leisure does not feature in the utility function. This implies that the household supplies $\bar{L}$ units of "raw" labour in the second period (There is no human capital investment in the second period, $I_2 = 0$, because there is no third period, so $L_2 = \bar{L}$.).

At the beginning of the first period, the household possesses an exogenously given amount of human capital, H_1. This could, for example, include the basic skills that any human possesses. By engaging in education, the human capital stock in the second period is augmented according to:

$$H_2 = G\left(I_1\right) H_1, \tag{3.71}$$

where $G\left(\cdot\right)$ is the human capital production function representing the training technology ($G\left(\cdot\right) > 0$ for $I_1 \geq 0$). It is assumed that $G\left(\cdot\right)$ features positive but diminishing marginal productivity of training hours, i.e. $G' > 0 > G''$. To rule out a zero-training corner solution, it is assumed that $\lim_{I_1 \to 0} G'\left(I_1\right) = \infty$. Human capital is valuable in the market so if the household possesses H_t units of human capital and works L_t raw hours, then the labour effort in *efficiency units* is $H_t L_t$ and gross wage income is $w_t H_t L_t$ (for $t = 1, 2$).

Ignoring initial financial assets ($a_0 = 0$), the budget constraints can be written as follows:

$$a_1 = \left(1 - t_{L1}\right) w_1 H_1 L_1 - C_1, \tag{3.72}$$

$$0 = \left[1 + r_1\left(1 - t_R\right)\right] a_1 + \left(1 - t_{L2}\right) w_2 H_2 \bar{L} - C_2, \tag{3.73}$$

where t_R is a proportional tax on interest income and t_{Lt} is the labour income tax. The lifetime budget constraint is thus given by:

$$C_1 + \frac{C_2}{1 + r_1\left(1 - t_R\right)} = \left(1 - t_{L1}\right) w_1 H_1 L_1 + \frac{\left(1 - t_{L2}\right) w_2 H_2 \bar{L}}{1 + r_1\left(1 - t_R\right)}. \tag{3.74}$$

The household chooses C_1, C_2, L_1, and I_1 in order to maximize lifetime utility (3.70) subject to the training technology (3.71) and the lifetime budget constraint (3.74). The first-order conditions associated with an interior solution to this maximization problem are:

$$\frac{U_1}{U_2} = 1 + r_1 (1 - t_R), \tag{3.75}$$

$$\frac{U_3}{U_1} = (1 - t_{L1}) w_1 H_1, \tag{3.76}$$

$$\frac{U_3}{U_1} = \frac{(1 - t_{L2}) w_2 G' (I_1) H_1}{1 + r_1 (1 - t_R)}, \tag{3.77}$$

where $U_t \equiv \partial U / \partial C_t$ is the marginal utility of consumption in period t (for $t = 1, 2$) and $U_3 \equiv \partial U / \partial (\bar{L} - L_1 - I_1)$ is the marginal utility of current leisure. The first two equations are familiar from the previous discussion. Equation (3.75) is the consumption Euler equation calling for an equalization of the marginal rate of substitution between present and future consumption to the after-tax interest factor. Similarly, equation (3.76) is the condition for optimal labour supply in the first period. It equates the marginal rate of substitution between present consumption and leisure to the after-tax wage rate.

Equation (3.77) determines the optimal amount of human capital investment. It calls for an equalization of the marginal rate of substitution between present consumption and leisure (left-hand side) to the present value of future additional after-tax wage income (right-hand side). By substituting (3.76) into (3.77) we obtain a simple condition characterizing the optimal investment decision:

$$G'(I_1) = \frac{1 - t_{L1}}{1 - t_{L2}} \frac{w_1}{w_2} [1 + r_1 (1 - t_R)]. \tag{3.78}$$

Several things are worth noting about this expression. First, the initial stock of human capital, H_1, does not affect the optimal investment decision since it appears on the right-hand side of both (3.76) and (3.77). Second, optimal investment does not depend on the utility function because working time and training time are perfect substitutes and labour supply in the second period is exogenous. Third, an increase in the current labour income tax or the interest income tax, and a decrease in the future labour income tax, all lead to a reduction in $G'(I_1)$, that is an increase in the optimal training effort. Whilst the effects of t_{L1} and t_{L2} should be obvious, the effect of the interest income tax warrants some further comment. Intuitively, as Eaton and Rosen (1980, p. 707) point out, an increase in t_R reduces the cost of borrowing (since interest payments are tax deductible) and allows the household to work less (and train more) in the current period. Fourth, if the labour income tax is time-invariant ($t_{L1} = t_{L2}$) then it drops out of (3.78) altogether. It follows that a general income tax, with $t_{L1} = t_{L2} = t_R$, has a positive effect on optimal human capital investment.

3.3.2 Borrowing constraints

As was pointed out by Sandmo (1985), the two-period model suffers from at least two potentially serious shortcomings. First, the model is based on the assumption that interest rates for lending and borrowing

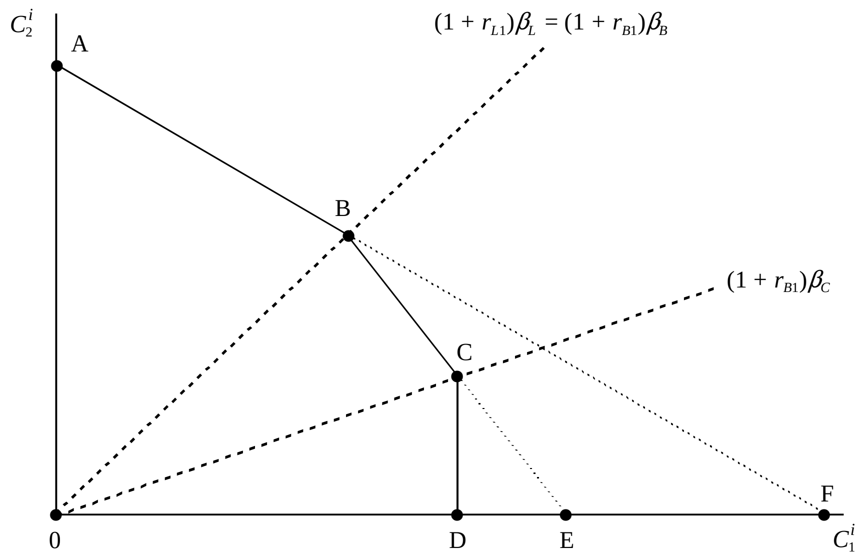

Figure 3.5: Capital market constraints and the choice set

are the same. Second, it is assumed that the household faces no quantity constraints on borrowing. In reality both assumptions are not likely to hold (for most households).

The consequences of differential lending and borrowing interest rates can easily be studied in the two-period model. Assume that the borrowing rate in period 1 is r_{B1}, the lending rate is r_{L1}, and let $r_{B1} > r_{L1}$. Then, ignoring initial assets at the beginning of period 1 ($a_0 = 0$), the two budget identities for household i are given by:

$$a_1^i = w_1^* \bar{L} - C_1^i, \tag{3.79}$$

$$C_2^i = \begin{cases} (1 + r_{L1})\, a_1^i + w_2^* \bar{L} & \text{for } a_1^i > 0 \\ \\ (1 + r_{B1})\, a_1^i + w_2^* \bar{L} & \text{for } a_1^i \leq 0 \end{cases}, \tag{3.80}$$

where $w_t^* \equiv w_t \,(1 - t_{Lt})$ is the after-tax wage rate in period t, assumed to be the same for all households. According to (3.80), the household receives interest payments $r_{L1} a_1^i$ if it saves in the first period ($a_1^i > 0$), but it must pay interest on loans equal to $r_{B1} a_1^i$ if it decides to borrow in the first period. Because $r_{B1} > r_{L1}$, the life-time budget constraint features a kink at the point where a_1^i changes sign. In terms of Figure 3.5, the kink is at point B and the choice set shrinks from 0AF (if $r_{L1} = r_{B1}$) to 0ABE.

If there is, in addition, a constraint on the maximum amount that can be borrowed, say $a_1^i \geq -a_{MAX}$ (with $a_{MAX} > 0$), then the lifetime budget constraint features an additional kink to the right of point B; say at point C in Figure 3.5. As a result of this credit constraint the choice set shrinks even further, say from 0ABE to 0ABCD.

As a consequence of the two types of capital market imperfections, it is likely that for many households the optimum (restricted) choice will be at one of the two kinks. For such households, many substitution effects due to changes in tax rates will no longer be relevant. Econometric research aimed at quantifying the effects of taxes on consumption, saving, and labour supply must take such features of the capital market into account—see also Section 3.4 below.

To illustrate some of these points, consider the following simple model in which households differ only in their rate of time preference. The lifetime utility function of household i is logarithmic:

$$\Lambda^i = \ln C_1^i + \beta_i \ln C_2^i, \tag{3.81}$$

where β_i is the discount factor due to pure time preference ($\beta_i \geq 0$). In this formulation, patient households feature a relatively high value of β_i whilst impatient households have a low β_i. Household i chooses C_1^i, C_2^i, and a_1^i in order to maximize Λ^i subject to the budget constraints (3.79)-(3.80).

A household which lends in the first period (a net saver) faces the lifetime budget constraint:

$$C_2^i = (1 + r_{L1}) \left(w_1^* \bar{L} - C_1^i \right) + w_2^* \bar{L}, \tag{3.82}$$

and chooses the consumption profile according to:

$$\frac{C_2^i}{C_1^i} = \beta_i (1 + r_{L1}), \qquad \text{(for } a_1^i > 0\text{)}. \tag{3.83}$$

At point B in Figure 3.5, $a_1^i = 0$, and the consumption profile is given by:

$$\frac{C_2^i}{C_1^i} = \frac{w_2^*}{w_1^*}. \tag{3.84}$$

Since (by assumption) after-tax wages are the same for all households, the kink is at the same location for all households. Saving households are located somewhere along the line segment AB in Figure 3.5. It thus follows from the comparison of (3.83) and (3.84) that:

$$a_1^i > 0 \qquad \Leftrightarrow \qquad \beta_i > \frac{w_2^*}{w_1^*} \frac{1}{1 + r_{L1}} \equiv \beta_L. \tag{3.85}$$

Any household whose β_i exceeds the critical lending value β_L will end up saving in the first period.

People who borrow in the first period (but are not credit constrained) face the lifetime budget constraint:

$$C_2^i = (1 + r_{B1}) \left(w_1^* \bar{L} - C_1^i \right) + w_2^* \bar{L}, \tag{3.86}$$

and choose the consumption profile:

$$\frac{C_2^i}{C_1^i} = \beta_i (1 + r_{B1}), \qquad \text{(for } -a_{MAX} < a_1^i < 0\text{)}. \tag{3.87}$$

Such households consume somewhere along the line segment BC in Figure 3.5. The comparison between (3.87) and (3.84) furthermore reveals that:

$$-a_{MAX} < a_1^i < 0 \qquad \Leftrightarrow \qquad \beta_i < \frac{w_2^*}{w_1^*} \frac{1}{1+r_{B1}} \equiv \beta_B. \tag{3.88}$$

Note that both (3.85) and (3.88) describe *interior solutions* for which $a_1^i \neq 0$. But since the borrowing rate is higher than the lending rate ($r_{B1} > r_{L1}$), it follows that β_L is larger than β_B, i.e. there are potentially many households who neither lend nor borrow (see also below). Such households have a time preference parameter, β_i, such that:

$$a_1^i = 0 \quad \Leftrightarrow \qquad \beta_B \leq \beta_i \leq \beta_L. \tag{3.89}$$

It remains to be determined what happens at the second kink (at point C) in Figure 3.5. At that point borrowers face a binding credit constraint, $a_1^i = -a_{MAX}$, and the consumption profile is given by:

$$\frac{C_2^i}{C_1^i} = \frac{w_2^* - (1+r_{B1})\, a_{MAX}/\bar{L}}{w_1^* + a_{MAX}/\bar{L}} \equiv \beta_C \,(1+r_{B1})\,. \tag{3.90}$$

Hence, any household whose β_i falls short of the critical value at which the credit constraint becomes binding will borrow up to the hilt and consume at point C:

$$a_1^i = -a_{MAX} \qquad \Leftrightarrow \qquad \beta_i < \beta_C. \tag{3.91}$$

In Figure 3.6 we visualize the different cases described by the model. It is assumed that the β_i parameters are distributed across the population according to the density function $f(\beta_i)$. The cumulative density function is defined as:

$$F(x) \equiv \int_0^x f(\beta_i)\, d\beta_i, \tag{3.92}$$

and has the usual properties $F(0) = 0$, $\lim_{x\to\infty} F(x) = 1$, and $F'(x) \equiv f(x)$. Intuitively, $F(x)$ is the area under the density function from 0 to x and it represents the fraction of the population for which $0 < \beta_i < x$. In Figure 3.6 four different population groups can be distinguished:

1. A fraction $F(\beta_C)$ of the population consists of credit constrained borrowers ($a_1^i = -a_{MAX}$);
2. A fraction $F(\beta_B) - F(\beta_C)$ of the population consists of borrowers who do not face a credit constraint ($-a_{MAX} < a_1^i < 0$);
3. A fraction $F(\beta_L) - F(\beta_B)$ of the population consists of households who neither borrow nor lend ($a_1^i = 0$);
4. A fraction $1 - F(\beta_L)$ of the population consists of households who are net savers ($a_1^i > 0$).

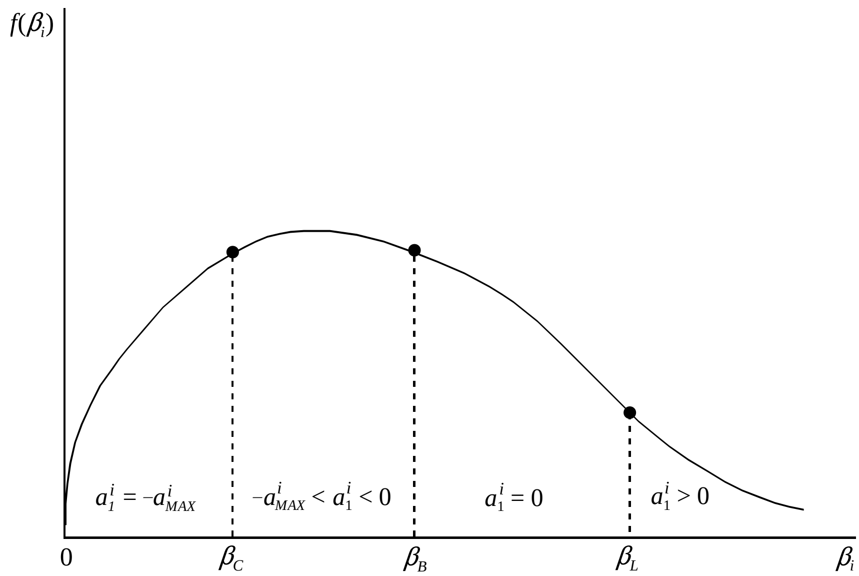

Figure 3.6: Capital market constraints and household patience

Provided a tax change does not shift them into another group, households in groups 2 and 4 behave as in the basic Fisherian model. Matters are drastically different, however, for households in groups 1 and 3. Consider the following example. The government taxes interest *receipts* so that the after-tax lending rate is $r_{L1}(1-t_R)$ and β_L is given by:

$$\beta_L \equiv \frac{w_2^*}{w_1^*}\frac{1}{1+r_{L1}(1-t_R)}. \tag{3.93}$$

Interest payments are not tax deductible, the borrowing rate continues to be r_B, and both β_B and β_C are unaffected. Next consider the effect of an increase in the interest income tax rate, t_R, on a typical non-saving household (in group 3). In view of (3.93), β_L will increase as a result of the tax increase:

$$\frac{\partial \beta_L}{\partial t_{L2}} \equiv \frac{r_{L1}\beta_L}{1+r_{L1}(1-t_R)} > 0. \tag{3.94}$$

In terms of Figure 3.6, β_L shifts to the right and some marginal savers turn into non-savers. Since nothing happens to β_B (see (3.88)), the population fraction of non-saving households is increased. As non-saving households simply consume their after-tax labour income in each period ($C_t^i = w_t^*\bar{L}$), it follows that the interest income tax increase has no effect at all on current and future consumption. The usual substitution effects associated with a change in the lending rate are simply not relevant for this group of households. (Of course, since β_C is also unaffected by the change in t_R, the behaviour of households in group 1 is also unchanged.)

3.3.3 Bequests

The third extension to the basic two-period consumption model recognizes the possibility that finitely-lived households may be linked to (certain members of) future generations. The typical example that comes to mind is that parents love their offspring, i.e. the child's welfare (or income) level features in the parents' utility function. We say that there exists one-sided or *unidirectional altruism* that runs from the parents to the offspring (but not vice versa).[9] We show how the basic consumption model can be extended to allow for intergenerational linkages operating via "altruistic" bequests. With this extended model it is possible to study the impact of inheritance taxation on consumption and savings decisions by households. The basic insights that we employ are due to Barro (1974), who used the model to study the role of public debt in an economy.

In the altruistic consumption model we continue to assume that all households supply a fixed amount of labour but we distinguish between two types of agents, namely "parents" and "children." To keep notation simple, we assume that each agent gets one child. Each agent lives for two periods, namely youth (superscript "Y") and old-age (superscript "O"). At the start of his life, the agent may receive a transfer (or bequest) from his parent. At the end of youth, the agent gets one child to whom it may in turn leave a bequest. Restricting attention to the case of unidirectional altruism, the lifetime utility function as of time period t of a young household is given by:

$$\Lambda_t^Y(\cdot) = U\left(C_t^Y, C_{t+1}^O\right) + \frac{1}{1+\delta}\Lambda_{t+1}^Y(\cdot), \quad \delta > 0, \tag{3.95}$$

where Λ_t^Y is lifetime utility of a young household at time t (the "parent"), C_t^Y is consumption by a young household in period t, C_t^O is consumption by an old household in period t, and $U(\cdot)$ is the non-altruistic part of the utility function. Incorporated in (3.95) is the notion that a household derives utility not only from *own* goods consumption (C_t^Y and C_{t+1}^O) but also from the utility level of a future young agent (the "child" whose welfare is denoted by $\Lambda_{t+1}^Y(\cdot)$). The parameter δ is assumed to be positive and it regulates the severity of the altruistic effect in the utility function. If δ is very large, then the weight of $\Lambda_{t+1}^Y(\cdot)$ in the parent's utility function is close to zero, and the altruistic effect is very weak. Conversely, if δ is close to zero, then the altruistic effect is very strong.[10]

The budget identities of a young agent at time t are given by:

$$C_t^Y + a_t + b_t = (1 + r_{t-1})b_{t-1} + w_t\bar{L}, \tag{3.96}$$

[9]The more complex case in which an agent cares for both his/her children and parents is referred to as one of two-sided or *bi-directional* altruism. To keep things simple we concentrate on unidirectional altruism in the text. The reader is referred to Kimball (1987) for a study of the model with two-sided altruism.

[10]It must be stressed that δ plays a completely different role than the household's own rate of time preference, ρ. To illustrate this point, assume that $U(C_t^Y, C_{t+1}^O)$ takes an additively separable form as in (3.40):

$$U(C_t^Y, C_{t+1}^O) = U(C_t^Y) + \frac{1}{1+\rho}U(C_{t+1}^O).$$

In this formulation, ρ affects the weight attached to own future felicity. In equation (3.95), δ affects the weight attached to the offspring's lifetime utility.

$$C^O_{t+1} = (1 + r_t)a_t + w_{t+1}\bar{L}, \tag{3.97}$$

where a_t is financial assets (excluding bequests received) at the end of period t, b_{t-1} is the gross bequest this agent received at the beginning of life, b_t is the inheritance left to the agent's child at the end of period t, r_t is the interest rate, w_t is the wage rate, and $\bar{L}$ is exogenous labour supply. Obviously, the child of this agent (who is young in period $t + 1$ and old in period $t + 2$) will face the following budget identities:

$$C^Y_{t+1} + a_{t+1} + b_{t+1} = (1 + r_t)b_t + w_{t+1}\bar{L}, \tag{3.98}$$
$$C^O_{t+2} = (1 + r_{t+1})a_{t+1} + w_{t+2}\bar{L}, \tag{3.99}$$

and similar expressions can be deduced for the "grandchild", the "great-grandchild", etcetera of the original agent. The thing to note about (3.96) and (3.98) is that in each case the offspring receives the interest-inclusive bequest (denoted by $(1 + r_{t-1})\, b_{t-1}$ and $(1 + r_t)b_t$, respectively) at the beginning of its first period of life.

With perfect capital markets, the agent can lend or borrow as he pleases, a_t can have either sign, and the consolidated budget constraint of a young agent at time t is given by:

$$C^Y_t + b_t + \frac{C^O_{t+1}}{1 + r_t} = (1 + r_{t-1})b_{t-1} + h_t, \tag{3.100}$$

where h_t is human wealth, i.e. the present value of lifetime wage income:

$$h_t \equiv w_t\bar{L} + \frac{w_{t+1}\bar{L}}{1 + r_t}. \tag{3.101}$$

By forward iteration of (3.95) we find that the *effective* objective function for the current parent depends on the consumption levels of all present and future generations:

$$\begin{aligned}
\Lambda^Y_t(\cdot) &= U\left(C^Y_t, C^O_{t+1}\right) + \frac{1}{1+\delta}\left[U\left(C^Y_{t+1}, C^O_{t+2}\right) + \frac{1}{1+\delta}\Lambda^Y_{t+2}(\cdot)\right] \\
&= U\left(C^Y_t, C^O_{t+1}\right) + \frac{1}{1+\delta}U\left(C^Y_{t+1}, C^O_{t+2}\right) + \left(\frac{1}{1+\delta}\right)^2 U\left(C^Y_{t+2}, C^O_{t+3}\right) + \ldots \\
&= \sum_{\tau=0}^{\infty}\left(\frac{1}{1+\delta}\right)^{\tau} U\left(C^Y_{t+\tau}, C^O_{t+\tau+1}\right).
\end{aligned} \tag{3.102}$$

The key thing to note about (3.102) is that the altruism parameter δ acts as a discounting factor for future non-altruistic utility functions (i.e. the $U(\cdot)$ functions). Formally, the utility function resembles a utility function of a single infinite-lived representative agent.

In the absence of constraints on bequests,[11] consolidation over the dynastic family can be achieved.

[11] Either we assume (counterfactually) that b_t can have either sign or we postulate that household decisions are such that the non-negativity constraint on bequests ($b_t \geq 0$) is never binding.

Indeed, by iterating (3.100) forward in time we can eliminate b_t, b_{t+1}, b_{t+2}, etcetera and obtain the dynastic budget constraint:

$$C_t^Y + \frac{C_{t+1}^O + C_{t+1}^Y}{1+r_t} + \frac{C_{t+2}^O + C_{t+2}^Y}{(1+r_t)(1+r_{t+1})} + \cdots = (1+r_{t-1})b_{t-1} + h_t + \frac{h_{t+1}}{1+r_t} + \cdots \tag{3.103}$$

The left-hand side of (3.103) represents the present value of consumption by all members of the household dynasty whilst the right-hand side is the present value of resources of all members of the household dynasty.

The dynastic decision problem is a multi-period generalization of our earlier problems. Indeed, the current dynastic head (the agent who is young at time t) effectively chooses a sequence for $C_{t+\tau}^Y$ and $C_{t+\tau+1}^O$ (for $\tau = 0, 1, \cdots, \infty$) in order to maximize the dynastic objective function (3.102) subject to the dynasty budget constraint (3.103). Although the dynastic head will no longer be alive after period $t+1$, he can nevertheless engineer the path of bequests in such a way that the optimal plan will actually be chosen by his descendants.

In closing this subsection, a number of remarks are in place. First, it can be shown that a (wealth) tax on bequests acts like an interest income tax in this framework. Second, if negative bequests are not taxed then the multidimensional choice set of the dynasty is kinked. It is likely that a lot of households will be at the kink (and leave zero bequests) in that case. Third, if the constraint on negative bequests is relevant and $b_t \geq 0$ becomes binding, then the chain connecting generations is broken and estate taxation will not have a substitution effect at all. The interested reader is invited to verify these results.

3.4 Empirical evidence

We have shown in this chapter that, in a dynamic setting, tax changes typically affect consumption, saving, and labour supply along two margins, namely the *intra*temporal (static) margin and the *inter*temporal (dynamic) margin. In the context of our simple two-period models, both the intratemporal substitution elasticity (e.g. our η in Section 3.2) and the intertemporal substitution elasticity (our σ) are thus crucial parameters. It is therefore not surprising that there exists a huge volume of literature attempting to obtain reliable estimates for the various substitution elasticities by means of econometric methods. In the remainder of this section we present a brief and incomplete survey of the available evidence. In the first subsection we discuss the evidence regarding the interest elasticity of household saving. In the second subsection we review the literature on intertemporal substitution effects in labour supply.

3.4.1 Interest elasticity of saving

Before turning to the empirical evidence, it is worth noting that there is no theoretical reason for the interest elasticity of private saving to be positive. This important point can be illustrated with the aid of the basic two-period model with exogenous labour supply discussed in Section 3.1 above. It is not difficult to show that the Slutsky equation for consumption in the two periods can be written as:

$$\frac{\partial C_1^M}{\partial r_1} = \frac{\partial C_1^H}{\partial r_1} + \frac{a_1}{1+r_1}\frac{\partial C_1^M}{\partial \Omega}, \tag{3.104}$$

$$\frac{\partial C_2^M}{\partial r_1} = \frac{\partial C_2^H}{\partial r_1} + a_1\left[1 - \frac{\partial C_1^M}{\partial \Omega}\right], \tag{3.105}$$

where C_t^M and C_t^H are, respectively, the Marshallian and Hicksian demands for C_t ($t = 1,2$). Clearly, the pure "own" substitution effect is positive, i.e. $\partial C_2^H/\partial r_1 > 0$ and thus (since there are only two goods) $\partial C_1^H/\partial r_1 < 0$. In Figure 3.1, the pure substitution effect is the move from E_0 to E'. The second terms on the right-hand sides of (3.104)-(3.105) represent the wealth effects, i.e. the combination of the income effect and the human wealth effect (IE and HWE in Figure 3.1). If both goods are normal (so that $0 < \partial C_1^M/\partial\Omega < 0$), it follows from (3.104)-(3.105) that the wealth effect is positive (negative) for a household that saves (borrows) in the first period. Not surprisingly, therefore, the uncompensated effects of a change in the interest rate are ambiguous. In the absence of future labour income and with homothetic preferences, we find from (3.24)-(3.25) that $\partial C_1^M/\partial r_1 \gtreqless 0$ for $\sigma \lesseqgtr 1$ and $\partial C_2^M/\partial r_1 > 0$ for all $\sigma \geq 0$.

The ambiguity of the uncompensated first-period consumption effect carries over into the uncompensated savings effect. Indeed, by using (3.4) and (3.104) we find:

$$\frac{\partial a_1^M}{\partial r_1}\left[= -\frac{\partial C_1^M}{\partial r_1}\right] = -\frac{\partial C_1^H}{\partial r_1} - \frac{a_1}{1+r_1}\frac{\partial C_1^M}{\partial \Omega}. \tag{3.106}$$

The first term on the right-hand side is positive but the second term is ambiguous in general. In the Cobb-Douglas case without future labour income the savings elasticity is zero. Interestingly, as can be seen from (3.24), the human wealth effect itself leads to a decrease in $\partial C_1^M/\partial r_1$ and thus an increase in $\partial a_1^M/\partial r_1$. This point was stressed by Summers (1981a) in the context of a calibrated general equilibrium model. The human wealth effect can ensure a positive interest elasticity of saving even in the case of Leontief preferences (if $\sigma = 0$).[12]

The empirical literature on the interest elasticity of private saving was recently surveyed by Elmendorf (1996) and Bernheim (2002). During the sixties and seventies, the empirical literature typically employed an ad hoc consumption function (or savings function) and attempted to measure the interest rate effect by including the interest rate as an explanatory variable. Wright (1967), for example, estimates

[12]See also Bernheim (2002, p. 1181). Evans (1983) and Starrett (1988a) cast doubt on the general validity of the Summers result.

a consumption function of the following type:

$$C_t = \alpha_0 + \alpha_1 Y_t^* + \alpha_2 \Omega_t + \alpha_3 r_t + u_t, \tag{3.107}$$

where C_t is consumption, Y_t^* is "normal income" (an exponentially weighted average of current and past actual disposable income), Ω_t is net worth, r_t is the after-tax interest rate, and u_t is the stochastic error term. He finds a significant negative estimate for α_3 which implies a savings elasticity in the range of 0.19 to 0.24. Blinder (1975) allows for changes in the income distribution and finds much lower estimates.[13]

Following the rational expectations revolution of the seventies and eighties, the use of ad hoc non-structural models (such as (3.107)) has fallen into disrepute. Instead, modern empirical work makes use of the consumption Euler equation—see Deaton (1992) and Attanasio (1999) for excellent surveys on this approach. In the context of our simple two-period model of Section 3.1, the Euler equation is given by (3.46). Making the obvious substitutions for adjacent time periods, this expression can be approximated by:[14]

$$\ln C_{t+1} - \ln C_t \approx \sigma \left[r_t - \rho - \frac{t_{Ct+1} - t_{Ct}}{1 + t_{Ct+1}} \right]. \tag{3.108}$$

This equation for consumption is called *structural* because it is derived from a microeconomic maximization problem. In (3.108), the conditioning variables are the interest rate and the current and future consumption tax rates. (Of course, in the presence of interest income taxation, the after-tax interest rate features in the Euler equation.) Following the pioneering contribution by Hall (1978), a large body of literature has emerged trying to estimate a stochastic version of equation (3.108) by econometric means. As is pointed out by Bernheim (2002, pp. 1210-1211), when applied to aggregate consumption data, the Euler equation approach typically yields very low estimates for the intertemporal substitution elasticity, σ. In contrast, when the method is applied at the level of individual households, the estimates for σ are typically larger than zero but less than unity. On the basis of an extensive survey of empirical studies, Elmendorf (1996, p. 19) suggests that a value of σ of about 0.37 may not be too far off the mark.

In a deterministic model, knowledge of σ is sufficient to compute the marginal propensity to consume out of total wealth and thus the interest elasticity of saving.[15] Indeed, in the two-period model,

[13]Other early studies include Wright (1969), Friend and Hasbrouck (1983), Boskin (1978), and Boskin and Lau (1982). The last two studies find a high elasticity of about 0.4. As is pointed out by Bernheim (2002, p. 1208), however, the typical finding in the early literature was a near-zero interest elasticity of saving.

[14]The Euler equation is given by:

$$\frac{C_{t+1}}{C_t} = \left(\frac{1+r_t}{1+\rho} \frac{1+t_{Ct}}{1+t_{Ct+1}} \right)^{\sigma}.$$

By taking logarithms and noting that $\ln(1+r_t) \approx r_t$, $\ln(1+\rho) \approx \rho$, and $\ln x \approx x - 1$ (around $x = 1$) we find:

$$\ln C_{t+1} - \ln C_t = \sigma \ln \left(\frac{1+r_t}{1+\rho} \right) + \sigma \ln \left(\frac{1+t_{Ct}}{1+t_{Ct+1}} \right) \approx \sigma \left[r_t - \rho + \frac{t_{Ct} - t_{Ct+1}}{1+t_{Ct+1}} \right].$$

[15]As Bernheim (2002, p. 1211) points out, however, in the case with uncertainty one cannot infer the interest elasticity of saving

current consumption equals $C_1 = \omega_1 \Omega$, where Ω is defined in (3.42) and ω_1 is defined as:

$$\omega_1 \equiv \left[1 + t_{C1} + \left(\frac{1+t_{C1}}{1+\rho} \right)^{\sigma} \left(\frac{1+t_{C2}}{1+r_1} \right)^{1-\sigma} \right]^{-1}. \tag{3.109}$$

For given values of r_1, ρ, t_{C1}, and t_{C2}, and σ, an implied value for ω_1 can be computed. For given values of initial wealth ($(1+r_0)\,a_0$) and present and future labour incomes ($w_1\bar{L}$ and $w_2\bar{L}$), total wealth can be computed as well as the strength of the human wealth effect. Using the expression for $\partial C_1/\partial r_1$ from (3.24) and noting that $\partial a_1/\partial r_1 = -\partial C_1/\partial r_1$ it is possible to compute the interest elasticity of saving. With an estimate for σ of 0.37, it is clear from (3.24) that the sum of the income effect and the substitution effect on consumption is positive. Hence, the interest effect on saving is negative on that account. Only if the human wealth effect is sufficiently strong, will the interest elasticity of saving be positive. Summers (1981) uses a multiperiod consumption-saving model and shows that the interest elasticity of saving can be quite large even if σ is low. Indeed, for the case with $\sigma = 1/3$, an interest rate of 4% per annum, and a rate of time preference of 3% per annum, Summers computes an interest elasticity of 2.38 (1981, p. 536).

3.4.2 Intertemporal substitution elasticity in labour supply

The modern empirical literature on intertemporal labour supply behaviour derives its inspiration from the pioneering contribution by Lucas and Rapping (1969). Before turning to the empirical evidence, we first sketch a prototypical structural model of life-cycle labour supply behaviour along the lines of MaCurdy (1981). In doing so we generalize the two-period model (used in Section 3.2 above) and assume that there exists a representative household with an infinite planning horizon. The lifetime utility function of this household is given by:

$$\Lambda_t \equiv \sum_{s=0}^{\infty} \left(\frac{1}{1+\rho} \right)^{s} U\left(C_{t+s}, \bar{L} - L_{t+s}\right), \tag{3.110}$$

where t is the planning period, ρ is the pure rate of time preferences, $U(\cdot)$ is the felicity function (with the usual properties that were mentioned below equation (2.1)), and C_{t+s} and $\bar{L} - L_{t+s}$ are, respectively, consumption and leisure at time $t+s$. The budget identity at time $t+s$ is given by:

$$a_{t+s} = (1 + r_{t+s-1})\, a_{t+s-1} + w^*_{t+s} L_{t+s} - C_{t+s}, \tag{3.111}$$

where a_{t+s} are financial assets at the end of period $t+s$, r_{t+s} is the rate of interest, $w^*_{t+s} \equiv (1 - t_{Lt+s})\, w_{t+s}$ is the after-tax real wage rate, and C_{t+s} is consumption. In the planning period the initial stock of financial assets, a_{t-1}, is given (determined in the past).[16]

The household chooses paths for consumption, labour supply, and financial assets in order to maxi-

from an estimate of σ. We return to this issue in Chapter 4.

[16]Note that equations (3.110) and (3.111) are the multiperiod counterparts to, respectively, equations (3.54) and (3.55) that appear in the two-period model.

mize lifetime utility (3.110) subject to a sequence of budget identities of the form (3.111), taking as given a_{t-1} and a solvency condition (which needs not concern us here). The Lagrangian associated with this maximization problem is given by:

$$\mathcal{L}_t \equiv \sum_{s=0}^{\infty} \left(\frac{1}{1+\rho}\right)^s U\left(C_{t+s}, \bar{L} - L_{t+s}\right) + \sum_{s=0}^{\infty} \lambda_{t+s} \left(\frac{1}{1+\rho}\right)^s \left[\left(1 + r_{t+s-1}\right) a_{t+s-1} + w^*_{t+s} L_{t+s} - C_{t+s} - a_{t+s}\right],$$

where $\lambda_{t+s}\left(1+\rho\right)^{-s}$ is the (scaled) Lagrange multiplier associated with the period $t+s$ constraint (3.111). Since s runs from 0 to ∞, a whole path of Lagrange multipliers must be postulated. The first-order conditions for an interior optimum are given by:

$$\frac{\partial \mathcal{L}_t}{\partial C_{t+s}} = \left(\frac{1}{1+\rho}\right)^s \left[U_C\left(C_{t+s}, \bar{L} - L_{t+s}\right) - \lambda_{t+s}\right] = 0, \tag{3.112}$$

$$\frac{\partial \mathcal{L}_t}{\partial L_{t+s}} = -\left(\frac{1}{1+\rho}\right)^s \left[U_{\bar{L}-L}\left(C_{t+s}, \bar{L} - L_{t+s}\right) - \lambda_{t+s} w^*_{t+s}\right] = 0, \tag{3.113}$$

$$\frac{\partial \mathcal{L}_t}{\partial a_{t+s}} = -\lambda_{t+s}\left(\frac{1}{1+\rho}\right)^s + \left(1 + r_{t+s}\right)\lambda_{t+s+1}\left(\frac{1}{1+\rho}\right)^{s+1} = 0. \tag{3.114}$$

The derivation of first two conditions is straightforward because they involve intratemporal comparisons. The condition for assets is, however, a little more complicated because it compares marginal costs and benefits through time. Additional saving in period $t+s$ is costly because it detracts from consumption, but it yields benefits in the form of higher interest income in the next period which can then be used for higher consumption. The marginal costs and benefits are represented by, respectively, the first and second term on the right-hand side of (3.114), and in the optimum they are equalized.

The first-order conditions for the planning period can be written as:

$$U_C\left(C_t, \bar{L} - L_t\right) = \lambda_t, \tag{3.115}$$

$$U_{\bar{L}-L}\left(C_t, \bar{L} - L_t\right) = \lambda_t w^*_t, \tag{3.116}$$

$$\lambda_t = \frac{1+r_t}{1+\rho}\lambda_{t+1} \tag{3.117}$$

Equations (3.115) and (3.116) implicitly define so-called λ-constant or *Frisch demand equations* for, respectively, goods consumption and leisure consumption.[17] The conditioning variable, λ_t, appearing in these demand equations is the marginal utility of initial wealth, i.e. it measures the change in (maximized) lifetime utility resulting from an infinitesimal change in initial wealth ($\lambda_t = \Lambda_t / \partial a_{t-1}$). The optimal lifetime decision problem thus consists of two components. First, the savings decision is solved in such a manner that the marginal utility of wealth evolves optimally according to (3.117). Second, conditional

[17] Together with Jan Tinbergen, Ragnar Frisch was awarded the first Nobel Prize in Economics in 1969 for his contributions to econometrics. He introduced the notion of Frisch demands in Frisch (1932). See also Frisch (1959) Heckman (1974, 1976), and Browning, Deaton, and Irish (1985, p. 507).

upon the value of λ_t, the period-t choices regarding consumption and labour supply are made according to (3.115)-(3.116). All the dynamic information (including the lifetime budget constraint) is thus incorporated in the variable λ_t.[18]

In order to illustrate the type of Frisch labour supply function that has been used in the empirical literature, we follow MaCurdy (1981) by assuming the felicity function to take the following (additively separable) form:

$$U\left(C_t, L_t\right) \equiv \alpha_t \frac{C_t^{1-1/\sigma_C} - 1}{1 - 1/\sigma_C} - \beta_t \frac{L_t^{1+1/\sigma_L}}{1 + 1/\sigma_L}, \tag{3.118}$$

where $\sigma_C > 0$ and $\sigma_L > 0$. In this specification, α_t and β_t are exogenously given taste parameters. These shift parameters can be used to capture certain life-cycle effects. For example, if β_t is postulated to rise over time, then ceteris paribus the household finds it more and more disagreeable to supply a given amount of labour as it gets older.

With the felicity function expressed directly in terms of labour supply (rather than leisure), the first-order conditions (3.115)-(3.116) change to, respectively, $U_C\left(C_t\right) = \lambda_t$, and $-U_L\left(L_t\right) = \lambda_t w_t^*$. The resulting Frisch labour supply is thus loglinear:

$$\ln L_t = -\sigma_L \ln \beta_t + \sigma_L \ln w_t^* + \sigma_L \ln \lambda_t. \tag{3.119}$$

Using this labour supply equation for adjacent time periods t and $t-1$ we obtain:

$$\Delta \ln L_t = -\sigma_L \Delta \ln \beta_t + \sigma_L \Delta \ln w_t^* - \sigma_L \left(r_{t-1} - \rho\right), \tag{3.120}$$

where $\Delta x_t \equiv x_t - x_{t-1}$ and we have used (3.117) lagged one period to deduce that $\Delta \ln \lambda_t = \ln\left(1+\rho\right) - \ln\left(1 + r_{t-1}\right) \approx \left(\rho - r_{t-1}\right)$. The expression in (3.120) is useful because it identifies the three separate reasons for a household to change its labour supply over time. First, if tastes change—for example β_t rises over time—then ceteris paribus this will lead to a decrease in labour supply. The *taste effect* is represented by the first term on the right-hand side of (3.120). Second, if the after-tax wage rises over time, then, holding constant $\Delta \ln \beta_t$ and r_{t-1}, labour supply will rise. This *intertemporal labour supply* effect is represented by the second term on the right-hand side of (3.120). Finally, the change in labour supply also depends on the gap between the interest rate and the pure rate of time preference because it affects the steepness of the Euler equation for labour supply.[19]

If households operate under conditions of uncertainty, the resulting labour supply equation becomes

[18] A loglinear *approximation* to the Frisch labour supply equation for the general model can be obtained from (3.115)-(3.116) by using the generalized implicit function theorem. As is shown by Card (1994, p. 51), such an equation takes the following form:

$$\ln L_t = \eta_{Lw} \ln w_t^* + \eta_{L\lambda} \ln \lambda_t,$$

where $\eta_{Lw} > 0$. If leisure is a normal good then $\eta_{L\lambda} > 0$. If felicity is separable in consumption and leisure ($U_{C,\bar{L}-L} = 0$, as is the case in (3.118)), then $\eta_{Lw} = \eta_{L\lambda}$.

[19] The Euler equation for labour supply can be deduced by noting that $-U_L\left(L_{t+1}\right) = \lambda_{t+1} w_{t+1}^*$ and $-U_L\left(L_t\right) = \lambda_t w_t^*$. Using

a little more complicated—see for example Card (1994, pp. 51-53). Here we ignore the formidable econometric complications involved in estimating a stochastic version of (3.120) and refer the interested reader to MaCurdy (1985), Pencavel (1986, pp. 44-51, 83-94), and Blundell and MaCurdy (1999, pp. 1591-1607) for details.

One of the main objectives of the empirical literature has been to obtain a reliable estimate of the *intertemporal substitution elasticity* in labour supply, that is the elasticity of employment with respect to the after-tax wage rate (the coefficient for $\Delta \ln w_t^*$ in (3.120)). Pencavel (1986, p. 85) presents a useful overview of the estimates of σ_L for prime-age males. Most estimates are positive (as theory predicts) but quite small. In a later critical survey article, Card (1994, p. 63) argues that "...the elasticity of intertemporal substitution is surely no higher than 0.5, and probably no higher than 0.2." In a recent study, Lee (2001) finds an estimate of about 0.5. He argues that previous studies found much lower estimates as a result of technical econometric problems (small sample bias and weak instrumental variables).

3.5 Punchlines

In this chapter we study the intertemporal decisions of a representative household. In the basic two-period Fisherian model, labour supply is exogenous and the household chooses present and future consumption in order to maximize its lifetime utility. By saving or borrowing the household can transfer resources intertemporally and achieve an optimal life-cycle pattern of consumption. The interest rate determines the relative price of future consumption. A change in the interest rate has three effects on present and future consumption, namely a pure substitution effect (SE), an income effect (IE), and a human wealth effect (HWE). Whilst the first two effects are well-known from static models, the third effect is specific to dynamic models. The HWE results from the change in the present value of future non-interest income. Even if the utility function features a low degree of substitutability between current and future consumption, the HWE can nevertheless explain a relatively high elasticity of saving with respect to the interest rate.

The two-period model is applied to a number of issues. First, several tax equivalencies are demonstrated. Tax systems are deemed to be equivalent if they give rise to the same household choices of present and future consumption. It is shown that a proportional tax on labour income plus initial assets is equivalent to a consumption tax. Furthermore, an interest income tax is equivalent to a particular type of wealth tax.

The second application of the two-period model concerns a quantitative-analytical study of consumption taxation. If the consumption tax is time-invariant then it does not affect the intertemporal

(3.117) we obtain:

$$\frac{L_{t+1}}{L_t} = \left(\frac{\beta_t}{\beta_{t+1}} \frac{1+\rho}{1+r_t} \frac{w_{t+1}^*}{w_t^*} \right)^{\eta}.$$

Taking logarithms and lagging once yields (3.120).

trade-off between present and future consumption and it only has an income effect. In contrast, if the present or future consumption tax is changed, there are both income and pure substitution effects and the consumption tax has features not unlike an interest income tax. Unlike for an interest income tax, however, the human wealth effect is absent under the consumption tax because the interest rate itself is not affected.

In the second section of this chapter we study an extended version of the Fisherian model in which labour supply is endogenous. Here the household must make both dynamic decisions (now versus the future) and static decisions (consumption versus leisure). In the extended model, tax changes in general have both *inter*temporal and *intra*temporal effects on present and future consumption and labour supply. Provided preferences take a particular form, a very useful analytical method can be used to solve the model. This *two-stage budgeting method* is explained in a detailed example.

In the third section of this chapter we study three further extensions to the two-period model. The first extension deals with the human capital accumulation decision. By investing in education and training in the current period, the household can increase its labour productivity and future wage income. Even though labour supply in the current period is endogenous and training is time-consuming, a time-invariant labour income tax does not affect the optimal training decision. An increase in the interest income tax, however, decreases the cost of borrowing and leads to an increase in the optimal training effort! The same result holds for a general income tax in which all income is taxed at the same proportional rate.

In the second extension we consider the implications for optimal consumption and savings choices of (i) a difference between the interest rate on lending and on borrowing, and (ii) a maximum borrowing limit (credit constraint). Each of these features gives rise to a kink in the household's lifetime budget constraint. If households differ in their degree of patience, it is demonstrated that potentially large groups of people will be located at one of these kinks. A tax on interest income received does not affect such households at all.

In the third extension we provide a brief discussion of a dynastic model in which members of different generations are linked via preferences (unidirectional altruism) and bequests (intergenerational transfers). Under certain (rather strong) conditions, parents leave bequests to their offspring and, in doing so, act *as if* they are maximizing an infinite-horizon dynastic utility function. Finitely-lived generations are linked to distant relatives via the perpetual chain of bequests. In such a dynastic model, a wealth tax on bequests is equivalent to an interest income tax. Both an asymmetric tax treatment of bequests and a prohibition on negative bequests may cause the intergenerational chain to snap.

In the final section of this chapter we briefly review the empirical evidence on the intertemporal substitution hypothesis. The results obtained thus far do not lend very significant support to this hypothesis. For the consumption-saving model a very low intertemporal substitution elasticity is typically found. This implies that agents are rather unwilling to substitute consumption across time and that the interest elasticity of saving is small. Furthermore, many studies find that current income is an important conditioning variable for current consumption. One interpretation of this result is that a significant proportion of the population faces capital market restrictions. For the intertemporal labour supply model matters are also not very encouraging. Typically, researchers find a very low estimate for the intertemporal substitution elasticity in labour supply (which sometimes even has the wrong sign). Of course, the jury is still out on the issue. Better data and better estimation methods may unearth more support for the intertemporal substitution hypothesis in the future. From a theoretical perspective, the life-cycle approach to consumption and labour supply remains one of the workhorse models of public economics.

Further reading

Basic two-period model. Atkinson and Stiglitz (1980, lecture 3) and Sandmo (1985) cover much of the same material as we do. On the altruistic model, see Barro (1974) and Bernheim and Bagwell (1988). Blomquist (1985) studies rate-progressive labour income taxes in a two-period model with endogenous labour supply. Diamond (1970) studies the incidence of an interest income tax.

Empirical evidence. Early papers on the interest elasticity of household saving include Wright (1967, 1969), Blinder (1975), Friend and Hasbrouck (1983), Boskin (1978), and Boskin and Lau (1982). Except for the last two studies, the typical result is a low interest elasticity of saving. See Sandmo (1985, pp. 280-283), Elmendorf (1996), and Bernheim (2002, pp. 1203-1211) for surveys. Deaton (1992) and Attanasio (1999) are up-to-date surveys on the intertemporal consumption-saving model. Some good references on intertemporal labour supply are MaCurdy (1981, 1985), Browning, Deaton, and Irish (1985), Pencavel (1986, pp. 44-51), Card (1994), and Blundell and MaCurdy (1999, pp. 1591-1607). Ham and Reilly (2002) test three theories of the labour market and consistently reject the intertemporal substitution model.

Calibrated models. There is a huge literature on calibrated simulation models with life-cycle features. Key references are Summers (1981a) and Auerbach and Kotlikoff (1987). See also Evans (1983) and Starrett (1988a) for critical remarks on the Summers approach as well as the rejoinder by Summers (1984).

Chapter 4

Taxation and choices under risk

The purpose of this chapter is to discuss the following topics:

- How can we model household consumption and savings behaviour in the face of uncertainty about either the yield on its savings (*capital risk*) or about its future non-interest income (*income risk*)?
- What is the role of *risk aversion* and how is this phenomenon captured by some often-used functional forms for the utility function?
- What is the effect of different taxes on risk taking behaviour by the household both with and without the loss-offset provision?
- How can we apply the basic stochastic model to study phenomena such as *precautionary savings* under income uncertainty, labour supply in the face of a risky wage rate, and *income tax evasion* with stochastic monitoring by the tax authority?
- How does the household react to additional riskiness and what is the role of *prudence*?

4.1 A basic stochastic model of consumption and saving

Up to this point we have assumed that households are endowed with perfect foresight regarding the future. Indeed, in the basic consumption model studied in the previous chapter, the household was assumed to know the future price level, real wage rate, and tax rates with perfect certainty. The objective of this chapter is to relax this assumption of perfect foresight and to study the effects of risk and uncertainty on household behaviour.

It is relatively straightforward to introduce risk and uncertainty into the two-period consumption-saving model. In order to keep the model as simple as possible, we make the following assumptions. First, just as in the previous chapter we consider only two time periods: period 1 is the present and period 2 is the remaining future (obviously, by construction, there is no period 3). Second, we continue to assume that households have perfect foresight regarding wages, taxes, and prices. Third, we assume that household labour supply is exogenous. Together with the second assumption this implies that there is no *income risk*, i.e. future non-interest income is non-stochastic (in Section 4.3 we study the topic of income risk). Fourth, we abstract from intergenerational bequests and assume that initial assets are zero. Fifth, we abstract from capital market imperfections such as differential lending and borrowing rates and quantity restrictions. As a result the household faces 'almost' no constraints on borrowing or lending.

The only kind of uncertainty the household faces is so-called *capital risk*. There are two assets of which one has an uncertain return. The safe asset (say "money") carries a certain yield of r (where $r \geq -1$). In addition there is a single risky asset (say "bonds" or "stocks") which carries an uncertain yield of $\tilde{x}$, where x is a stochastic variable with a known probability distribution. To keep matters simple it is assumed that at worst the household can lose its entire risky investment, i.e. $\Pr\{\tilde{x} \geq -1\} = 1$, where $\Pr\{\textbf{event}\}$ denotes the probability of **event** occurring, and a tilde above a variable denotes that the variable in question is stochastic.[1]

The household determines the optimal allocation of its saving over the two assets. Since utility will in general be stochastic, the household is assumed to maximize its *expected utility*. The real budget constraints facing the household can be written as:

$$b + m^* + C_1 = w_1 \bar{L}, \tag{4.1}$$

$$\tilde{C}_2 = (1+r)m^* + (1+\tilde{x})\, b + w_2 \bar{L}, \tag{4.2}$$

where b is the amount of risky assets bought ($b \geq 0$, i.e. the household is a demander of risky assets), m^* is the amount of safe assets bought or borrowed, $\bar{L}$ is exogenous labour supply, w_t is the exogenous real wage rate in period t ($t = 1, 2$), C_1 is non-stochastic consumption in period 1, and $\tilde{C}_2$ is stochastic consumption in period 2. Investing in the risky asset is an *temporal uncertain prospect* because time must elapse before uncertainty is resolved and "nature" reveals the realized value of $\tilde{x}$. As a consequence, consolidation of the budget identities (4.1)-(4.2) is not convenient because it hides the sequence of events.

To simplify the notation it is useful to rewrite the budget identities somewhat by defining $m \equiv m^* + w_2\bar{L}/\,(1+r)$:

$$b + m + C_1 = h_0, \tag{4.3}$$

$$\tilde{C}_2 = (1+r)m + (1+\tilde{x})\, b, \tag{4.4}$$

where $h_0 \equiv w_1\bar{L} + w_2\bar{L}/\,(1+r)$ is human wealth, i.e. the present value of non-interest income capital-

[1] We follow convention in this literature by using the superscript tilde to denote stochastic variables. (Elsewhere in the book a tilde denotes a rate of change.) See, for example, Eeckhoudt, Gollier, and Schlesinger (2005) for an accessible textbook on economic decision making in a risky environment.

ized at the risk-free rate, and m represents *gross* risk-free assets.

The representative household's lifetime expected utility function is given by:

$$E(\tilde{\Lambda}) = U(C_1) + \frac{1}{1+\rho} E(U(\tilde{C}_2)), \tag{4.5}$$

where $E(\cdot)$ is the expectations operator (the expectation based on the known probability density function of $\tilde{x}$), ρ is the constant rate of pure time preference ($\rho > 0$), and $U(\cdot)$ is the felicity function. This function features a positive but diminishing marginal felicity, i.e. $U'(\cdot) > 0 > U''(\cdot)$. (Below we assume a specific functional form for the felicity function.) No expectations operator is needed in front of the first term on the right-hand side of (4.5) because C_1 and hence $U(C_1)$ are non-stochastic.

The stochastic assumption regarding $\tilde{x}$ completes the model. The probability density function for $\tilde{x}$ is denoted by $f(\tilde{x})$ and is defined over the interval $\tilde{x} \in [-1, \infty)$. If the realization of $\tilde{x}$ equals the lower bound ($x = -1$) the investor "loses his entire investment principal and all" and if the realization of $\tilde{x}$ equals the upper bound ($x \to \infty$) the investor "strikes it lucky by hitting the jackpot". Given the stochastic process for $\tilde{x}$, the expression for expected lifetime utility is given by:

$$E(\tilde{\Lambda}) = U(C_1) + \frac{1}{1+\rho} \int_{-1}^{\infty} f(\tilde{x})\, U(\tilde{C}_2) d\tilde{x}, \tag{4.6}$$

where we have used the fact that $E(U(\tilde{C}_2)) \equiv \int_{-1}^{\infty} f(\tilde{x})\, U(\tilde{C}_2) d\tilde{x}$.

By substituting (4.3) and (4.4) into (4.6) we find the following expression for expected lifetime utility:

$$\begin{aligned} E(\tilde{\Lambda}) &= U\Big(h_0 - (b+m)\Big) \\ &\equiv + \frac{1}{1+\rho} \int_{-1}^{\infty} f(\tilde{x})\, U\Big((b+m) \left[(1+r)\omega + (1+\tilde{x})(1-\omega) \right] \Big) d\tilde{x}, \end{aligned} \tag{4.7}$$

where $\omega \equiv m/(m+b)$ is the portfolio *share* of the risk-free asset (and $1-\omega \equiv b/(b+m)$ is the share of the risky asset), and we have used the fact that $\tilde{C}_2$ can be written as:

$$\tilde{C}_2 = (b+m) \left[(1+r)\omega + (1+\tilde{x})(1-\omega) \right]. \tag{4.8}$$

The household's choice problem involves two types of decisions. The *savings decision* concerns the optimal choice of $b+m$ whilst the *portfolio decision* involves the optimal choice of ω. The variables that are exogenous to the household are the risk-free interest rate (r), human wealth (h_0), and the (parameters of the) stochastic distribution of $\tilde{x}$.

The first-order necessary condition for the optimal savings decision is:

$$\frac{\partial E(\tilde{\Lambda})}{\partial (b+m)} = -U'(C_1) + \frac{1}{1+\rho} \int_{-1}^{\infty} \left[(1+r)\omega + (1+\tilde{x})(1-\omega) \right] f(\tilde{x})\, U'(\tilde{C}_2) d\tilde{x} = 0. \tag{4.9}$$

By rewriting this expression slightly we obtain the consumption Euler equation in a stochastic setting:

$$U'(C_1) = \frac{1}{1+\rho} E\left(\left[(1+r)\omega + (1+\tilde{x})(1-\omega)\right] U'(\tilde{C}_2)\right). \tag{4.10}$$

The first-order necessary condition for the optimal portfolio decision is:

$$\frac{\partial E(\tilde{\Lambda})}{\partial \omega} = \frac{b+m}{1+\rho} \int_{-1}^{\infty} (r-\tilde{x}) f(\tilde{x}) U'(\tilde{C}_2) d\tilde{x} = 0, \tag{4.11}$$

where we have used the fact that $(b+m)$ is non-stochastic. By simplifying (4.11) we obtain:

$$0 = E\left(U'(\tilde{C}_2)(\tilde{x}-r)\right). \tag{4.12}$$

Equation (4.12) implicitly determines the *optimum portfolio*, i.e. the optimal division of saving over the safe and the risky asset. Intuitively it says that the expected marginal utility per Euro invested should be equated for the two assets.

4.1.1 A special case: Iso-elastic felicity

In principle one can deduce comparative static effects directly from equations (4.10) and (4.12). In order to simplify the discussion, however, we postulate a specific (iso-elastic) functional form for the felicity function:

$$U(C_t) = \begin{cases} (1/\gamma_R)\left[C_t^{\gamma_R} - 1\right] & \text{if } \gamma_R \neq 0 \\ \\ \ln C_t & \text{if } \gamma_R = 0, \end{cases} \tag{4.13}$$

where γ_R (< 1) characterizes the *degree of risk aversion* exhibited by the agent (see below). We shall call $1-\gamma_R$ the coefficient of *relative* risk aversion.[2]

Armed with the functional form (4.13), the optimal portfolio condition (4.12) can be simplified. We obtain after some steps:

$$\begin{aligned} 0 &= E\left(U'(\tilde{C}_2)(\tilde{x}-r)\right) \\ &= E\left(\tilde{C}_2^{\gamma_R-1}(\tilde{x}-r)\right) \\ &= E\left((b+m)^{\gamma_R-1}\left[(1+r)\omega + (1+\tilde{x})(1-\omega)\right]^{\gamma_R-1}(\tilde{x}-r)\right) \\ &= E\left(\left[(1+r)\omega + (1+\tilde{x})(1-\omega)\right]^{\gamma_R-1}(\tilde{x}-r)\right), \end{aligned} \tag{4.14}$$

where we have used the expression for $\tilde{C}_2$ from (4.8) in going from the second to the third line, and have

[2]Note the close correspondence with the iso-elastic felicity function that was used in the deterministic context in the previous chapter (see Section 3.1.2). The relationship between the intertemporal substitution elasticity, σ, and the risk-aversion parameter, γ_R, is $\sigma = 1/(1-\gamma_R)$.

noted that m and b are non-stochastic in going from the third to the fourth line. Equation (4.14) implicitly determines the optimal portfolio share, ω^*, as a function of r, γ_R, and the parameters characterizing the probability distribution of $\tilde{x}$. As was pointed out by Samuelson (1969), this optimal portfolio share maximizes the so-called *subjective mean return* on the portfolio, x^*, which is implicitly defined as:[3]

$$\begin{aligned}(1+x^*)^{\gamma_R} &\equiv \max_{\omega} E\left((1+r)\omega + (1+\tilde{x})(1-\omega)\right)^{\gamma_R} \\ &= E\left((1+r)\omega^* + (1+\tilde{x})(1-\omega^*)\right)^{\gamma_R}. \end{aligned} \tag{4.15}$$

The stochastic consumption Euler equation can also be simplified with the aid of (4.13). Indeed, by using (4.13) in (4.10) and simplifying we obtain:

$$\begin{aligned}C_1^{\gamma_R-1} &= \frac{1}{1+\rho} E\left(\tilde{C}_2^{\gamma_R-1}\left[(1+r)\omega + (1+\tilde{x})(1-\omega)\right]\right) \\ &= \frac{1}{1+\rho}(b+m)^{\gamma_R-1} E\left((1+r)\omega + (1+\tilde{x})(1-\omega)\right)^{\gamma_R}, \end{aligned} \tag{4.16}$$

where we have used (4.8) to get from the first to the second line. By using (4.3) and (4.15) we can further simplify (4.16) to obtain a simple expression for consumption in the current period:

$$C_1 = \xi_1 h_0, \tag{4.17}$$

where ξ_1 is the marginal propensity to consume out of total wealth:

$$\xi_1 \equiv \frac{(1+x^*)^{\gamma_R/(\gamma_R-1)}}{(1+\rho)^{1/(\gamma_R-1)} + (1+x^*)^{\gamma_R/(\gamma_R-1)}}. \tag{4.18}$$

Interestingly, the optimal consumption plan for the current period looks very much like the solution that would be obtained under certainty. Indeed, in the absence of uncertainty about the bond yield, maximization of lifetime utility would give rise to the same expression for ξ_1 but with the subjective mean return, x^*, replaced by $\max[\bar{x}, r]$, where $\bar{x}$ is the certain return on the (not so) "risky" asset. Note furthermore that in the case of logarithmic felicity ($\gamma_R = 0$), x^* drops out of (4.18) altogether, $\xi_1 = (1+\rho)\,/\,(2+\rho)$ and the existence of capital risk does not affect the level of present consumption at all. Finally, we have now established that with iso-elastic felicity functions, there exists a so-called *separability property* between the savings problem (choosing when to consume) and the portfolio problem (choosing what to use as a savings instrument).[4]

[3]The proof of Samuelson's claim runs as follows. The first-order condition associated with the maximization problem indicated in (4.15) is given by:

$$\frac{dE\left((1+r)\omega + (1+\tilde{x})(1-\omega)\right)^{\gamma_R}}{d\omega} = E\left(\left((1+r)\omega + (1+\tilde{x})(1-\omega)\right)^{\gamma_R-1}(r-\tilde{x})\right) = 0,$$

which coincides with (4.14).

[4]Samuelson derives this result in a multi-period consumption saving model (1969a, p. 244). Hagen (1971) studies such a multi-period model in the presence of proportional income taxes.

4.1.2 Taxation and risk taking

How do taxes affect risk taking by the household? An often expressed view is that taxation of asset income discriminates against risk taking because it lowers the expected rate of return on risky assets. An alternative view was expressed six decades ago by Domar and Musgrave (1944) who argued that taxation may increase risk taking. On the one hand, the government takes a share of the expected return but on the other hand with *perfect loss offsets* the government also shares in the risk of losses. As a result it may well be the case that asset income taxation increases risk taking by the household.

In order to study the effect of taxation on risk taking we now extend the model by introducing various types of tax systems. These systems differ only in their definition of the tax base. In the so-called *net taxation* case the tax is levied on the excess return on the risky asset only, whereas in the *gross taxation* case both assets are taxed at the same rate. We study the general version of the model, i.e. the felicity function is not restricted to be iso-elastic.

4.1.2.1 Net taxation case

In the net taxation case we assume that there is a tax, t_A, on the *excess return* on the risky asset, $\tilde{x} - r$. This tax rate is non-stochastic and the tax base is equal to $b(\tilde{x} - r)$. The current budget constraint is still given by (4.3) but (4.4) is modified to:

$$\tilde{C}_2 = (1+r)m + (1+\tilde{x})\,b - t_A b\,(\tilde{x} - r)\,, \tag{4.19}$$

where m is gross risk-free assets (defined directly above equation (4.3)). Using the same steps as before, we can rewrite $\tilde{C}_2$ as follows:

$$\begin{aligned}\tilde{C}_2 &= (1+r)m + (1+r+\tilde{x}-r)\,b - t_A b\,(\tilde{x}-r) \\ &= (1+r)\,(m+b) + (1-t_A)\,(\tilde{x}-r)\,b \\ &= (1+r)\,(h_0 - C_1) + (1-t_A)\,(\tilde{x}-r)\,b,\end{aligned} \tag{4.20}$$

where h_0 is human wealth and we have used (4.3) in the final step.

Using (4.20) in (4.6), expected lifetime utility can now be written as:

$$E(\tilde{\Lambda}) = U(C_1) + \frac{1}{1+\rho}\int_{-1}^{\infty} f(\tilde{x})\,U\Big((1+r)\,(h_0 - C_1) + (1-t_A)\,(\tilde{x}-r)\,b\Big)d\tilde{x}. \tag{4.21}$$

Expressed in this manner, the *savings decision* is the choice of current consumption (which implies the choice of $b+m$) and the *portfolio decision* is the choice of b (which then implies the choice of the portfolio share of the safe asset, ω). The exogenous variables to the household are r, h_0, t_A, and (the parameters of) the distribution of $\tilde{x}$.

The first-order necessary conditions for the optimal savings and portfolio decisions are, respectively,

$\partial E(\tilde{\Lambda})/\partial C_1 = 0$ and $\partial E(\tilde{\Lambda})/\partial b = 0$ or:

$$U'(C_1) = \frac{1+r}{1+\rho} E\left(U'(\tilde{C}_2)\right) \tag{4.22}$$

$$0 = E\left(U'(\tilde{C}_2)(\tilde{x} - r)\right), \tag{4.23}$$

where we have used the fact that ρ and $1 - t_A$ are non-stochastic in deriving (4.23).

Armed with this simple model we can determine how the tax rate affects current consumption and the demand for the risky asset. The key thing to note is that the tax rate, t_A, only enters the first-order conditions (4.22) and (4.23) via it effect on future (stochastic) consumption, $\tilde{C}_2$, which is defined in (4.20) above. By partially differentiating (4.22)-(4.23) and (4.20) with respect to the tax rate we find:

$$U''(C_1)\frac{\partial C_1}{\partial t_A} = \frac{1+r}{1+\rho} E\left(U''(\tilde{C}_2)\frac{\partial \tilde{C}_2}{\partial t_A}\right), \tag{4.24}$$

$$0 = E\left(U''(\tilde{C}_2)(\tilde{x} - r)\frac{\partial \tilde{C}_2}{\partial t_A}\right), \tag{4.25}$$

$$\frac{\partial \tilde{C}_2}{\partial t_A} = -(1+r)\frac{\partial C_1}{\partial t_A} + \left[(1 - t_A)\frac{\partial b}{\partial t_A} - b\right](\tilde{x} - r). \tag{4.26}$$

Given that $U''(\tilde{C}_2) < 0$, the only possible way for both (4.24) and (4.25) to hold simultaneously for all realizations of $\tilde{x}$, is if $\partial \tilde{C}_2/\partial t_A = 0$. It follows from (4.26) that this is the case if and only if the following conditions are both satisfied:

$$\frac{\partial C_1}{\partial t_A} = 0, \tag{4.27}$$

$$\frac{\partial b}{\partial t_A} = \frac{b}{1 - t_A} > 0. \tag{4.28}$$

Equation (4.27) shows that present consumption (and thus saving) is not changed if the tax is changed, whereas (4.28) shows that an increase in the tax leads to an *increase* in the demand for the risky asset! The intuition behind this Domar-Musgrave result is as follows. By adopting the portfolio rule (4.28), the household is able to hold constant the *probability distribution of* consumption in the second period ($\tilde{C}_2$). To show that this is indeed the case, consider the mean of future consumption:

$$E(\tilde{C}_2) \equiv E\left((1+r)(h_0 - C_1) + (1 - t_A)(\tilde{x} - r)b\right) \quad \Rightarrow$$

$$\frac{\partial E(\tilde{C}_2)}{\partial t_A} = E\left(\left[(1 - t_A)\frac{\partial b}{\partial t_A} - b\right](\tilde{x} - r)\right) = 0. \tag{4.29}$$

Similarly, for the variance of future consumption we find:

$$V(\tilde{C}_2) \equiv E\left(\tilde{C}_2 - E(\tilde{C}_2)\right)^2 \quad \Rightarrow$$

$$\frac{\partial V(\tilde{C}_2)}{\partial t_A} = E\left(2(\tilde{C}_2 - E(\tilde{C}_2))\left[\frac{\partial \tilde{C}_2}{\partial t_A} - \frac{\partial E(\tilde{C}_2)}{\partial t_A}\right]\right) = 0, \tag{4.30}$$

and similarly for all higher-order moments of the distribution. Given that the originally chosen distribution was optimal to start with, the agent continues to select it given that it is available even after the tax is changed (Sandmo, 1985, p. 295). Note that this conclusion is independent of the shape of the agent's preferences. Furthermore, expected utility does not change as a result of the tax change (see also Figure 4.5 below for an illustration).

4.1.2.2 Gross taxation case

In the gross taxation case we assume that there is a non-stochastic tax, t_A, on the gross return on both assets. The tax base is thus equal to $b\tilde{x} + rm$ and the real budget constraints consist of equation (4.3) and:

$$\tilde{C}_2 = [1 + r(1 - t_A)]\, m + [1 + \tilde{x}(1 - t_A)]\, b, \tag{4.31}$$

where $r(1 - t_A)$ is the (deterministic) after-tax yield on the safe asset, and $\tilde{x}(1 - t_A)$ is the (stochastic) after-tax yield on the risky asset. We can rewrite $\tilde{C}_2$ as follows:

$$\begin{aligned} \tilde{C}_2 &= [1 + r(1 - t_A)]\,(m + b) + (1 - t_A)\,(\tilde{x} - r)\, b \\ &= [1 + r(1 - t_A)]\,(h_0 - C_1) + (1 - t_A)\,(\tilde{x} - r)\, b, \end{aligned} \tag{4.32}$$

where we have used (4.3) in the final step. By substituting (4.32) into (4.6) we obtain the expression for lifetime utility:

$$E(\tilde{\Lambda}) = U(C_1) + \frac{1}{1+\rho} \int_{-1}^{\infty} f(\tilde{x})\, U\Big([1 + r(1 - t_A)]\,(h_0 - C_1) + (1 - t_A)\,(\tilde{x} - r)\, b \Big) d\tilde{x}. \tag{4.33}$$

The first-order necessary conditions for, respectively, the optimal savings and portfolio decisions are:

$$U'(C_1) = \frac{1 + r(1 - t_A)}{1+\rho} E\left(U'(\tilde{C}_2)\right), \tag{4.34}$$

$$0 = E\left(U'(\tilde{C}_2)\,(\tilde{x} - r)\right). \tag{4.35}$$

How does the tax rate affect current consumption and the demand for the risky asset? As for the net taxation case, the tax rate enters the first-order conditions (4.34)-(4.35) via the expression for $\tilde{C}_2$ (given in (4.32) above). Unlike for the net taxation case, however, the tax also distorts the Euler equation (4.34). Since there are competing wealth and substitution effects, no unambiguous conclusion is possible about the relationship between risk taking and the tax rate in the gross taxation case. Indeed, by differentiating (4.34)-(4.35) and (4.32) with respect to the tax rate we find:

$$U''(C_1) \frac{\partial C_1}{\partial t_A} = \frac{1 + r(1 - t_A)}{1+\rho} E\left(U''(\tilde{C}_2) \frac{\partial \tilde{C}_2}{\partial t_A}\right) - \frac{r}{1+\rho} E\left(U'(\tilde{C}_2) \frac{\partial \tilde{C}_2}{\partial t_A}\right), \tag{4.36}$$

$$0 = E\left(U''(\tilde{C}_2)\,(\tilde{x} - r)\, \frac{\partial \tilde{C}_2}{\partial t_A}\right), \tag{4.37}$$

$$\frac{\partial \tilde{C}_2}{\partial t_A} = -\left[1 + r\left(1 - t_A\right)\right]\frac{\partial C_1}{\partial t_A} - r\left(h_0 - C_1\right) + \left[\left(1 - t_A\right)\frac{\partial b}{\partial t_A} - b\right]\left(\tilde{x} - r\right). \tag{4.38}$$

Comparing these expressions to (4.24)-(4.26), we find that there are additional terms in both (4.36) and (4.38). As a result, it is no longer optimal for the household to adopt (4.27)-(4.28). In the interest of space, the full implications of (4.36)-(4.37) are studied in an exercise for this chapter. Here it suffices to note that the solution is much more complicated than (4.27)-(4.28) because the optimal rule now depends on the shape of the agent's preferences. Although no firm general conclusion are possible, Sandmo (1985, p. 297) argues that for reasonable parameter values the substitution effect is likely to dominate the wealth effect so that the Domar-Musgrave view continues to hold even in the gross taxation case.

4.2 Portfolio decision in isolation

In this section we study the portfolio decision in isolation, i.e. we step back from the two-period consumption-saving model and focus only on the household's choice of portfolio investment. One advantage of doing so is that it allows us to investigate in more detail the role of preferences in the portfolio decision.

The key assumptions we employ in this section are the following. The individual household is assumed to maximize expected utility of wealth at the end of the period (this is not unlike our $\tilde{C}_2$ in the model of Section 4.1). The household is risk-averse and can invest in two assets, namely a safe asset with a certain yield of r, and a risky asset with an uncertain yield $\tilde{x}$. To keep the analysis as simple as possible, a very simple probability distribution for $\tilde{x}$ is postulated. Indeed, it is assumed that only two *states* are possible for the yield on the risky asset, namely a state in which the yield is low (subscript "L") and one in which the yield is high (subscript "H"). The probability of a low yield ($\tilde{x} = x_L < r$) is π_L and the probability of a high yield ($\tilde{x} = x_H > r$) is $\pi_H \equiv 1 - \pi_L$. The analytical advantage of this two-state model lies in the fact that a simple graphical treatment is possible (see also Stiglitz (1969) and Atkinson and Stiglitz (1980, lecture 4)).

The amount of wealth to be invested is given by:

$$A_0 \equiv b + m, \tag{4.39}$$

where A_0 is initial wealth (exogenously given), b is the amount of risky assets, and m is the amount of safe assets. Depending on the household's investment decision, final wealth is stochastic:

$$\tilde{A}_1 = \begin{cases} A_L \equiv m\left(1 + r\right) + b\left(1 + x_L\right) & \text{with probability } \pi_L \\ \\ A_H \equiv m\left(1 + r\right) + b\left(1 + x_H\right) & \text{with probability } 1 - \pi_L \end{cases}, \tag{4.40}$$

where A_L and A_H denote final wealth in, respectively, the low-yield and the high-yield case. Note that

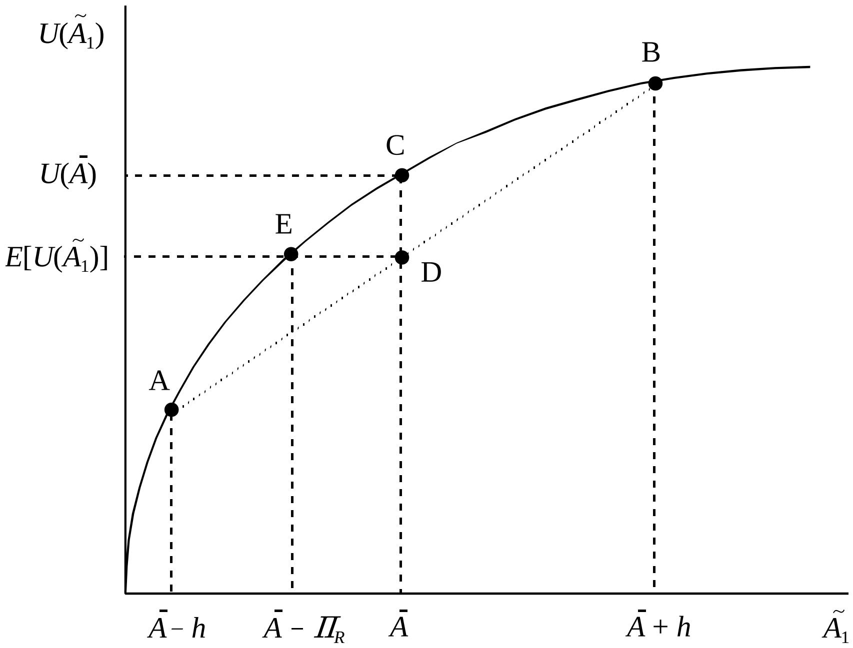

Figure 4.1: Risk aversion and the risk premium

the household can influence A_L and A_H by choice of b.

Utility is stochastic and depends on final wealth, i.e. utility is $U(\tilde{A}_1)$. Risk aversion implies that the utility function features the following derivatives:

$$U'(\tilde{A}_1) > 0, \qquad U''(\tilde{A}_1) < 0. \tag{4.41}$$

Intuitively, a risk averse household prefers a safe final wealth of $\bar{A}$ over a random distribution with the mean of final wealth equal to $\bar{A}$. Graphically this is shown in Figure 4.1. The utility of the certain prospect is $U(\bar{A})$. The stochastic prospect consists of final wealth $\tilde{A}_1 = \bar{A} - h$ with probability $\frac{1}{2}$ or $\tilde{A}_1 = \bar{A} + h$ with probability $\frac{1}{2}$. The expected value of the stochastic prospect is thus by construction equal to the certain prospect, i.e. $E(\tilde{A}_1) = \bar{A}$. The expected utility of the stochastic prospect is $E(U(\tilde{A}_1))$, which is at point D. The so-called *risk premium* is equal to Π_R, where Π_R is such that $U\left(\bar{A} - \Pi_R\right) = E(U(\tilde{A}_1))$. Clearly, Π_R is positive for a risk-averse household because such a household needs to be compensated for accepting some risk. Note also that Π_R in general depends on both the distribution of $\tilde{x}$ and on the shape of the utility function (see the classic analysis by Pratt (1964)).

Expected utility can be written as:

$$E(U(\tilde{A}_1)) \equiv \pi_L U\left(A_L\right) + \left(1 - \pi_L\right) U\left(A_H\right), \tag{4.42}$$

where A_L and A_H are defined in (4.40) above. The household chooses b and m in order to maximize (4.42) subject to the wealth constraint (4.39). In Figure 4.2 the optimum portfolio choice is illustrated.

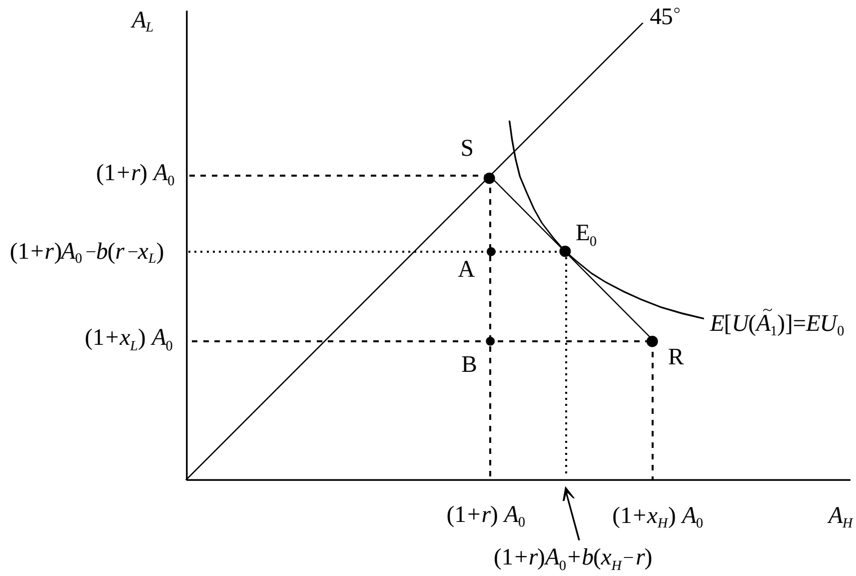

Figure 4.2: The optimal portfolio decision

Final wealth in the "bad" and the "good" state is measured on, respectively, the vertical axis and the horizontal axis. The perfectly safe portfolio ($b = 0$, $m = A_0$) is represented by point S. This point lies along the 45-degree line because final wealth is the same in both states (no risk is taken by the household at all!). Point R represents the perfectly risky portfolio ($m = 0$, $b = A_0$). For that portfolio, final wealth equals $A_L = (1 + x_L)\,A_0$ with probability π_L and $A_H = (1 + x_H)\,A_0$ with probability $1 - \pi_L$. By varying b (and m, such that $b + m = A_0$) any portfolio between points S and R is attainable.

The mathematical expression for the budget line SR is obtained as follows. First, we note from (4.39)-(4.40) that:

$$A_L = (1 + r)\,A_0 - b\,(r - x_L)\,, \tag{4.43}$$

$$A_H = (1 + r)\,A_0 + b\,(x_H - r)\,. \tag{4.44}$$

Second, by solving (4.44) for b and substituting the resulting expression in (4.43) we find:

$$A_L = -\frac{r - x_L}{x_H - r} A_H + \frac{x_H - x_L}{x_H - r}\,(1 + r)\,A_0. \tag{4.45}$$

The budget line is downward sloping and an increase in the amount to be invested shifts it in a parallel fashion to the right (these results follow from the assumption $x_L < r < x_H$).

The slope of the indifference curve is given by:

$$\begin{aligned} dE(U(\tilde{A}_1)) &\equiv \pi_L U'\,(A_L)\,dA_L + (1 - \pi_L)\,U'\,(A_H)\,dA_H = 0 \qquad \Leftrightarrow \\ \frac{dA_L}{dA_H} &= -\frac{1 - \pi_L}{\pi_L}\frac{U'\,(A_H)}{U'\,(A_L)} < 0. \end{aligned} \tag{4.46}$$

In an interior optimum, the indifference curve is tangent to the budget line. It follows from (4.45) and (4.46) that b is chosen such that:

$$\frac{1-\pi_L}{\pi_L}\frac{U'(A_H)}{U'(A_L)} = \frac{r-x_L}{x_H-r}. \tag{4.47}$$

The interior optimum is at point E_0 in Figure 4.2.

An ancient theorem from Greek antiquity can be used to deduce something about the optimal choice of risky assets. We know from the Theorem of Pythagoras that the length of the line segment SE_0 is equal to:

$$\begin{aligned}\text{(length of) } SE_0 &= \sqrt{(SA)^2 + (AE_0)^2} = \sqrt{[b(r-x_L)]^2 + [b(x_H-r)]^2} \\ &= b\sqrt{(r-x_L)^2 + (x_H-r)^2}.\end{aligned} \tag{4.48}$$

Similarly, we know that the length of line segment SR is:

$$\begin{aligned}\text{(length of) } SR &= \sqrt{(SB)^2 + (BR)^2} = \sqrt{[A_0(r-x_L)]^2 + [A_0(x_H-r)]^2} \\ &= A_0\sqrt{(r-x_L)^2 + (x_H-r)^2}.\end{aligned} \tag{4.49}$$

By combining the expressions in (4.48)-(4.49), we find that the portfolio *share* of the risky asset is represented geometrically by:

$$\frac{b}{A_0} = \frac{\text{(length of) } SE_0}{\text{(length of) } SR}. \tag{4.50}$$

This result is intuitively obvious. The closer E_0 is to the perfectly safe point S, the smaller is the portfolio share of risky assets.

4.2.1 Wealth effects in the portfolio decision

How does the optimum portfolio depend on the household's initial wealth level, A_0? It is clear from (4.45) that an increase in A_0 shifts the budget line in a parallel fashion to the right, say from S_0R_0 to S_1R_1 in Figure 4.3. Not surprisingly, the effect on the optimum portfolio depends on the form of the utility function. We consider two often-used functional forms.

4.2.1.1 CRRA preferences

In Figure 4.3 we plot the case for iso-elastic (or CRRA) preferences over terminal wealth:

$$U(\tilde{A}_1) = \begin{cases} (1/\gamma_R)\left[\tilde{A}_1^{\gamma_R} - 1\right] & \text{if } \gamma_R \neq 0,\ \gamma_R < 1 \\ \\ \ln \tilde{A}_1 & \text{if } \gamma_R = 0, \end{cases} \tag{4.51}$$

where $1 - \gamma_R$ is the (constant) coefficient of *relative risk aversion*.[5] Using (4.51), the first-order condition (4.47) simplifies to:

$$\begin{aligned} \frac{1-\pi_L}{\pi_L}\frac{U'(A_H)}{U'(A_L)} &= \frac{r - x_L}{x_H - r} \quad \Leftrightarrow \\ \frac{1-\pi_L}{\pi_L}\frac{A_L^{1-\gamma_R}}{A_H^{1-\gamma_R}} &= \frac{r - x_L}{x_H - r} \quad \Leftrightarrow \\ \frac{A_L}{A_H} &= \left(\frac{\pi_L}{1-\pi_L}\frac{r - x_L}{x_H - r}\right)^{1/(1-\gamma_R)} \equiv \theta_R. \end{aligned} \tag{4.52}$$

In the optimum, $A_L = \theta_R A_H$ (with θ_R a positive constant) so the wealth expansion path is a straight line from the origin–see the line WEP in Figure 4.3. It follows that the same portfolio composition is chosen at all wealth levels and that the wealth elasticity of the demand for the risky asset is unity.

We can find the level solutions for b and m as follows. First, by substituting (4.52) into the budget line (4.45) we obtain the solutions for A_H and A_L:

$$A_H = \frac{(x_H - x_L)(1+r)A_0}{\theta_R(x_H - r) + (r - x_L)}, \tag{4.53}$$

$$A_L = \frac{\theta_R(x_H - x_L)(1+r)A_0}{\theta_R(x_H - r) + (r - x_L)}. \tag{4.54}$$

Next, we note that the optimal value for b can be recovered by substituting (4.53) into (4.44):

$$b = \frac{(1-\theta_R)(1+r)A_0}{\theta_R(x_H - r) + (r - x_L)} \tag{4.55}$$

Finally, by using (4.55) in (4.39) we obtain the optimal value for m:

$$\begin{aligned} m &= A_0 - b \\ &= \frac{\left[\theta_R(1+x_H) - (1+x_L)\right]A_0}{\theta_R(x_H - r) + (r - x_L)}. \end{aligned} \tag{4.56}$$

[5] The coefficient of relative risk aversion is defined as:

$$R_R(z) \equiv -\frac{zU''(z)}{U'(z)}.$$

For iso-elastic preferences, $R_R(z) = 1 - \gamma_R$, a constant. Such preferences are therefore often referred to as **C**onstant **R**elative **R**isk **A**version or CRRA preferences. See Gollier (2001, p. 27).

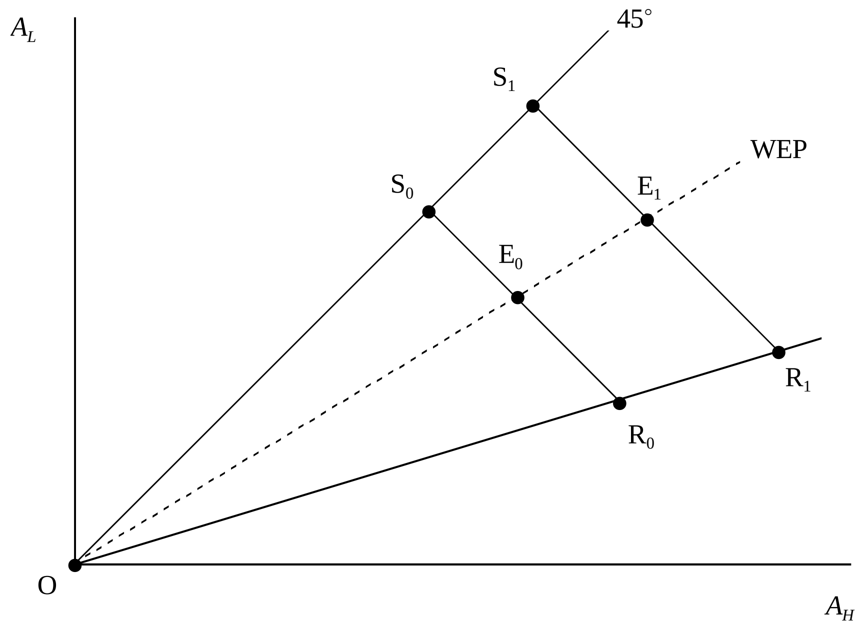

Figure 4.3: Wealth expansion path for CRRA preferences

A very risk averse investor has a very high value of $1-\gamma_R$ and will choose to stay close to point S. For such an investor, θ_R will be close to unity (see (4.52)), b will be close to zero (see (4.55)), and m will be close to A_0 (see (4.56)). Conversely, an investor who is only mildly risk averse will be characterized by a low (but positive) value of $1-\gamma_R$ and will choose a value of θ_R close to unity.

4.2.1.2 CARA preferences

In Figure 4.4 we illustrate the case for exponential (or CARA) preferences:

$$U(\tilde{A}_1) = -\frac{\exp[-\gamma_A \tilde{A}_1]}{\gamma_A}, \tag{4.57}$$

where γ_A (> 0) is the (constant) measure of *absolute risk aversion*.[6] The first-order condition (4.47) for such preferences simplifies to:

$$\begin{aligned}
\frac{1-\pi_L}{\pi_L}\frac{U'(A_H)}{U'(A_L)} &= \frac{r-x_L}{x_H-r} \quad \Leftrightarrow \\
\frac{1-\pi_L}{\pi_L}\frac{\exp[-\gamma_A A_H]}{\exp[-\gamma_A A_L]} &= \frac{r-x_L}{x_H-r} \quad \Leftrightarrow \\
A_L - A_H &= \frac{1}{\gamma_A}\ln\left(\frac{\pi_L}{1-\pi_L}\frac{r-x_L}{x_H-r}\right) \equiv -\theta_A.
\end{aligned} \tag{4.58}$$

[6]The coefficient of absolute risk aversion is defined as:

$$R_A(z) \equiv -\frac{U''(z)}{U'(z)}.$$

For exponential preferences, $R_A(z) = 1-\gamma_A$, a constant. Such preferences are therefore often referred to as **C**onstant **A**bsolute **R**isk **A**version or CARA preferences. See Gollier (2001, p. 27).

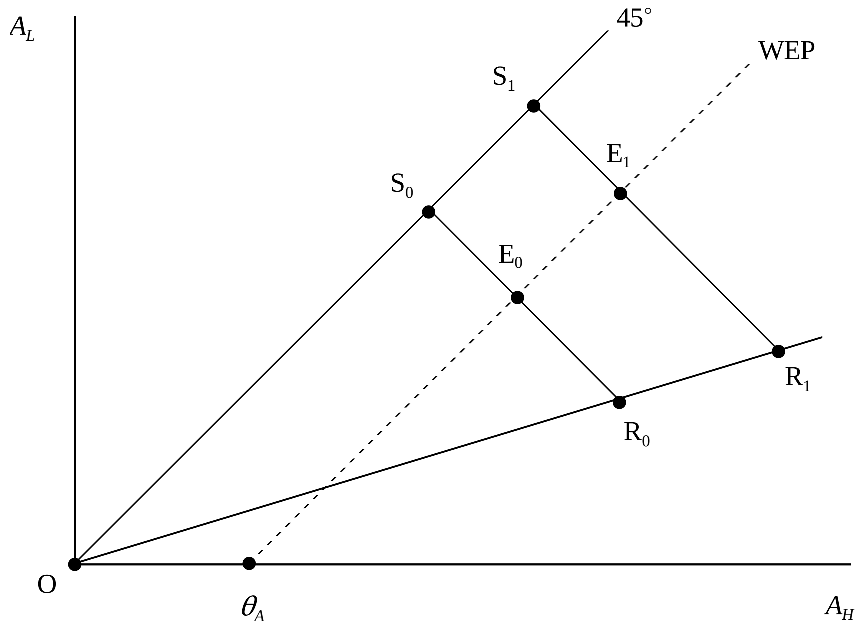

Figure 4.4: Wealth expansion path for CARA preferences

In the optimum, $A_L = A_H - \theta_A$ (with θ_A a positive constant), so the wealth expansion path is a straight line parallel to the 45° line–see the line WEP in Figure 4.4. In this case the wealth elasticity of the demand for the risky asset is zero, i.e. the same amount of risky assets is chosen for all wealth levels.

The level solutions are found as follows. By substituting (4.58) into the budget line (4.45) we obtain:

$$A_H = \frac{\theta_A (x_H - r) + (x_H - x_L)(1+r) A_0}{x_H - x_L}, \tag{4.59}$$

$$A_L = \frac{\theta_A (x_L - r) + (x_H - x_L)(1+r) A_0}{x_H - x_L}. \tag{4.60}$$

We infer from (4.43)-(4.44) that $A_H - A_L = b(x_H - x_L)$ so that b is equal to:

$$b = \frac{\theta_A}{x_H - x_L}, \tag{4.61}$$

which is independent of A_0 as was asserted above. Finally, the optimal value for m is found by using (4.61) and (4.39):

$$m \equiv A_0 - b = \frac{A_0 (x_H - x_L) - \theta_A}{x_H - x_L}. \tag{4.62}$$

In this case, a very risk averse investor has a very high value of γ_A and a very low value for θ_A (see (4.58)), so that the wealth expansion path will be close to the "safe haven" of the 45-degree line.

4.2.1.3 Wealth elasticity

We have established that the wealth elasticity of the demand for the risky asset is unity for the case of CRRA preferences (see (4.55)) and zero for the case of CARA preferences (see (4.61)). The remainder of this subsection studies how the wealth elasticity is affected by the investor's risk attitude in the general case. Equation (4.47) defines an implicit function relating the optimal demand for risky assets, b, to the initial wealth level, A_0. For convenience, we restate (4.47) in a lightly rewritten format as:

$$(r - x_L)\,\pi_L U'(A_L) = (x_H - r)(1 - \pi_L)\,U'(A_H). \tag{4.63}$$

By differentiating this expression with respect to b and A_0, and noting (4.43)-(4.44), we obtain:

$$\begin{aligned}(r - x_L)\,\pi_L U''(A_L)\,[(1+r)\,dA_0 - (r - x_L)\,db] &= (x_H - r)(1 - \pi_L)\,U''(A_H) \\ &\quad \times [(1+r)\,dA_0 + (x_H - r)\,db], \end{aligned} \tag{4.64}$$

or:

$$\frac{1}{1+r}\frac{db}{dA_0} = \frac{[-(r - x_L)\,\pi_L U''(A_L) + (x_H - r)(1 - \pi_L)\,U''(A_H)]}{-\left[(r - x_L)^2\,\pi_L U''(A_L) + (x_H - r)^2 (1 - \pi_L)\,U''(A_H)\right]}. \tag{4.65}$$

Clearly, the denominator of (4.65) is positive (because $U''(\cdot) < 0$, $(r - x_L)^2 > 0$, and $(x_H - r)^2 > 0$). The sign of db/dA_0 is thus determined by the sign of the numerator, which can be written in short-hand notation as $E((\tilde{x} - r)\,U''(\tilde{A}_1))$. Since $\tilde{x} - r$ is negative in the bad state and positive in the good state, the sign of the numerator depends on the curvature of the utility function. It is thus possible to relate the sign of db/dA_0 to a measure of the investor's risk attitude.

By using the coefficient of *absolute* risk aversion, $R_A(z) \equiv -U''(z)/U'(z)$, it is possible to write (4.65) as:

$$\frac{1}{1+r}\frac{db}{dA_0} = \frac{R_A(A_L) - R_A(A_H)}{(r - x_L)\,R_A(A_L) + (x_H - r)\,R_A(A_H)}, \tag{4.66}$$

where we have used the first-order condition (4.63) to simplify the expression. Whilst the denominator continues to be positive (as $R_A(\cdot) > 0$ and $x_L < r < x_H$), the sign of the numerator hinges only on the difference between $R_A(A_L)$ and $R_A(A_H)$. With CARA preferences, $R_A(A_L) = R_A(A_H) = 1 - \gamma_A$ (a constant) and it follows that $db/dA_0 = 0$. As Arrow (1971, p. 96) has pointed out, however, everyday observation seems to suggests that the degree of absolute risk aversion is decreasing in wealth, i.e. $R'_A(z) < 0$. Preferences for which this holds are called *DARA preferences* (where DARA stands for **D**iminishing **A**bsolute **R**isk **A**version). In the context of our two-state model, with DARA preferences $R_A(A_L)$ exceeds $R_A(A_H)$ and it follows from (4.66) that the wealth elasticity of the demand for the risky asset is positive.

In a similar fashion we can use the coefficient of *relative* risk aversion, $R_R(z) \equiv -zU''(z)/U'(z)$, and write (4.65) in elasticity format as:

$$
\begin{aligned}
\frac{A_0}{b}\frac{db}{dA_0} &= \frac{(1+r)A_0}{b}\frac{\left[\frac{R_R(A_L)}{A_L}-\frac{R_R(A_H)}{A_H}\right]}{(r-x_L)\frac{R_R(A_L)}{A_L}+(x_H-r)\frac{R_R(A_H)}{A_H}} \\
&= \frac{(1+r)A_0}{b}\frac{[A_H R_R(A_L)-A_L R_R(A_H)]}{(r-x_L)A_H R_R(A_L)+(x_H-r)A_L R_R(A_H)} \\
&= \frac{(1+r)A_0}{b}\frac{b+\frac{\phi A_L}{x_H-x_L}}{(1+r)A_0-\frac{\phi A_L(x_H-r)}{x_H-x_L}},
\end{aligned}
\tag{4.67}
$$

where $\phi \equiv [R_R(A_L)-R_R(A_H)]/R_R(A_L)$.[7] We reach a similar conclusion as for the DARA case. If the utility function features **D**iminishing **R**elative **R**isk **A**version (DRRA), then $R_R'(z) < 0$ and in the two-state model $\phi > 0$. It follows from (4.67) that the wealth elasticity of the demand for the risky asset is greater than unity.

As is pointed out by Gollier (2001, p. 25), assuming preferences to exhibit DRRA constitutes a much stronger requirement than assuming the DARA property. This can be readily seen by noting that the two measures of risk aversion can be related according to $R_R(z) \equiv zR_A(z)$, so that $R_R'(z) \equiv R_A(z) + zR_A'(z)$. DARA only requires $R_A'(z)$ to be negative. DRRA requires $R_A'(z)$ to be sufficiently negative (to offset the positive term $R_A(z)$ in the expression for $R_R'(z)$). He also suggests that there is no clear evidence in favour of the DRRA assumption, whereas there is for the DARA property.

[7] Once again we have made use of the first-order condition (4.63) to obtain the expression in the first line. In going from the second to the third line, it is useful to note that (4.43) and (4.44) imply:

$$(r-x_L)A_H + (x_H-r)A_L = (x_H-x_L)(1+r)A_0.$$

Intermezzo 4.1

Introspective estimate for the degree of risk aversion. In a recent monograph on the economics of risk, Gollier (2001, pp. 30-31) presents an interesting way to estimate a person's coefficient of relative risk aversion by introspection. Suppose you are the person we are studying today. The maintained assumption in this approach is that your preferences are of the CRRA type:

$$U(\tilde{A}_1) = \begin{cases} (1/\gamma_R)\left[\tilde{A}_1^{\gamma_R} - 1\right] & \text{if } \gamma_R \neq 0,\ \gamma_R < 1 \\ \\ \ln \tilde{A}_1 & \text{if } \gamma_R = 0, \end{cases} .$$

The objective is to find the magnitude of your coefficient of relative risk aversion, $R_R \equiv 1 - \gamma_R$.

Thought experiment. Answer the following question for yourself: What is the share of your wealth that you are willing to pay in order to escape the risk of gaining or losing a share α of it with equal probability? Denote the share that you are willing to pay $\hat{\pi}$. Then your answer implies that you are indifferent between the following prospects:

$$\frac{1}{2}\left[\frac{[(1-\alpha)\,A]^{\gamma_R}}{\gamma_R}\right] + \frac{1}{2}\left[\frac{[(1+\alpha)\,A]^{\gamma_R}}{\gamma_R}\right] = \frac{[(1-\hat{\pi})\,A]^{\gamma_R}}{\gamma_R} \quad \Leftrightarrow$$

$$\frac{(1-\alpha)^{\gamma_R}}{2} + \frac{(1+\alpha)^{\gamma_R}}{2} = (1-\hat{\pi})^{\gamma_R}\,. \tag{I.1}$$

We can use (I.1) to compute γ_R once we know $\hat{\pi}$ (from you) and fix a value of α in the thought experiment.

Implied estimate	$\alpha = 10\%$	$\alpha = 30\%$
$1 - \gamma_R = 0.5$	0.3	2.3
$1 - \gamma_R = 1$	0.5	4.6
$1 - \gamma_R = 4$	2.0	16.0
$1 - \gamma_R = 10$	4.4	24.4
$1 - \gamma_R = 40$	8.4	28.7

The above Table (which is taken from Gollier) shows implied estimates of $1 - \gamma_R$ for $\alpha = 10\%$ and $\alpha = 30\%$. To interpret these results, consider the column labeled $\alpha = 10\%$. Reasonable answers for $\hat{\pi}$ are in the range 0.5% to 2%, which implies a value for $1 - \gamma_R$ in the range of 1 to 4. Conversely, note that $1 - \gamma_R = 40$ implies (rather unreasonably) that you would pay 8.4% of your wealth to avoid the risk of gaining or losing 10% with equal probability. People are not that risk averse!

****

4.2.2 Effects of taxation

In this subsection we return to the theme of this chapter, namely the effects of taxation on risk taking. We consider two types of taxes, namely a proportional wealth tax (t_W), and a proportional tax (t_A) on either (i) the excess return on the risky asset or (ii) on both assets. Note that the latter two cases have also been studied above in the two-period consumption-savings model. In the first instance we assume *full loss offset*, i.e. losses incurred in the bad state can be deducted in full from the tax base.

The proportional wealth tax affects terminal wealth according to:

$$\tilde{A}_1 = \begin{cases} A_L \equiv (1-t_W)\left[A_0(1+r)+b(x_L-r)\right] & \text{with probability } \pi_L \\ \\ A_H \equiv (1-t_W)\left[A_0(1+r)+b(x_H-r)\right] & \text{with probability } \pi_H \end{cases}. \tag{4.68}$$

It follows from (4.68) that the budget line in (A_L, A_H)-space is given by:

$$A_L = -\frac{r-x_L}{x_H-r}A_H + \frac{x_H-x_L}{x_H-r}(1+r)(1-t_W)A_0. \tag{4.69}$$

As can be seen from (4.69), the wealth tax has no effect on the slope of the budget line: the term $(1-t_W)$ simply reduces initial wealth. The effects of an *decrease* in the wealth tax are illustrated in Figure 4.3 (for CRRA preferences) and in Figure 4.4 (for CARA preferences). In both diagrams the equilibrium shifts from E_0 to E_1. Since the slope of the budget line is not affected, the optimal point is located on the initial wealth expansion path WEP in both figures. For the CRRA case, the demand for the risky asset is increased as a result of the increase in after-tax wealth (see (4.55)). In contrast, for the CARA case, the demand for the risky asset is unaffected by the wealth increase (see (4.61)).

A proportional income tax on the *excess return* on the risky asset affects terminal wealth according to:

$$\tilde{A}_1 = \begin{cases} A_L \equiv A_0(1+r)+b(1-t_A)(x_L-r) & \text{with probability } \pi_L \\ \\ A_H \equiv A_0(1+r)+b(1-t_A)(x_H-r) & \text{with probability } \pi_H \end{cases}, \tag{4.70}$$

so that the budget line in (A_L, A_H)-space is still as given in (4.45) above:

$$A_L = -\frac{r-x_L}{x_H-r}A_H + \frac{x_H-x_L}{x_H-r}(1+r)A_0. \tag{4.71}$$

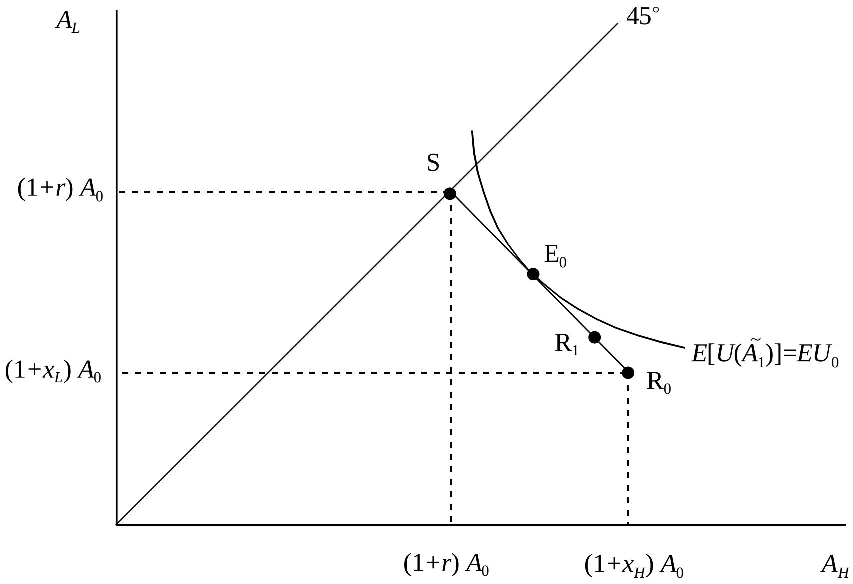

Figure 4.5: Proportional income tax on excess return (CRRA preferences)

Obviously, the tax affects neither the slope nor the position of the budget line. The only thing that it does affect is the *length* of the budget line. Indeed, it follows from (4.70) that under the perfectly risky portfolio (for which $a = 0$ and $b = A_0$), final wealth is equal to $A_L = [1 + r + (1 - t_A)(x_L - r)]$ is the bad state and $A_H = [1 + r + (1 - t_A)(x_H - r)]$ in the good state. In terms of Figure 4.5, an increase in t_A thus moves the perfectly risky point from R_0 to R_1. Barring corner solutions, the same optimal point can still be attained (at point E_0). The portfolio *share* of the risky asset rises, from SE_0/SR_0 to SE_0/SR_1. (Note that SR_1 is $(1 - t_A)$ times SR_0.)

A proportional income tax on the return on both assets affects terminal wealth according to:

$$\tilde{A}_1 = \begin{cases} A_L \equiv A_0\,[1 + r\,(1 - t_A)] + b\,(1 - t_A)\,(x_L - r) & \text{prob. } \pi_L \\ \\ A_H \equiv A_0\,[1 + r\,(1 - t_A)] + b\,(1 - t_A)\,(x_H - r) & \text{prob. } \pi_H \end{cases}, \tag{4.72}$$

so that the budget line in (A_L, A_H)-space becomes:

$$A_L = -\frac{r - x_L}{x_H - r} A_H + \frac{x_H - x_L}{x_H - r}\,[1 + r\,(1 - t_A)]\,A_0. \tag{4.73}$$

The slope of the budget line is unaffected by the tax, and t_A acts as a tax on initial wealth. An increase in t_A shifts the budget line in a parallel fashion towards the origin. In Figure 4.6 the case of CRRA preferences is illustrated. The equilibrium shifts from E_0 to E_1. The EPL line represents the locus for which the portfolio composition is the same as at point E_0 (b/A_0 constant). Since point E_1 lies to the right of the EPL line, the portfolio share of the risky asset increases as a result of the tax change (i.e.

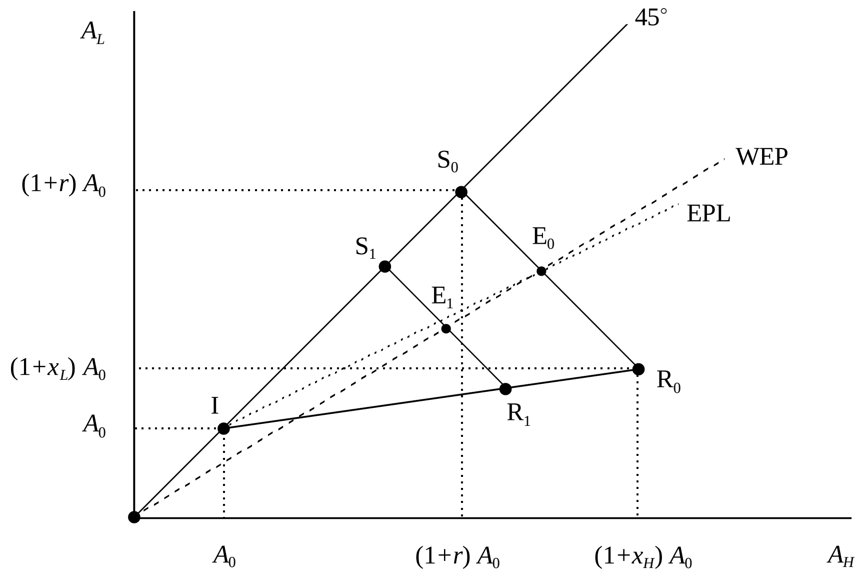

Figure 4.6: Proportional income tax (CRRA preferences)

S_1E_1/S_1R_1 is larger than S_0E_0/S_0R_0).[8]

4.2.2.1 No loss offsets

Matters are much less clear-cut if the government is a fair-weather friend and there is no loss offset provision. To show why this is so, consider the special case in which there is a zero return on the safe asset ($r = 0$), a negative return on the risky asset in the bad state ($x_L < 0$), and preferences are of the CRRA type. A proportional interest income tax now affects terminal wealth according to:

$$\tilde{A}_1 = \begin{cases} A_L \equiv A_0 + bx_L & \text{with probability } \pi_L \\ \\ A_H \equiv A_0 + b(1-t_A)x_H & \text{with probability } \pi_H \end{cases}. \tag{4.74}$$

In the bad state the investor loses $-bx_L$ (> 0) which he cannot deduct against tax liabilities. In the good state the gains are bx_H which are taxed at the proportional rate t_A. The budget line is now given by:

$$A_L = \frac{x_L}{x_H(1-t_A)}A_H + \frac{[x_H(1-t_A) - x_L]A_0}{x_H(1-t_A)}. \tag{4.75}$$

[8] In Figure 4.6, the line IR_1R_0 connects all purely risky points (for which $b = A_0$). The mathematical expression for this line is obtained as follows. Substituting $b = A_0$ into (4.72) we find:

$$A_L = [1 + x_L(1-t_A)]A_0, \; A_H = [1 + x_H(1-t_A)]A_0.$$

Eliminating $(1-t_A)$ we find:

$$A_L - A_0 = \frac{x_L}{x_H}(A_H - A_0),$$

which passes through point I and has a positive slope less than unity (because $x_H > x_L$).

In terms of Figure 4.7, the safe point S is unaffected by the tax because $r = 0$ so no taxable income is generated at that point. An increase in the tax rotates the budget line clockwise and moves the pure risky point from R_0 to R_1. As a result of the tax change, the equilibrium shifts from E_0 to E_1, there are both substitution and wealth effects, and the usual Hicksian decomposition can be made. Indeed, in terms of the optimal choices for A_H and A_L, the pure substitution effect is represented by the move from E_0 to E' whilst the income effect is the move from E' to E_1.

The effect on the demand for the risky asset can be determined graphically by noting that the vertical distance between the safe point S and the initially optimal point E_0 is equal to $-b_0 x_L$, where b_0 is the initial demand for the risky asset. Similarly, the vertical distance between S and E_1 is equal to $-b_1 x_L$, where b_1 is the new optimal choice for b. Since E_1 lies to the north of E_0, it follows that for the case drawn in Figure 4.7, the demand for the risky asset decreases as a result of the tax increase. If compensation is made only if the good state occurs (and the investor has to pay the tax), then the vertical distance between S and E' measures $-b' x_L$, where b' is the compensated (Hicksian) demand for the risky asset. Formally, the Slutsky equation for the demand for the risky asset is given by:

$$\frac{\partial b}{\partial x_H^*} = \left(\frac{\partial b}{\partial x_H^*}\right)_{EU_0} + b\frac{\partial b}{\partial Z_H}, \tag{4.76}$$

where $x_H^* \equiv x_H (1 - t_A)$ is the after-tax return on the risky asset in the good state and Z_H is the (hypothetical) compensation in the good state.[9] The first term on the right-hand side of (4.76) is the pure substitution effect, which is positive. The second term is a negative income effect due to compensation. It follows that the slope of the uncompensated demand for the risky asset is ambiguous. An increase in the tax t_A leads to a decrease in x_H^* and thus to a decrease in the Hicksian demand for the risky asset. Since the income effect is negative, the uncompensated effect of the tax increase is also ambiguous (i.e. $\partial b/\partial t_A = [\partial b/\partial t_A]_{EU_0} - b x_H (\partial b/\partial Z_H) \gtreqless 0$).

4.2.3 Many risky assets

Up to this point attention has been restricted to the case of two assets, of which one is safe and one is risky. In this subsection we briefly discuss how the portfolio model can be extended to the case of many risky assets. The approach adopted here was originally suggested by Sandmo (1977). Just as in the basic model, the utility function depends on final wealth, $\tilde{A}_1$, as in equation (4.41) above, and there is one safe

[9]See Diamond and Yaari (1972), Fischer (1972), and Sandmo (1977) on the derivation of comparative static effects (including Hicksian decomposition) in two-period models with many risky assets. Sandmo (1977, p. 373) computes a Slutsky equation for the demand for the risky asset with respect to a change in the expected rate of return $\bar{x} \equiv \pi_L x_L + (1 - \pi_L) x_H^*$ and finds:

$$\frac{\partial b}{\partial \bar{x}} = \left(\frac{\partial b}{\partial \bar{x}}\right)_{EU_0} + \frac{b}{1+r}\frac{\partial b}{\partial A_0}.$$

The own substitution effect is positive. The income effect is also positive if the risky asset is normal (as is the case with CRRA preferences). Hence, the uncompensated demand slopes upward. This form of the Slutsky equation differs from (4.76) because there only x_H^* is changed, so that both the mean return on the risky asset and its spread around the mean are changed. See also below.

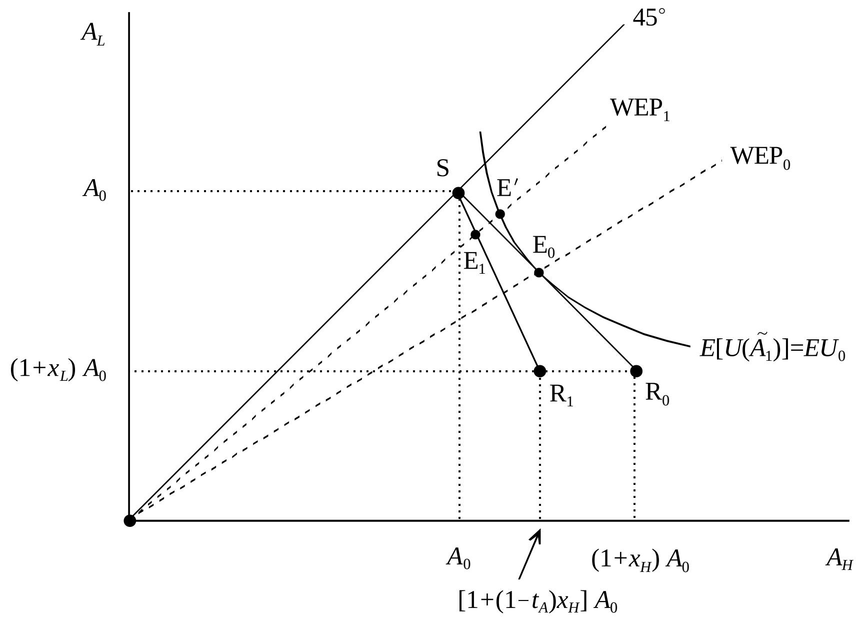

Figure 4.7: No loss offset and the proportional income tax (CRRA preferences)

asset, denoted by m, which carries a certain rate of return r. In contrast, there are now n risky assets, denoted by b_i, with different stochastic return characteristics, i.e. no two assets have perfectly correlated returns. The investor's budget constraint (4.39) is thus generalized to:

$$A_0 = m + \sum_{i=1}^{n} b_i, \tag{4.77}$$

where m and b_i are allowed to be negative (Note that a negative holding of any asset is interpreted as a short sale or debt). The expression for final wealth is:

$$\tilde{A}_1 \equiv (1+r)\,m + \sum_{i=1}^{n} b_i\,(1+\tilde{x}_i)\,, \tag{4.78}$$

where $\tilde{x}_i$ is the stochastic return on risky asset b_i. To prepare for the comparative static exercises to come, the stochastic return on asset b_i is written as:

$$\tilde{x}_i - r = \mu_i + \gamma_i \varepsilon_i, \tag{4.79}$$

where μ_i is a parameter representing the expected (mean) *excess* return on the risky asset, γ_i is a shift parameter (set equal to unity initially), and ε_i is a stochastic term with mean zero, i.e. $E(\varepsilon_i) = 0$ so that it follows from (4.79) that the expected return on the risky asset is $E(\tilde{x}_i) = r + \mu_i$. The advantage of this formulation is that it allows for two interesting types of comparative static experiments. First, by holding constant γ_i $(= 1)$ and changing μ_i, the effects on asset demands of a change in the *mean return* on risky asset b_i can be computed. Second, by holding constant μ_i and changing γ_i the effects on asset demands of an increase in the *riskiness* of b_i can be studied.

The key insight of Sandmo (1977) is that the derivation of comparative static results is straightforward provided the model is rewritten in the standard format (employed, for example, in Chapter 2 above) of (expected) utility maximization subject to a single budget constraint. The transformed model make use of the following auxiliary variables:

$$A_1^C \equiv (1+r)\, m + \sum_{i=1}^{n} b_i \,(1+r+\mu_i)\,, \tag{4.80}$$

$$b_i^* \equiv -\gamma_i b_i, \tag{4.81}$$

$$P_i \equiv \frac{\mu_i}{(1+r)\,\gamma_i}. \tag{4.82}$$

In equation (4.80), A_1^C represents terminal wealth that would be obtained under the investment portfolio (m, b_i) if all assets carried a certain return (i.e. if $\varepsilon_i \equiv 0$ for all i). In the maximization problem, the investor chooses the "commodities" A_1^C and b_i^*, rather than m and b_i. Note that P_i is (4.82) is interpreted as the price of commodity b_i^*.

Using the transformed variables, the expression for final wealth, given in (4.78) above, can be rewritten as follows:

$$\tilde{A}_1 \equiv A_1^C - \sum_{i=1}^{n} b_i^* \varepsilon_i. \tag{4.83}$$

Similarly, the budget constraint (4.77) can be rewritten as:

$$A_0 = \frac{A_1^C}{1+r} + \sum_{i=1}^{n} P_i b_i^*, \tag{4.84}$$

where $1/\,(1+r)$ can thus be seen as the price of commodity A_1^C. The investor chooses A_1^C and b_i^* (for $i = 1, \cdots, n$) in order to maximize expected utility $E(U(\tilde{A}_1))$, subject to the budget constraint (4.84) and the definition of terminal wealth (4.83). Formally this optimization problem has exactly the same structure as the standard consumption choice model. Hence, it follows that the Marshallian demand for b_i^* can be written as $b_i^*\,(1/\,(1+r)\,, P_1, \cdots, P_n, A_0)$ and that the Slutsky equation for b_i^* takes the following form:[10]

$$\frac{\partial b_i^*}{\partial P_j} = \left(\frac{\partial b_i^*}{\partial P_j}\right)_{EU_0} - b_j^* \frac{\partial b_i^*}{\partial A_0}, \tag{4.85}$$

where the compensated derivatives feature the usual properties of (i) negative "own" effects ($(\partial b_i^*/\partial P_j)_{EU_0} < 0$) and Slutsky symmetry ($(\partial b_i^*/\partial P_j)_{EU_0} = (\partial b_j^*/\partial P_i)_{EU_0}$).

By using (4.81) and retransforming, we find that the Marshallian demand for risky asset b_i can be

[10]The Slutsky equation for the standard consumption model is formally derived in an Intermezzo in Chapter 2 above.

written as:

$$b_i = -\frac{1}{\gamma_i} b_i^* \left(1/\left(1+r\right), P_1, \cdots, P_n, A_0\right). \tag{4.86}$$

Armed with this expression we can investigate the comparative static effects of changes in μ_i and/or γ_i. Consider first an increase in the mean return of a risky asset b_j. By differentiating (4.86) with respect to μ_j (keeping $\gamma_i = 1$ for all i), and noting (4.82) and (4.85)-(4.86) we find:

$$\frac{\partial b_i}{\partial \mu_j} = \left(\frac{\partial b_i}{\partial \mu_j}\right)_{EU_0} + \frac{b_j}{1+r}\frac{\partial b_i}{\partial A_0}. \tag{4.87}$$

In the remainder of this subsection we assume (for convenience of exposition) that there are no short positions in risky assets ($b_i > 0$) and that mean excess returns are positive ($\mu_i > 0$ for all $i = 1, \cdots, n$). Equation (4.87) shows that for normal assets ($\partial b_i/\partial A_0 > 0$), the own effect of an increase in the mean return is positive ($\partial b_i/\partial \mu_i > 0$) because both the pure substitution effect (first term on the right-hand side of (4.87)) and the income effect (second term on the right-hand side) are positive.

Next, we consider an increase in the riskiness of asset b_j. By differentiating (4.86) with respect to γ_j (keeping constant μ_i for all i), and noting (4.82) we obtain:

$$\frac{\partial b_i}{\partial \gamma_j} = -\mu_j \frac{\partial b_i}{\partial \mu_j} \qquad (\text{for } i \neq j), \tag{4.88}$$

$$\frac{\partial b_i}{\partial \gamma_i} = -\mu_i \frac{\partial b_i}{\partial \mu_i} - b_i \qquad (\text{for all } i). \tag{4.89}$$

According to (4.88), for assets b_j with a positive expected excess return ($\mu_j > 0$), an increase in their riskiness has the opposite effect (on asset b_i) of a decrease in their average return. The own effect of an increase in riskiness is given in (4.89). For normal assets, an increase in riskiness leads to a decrease in the demand for that asset as both terms on the right-hand side of (4.89) are positive in that case.

4.3 Extensions and applications

In this section we study a number of extensions and applications of the models discussed so far. Throughout this section we abstract from capital risk and instead assume that the household faces *uninsurable income risk*, i.e. its (current or future) income is not known with certainty and there is no income insurance available. In the first subsection, labour supply is exogenous and the effect of income risk on the savings decision is studied. In the second subsection we endogenize the labour supply decision in the face of wage and non-wage income uncertainty. Finally, in the third subsection we briefly study the topic of tax evasion.

4.3.1 Income risk and precautionary saving

How does future income uncertainty affect the current consumption and savings decision? This question was first studied in the context of two-period models by Leland (1968), Sandmo (1970), and Drèze and Modigliani (1972). Leland (1968, p. 465) defines *precautionary saving* as the additional amount of saving that takes place because future income is stochastic rather than deterministic. In this subsection we study the issue of precautionary saving in the context of a simplified version of the model by Sandmo (1970).

To keep things as simple as possible, we assume that the household's lifetime utility function is intertemporally separable as in equation (4.5) above. We abstract from capital risk and assume that there is a single asset carrying a certain rate of return r. Labour supply in both periods is equal to the time endowment, $\bar{L}$. The household budget constraints in the two periods are given by:

$$C_1 + S_1 = w_1 \bar{L}, \tag{4.90}$$

$$\tilde{C}_2 = \tilde{w}_2 \bar{L} + (1+r) S_1, \tag{4.91}$$

where S_1 is saving in the current period, w_1 is the current wage rate (known with certainty), and $\tilde{w}_2$ is the stochastic future wage rate. As a result of wage uncertainty, and in the absence of insurance opportunities, non-interest income is stochastic, as is future consumption. For the comparative statics exercises we follow a similar approach as before, and write the future wage rate as:

$$\tilde{w}_2 \equiv \bar{w}_2 + \gamma \tilde{\varepsilon}, \tag{4.92}$$

where $\bar{w}$ is expected (mean) future wage rate, γ is a shift parameter (set equal to unity initially), and $\tilde{\varepsilon}$ is a stochastic term with mean zero, i.e. $E(\tilde{\varepsilon}) = 0$. The household knows the stochastic distribution of $\tilde{\varepsilon}$. Two cases can be considered. First, by changing $\bar{w}_2$ (holding constant γ), the *life-cycle effect* on current consumption and saving can be studied. Second, by holding constant the expected wage $\bar{w}_2$ and changing γ the *precautionary savings effect* can be analyzed.

Because there is no capital risk, the savings decision constitutes a certain prospect (rather than a temporal uncertain prospect as in Section 4.1 above) so that (4.90)-(4.91) can be consolidated into a single lifetime budget constraint:

$$\tilde{C}_2 = \tilde{w}_2 \bar{L} + \left(w_1 \bar{L} - C_1 \right) (1+r). \tag{4.93}$$

The household chooses current consumption, C_1 (and thus saving), in order to maximize lifetime utility (4.5) subject to the constraint (4.93). The first-order condition for this maximization problem is given by:

$$\frac{dE(\tilde{\Lambda})}{dC_1} = U'(C_1) - \frac{1+r}{1+\rho} E(U'(\tilde{C}_2)) = 0. \tag{4.94}$$

Evaluated at the optimum choice for C_1, the following second-order condition must be satisfied:

$$\frac{d^2E(\tilde{\Lambda})}{dC_1^2} \equiv |\Delta_0| \equiv U''(C_1) + \frac{(1+r)^2}{1+\rho}E(U''(\tilde{C}_2)) < 0. \tag{4.95}$$

For a risk averse household, $U''(z) < 0$ for all z, $|\Delta_0| < 0$, and the solution for C_1 determined by (4.94) is indeed the optimum. Suppressing all variables that are kept constant, we write the solution for current consumption in general terms as $C_1 = C_1(\bar{w}_2, \gamma)$ and determine the partial derivatives in the usual fashion.

4.3.1.1 Life-cycle effect

The life-cycle savings effect can be computed by differentiating the first-order condition (4.94) with respect to C_1 and $\bar{w}_2$ and noting (4.92)-(4.93). After some straightforward steps we obtain:

$$\frac{\partial C_1}{\partial \bar{w}_2} = \frac{(1+r)\,\bar{L}}{1+\rho}\frac{E(U''(\tilde{C}_2))}{|\Delta_0|} > 0, \tag{4.96}$$

where the sign follows from the fact that $|\Delta_0| < 0$ and $E(U''(\tilde{C}_2)) < 0$. Just as in the deterministic case of Chapter 3, current and future consumption are both normal goods. An increase in expected future wage income leads to an increase in current consumption, a decrease in current saving, and an increase in the *expected level* of future consumption. The latter effect can be demonstrated by substituting (4.92) into (4.93), taking expectations (noting $E(\tilde{\varepsilon}) = 0$), and differentiating with respect to $\bar{w}_2$:

$$\frac{\partial E(\tilde{C}_2)}{\partial \bar{w}_2} = \bar{L} - (1+r)\frac{\partial C_1}{\partial \bar{w}_2} = \bar{L}\frac{E(U''(\tilde{C}_2))}{|\Delta_0|} > 0, \tag{4.97}$$

where we have used (4.96) to arrive at the second expression. The property of risk aversion is all we need to ensure normality of both goods.

4.3.1.2 Precautionary savings effect

To compute the precautionary savings effect we substitute (4.92)-(4.93) into (4.94) and differentiate with respect to C_1 and γ. After some straightforward steps we obtain:

$$\frac{\partial C_1}{\partial \gamma} = \frac{(1+r)\,\bar{L}}{1+\rho}\frac{E(U''(\tilde{C}_2)\tilde{\varepsilon})}{|\Delta_0|} \lesseqgtr 0. \tag{4.98}$$

The effect of increased variability of future wage income is ambiguous in general. The sign of the effect is determined by $E(U''(\tilde{C}_2)\tilde{\varepsilon})$ which is ambiguous. It follows from (4.92)-(4.93) that $\tilde{C}_2$ is increasing in $\tilde{\varepsilon}$, but this is not enough information to determine how $U''(\tilde{C}_2)$ varies with $\tilde{\varepsilon}$. The assumption of risk aversion just ensures that $U''(z) < 0$ for all z, but it is silent about the *third* derivative of the felicity function. Clearly, if $U'''(z) > 0$ for all z, then $U''(\tilde{C}_2)$ is increasing in $\tilde{\varepsilon}$ ($U''(\tilde{C}_2)$ and $\tilde{\varepsilon}$ feature a positive correlation), so that $E(U''(\tilde{C}_2)\tilde{\varepsilon}) > 0$ and $\partial C_1/\partial\gamma < 0$ and thus $\partial S_1/\partial\gamma > 0$ (precautionary saving).

In a recent paper, Kimball (1990) formally proves that $U'''(z) > 0$ is a both a necessary and a sufficient condition for the risk-averse household to exhibit a precautionary savings motive. He coins the term *prudence* which is meant to capture the intensity of the precautionary savings motive (1990, p. 54). Defining the *index of absolute prudence* by $P_A(z) \equiv -U'''(z)/U''(z)$, he furthermore demonstrates that, in a model like ours, precautionary saving occurs if and only if $P_A(z) > 0$.[11] He concludes by noting that "...the sign of the third derivative of the utility function governs the presence or absence of a precautionary saving motive just as the sign of the second derivative governs the presence or absence of risk aversion" (1990, p. 68).

The index of absolute prudence can be related to the index of absolute risk aversion, $R_A(z) \equiv -U''(z)/U'(z)$. Indeed, by differentiating $R_A(z)$ we find:

$$R'_A(z) = R_A(z)\left[R_A(z) - P_A(z)\right]. \tag{4.99}$$

Recall from the previous discussion that there are strong reasons to believe that preferences display declining absolute risk aversion (DARA), i.e. $R'_A(z) < 0$. It follows from (4.99) that the DARA property implies positive prudence, i.e. $P_A(z) > R_A(z) > 0$. So if DARA is a reasonable property then so is the property of positive prudence (Kimball, 1990, p. 65).

4.3.2 Labour supply and risk

Block and Heineke (1973) were the first to study the optimal (static) labour supply decision in an environment in which either non-labour income or the wage rate is risky. Their model was subsequently used by Eaton and Rosen (1980a) to study the effect of (optimal) income taxation with a risky wage rate. In this subsection we present a simple version of the labour supply model and study its key features.

To keep things simple, it is assumed that the household utility function is separable in consumption and leisure:

$$E(\tilde{\Lambda}) = E(U(\tilde{C})) + E(V(\bar{L} - L)), \tag{4.100}$$

where C is consumption, L is labour supply, and $\bar{L}$ is the time endowment. The felicity functions feature the usual properties, i.e. $U'(\cdot) > 0 > U''(\cdot)$ and $V'(\cdot) > 0 > V''(\cdot)$. The household chooses its optimal labour supply in the face of uncertainty about its various income sources. Consumption is therefore stochastic. Block and Heineke (1973) study two versions of the model. In the first version of the model, the wage rate is known in advance, but non-wage income is stochastic. For this model, the budget

[11] Note that for the CRRA utility function defined in equation (4.51) we find $P_A(z) = (2 - \gamma_R)/z > 0$. For for the CARA utility function given in (4.57) we find $P_A(z) = \gamma_A > 0$.

constraint can be written as follows:

$$\tilde{C} = wL + \tilde{m}, \tag{4.101}$$

where $\tilde{m}$ is stochastic income from sources other than the labour market.

By substituting (4.101) into (4.100), expected utility depends on the single choice variable, L. The first-order condition is thus given by:

$$\frac{dE(\tilde{\Lambda})}{dL} = wE(U'(wL + \tilde{m})) - V'(\bar{L} - L) = 0, \tag{4.102}$$

where we have used the fact that $E(V'(\bar{L} - L)) = V'(\bar{L} - L)$ and that the wage is known with certainty. According to (4.102), the expected marginal utility of working a little more is equated to the marginal disutility of supplying that additional unit of labour. At the optimum, the following second-order condition is satisfied:

$$\frac{d^2E(\tilde{\Lambda})}{dL^2} \equiv |\Delta_1| = w^2 E(U''(wL + \tilde{m})) + V''(\bar{L} - L) < 0, \tag{4.103}$$

where the sign follows from the assumptions made regarding the felicity functions.

In order to investigate the effect of uncertainty on labour supply, we write non-labour income as:

$$\tilde{m} \equiv \bar{m} + \gamma\tilde{\varepsilon}, \tag{4.104}$$

where $\bar{m}$ is the expected value of $\tilde{m}$, γ is a shift parameter (set equal to unity initially), and $\tilde{\varepsilon}$ is a stochastic term with mean zero, i.e. $E(\tilde{\varepsilon}) = 0$. An increase in expected non-labour income affects labour supply as follows:

$$\frac{\partial L}{\partial \bar{m}} = \frac{wE(U''(\tilde{C}))}{-|\Delta_1|} < 0, \tag{4.105}$$

where the sign follows from the fact that $|\Delta_1| < 0$ and $U''(\tilde{C}) < 0$. For a given value of L, an increase in $\bar{m}$ shifts the distribution of $\tilde{C}$ to the right, and the distribution of $U'(\tilde{C})$ to the left. The expected marginal utility of consumption falls and the household cuts back its labour supply as a result. Leisure is thus a normal good.

An increase in the variability of non-labour income can be studied by holding constant $\bar{m}$ and increasing γ marginally. The effect on labour supply is:

$$\frac{\partial L}{\partial \gamma} = \frac{wE(U''(\tilde{C})\tilde{\varepsilon})}{-|\Delta_1|} \lesseqgtr 0, \tag{4.106}$$

where we have used the fact that $\gamma = 1$ initially. Just as in the case of precautionary savings, the effect of increased riskiness is ambiguous in general. Again, if the third derivative of $U(\tilde{C})$ is positive, and the household displays positive prudence (now with respect to non-labour income fluctuations), then $U''(\tilde{C})$ is increasing in $\tilde{\varepsilon}$, $E(U''(\tilde{C})\tilde{\varepsilon}) > 0$, and labour supply increases as non-labour income becomes more risky, $\partial L/\partial \gamma > 0$.

In the second version of their model, Block and Heineke (1973) assume that non-labour income is constant and the wage rate itself is stochastic. In that case, the household budget constraint is given by:

$$\tilde{C} = \tilde{w}L + m, \tag{4.107}$$

where $\tilde{w}$ is stochastic. The first- and second-order conditions for expected utility maximization are:

$$\frac{dE(\tilde{\Lambda})}{dL} = E(U'(\tilde{w}L + m)\tilde{w}) - V'(\bar{L} - L) = 0, \tag{4.108}$$

$$\frac{d^2E(\tilde{\Lambda})}{dL^2} \equiv |\Delta_2| = E(U''(\tilde{w}L + \tilde{m})\tilde{w}^2) + V''(\bar{L} - L) < 0. \tag{4.109}$$

Equation (4.108) differs substantially from its deterministic counterpart (4.102). Indeed, with a stochastic wage, the optimal labour supply decision depends on a measure of the *covariance* between the wage rate and the marginal utility of consumption, i.e. on $E(U'(\tilde{w}L + m)\tilde{w})$, rather than on the marginal utility of consumption itself. Clearly, the second-order condition is satisfied (given the assumptions made about preferences) because the covariance between $U''(\tilde{C})$ and $\tilde{w}^2$ must be negative, i.e. $E(U''(\tilde{w}L + \tilde{m})\tilde{w}^2) < 0$.

In order to study the effects on labour supply of wage uncertainty, we employ our usual trick and write the wage rate as:

$$\tilde{w} = \bar{w} + \gamma\tilde{\varepsilon}, \tag{4.110}$$

with $E(\tilde{w}) = \bar{w}$ (a constant), $\gamma = 1$ initially, and $E(\tilde{\varepsilon}) = 0$. Recall from the deterministic labour supply model discussed in Chapter 2 that the slope of the uncompensated (Marshallian) labour supply curve is ambiguous due to offsetting income and substitution effects. How does this work in a setting with risk? An increase in the expected wage rate (holding constant $\gamma = 1$) changes labour supply according to:

$$\frac{\partial L}{\partial \bar{w}} = \frac{E(U'(\tilde{C})) + LE(\tilde{w}U''(\tilde{C}))}{-|\Delta_2|} \lesseqgtr 0. \tag{4.111}$$

The ambiguity arises because $E(U'(\tilde{C}))$ is positive whereas the sign of $E(\tilde{w}U''(\tilde{C}))$ is ambiguous. With positive prudence, however, $U''(\tilde{C})$ is increasing in $\tilde{w}$, $E(U''(\tilde{C})\tilde{w}) > 0$, and the uncompensated labour supply curve is an upward sloping function of the expected wage rate, $\partial L/\partial \bar{w} > 0$.[12]

[12]It is not difficult to show that the compensated effect, keeping expected utility constant (by means of a change in non-labour

Finally, an increase in γ (holding constant $\bar{w}$) has the following effect on optimal labour supply:

$$\frac{\partial L}{\partial \gamma} = \frac{E(U'(\tilde{C})\tilde{\varepsilon})}{-|\Delta_2|} + \frac{LE(\tilde{w}U''(\tilde{C})\tilde{\varepsilon})}{-|\Delta_2|} \lesseqgtr 0. \tag{4.112}$$

Block and Heineke (1973, p. 383) call the first and second term on the right-hand side of (4.112), respectively, the *uncertainty substitution effect* and the *income uncertainty effect*. Although it is difficult to provide a simple intuitive interpretation[13] of the two effects, their respective signs can be determined. The uncertainty substitution effect is negative because $\tilde{C}$ depends negatively on $\tilde{\varepsilon}$, so that, since $U''(\tilde{C}) < 0$ it follows that $E(U'(\tilde{C})\tilde{\varepsilon}) < 0$. Under positive prudence, the income uncertainty effect is positive because $U''(\tilde{C})$ is increasing in $\tilde{\varepsilon}$, i.e. $E(\tilde{w}U''(\tilde{C})\tilde{\varepsilon}) > 0$. Hence, unlike the previous case of non-labour income uncertainty, here the prudence property does not automatically imply a positive relationship between labour supply and riskiness.

Inspired by Eaton and Rosen (1980a, pp. 367-368), we use the approach of Rothschild and Stiglitz (1971, p. 67) in order to further investigate the conditions under which $\partial L/\partial \gamma$ is likely to be positive. (This method is discussed in greater detail in the Intermezzo below.) First we define the following function:

$$\Omega_L(L, \tilde{\varepsilon}) \equiv \tilde{w}U'(\tilde{w}L + m) - V'(\bar{L} - L), \tag{4.113}$$

and we note that the first-order condition (4.108) can be written as $E(\Omega_L(L, \tilde{\varepsilon})) = 0$. Note that $\Omega_L(L, \tilde{\varepsilon})$ is monotonically decreasing in L. Rothschild and Stiglitz have shown that a mean-preserving spread in wage income (of which an increase in γ is just a simple example), leads to an increase (decrease) in labour supply if $\Omega_L(L, \tilde{\varepsilon})$ is convex (concave) in $\tilde{\varepsilon}$ (1971, p. 67). Using (4.113) and (4.110) we find the partial derivatives of the $\Omega_L(L, \tilde{\varepsilon})$ function:

$$\frac{\partial \Omega_L(L, \tilde{\varepsilon})}{\partial \tilde{\varepsilon}} = \gamma U'(\tilde{w}L + m) + \gamma L\tilde{w}U''(\tilde{w}L + m), \tag{4.114}$$

$$\frac{\partial^2 \Omega_L(L, \tilde{\varepsilon})}{\partial \tilde{\varepsilon}^2} = \gamma^2 L \left[2U''(\tilde{w}L + m) + \tilde{w}LU'''(\tilde{w}L + m)\right]. \tag{4.115}$$

If $\partial^2 \Omega_L(L, \tilde{\varepsilon})/\partial \tilde{\varepsilon}^2 > 0$ (< 0) *over the entire domain of* $\tilde{w}$, then $\Omega_L(L, \tilde{\varepsilon})$ is strictly convex (concave) and $\partial L/\partial \gamma > 0$ (< 0).[14] Using the index of absolute prudence ($P_A(z) \equiv -U'''(z)/U''(z)$) in (4.115), it is

income m), is given by:

$$\left(\frac{\partial L}{\partial \bar{w}}\right)_{E\Lambda_0} = \frac{E(U'(\tilde{C}))}{-|\Delta_2|} > 0.$$

Menezes and Wang (2005) discuss how duality methods can be employed in the context of a model with wage rate uncertainty. See also Menezes et al. (2005).

[13] Indeed, as was pointed out by Davis (1989, p. 135), the very concept of a pure substitution effect is ambiguous in the context of an increase in risk. This is because there are different compensation methods that can be used, each giving different compensated effects. See also Eaton and Rosen (1980a, p. 369), Dardanoni (1988, p. 438) and Menezes and Wang (2005) for attempts to decompose income and substitution effects.

[14] If $\Omega_L(L, \varepsilon)$ is neither convex nor concave, then the effect of an increase in risk is ambiguous. See Rothschild and Stiglitz (1971, p. 67). Of course, for a given distribution for $\tilde{\varepsilon}$ and a specific utility function, the effect can be computed.

thus possible to derive the following condition for the labour supply response:

$$\frac{\partial L}{\partial \gamma} \gtreqless 0 \qquad \Leftrightarrow \qquad \tilde{w} L P_A(\tilde{C}) \gtreqless 2. \tag{4.116}$$

Eaton and Rosen (1980a, p. 368) derive a similar condition but state it in terms of the coefficient of relative risk aversion, R_R, which they assume to be constant. As (4.116) shows, however, the key determinant in the condition is not so much the attitude toward risk (as captured by the curvature of $U(\tilde{C})$) but rather prudence itself (i.e. the curvature of the marginal felicity function, $U'(\tilde{C})$). If prudence is zero or low, then the household reacts to increased wage risk by cutting back labour supply (dominant uncertainty substitution effect). Conversely, if prudence is high enough (for example, because preferences display DARA and the coefficient of absolute risk aversion is very high), then the income uncertainty effect may dominate and labour supply may increase.

Eaton and Rosen (1980a) consider the effects of proportional labour income taxation on the optimal supply of labour. With such a tax, the household budget constraint (4.101) is modified to:

$$\tilde{C} = \tilde{w}\,(1 - t_L)\,L + m, \tag{4.117}$$

where t_L is the tax rate and we have assumed that other income is untaxed. By substituting (4.110) into (4.117) we find:

$$\tilde{C} = \bar{w}^* L + \gamma^* \tilde{\varepsilon} L + m, \tag{4.118}$$

where $\bar{w}^* \equiv (1 - t_L)\,\bar{w}$ is the expected after-tax wage and $\gamma^* \equiv (1 - t_L)\,\gamma$ is the tax-inclusive risk parameter. An increase in the tax rate thus has two separate effects. On the one hand it reduces the expected after-tax wage rate. This affects labour supply as in equation (4.111) above. On the other hand, the tax increase also reduces the variability of the after-tax wage rate, i.e. the tax partially insures the individual against wage fluctuations. This is of course the Domar-Musgrave result popping up in the context of labour supply. The effects of a decrease in γ^* are covered by equation (4.112) above. Not surprisingly, since $\partial L/\partial \bar{w}$ and $\partial L/\partial \gamma$ are both ambiguous, the overall effect on labour supply of the tax increase is ambiguous too.

Intermezzo 4.2

Economic effects of increased risk. In the text we use an intuitive approach to examine the effects on optimal decision variables when riskiness increases. See, for example, the discussions surrounding equations (4.98), (4.106), and (4.112). Here we briefly discuss the more formal approach of Rothschild and Stiglitz (1970, 1971) and Diamond and Stiglitz (1974). The agent has a control variable X and an objective function $\Omega(X, \tilde{p})$ which depends on both X and some stochastic variable $\tilde{p}$. To keep things simple, $\tilde{p}$ is written as follows:

$$\tilde{p} \equiv \bar{p} + \gamma\tilde{\varepsilon}, \tag{I.1}$$

where $\bar{p}$ is a constant and $\tilde{\varepsilon}$ is a stochastic variable in the interval $[\varepsilon_{MIN}, \varepsilon_{MAX}]$. The density function is $f(\tilde{\varepsilon})$ and the cumulative density function is $F(y)$:

$$F(y) \equiv \int_{\varepsilon_{MIN}}^{y} f(\tilde{\varepsilon})\, d\tilde{\varepsilon}. \tag{I.2}$$

The expected value of $\tilde{\varepsilon}$ is assumed to be zero:

$$E(\tilde{\varepsilon}) \equiv \int_{\varepsilon_{MIN}}^{\varepsilon_{MAX}} \tilde{\varepsilon} f(\tilde{\varepsilon})\, d\tilde{\varepsilon} = 0, \tag{I.3}$$

so that it follows from (I.1) that $E(p) = \bar{p}$. An increase in γ constitutes a mean-preserving spread in $\tilde{p}$.

We write the expected value of the objective function as:

$$\Psi(X) \equiv \int_{\varepsilon_{MIN}}^{\varepsilon_{MAX}} \Omega(X, \tilde{p})\, f(\tilde{\varepsilon})\, d\tilde{\varepsilon}, \tag{I.4}$$

with $\tilde{p}$ given in (I.1) above. The agent chooses X in order to maximize $\Psi(X)$. The first-order condition is given by $\Psi'(X^*) = 0$ or:

$$0 = \int_{\varepsilon_{MIN}}^{\varepsilon_{MAX}} \Omega_X(X^*, \tilde{p})\, f(\tilde{\varepsilon})\, d\tilde{\varepsilon}, \tag{I.5}$$

where $\Omega_X \equiv \partial\Omega(X, \tilde{p})/\partial X$, and X^* is the optimal choice for the control variable (which we assume to be unique). The second-order condition for a maximum is that $\Omega_{XX}(X, \tilde{p}) < 0$ for $X = X^*$. Assume that $\Omega_{XX}(X, \tilde{p}) < 0$ also holds in the neighbourhood of X^*. Rothschild and Stiglitz (1971, p. 67) state the following results.

(P1) If $\Omega_X(X^*, \tilde{p})$ is a concave (convex) function of $\tilde{\varepsilon}$, i.e. $\partial\Omega_X(X^*, \tilde{p})/\partial\tilde{\varepsilon}^2 < 0\ (> 0)$, then an increase in riskiness will decrease (increase) X^*.

(P2) If $\Omega_X(X^*, \tilde{p})$ is neither convex nor concave function in $\tilde{\varepsilon}$, then the effect of an increase in riskiness on X^* is ambiguous.

In the text we have used (P1) in detail in the discussion surrounding equation (4.112), where L is the control variable and $\tilde{p}$ is the wage rate. To determine the sign of the precautionary savings effect in (4.98), we note that the control variable is now current consumption ($X = C_1$), and that the first-order condition (4.94) can be written as $E(\Omega_X(X^*, \tilde{p})) = 0$, where $\tilde{p} = \tilde{w}_2$ the future wage rate, and $\Omega_X(X^*, \tilde{p})$ is defined as:

$$\Omega_X(X^*, \tilde{p}) \equiv U'(X^*) - \frac{1+r}{1+\rho} U'(\tilde{p}\bar{L} + (w_1\bar{L} - X^*)(1+r)). \tag{I.6}$$

Concavity or convexity of $\Omega_X(X^*, \tilde{p})$ in $\tilde{p}$ clearly depends on the sign of $U'''(\cdot)$ as is argued intuitively below (4.98). The discussion below equation (4.106) can be similarly understood in terms of (P1).

****

4.3.3 Tax evasion

In an interesting application of the economic theory of decision making under risk, Allingham and Sandmo (1972) study the issue of income tax evasion. Their paper makes use of insights from the economics of crime, a field of study that was pioneered by Becker (1968). How much of its income will a risk averse household declare for tax purposes if the tax authority is unable to perfectly monitor this household's income? On the one hand, by under-reporting its income, the household evades some taxes it would otherwise have had to pay. This is the *crime*. On the other hand, the tax authority may find out about the household's undeclared income and levy a penalty exceeding the evaded tax. This is the *punishment*.

Even though the household's before-tax income is deterministic, the tax declaration decision is a decision under risk because the probability of getting caught is less than one. To study the key trade-off facing the household, consider the following simple model. The household's actual before-tax income is exogenously given and equal to W. The household declares an amount X to the tax agency and pays a proportional tax, t_Y, on this declared income. With probability π the household is investigated by the tax agency who will then observe W. If W exceeds X then the household pays tax on the undeclared part of income, $W - X$ (≥ 0), at the penalty rate t_P, which exceeds t_Y.[15] With probability $1 - \pi$ the household is not investigated (and thus gets way with the crime). Consumption of the household is thus stochastic and given by:

$$\tilde{C} = \begin{cases} C_L \equiv W - t_Y X - t_P (W - X) & \text{with probability } \pi \\ \\ C_H \equiv W - t_Y X & \text{with probability } 1 - \pi \end{cases}. \tag{4.119}$$

[15]If the household reports too much income ($W - X < 0$), then it receives a rebate at the statutory tax rate, t_Y, if it is monitored by the tax authority.

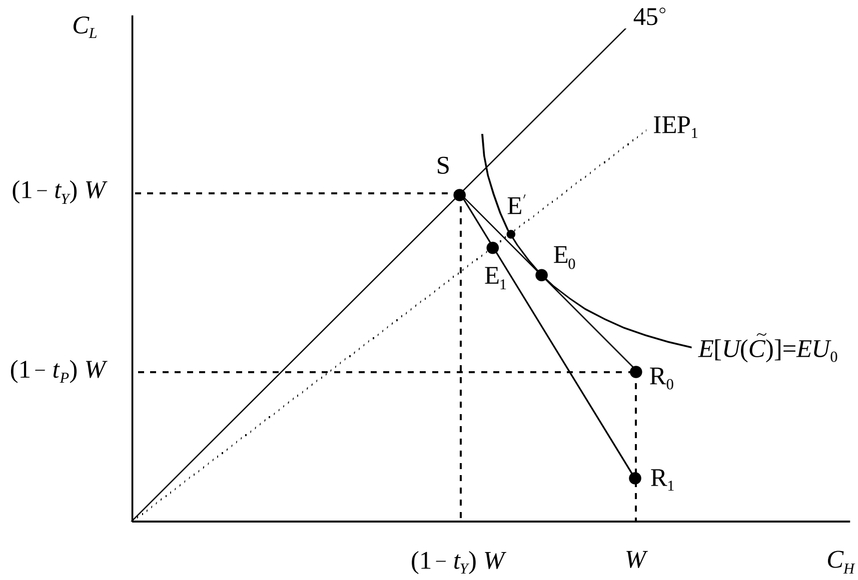

Figure 4.8: Income tax evasion and the penalty rate

The household is risk averse and chooses its control variable, X, in order to maximize its expected utility:

$$E(U(\tilde{C})) \equiv \pi U(C_L) + (1-\pi)\,U(C_H), \tag{4.120}$$

where $U'(\cdot) > 0 > U''(\cdot)$. The first-order condition for an internal solution to the maximization problem is given by:

$$\begin{aligned}\frac{dE(U(\tilde{C}))}{dX} &= \pi U'(C_L)\frac{dC_L}{dX} + (1-\pi)\,U'(C_H)\frac{dC_H}{dX} \\ &= \pi U'(C_L)\,(t_P - t_Y) - (1-\pi)\,U'(C_H)t_Y = 0.\end{aligned} \tag{4.121}$$

At the optimum, the following second-order condition should also be satisfied:

$$\frac{d^2E(U(\tilde{C}))}{dX^2} \equiv |\Delta_3| = \pi\,(t_P - t_Y)^2\,U''(C_L) + (1-\pi)\,t_Y^2U''(C_H) < 0, \tag{4.122}$$

where the sign follows from the fact that $U''(\cdot) < 0$.

The optimal choice for X can be illustrated with the aid of Figure 4.8. Consumption in the two states is measured on the two axes and the budget line is given by SR$_0$. The perfectly safe point S is attained if the household declares its total pre-tax income ($X = W$). Regardless of whether it will be monitored, it will have a certain consumption level equal to after-tax income, $(1-t_Y)\,W$. The perfectly risky point is at point R$_0$. At that point the household does not declare any income for tax purposes at all, i.e. $X = 0$ and consumption is either W if it is not monitored (with probability $1-\pi$) or $(1-t_P)\,W$ if it is "caught red-handed" (with probability π).

By using equation (4.119), and eliminating X, the budget equation can be written as:

$$C_L = \frac{t_P (1 - t_Y) W}{t_Y} - \frac{t_P - t_Y}{t_Y} C_H. \tag{4.123}$$

The slope of the budget line passing through points S and R_0 is thus given by $-(t_P - t_Y)/t_Y$. By using (4.121) we find that the optimum occurs at point E_0, where there is a tangency between an indifference curve and the budget line SR_0. At point E_0 the following condition is satisfied:

$$\frac{1-\pi}{\pi} \frac{U'(W - t_Y X)}{U'((1-t_P) W + (t_P - t_Y) X)} = \frac{t_P - t_Y}{t_Y}, \tag{4.124}$$

where we have substituted the definitions of C_L and C_H from (4.119). We note for future reference that the horizontal distance between point R_0 and point E_0 is equal to $t_Y X$.

The effect of the penalty rate on the optimal amount of declared income can be investigated by differentiating the first-order condition (4.124) with respect to t_P:

$$\frac{\partial X}{\partial t_P} = \frac{\pi \left[U'(C_L) - (W - X)(t_P - t_Y) U''(C_L)\right]}{-|\Delta_3|} > 0. \tag{4.125}$$

An increase in the penalty rate thus increases the amount of declared income, as one would expect. In terms of Figure 4.8, the budget line rotates from SR_0 to SR_1 and there are both income and substitution effects. As (4.125) shows, however, the total effect is unambiguous because the two effects change the optimal X in the same direction. In Figure 4.8 we illustrate the Slutsky decomposition for the case of CRRA preferences. After the tax increase, the new income expansion path is given by IEP_1. The substitution effect is the move from E_0 to E', whilst the income effect is the move from E' to E_1. Since, for a given income tax rate, any horizontal move to the left represents an increase in X, it follows that both effects work in the direction of increased declared income.

4.4 Punchlines

In this chapter we study household decision making under conditions of risk. Two main types of risk are considered, namely *capital risk* and *income risk*. With capital risk the yield on an asset is unknown in advance, though the household typically knows its stochastic distribution. With income risk, asset yields are known but income itself is stochastic.

Faced with risky prospects, the household is assumed to possess a so-called von-Neumann-Morgenstern (vNM) utility function which depends on the possible outcomes. Throughout the chapter we employ the *expected utility hypothesis*, according to which the household acts in such a way as to maximize the mathematical expectation of this vNM utility function, using the probabilities of different outcomes (or *states*) for weighting purposes. Although the expected utility hypothesis is not without its detractors, it remains the dominant hypothesis for the analysis of economic decision making under risk.

In the first section of this chapter we study a simple two-period model in which there is no income risk but the household must make an optimal decision concerning its investment portfolio in the presence of one risk-free asset and one risky (but potentially high-yield) asset. In addition to the *portfolio decision*, the household must also make an intertemporal decision, i.e. the decision when and how much to consume during its lifetime (*savings decision*). We use this capital risk model to study the effects of interest income taxation. Depending on the details of the tax system, it is possible that risk taking increases if the tax on interest income is increased! This is the famous Domar-Musgrave result. On the one hand the government takes a share of the expected return but on the other hand (with perfect loss offsets) it also shares in the risk of losses.

In the second section we abstract from the savings decision and zoom in on the pure portfolio decision. In doing so we are able to study in greater detail the various aspects of often-used functional forms of the vNM utility function. Furthermore, by restricting the number of possible states to two, a simple graphical apparatus can be used to visualize the results and to emphasize the similarity to the deterministic case. Throughout this chapter attention is restricted to the case of *risk averse* decision makers. Such a decision maker prefers the certain outcome A over any risky prospect for which the mathematical expectation is equal to A, i.e. to take on risk he demands a *risk premium*. Mathematically this means that the risk averter's vNM utility function is a concave function of the stochastic outcome. We discuss the key features of the optimal portfolio decision for two often-used functional forms, namely the CRRA and CARA forms, which differ in the way in which the degree of concavity of the vNM utility function is modelled.

The third section of this chapter studies some extensions and applications which abstract from capital risk and instead focus on various types of income risk. The first application augments the two-period consumption-saving model (of the first section) by postulating the existence of a single asset carrying a certain return and by assuming that future wage income is risky. In this context the households has two motives for saving, namely a *life-cycle* motive (smoothing consumption over time) and a *precautionary* motive (accounting for the riskiness of future income). The assumption of risk aversion is sufficient to guarantee an operative life-cycle motive. In contrast, the strength of the precautionary motive is regulated by so-called *prudence*, which measures the curvature of the *marginal* utility function. A prudent household will react to increased riskiness by saving more.

Next we turn to the reaction of optimal labour supply when either non-labour income or the wage rate itself is risky. Again the degree of prudence determines the reaction of the decision variable (labour supply) to increased riskiness in either non-labour income or the wage rate. A prudent worker will react to both types of risk increases by supplying more hours.

Finally, we conclude section three of this chapter by applying the expected utility theory to the household's decision concerning its income tax declaration. If the tax authority cannot monitor the household's income perfectly, the probability of being found out when cheating is less than unity and the household faces a stochastic income unless it is so risk averse as to stick to the perfectly safe option of declaring its income truthfully. In the internal optimum, the moderately risk averse household takes on some risk, and evades some of the taxes it would otherwise have to pay. The amount of the tax that is thus evaded can, however, be reduced by the tax authority by increasing the penalty rate levied on undeclared income that is discovered (with non-zero probability) upon auditing.

Further reading

General treatments of decision making under risk. There are several excellent books on the economics of risk and uncertainty. See, for example, Hirshleifer and Riley (1992), Gollier (2001), and Eeckhoudt, Gollier, and Schlesinger (2005). The first two are rather advanced whilst the third is very accessible. Brief introductions to the expected utility hypothesis are Schmeidler and Wakker (1987) and Machina (1987). Kreps (1988) is an advanced introduction to the theory of individual decision making both with and without risk.

Two-period model with risk. Early papers using the two-period model with income and/or capital risk are Leland (1968), Sandmo (1969, 1970, 1974b), Drèze and Modigliani (1972), and Ahsan (1976). Recent contributions include Barsky, Mankiw, and Zeldes (1986), Kimball (1990), Elmendorf and Kimball (2000), and Aura, Diamond, and Geanakoplos (2002). Block and Heineke (1975) model the joint saving-labour supply decision.

Portfolio model. The key sources for the pure portfolio model are Markowitz (1952), Tobin (1958), Arrow (1965, 1971b), Mossin (1968), Stiglitz (1969), and Sandmo (1974b, 1977). See also Atkinson and Stiglitz (1980, lecture 4) and Sandmo (1985, pp. 293-299) for a further discussion of the portfolio approach. For an empirical study on the effects of income risk on the portfolio choice, see Guiso, Jappelli, and Terlizzese (1996).

Labour supply and risk. Early papers on the labour supply decision under risk are Block and Heineke (1973), Eaton and Rosen (1980a), Coyte (1986), Dardanoni (1988), and Bodie, Merton, and Samuelson (1992). Recent papers include Hartwick (2000), and Menezes and Wang (2005). Kanbur (1981) models the occupational choice as an example of risk taking (safe versus risky professions).

Human capital. Interesting papers on the household's human capital investment decision under conditions of risk are Levhari and Weiss (1974), Eaton and Rosen (1980b), Hamilton (1987), Judd (1998), and Anderberg and Andersson (2003).

Tax evasion. A classic paper on the economics of crime is Becker (1968). Applications to the issue of income tax evasion are by Allingham and Sandmo (1972), Yitzhaki (1974), and Baldry (1979). See also the recent survey by Slemrod and Yitzhaki (2002).

Chapter 5

Taxation and the firm

The purpose of this chapter is to discuss the following topics:

- What are the main taxes that are typically levied on firms?
- How do taxes affect the behaviour of perfectly competitive firms in a static setting?
- How do the results change if one allows for imperfectly competitive firms?
- What are the main determinants of the cost of capital if the firm determines its financial structure in an optimal forward-looking fashion?
- How do taxes affect the firm's optimal dynamic plans when investment is subject to adjustment costs and expectations are fulfilled (perfect foresight)?

5.1 A basic static model of firm behaviour

As is pointed out by Atkinson and Stiglitz (1980, pp. 129-130), a large number of different taxes (may) affect the firm.[1] Such taxes may be specific or ad valorem. A first group of taxes are those levied on factors of production. Examples include the payroll tax, where the tax base is the wage bill, and the corporate profits tax, which can be seen as a tax on the return on capital. In recent years, taxes on the use of energy have been imposed to induce firms to engage in cleaner (more fuel-efficient) production methods. Negative taxes (or subsidies) are also observed in many countries. Examples are wage subsidies, e.g. on unskilled labour, and the investment tax credit which is a kind of subsidy on capital accumulation by firms.

A second group of taxes affecting the firm are those levied on output. Examples are the value-added tax and the production tax. As we shall see in this chapter, a third group of taxes that may affect the firm

[1] We define the *firm* in the traditional way as a profit maximizing agent facing a known technology and operating subject to a well-defined market constraint. See Archibald (1987) for a discussion of alternative concepts. See also Roberts (2004) for a discussion of modern theories of the firm.

are actually taxes levied on households, e.g. the capital gains tax or the tax on dividend income. Such taxes may affect the cost of capital the firm is faced with in the market!

In this section we develop some basic *partial equilibrium* models of firm behaviour in a *static* setting. General equilibrium effects of firm taxation are studied in Chapter 6 below whilst some intertemporal considerations are discussed in later sections of this chapter. In the first subsection we study the behaviour of a representative perfectly competitive firm, focusing on factor substitution effects. In the second subsection we study tax effects in simple models of monopoly and oligopolistic competition.

5.1.1 Perfect competition

In this subsection we study the behaviour of a representative perfectly competitive firm. The basic assumptions that are employed in the model are the following. First, the firm produces a single homogeneous product under conditions of perfect competition on both its input markets and the market for its output (price taking behaviour). Second, technology available to the firm features constant returns to scale. Third, we abstract from adjustment costs and assume that all factors of production are freely adjustable to the firm. (In Section 5.3 below we introduce adjustment costs of investment which renders the firm's capital stock a quasi-fixed factor of production.)

In order to develop a simple graphical apparatus we first study the case in which there are only two factors of production, namely labour, L, and physical capital, K. Profit of the firm is defined as the difference between revenue and total costs:

$$\Pi \equiv PF(K,L) - WL - R^K K, \tag{5.1}$$

where Π is profit, P is the goods price, $F(\cdot)$ is a linear homogeneous production function, W is the wage rate, and R^K is the rental rate on capital.[2] The production function is strictly quasi-concave in its arguments, features positive but diminishing returns to each production factor, and possesses isoquants which bulge towards the origin. In technical terms, the properties are given by:

$$\begin{array}{ll} F_L \equiv \frac{\partial F}{\partial L} > 0, & F_K \equiv \frac{\partial F}{\partial K} > 0, \\ F_{LL} \equiv \frac{\partial^2 F}{\partial L^2} < 0, & F_{KK} \equiv \frac{\partial^2 F}{\partial K^2} < 0, \\ F_{KL} = F_{LK} \equiv \frac{\partial^2 F}{\partial L \partial K} > 0, & F_{LL}F_{KK} - F_{KL}^2 > 0. \end{array} \tag{5.2}$$

Note that in this particular case, with only two factors of production and constant returns to scale, capital and labour must be *cooperative factors* of production, i.e. the marginal product of one factor increases as the use of the other factor is increased.

The firm chooses its inputs of capital and labour in such a way that profit is maximized, taking as given the available technology, the output price, and the input prices. The first-order necessary condi-

[2] It is straightforward to distinguish multiple types of capital or labour with associated wage and rental rates.

tions are given by:

$$PF_K(K,L) = R^K, \tag{5.3}$$

$$PF_L(K,L) = W. \tag{5.4}$$

For each factor of production, the value of the marginal product is equated to the rental rate. Of course, with constant returns to scale, equations (5.3)-(5.4) do not pin down a unique profit maximizing output *level* for the firm. For any output level, the firm will choose the capital-labour ratio such that (5.3)-(5.4) hold. In terms of Figure 5.1, for a given output level Y_0, the firm chooses its input mix such that the marginal rate of technical substitution between the two factors (i.e. F_L/F_K) is equated to the relative price of those two factors (i.e. W/R^K). This occurs at point A.

An equivalent way to describe the firm's behaviour recognizes the fact that a profit maximizing firm must necessarily produce in a cost-minimizing fashion. To elaborate on this result, we write profit in the following fashion:

$$\Pi = PY - c\left(W, R^K\right) Y, \tag{5.5}$$

where $c\left(W, R^K\right)$ is the *unit cost function*, i.e. $c\left(W, R^K\right) \equiv \min WL + R^K K$ subject to $F(K,L) = 1$, and $c\left(W, R^K\right) Y$ is total cost. The *conditional input demand* functions are obtained by using Shephard's Lemma (see Intermezzo 5.1):

$$K = \frac{\partial c\left(W, R^K\right)}{\partial R^K} Y_0, \qquad L = \frac{\partial c\left(W, R^K\right)}{\partial W} Y_0. \tag{5.6}$$

These conditional demands also describe point A in Figure 5.1.[3]

Intermezzo 5.1

The cost function. The cost function is yet another very useful tool from duality theory which will be used time and again. The cost function is analogous to the *expenditure function* discussed in the household model of Chapter 2. There is a single output, Y, which is produced by using n production factors. Factor Z_i carries a rental rate of W_i (for $i = 1, \cdots, n$). Formally, the cost function represents the minimum level of factor costs needed to produce a given level of output, say Y_0, when faced with the rental rates W_i:

$$C(\mathbf{w}, Y_0) \equiv \min_{\{Z_i\}} \sum_{i=1}^{n} W_i Z_i \quad \text{subject to:} \quad F(Z_1, \cdots, Z_n) = Y_0,$$

[3] This can be seen by noting that the first-order conditions underlying cost minimization are $W = \lambda F_L$ and $R^K = \lambda F_K$, where λ is the Lagrange multiplier for the production constraint. Combining the two conditions yields $F_L/F_K = W/R^K$.

where $\mathbf{w} \equiv (W_1, \cdots, W_n)$ is the vector of factor prices, $\mathbf{Z} \equiv (Z_1, \cdots, Z_n)$ is the vector of inputs, and $F(\mathbf{Z}) \equiv F(Z_1, \cdots, Z_n)$ is the strictly quasi-concave production function.

The following key properties of the cost function can be established (McFadden, 1978, pp. 47-48).

(P1) $C(\mathbf{w}, Y_0)$ is homogeneous of degree one in rental rates ($C(\lambda \mathbf{w}, Y_0) = \lambda C(\mathbf{w}, Y_0)$ for $\lambda > 0$);

(P2) $C(\mathbf{w}, Y_0)$ is concave in rental rates ($C(\lambda \mathbf{w}^0 + (1-\lambda)\mathbf{w}^1, Y_0) \geq \lambda C(\mathbf{w}^0, Y_0) + (1-\lambda) C(\mathbf{w}^1, Y_0)$ for $0 \leq \lambda \leq 1$);

(P3) $C(\mathbf{w}, Y_0)$ is non-decreasing in $\mathbf{w}$ (if $\mathbf{w}^1 > \mathbf{w}^0$ then $C(\mathbf{w}^1, Y_0) > C(\mathbf{w}^0, Y_0)$);

(P4) $C(\mathbf{w}, Y_0)$ is continuous in $\mathbf{w}$;

(P5) If $F(\mathbf{Z})$ features CRTS then the cost function is linear in Y_0 and can be written as $C(\cdot) = c(\mathbf{w}) Y_0$ (where $c(\mathbf{w})$ is unit-cost);

(P6) The Allen-Uzawa substitution elasticity between factors i and j is defined as:

$$\sigma_{ij} \equiv \frac{C(\cdot) C_{ij}(\cdot)}{C_i(\cdot) C_j(\cdot)}, \qquad i \neq j, \tag{I.1}$$

where $C_i \equiv \partial C(\cdot)/\partial W_i$ and $C_{ij} \equiv \partial^2 C(\cdot)/\partial W_i \partial W_j$.

The *conditional factor demand curves* are given by Shephard's Lemma:

$$Z_i(\mathbf{w}, Y_0) = \frac{\partial C(\mathbf{w}, Y_0)}{\partial W_i}. \tag{I.2}$$

Properties of the conditional factor demands are:

(P7) $Z_i(\mathbf{w}, Y_0)$ is decreasing in W_i, $\partial Z_i(\mathbf{w}, Y_0)/\partial W_i = \partial^2 C(\mathbf{w}, Y_0)/\partial W_i^2 < 0$;

(P8) Cross-price effects are symmetric, $\partial Z_i(\mathbf{w}, Y_0)/\partial W_j = \partial Z_j(\mathbf{w}, Y_0)/\partial W_i$ for all i, j; Inputs i and j are substitutes (complements) if $\partial Z_i(\mathbf{w}, Y_0)/\partial W_j > 0$ (< 0);

(P9) $Z_i(\mathbf{w}, Y_0)$ is homogeneous of degree zero in rental rates $(W_1, \cdots, W_n)$;

(P10) If $F(\mathbf{Z})$ features constant returns to scale then, since $C(\cdot) = c(\mathbf{w}) Y_0$, the conditional factor demands simplify to:

$$Z_i(\mathbf{w}, Y_0) = \frac{\partial c(\mathbf{w})}{\partial W_i} Y_0. \tag{I.3}$$

****

How do taxes affect the firm's objective function and its first-order conditions? First consider the case of an ad valorem payroll tax, t_P. Using the direct approach, profit is amended as follows:

$$\Pi \equiv PF(K,L) - W(1+t_P)L - R^K K, \tag{5.7}$$

and the first-order conditions (5.3)-(5.4) are changed to:

$$PF_K(K,L) = R^K, \tag{5.8}$$
$$PF_L(K,L) = W(1+t_P). \tag{5.9}$$

Holding constant output, an increase in the payroll tax induces a factor substitution effect, i.e. the firm will want to adopt a higher capital-labour ratio. In terms of Figure 5.1, the relative rental rate on capital ($R^K/W(1+t_P)$) falls and the firm moves from point A to point B.

The same conclusion can also be derived by means of the dual approach. The payroll tax pushes up the rental rate on labour, and it follows from the properties of the conditional factor demand functions (5.6) that:

$$\frac{\partial L}{\partial W} = \frac{\partial^2 c(W,R^K)}{\partial W^2} Y_0 < 0, \tag{5.10}$$
$$\frac{\partial K}{\partial W} = \frac{\partial^2 c(W,R^K)}{\partial R^K \partial W} Y_0 = -\frac{W}{R^K}\frac{\partial^2 c(W,R^K)}{\partial W^2} Y_0 > 0, \tag{5.11}$$

where we have made use of the fact that factor demand functions are homogeneous of degree zero in input prices to arrive at the second expression in (5.11). In the two-factor case under consideration here, we reach the rather unsurprising conclusion that the factors must be substitutes.

As a second example, consider an ad valorem output tax, t_Y. Profit (5.1) is now given by:

$$\Pi \equiv (1+t_Y)PF(K,L) - WL - R^K K, \tag{5.12}$$

and the first-order conditions are modified to:

$$PF_K(K,L) = \frac{R^K}{1+t_Y}, \tag{5.13}$$
$$PF_L(K,L) = \frac{W}{1+t_Y}. \tag{5.14}$$

In this case there is no factor substitution effect. The relative rental rate on capital is unchanged so that, holding constant output, there is no effect on the capital labour ratio either. The output tax acts as a common tax on both factors of production. The same conclusion follows from the input demand functions (5.6) which are homogeneous of degree zero in the input prices.

The third example considers a specific subsidy on labour, s_L, for which the profit expression and the

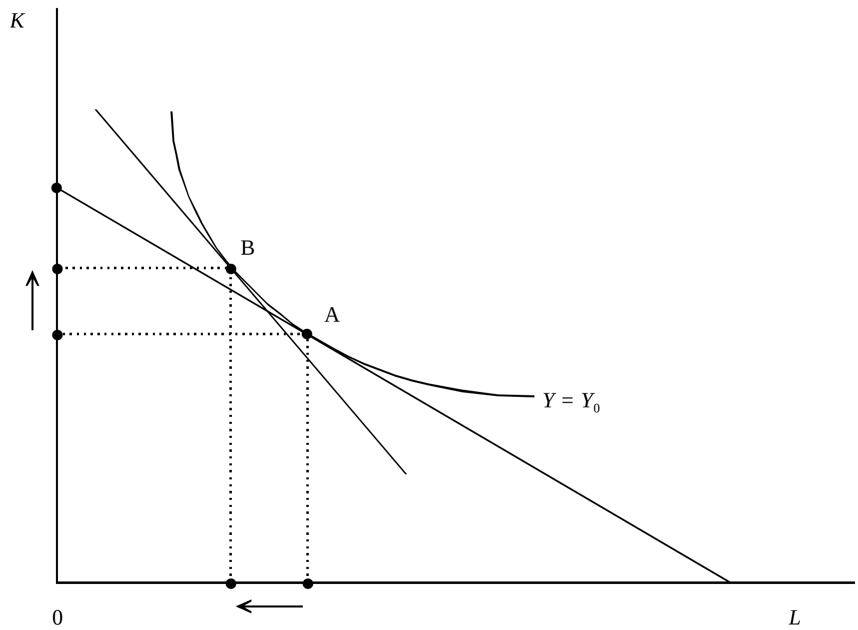

Figure 5.1: Factor Substitution Effect

first-order conditions are modified to:

$$\Pi \equiv PF(K,L) - (W - s_L)L - R^K K, \tag{5.15}$$

$$PF_K(K,L) = R^K, \tag{5.16}$$

$$PF_L(K,L) = W - s_L. \tag{5.17}$$

Two things are worth noting about this case. First, there is a factor substitution effect in the direction of a lower capital-labour ratio (labour becomes cheaper vis-a-vis capital). Second, the specific subsidy is equivalent to an ad valorem subsidy equal to $\bar{s}_L \equiv s_L/W$ (since $W - s_L$ equals $(1-\bar{s}_L)W$ in that case).

As a final example consider a tax on pure profit, t_K, which changes the profit definition to:

$$\Pi \equiv (1-t_K)\left[PF(K,L) - WL - R^K K\right]. \tag{5.18}$$

Clearly, the first-order conditions for profit maximization are still as in (5.3)-(5.4), i.e. there is no factor substitution effect. The relative rental rate on capital is unaffected so there is no effect on the capital-labour ratio either. If all firms are exactly the same, then they all make zero (excess) profits, the term in square brackets on the right-hand side of (5.18) is zero, and the tax raises no revenue at all. In the alternative (and more realistic) case where firms are heterogeneous, some firms may make positive profits whilst others may incur losses. In the long run, however, the loss-making firms will exit and the marginal firm will be the one making exactly zero profits. By definition, the profit tax will not affect that marginal firm at all. Hence, long-run output will not be affected under this scenario either. The profit tax falls on pure economic profit, i.e. on the "return on entrepreneurship". As Atkinson and Stiglitz (1980, p. 132) point out, however, the validity of this "Marshallian" view on the profit tax is critically dependent on the definition of the tax base. Indeed, as is clear from (5.18), the tax base excludes the cost of capital, $R^K K$, and thus only applies to pure profits. In reality matters are much more complex—see Sections 5.2 and 5.3 for a further study of the effects of corporate taxation.

How do changes in tax rates (or input prices) affect supply in a competitive market? We answer this question for the *symmetric case*, in which all firms possess the same technology and thus face the same cost function. Output of firm i is denoted by Y_i and profit of firm i is given by:

$$\Pi_i \equiv \left[P - c\left(W\left(1+t_P\right), R^K\right)\right] Y_i, \tag{5.19}$$

where t_P is the payroll tax (all other taxes are abstracted from) and $c(\cdot)\,Y_i$ is total cost of firm i. Profit maximization ($d\Pi_i/dY_i = 0$) yields the familiar condition equating price to marginal cost of production:

$$P = c\left(W^*, R^K\right), \tag{5.20}$$

where $W^* \equiv W\left(1+t_P\right)$. Note that the scale of each firm is irrelevant in this constant-returns-to-scale setting. Provided the firms act as perfect competitors, total output in the market can be produced either by one very large firm or by many small firms. The supply curve is horizontal at the price level stated in equation (5.20). Note furthermore, that all active firms will make zero excess profits under marginal cost pricing.

Armed with equation (5.20) we can investigate how supply reacts to a change in the payroll tax:

$$\frac{\partial P}{\partial t_P} = \frac{\partial c}{\partial t_P} = \frac{\partial c\left(W^*, R^K\right)}{\partial W^*}\frac{\partial W^*}{\partial t_P} = \frac{W L_i}{Y_i} > 0, \tag{5.21}$$

where we have used Shephard's Lemma in the final step. An increase in the payroll tax raises the cost of labour to producers and increases the marginal cost of production. As a result, the competitive supply curve is shifted up, more so the larger is L_i/Y_i, i.e. the more important labour is in the production process.

5.1.2 Imperfect competition

At the end of the previous subsection we demonstrated that in a competitive setting, any cost increase induces a one-for-one increase in the supply price of the commodity.[4] In this subsection we show that this result is no longer valid if the goods market is imperfectly competitive. Indeed, in such a setting, cost changes typically lead to more than one-for-one price changes, i.e. *price overshifting* occurs. To demonstrate this important phenomenon we consider two cases of imperfectly competitive behaviour,

[4] The general equilibrium repercussions of cost changes in a competitive model are studied in Chapter 6 below.

namely the *monopoly* case, in which there is a single supplier in the market, and the *oligopoly* case, in which there are a number of firms in the market reacting strategically to each other's output decisions. We continue to assume that factor prices are exogenous, i.e. the models are partial equilibrium in nature. To facilitate comparison with the competitive case we focus attention on the effects of a payroll tax.

5.1.2.1 Monopoly

Assume that firm i is the monopolistic producer of the good. Profit of firm i is given by:

$$\Pi_i \equiv P(Y_i) Y_i - c\left(W^*, R^K\right) Y_i, \tag{5.22}$$

where $P_i = P(Y_i)$ is the *inverse* demand function for the good and P_i is the price set by firm i. The demand curve is downward sloping, $P'(Y_i) < 0$, and the demand elasticity is greater than unity in absolute value:

$$\varepsilon(Y_i) \equiv \left|\frac{\partial Y_i}{\partial P_i}\frac{P_i}{Y_i}\right| = \left|\frac{P(Y_i)}{Y_i P'(Y_i)}\right| > 1. \tag{5.23}$$

Firm i chooses its output in order to maximize its profit. The first-order condition is given by:

$$\frac{d\Pi_i}{dY_i} \equiv P(Y_i) + Y_i P'(Y_i) - c\left(W^*, R^K\right) = 0, \tag{5.24}$$

whilst the second-order condition is:

$$\frac{d^2\Pi_i}{dY_i^2} \equiv 2P'(Y_i) + Y_i P''(Y_i) < 0. \tag{5.25}$$

By noting (5.23), it is easy to deduce that the profit maximizing output level induces an optimal price which is equal to a *markup* times marginal cost of production:

$$P_i = \frac{c\left(W^*, R^K\right)}{1 - 1/\varepsilon(Y_i)}. \tag{5.26}$$

Since, the demand elasticity is greater than unity, it follows that the markup also exceeds unity ($1/\left[1 - 1/\varepsilon(Y_i)\right] > 1$), i.e. price exceeds marginal cost ($P_i > c\left(W^*, R^K\right)$).

How does the monopolistic producer react to an increase in the payroll tax which leads to an increase in marginal cost? By using (5.24) and applying the implicit function theorem we find that optimal output decreases:

$$\frac{\partial Y_i}{\partial t_P} = \frac{1}{2P'(Y_i) + Y_i P''(Y_i)}\frac{\partial c}{\partial t_P} < 0, \tag{5.27}$$

where the sign follows from the second-order condition for profit maximization given in (5.25) above.

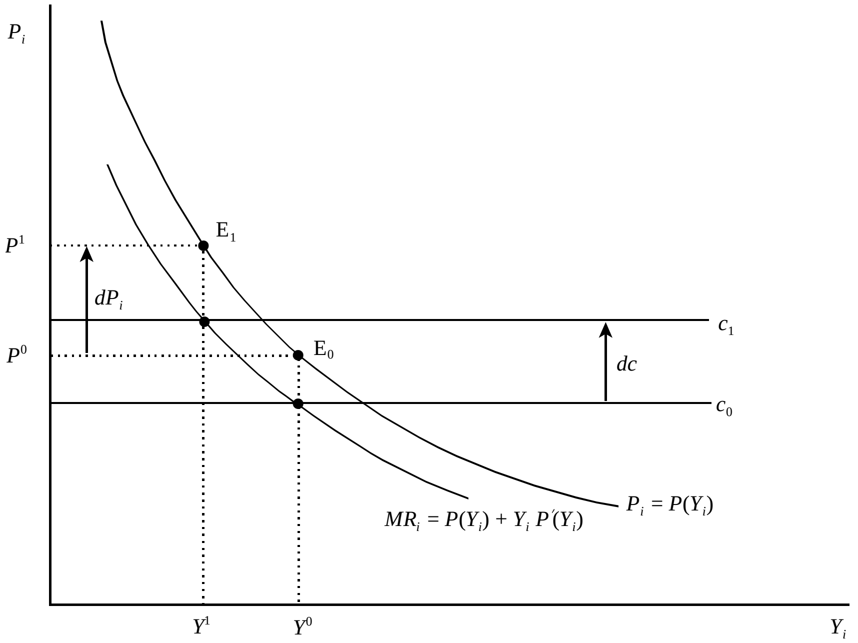

Figure 5.2: Overshifting of cost changes under monopoly

Since the demand curve facing the monopolist is downward sloping, the reduction in output leads to an increase in the price, i.e. some price shifting occurs. Indeed, as was stressed by Seade (1985), *over*shifting is quite likely to occur. Consider, for example, the case in which the demand elasticity is independent of the output level, i.e. $\varepsilon\,(Y_i) = \varepsilon$ (a constant). In that case the price setting rule (5.26) reduces to $P_i = c/\,(1 - 1/\varepsilon)$ so that:

$$\frac{\partial P_i}{\partial t_P} = \frac{1}{1-1/\varepsilon}\frac{\partial c}{\partial t_P} > \frac{\partial c}{\partial t_P}, \tag{5.28}$$

where the sign follows from the fact that $\varepsilon > 1$.

The overshifting result has been illustrated in Figure 5.2. Marginal cost is initially equal to c_0 and the profit maximizing quantity is set at that point where marginal cost equals marginal revenue, MR_i. The output level is Y^0 and the corresponding price level is equal to P^0; see point E_0 on the demand curve $P\,(Y_i)$. An increase in the payroll tax shifts the marginal cost curve to c_1. The monopolist reduces output to Y^1 and changes the price to P^1. As is apparent from the diagram, the increase in the price exceeds the original increase in marginal cost.

For the general case, with a non-constant demand elasticity, overshifting is also a distinct possibility. It follows from (5.24) that a change in marginal cost will change the monopolist's output and price according to:

$$\frac{dY_i}{dc} = \frac{1}{2P'\,(Y_i) + Y_iP''\,(Y_i)} < 0, \tag{5.29}$$

$$\frac{dP_i}{dc} = \frac{P'(Y_i)}{2P'(Y_i) + Y_i P''(Y_i)} \equiv S_i > 0, \tag{5.30}$$

where S_i is the *shifting parameter*. Since both the denominator and the numerator of (5.30) are negative it follows that $S_i > 0$. Price overshifting occurs if $S_i > 1$ or:

$$S_i > 1 \qquad \Leftrightarrow \qquad P'(Y_i) + Y_i P''(Y_i) > 0 \qquad \Leftrightarrow \qquad E_i \equiv -\frac{Y_i P''(Y_i)}{P'(Y_i)} > 1, \tag{5.31}$$

where E_i is a measure, first suggested by Seade (1985, p. 12), representing the elasticity of the *slope* of the inverse demand function. Provided this elasticity measure exceeds unity, price overshifting occurs.

Even though with price overshifting the price is increased by more than the cost change, it is impossible for profit to increase as a result of the cost change, i.e. there is no *profit overshifting* in the monopoly case. This result can be demonstrated by differentiating (5.22) with respect to c:

$$\frac{d\Pi_i}{dc} = \left[P(Y_i) + Y_i P'(Y_i) - c\left(W^*, R^K\right)\right]\frac{dY_i}{dc} - Y_i = -Y_i, \tag{5.32}$$

where we have used the first-order condition (5.24) to arrive at the final expression. The cost increase leads to a reduction in maximized profits. The intuition behind this result is provided by Seade (1985, p. 5). The monopolist could have imposed the higher cost on himself but chose not to do so. The fact that he did not choose to do so, implies that it is optimal (i.e. profit maximizing) not to do so. Hence, profit cannot increase as a result of cost changes.

5.1.2.2 Oligopoly

Both under perfect competition and under monopoly, it is impossible for cost changes to lead to higher profits. This result is somewhat disturbing because, as was argued by Seade (1985), profit shifting does seems to be a fact of life, i.e. it is not just a theoretical curiosity:

> ...casual-observation real-world examples seem to be easy to find, in taxation contexts or otherwise, of apparently profitable cost increases. A notable case in point is the world oil industry in the years 1973-4, when the operating profits (exclusive of stock-revaluation) of the large multinational oil companies underwent marked increases, in the wake of the steep increase in the price of their prime input, crude (Seade, 1985, p. 5).

Seade (1985) demonstrated that it is quite possible for cost increases to be profitable to each firm under conditions of oligopoly. In the remainder of this subsection we study a simple symmetric version of Seade's conjectural variations oligopoly model and derive the conditions under which price overshifting and profit overshifting occur. We continue to assume that factor prices are exogenous, i.e. the model describes a partial equilibrium. There is a fixed number (N) of identical firms each producing the homogeneous good. Profit of representative firm i is given by:

$$\Pi_i \equiv P(Y) Y_i - c\left(W^*, R^K\right) Y_i, \tag{5.33}$$

where $P(\cdot)$ is the (downward sloping) inverse demand function which depends on *total* output of the oligopolistic sector, Y:

$$Y \equiv \sum_{i=1}^{N} Y_i. \tag{5.34}$$

Firm i conjectures a relationship between aggregate output and its own production. Following Seade (1980a, p. 480) we assume that this conjectured relationship takes the form of $dY/dY_i = \eta$, where η is the *conjectural coefficient*; a positive constant which is the same for all firms in the symmetric version of the model. A natural upper bound for η is N, in which case all rival firms react by increasing output by the same amount as firm i (i.e. $dY_j/dY_i = 1$ for all $j \neq i$).

Firm i chooses its output level in order to maximize profit (5.33) taking into account the conjectured relationship $dY/dY_i = \eta$. The first-order condition is given by:

$$\frac{d\Pi_i}{dY_i} \equiv P(Y) + \eta Y_i P'(Y) - c\left(W^*, R^K\right) = 0, \tag{5.35}$$

and the second-order condition is:

$$\frac{d^2\Pi_i}{dY_i^2} \equiv 2\eta P'(Y) + \eta^2 Y_i P''(Y) < 0. \tag{5.36}$$

In addition to these optimality conditions, Seade (1980b, p. 24) also derives so-called *stability conditions* which ensure that, following a perturbation in one of the exogenous variables, the oligopolistic sector will eventually settle down in a new symmetric equilibrium. For the symmetric model used in this subsection, the key stability condition is given by:

$$(N+\eta) P'(N\bar{Y}) + \eta N\bar{Y} P''(N\bar{Y}) < 0. \tag{5.37}$$

By assumption all firms are the same, so in the symmetric equilibrium $Y_i = \bar{Y}$ and $P_i = \bar{P}$ (for all $i = 1, 2, \cdots, N$), aggregate output is equal to $Y = N\bar{Y}$, and the first-order condition (5.35) reduces to:

$$P(N\bar{Y}) + \eta \bar{Y} P'(N\bar{Y}) = c\left(W^*, R^K\right). \tag{5.38}$$

Armed with the expression (5.38), we can investigate the issue of price overshifting in the symmetric oligopoly equilibrium. By differentiating (5.38) with respect to c and $\bar{Y}$ we find:

$$\frac{d\bar{Y}}{dc} = \frac{1}{(N+\eta)\, P'\left(N\bar{Y}\right) + \eta N\bar{Y}P''\left(N\bar{Y}\right)} < 0, \tag{5.39}$$

where the sign follows from the stability condition (5.37). Just as in the monopoly case, output is reduced so some price shifting takes place. To derive the extent of price shifting, we note that in the symmetric equilibrium, $P_i = \bar{P} = P\left(N\bar{Y}\right)$, so that the change in the symmetric price following a cost change is given by:

$$\frac{d\bar{P}}{dc} = \frac{NP'\left(N\bar{Y}\right)}{(N+\eta)\, P'\left(N\bar{Y}\right) + \eta N\bar{Y}P''\left(N\bar{Y}\right)} \equiv \bar{S} > 0, \tag{5.40}$$

where $\bar{S}$ is the shifting coefficient for the oligopoly model. It is not difficult to derive the following condition for price overshifting from (5.40):

$$\bar{S} > 1 \quad \Leftrightarrow \quad P'\left(N\bar{Y}\right) + N\bar{Y}P''\left(N\bar{Y}\right) > 0 \quad \Leftrightarrow \quad \bar{E} \equiv -\frac{N\bar{Y}P''\left(N\bar{Y}\right)}{P'\left(N\bar{Y}\right)} > 1, \tag{5.41}$$

where $\bar{E}$ is again Seade's measure for the elasticity of the slope of the inverse demand function. Comparing the expressions (5.39)-(5.41) to the corresponding ones for the monopoly model (i.e. equations (5.29)-(5.31)) we find that the monopoly model is a special case of the oligopoly model with only one firm and a conjectural coefficient of unity ($N = \eta = 1$).

Unlike for the monopoly case, profit overshifting is a distinct possibility in the symmetric oligopoly model. In the symmetric equilibrium, profit of each firm is given by $\bar{\Pi} \equiv P\left(N\bar{Y}\right)\bar{Y} - c\bar{Y}$. By differentiating this expression we obtain:

$$\begin{aligned} \frac{d\bar{\Pi}}{dc} &= \left[P\left(N\bar{Y}\right) + N\bar{Y}P'\left(N\bar{Y}\right) - c\right]\frac{d\bar{Y}}{dc} - \bar{Y} \\ &= \frac{(N-\eta)\,\bar{Y}P'\left(N\bar{Y}\right)}{(N+\eta)\, P'\left(N\bar{Y}\right) + \eta N\bar{Y}P''\left(N\bar{Y}\right)} - \bar{Y}, \end{aligned} \tag{5.42}$$

where we have used (5.38) and (5.39) to arrive at the second line. Whereas the first term on the right-hand side of (5.42) is zero for the monopoly case (with $N = \eta = 1$), it is positive for the oligopoly model (recall that $0 < \eta < N$). Hence, in principle the first term can offset the negative second term on the right-hand side of (5.42), so that profit increases following a cost change.

By using the definition for $\bar{E}$ (given in (5.41) above), the expression in (5.42) can be simplified as follows:

$$\begin{aligned} \frac{d\bar{\Pi}}{dc} &= \frac{(N-\eta)\,\bar{Y}P'\left(N\bar{Y}\right) - \bar{Y}\left[(N+\eta)\, P'\left(N\bar{Y}\right) + \eta N\bar{Y}P''\left(N\bar{Y}\right)\right]}{(N+\eta)\, P'\left(N\bar{Y}\right) + \eta N\bar{Y}P''\left(N\bar{Y}\right)} \\ &= \frac{-\eta\bar{Y}\left[2P'\left(N\bar{Y}\right) + N\bar{Y}P''\left(N\bar{Y}\right)\right]}{(N+\eta)\, P'\left(N\bar{Y}\right) + \eta N\bar{Y}P''\left(N\bar{Y}\right)} \end{aligned}$$

$$= \frac{\eta \bar{Y} P'(N\bar{Y})}{(N+\eta) P'(N\bar{Y}) + \eta N \bar{Y} P''(N\bar{Y})} [\bar{E} - 2]. \tag{5.43}$$

In view of the stability condition (5.37) and the fact that demand slopes downward, it is obvious that the fraction appearing on the right-hand side of (5.43) is positive. Hence, profit overshifting occurs in the symmetric oligopoly model if and only if $\bar{E} > 2$. To relate this condition to the elasticity of the demand curve itself, consider the iso-elastic demand curve, $Y = P^{-\varepsilon}$, for which $\bar{E} = 1 + 1/\varepsilon$. Provided the demand curve is inelastic ($\varepsilon < 1$), an increase in cost *must* increase the profit level for each oligopolist (Seade, 1985, p. 16).

The intuition behind the profit overshifting result is provided by Seade (1985, p. 5). If an individual producer cuts back his production, then his profit will go down but the price set by all firms (and their profits) will go up. Since the costs are borne by the individual producers and the benefits accrue to all producers, too little "restraint" will be practiced, i.e. too much is produced by each firm.[5] The cost increase forces all firms to cut production and thus increases restraint. As Seade puts it, "the increase in cost can be seen as imposing on the producers some of the collusion they themselves had been unable to achieve."

5.2 Financial structure of the firm

Useful as they are for organizational purposes, the basic static models discussed in the previous section are somewhat unsatisfactory for at least two reasons. First, because the models are static, they cannot be used to study the inherently dynamic capital accumulation decision of the firm (i.e. its *real investment plans*). Second, such static models cannot say anything about the determination of the cost of capital facing a given firm (i.e. the firm's *optimal financial policy*). In this section we deal with the second of these deficiencies by studying the determinants of the cost of capital facing the firm. (In Section 5.3 we jointly study the firm's investment and financial decisions.) To keep things as simple as possible, we restrict attention to the case of a perfectly competitive firm. Central to our discussion is the classic deterministic infinite-horizon model by Auerbach (1979).

The key issue studied by Auerbach concerns the question how a *mature* firm should finance a *given* amount of investment. The representative firm is competitive, operates under perfect foresight, has an infinite planning horizon, and cannot go bankrupt. It faces three modes of financing, namely sales of common stock (share emissions), retention of earnings, and sales of corporate debt (single-period corporate bonds).[6]

The model is formulated in discrete time, and the timing of payments is as follows. At the beginning of each period firms distribute dividends to pre-existing shareholders, pay interest on corporate bonds,

[5] This is not unlike the public good problem studied in Chapter 12 below. See also Kurz (1985).

[6] In contrast, an *immature* (starting) firm may not have profits that are high enough to finance its investment projects and may also face restrictions regarding its ability to incur corporate debt. See Sinn (1991, pp. 39-42) on the cost of capital facing immature firms.

repay principal on debt outstanding from the previous period, and sell new shares *ex dividend*.

The stylized tax system is as follows. All tax rates are assumed to be constant over time. The corporate tax levied on the firm is denoted by t_K and by assumption interest payments on corporate debt are deductible from taxable profit. The dividend tax and capital gains tax levied on the household-investors are denoted by, respectively, t_D and t_G. It is assumed that capital gains are taxed upon accrual (not on realization), and that dividends are taxed more heavily, i.e. $t_D > t_G$.

Household-investors all face the same personal tax rates t_D and t_G, have a one-period discount rate of ρ_t, face an interest rate of r_t, and, like firms, are blessed with perfect foresight. Below we shall normalize the number of household-investors to unity and refer to the representative household where needed.

5.2.1 The firm

The ex-dividend value of the representative firm's equity at the beginning of time t is denoted by V_t:

$$V_t = V_t^O + V_t^N, \tag{5.44}$$

where V_t^O is the ex-dividend value of *pre-existing* shares at the beginning of time t and V_t^N is the ex-dividend value of *new* shares sold at the beginning of time t. Pre-existing equity may be *diluted* by new emissions (see below) and the dilution parameter is defined as follows:

$$\delta_t \equiv \frac{V_t^N}{V_t}. \tag{5.45}$$

The stock of corporate debt outstanding at the beginning of period t is B_t and the firm's *leverage* is thus represented by:

$$\beta_t \equiv \frac{B_t}{B_t + V_t}. \tag{5.46}$$

Dividends, D_t, paid at the end of period t (i.e. at beginning of period $t+1$) are defined as follows:

$$D_t = x_{t+1} + B_{t+1} + V_{t+1}^N - [1 + r_t(1 - t_K)] B_t, \tag{5.47}$$

where dividends must be non-negative (by definition), and x_t is the corporate cash flow of the firm at the beginning of period t. This cash flow is net of the corporate tax but before account is taken of the net sales of debt and equity, interest payments, and tax savings on interest payments.[7]

[7]Cash flow is thus given by $x_t \equiv (1 - t_K)\Pi(K_t) - I_t$, where I_t is investment spending, and $\Pi(K_t)$ is gross operating profit expressed as a function of the capital stock at the beginning of period t :

$$\Pi(K_t) \equiv \max_{\{L_t\}} F(K_t, L_t) - W_t L_t = F_K(K_t, L_t^*) K_t,$$

where W_t is the wage rate and $F(K_t, L_t)$ is the production function (featuring constant returns to scale). The second expression results from the fact that the firm employs labour up to the point at which $W_t = F_L(K_t, L_t^*)$, where L_t^* is the optimal amount of labour used.

It is assumed that the firm's objective is to maximize the wealth of *existing* shareholders, employing the following three instruments. First, it can choose an *investment policy*, i.e. a time path for cash flows, which we denote by $I_X \equiv \{x_t, x_{t+1}, x_{t+2} \cdots\}$. Second, it can choose a *debt policy*, i.e. a time path for corporate borrowing, which is denoted by $I_B \equiv \{B_t, B_{t+1}, B_{t+2} \cdots\}$. Third, it can choose an *equity policy*, i.e. a time path for equity emissions, which is denoted by $I_V \equiv \{V_t^N, V_{t+1}^N, V_{t+2}^N \cdots\}$. In view of (5.47), once I_X, I_B, and I_V are chosen, the path of dividends, $I_D \equiv \{D_{t-1}, D_t, D_{t+1} \cdots\}$, is determined also.

It is assumed that the value of shares is equal to the present value of the net distributions received by the owners (the no-bubble, *fundamental* equity value). The net distribution (received by the group of share holders) at the end of period t is equal to:

$$E_t \equiv (1 - t_D) D_t - t_G \left(V_{t+1}^O - V_t \right), \tag{5.48}$$

where t_D is the dividend tax, $(1 - t_D) D_t$ is after-tax dividend payments, t_G is the capital gains tax, and $V_{t+1}^O - V_t$ is the accrued capital gain (or loss, if this term is negative). (Recall that V_{t+1}^O is the ex-dividend value of pre-existing shares at the beginning of time $t + 1$).

Consider two time periods t and s, where $s > t$. Clearly, if new equity is issued between times t and s, then not all of E_s will go to shares that were held as of period t. Some of these distributions will go to shares that were issued between t and s. Define μ_t^s as the fraction of shares held during period s that were in existence before the start of period t:

$$\begin{aligned}\mu_t^s &\equiv \frac{V_t^O}{V_t} \times \frac{V_{t+1}^O}{V_{t+1}} \times \cdots \frac{V_s^O}{V_s} \\ &= (1 - \delta_t)(1 - \delta_{t+1}) \cdots (1 - \delta_s), \end{aligned} \tag{5.49}$$

where δ_t is defined in (5.45) above. Obviously, if no new shares are issued between t and s, there is no dilution and $\mu_t^s = 1$. The original shareholders receive the distributions. In all other cases, the more dilution there is, the smaller is the share of the "original" shareholders in the distributions.

Since the discount rate of equity owners in period s is ρ_s we find that the fundamental value of shares owned at the beginning of period t is given by the discounted present value of receipts:

$$V_t^O = \sum_{s=t}^{\infty} \left[\prod_{z=t}^{s} \frac{1}{1 + \rho_z} \right] \mu_t^s E_s. \tag{5.50}$$

According to (5.50), the "original" shareholders receive $\mu_t^s E_s$ (in period s), have a discount rate of ρ_z in period z, and feature an infinite planning horizon. Using (5.44), (5.45), and (5.50), it is possible to derive the following arbitrage-like equation for V_t (see Intermezzo 5.2):

$$\rho_t = \frac{E_t + V_{t+1}^O - V_t}{V_t}. \tag{5.51}$$

Intuitively, the one-period holding yield on shares, equalling dividend plus capital gain over initial value, must equal the required one-period discount rate of the household-investors.

Intermezzo 5.2

Deriving the arbitrage equation. The derivation of the arbitrage equation (5.51) is far from straightforward. Two intermediate results are useful.

Result 1: Difference equation for V_t^O:

$$V_t^O = \frac{1-\delta_t}{1+\rho_t}\left[E_t + V_{t+1}^O\right]. \tag{I.1}$$

Result 2: Relationship between V_t^O and V_t:

$$\frac{1+\rho_t}{1-\delta_t}V_t^O = (1+\rho_t)\,V_t. \tag{I.2}$$

By combining (I.1) and (I.2) we obtain (5.51). The proof of Result 2 is obvious: it follows readily from (5.44) by using the definition of δ_t in (5.45). Result 1 is proved as follows. First, by using (5.49)-(5.50) we can write V_t^O as:

$$V_t^O = \frac{1-\delta_t}{1+\rho_t}E_t + \frac{1-\delta_t}{1+\rho_t}\frac{1-\delta_{t+1}}{1+\rho_{t+1}}E_{t+1} + \cdots \tag{I.3}$$

and V_{t+1}^O as:

$$\begin{aligned} V_{t+1}^O &= \sum_{s=t+1}^{\infty}\left[\prod_{z=t+1}^{s}\frac{1}{1+\rho_z}\right]\mu_{t+1}^s E_s \\ &= \frac{1-\delta_{t+1}}{1+\rho_{t+1}}E_{t+1} + \frac{1-\delta_{t+1}}{1+\rho_{t+1}}\frac{1-\delta_{t+2}}{1+\rho_{t+2}}E_{t+2} + \cdots \\ &= \frac{1+\rho_t}{1-\delta_t}\left[\frac{1-\delta_t}{1+\rho_t}\frac{1-\delta_{t+1}}{1+\rho_{t+1}}E_{t+1} + \frac{1-\delta_t}{1+\rho_t}\frac{1-\delta_{t+1}}{1+\rho_{t+1}}\right. \\ &\quad \left. \times\frac{1-\delta_{t+2}}{1+\rho_{t+2}}E_{t+2} + \cdots\right]. \end{aligned} \tag{I.4}$$

But, in view of (I.3), the term in square brackets on the right-hand side of (I.4) can be written as $V_t^O - \frac{1-\delta_t}{1+\rho_t}E_t$ so V_{t+1}^O itself can be written as:

$$\begin{aligned} V_{t+1}^O &= \frac{1+\rho_t}{1-\delta_t}\left[V_t^O - \frac{1-\delta_t}{1+\rho_t}E_t\right] \\ &= \frac{1+\rho_t}{1-\delta_t}V_t^O - E_t, \end{aligned} \tag{I.4}$$

which is a slightly rewritten version of (I.1).

****

By assumption the firm maximizes the wealth of existing stock holders at time t, which is defined as follows:

$$\Omega_t^O \equiv V_t^O + E_{t-1}, \tag{5.52}$$

where E_{t-1} is the net distribution of dividends at the end of period $t-1$ (i.e. at the beginning of period t). By using (5.44), (5.47), and (5.48) we find that (5.52) can be rewritten as follows:

$$\Omega_t^O = \Omega_t + \omega_0, \tag{5.53}$$

where Ω_t and ω_0 are defined as follows:

$$\Omega_t \equiv (1-t_G)\,V_t - (t_D - t_G)\,V_t^N + (1-t_D)\,[B_t + x_t]\,, \tag{5.54}$$

$$\omega_0 \equiv t_G V_{t-1} - (1-t_D)\,[1 + r_{t-1}\,(1-t_K)]\,B_{t-1}. \tag{5.55}$$

The key thing to note about ω_0 is that it is predetermined at the beginning of period t and is thus irrelevant for the maximization of Ω_t^O (the past cannot be undone). The effective objective function of the firm is thus given by (5.54).

5.2.2 Wealth maximization: no personal taxes

In order to develop the intuition behind the results to come, we first look at the case where all personal taxes are zero, i.e. $t_G = t_D = 0$. The firm's objective function (5.54) simplifies to:

$$\Omega_t = V_t + B_t + x_t, \tag{5.56}$$

and we find the *standard result* that the wealth maximizing firm maximizes the sum of the market value of its securities ($V_t + B_t$) and its current cash flow (x_t).

Furthermore, using (5.44), (5.48), and (5.51) it is now possible to derive the famous *Modigliani-Miller* result (see Intermezzo 5.3):

$$V_t = \sum_{s=t}^{\infty}\left[\prod_{z=t}^{s}\frac{1}{1+\rho_z}\right]\left[D_s - V_{s+1}^N\right]. \tag{5.57}$$

According to this expression, the particular source of equity funds used to finance a given investment policy (I_X) has no impact on the firm's valuation. Once I_X and I_B are determined so is I_V, since $D_s - V_{s+1}^N$

is (by definition) independent of V^N.[8] Put differently, a unit increase (decrease) in dividends financed by a unit increase (decrease) in the value of new equity has no effect on V_t (which is independent of the *levels* of D_s and V_{s+1}^N).

Shareholder wealth can now be written as follows:

$$\Omega_t = x_t + \sum_{s=t+1}^{\infty} \left[\prod_{z=t}^{s-1} \frac{1}{1+\theta_z} \right] x_s, \tag{5.58}$$

where θ_z is the cost of capital in period z:

$$\theta_z \equiv \beta_z r_z (1 - t_K) + (1 - \beta_z) \rho_z, \tag{5.59}$$

and where β_z is defined in (5.46) above. According to (5.58), shareholder wealth is equal to the present value of after-corporate-tax cash flows, using the discount rate θ_z which is determined by the financial policy of the firm. As (5.59) shows, the cost of capital in period t is a weighted average of the costs associated with debt and equity holdings, with the weights being $\beta_t \equiv B_t / (B_t + V_t)$ and $1 - \beta_t \equiv V_t / (B_t + V_t)$ (see (5.46) above). Note furthermore that the interest paid on corporate debt is deductible from the corporate tax base so $r_z (1 - t_K)$ is the net cost of debt to the firm.

In view of (5.58)-(5.59) it is now possible to say something about the optimal *financial policy* of the firm, which amounts to the optimal choice of the leverage parameter β_t. Clearly, since β_t does not affect the stream of cash flows (x_t), the firm should choose β_t in order to minimize the cost of capital, θ_t, by finding the cheapest source of funding. Since equation (5.59) is linear in β_t, it follows that there are three cases:[9]

- If $r_t (1 - t_K) > \rho_t$, then debt is relatively expensive and it is optimal to choose $\beta_t = 0$: finance by equity alone;
- If $r_t (1 - t_K) = \rho_t$, then debt and equity are equally expensive and the firm is indifferent about its choice of β_t;
- If $r_t (1 - t_K) < \rho_t$, then debt is relatively cheap and it is optimal to choose $\beta_t = 1$: finance by bonds alone.

Once the optimal cost of capital has been determined, the firm makes its optimal real (production and investment) plans in such a way as to maximize its stock market value (sequential decision making). The cost of capital thus influences the firm's real activities but not vice versa. The interaction between the cost of capital and the output-investment plans is studied in more detail in Section 5.3 below.

[8] To see why this is the case, recall that:

$$D_s - V_{s+1}^N = x_{s+1} + B_{s+1} - (1 + r_s) B_s.$$

[9] In the special case with $\rho_t = r_t$, bonds are preferred if there is a positive corporate tax ($t_K > 0$) and the firm is indifferent about its source of financing if the corporate tax is zero ($t_K = 0$).

Intermezzo 5.3

Derivations of equations (5.57) and (5.58)-(5.59). The derivation of (5.57) proceeds as follows. In the first step we use (5.44), (5.48) and (5.51) to get:

$$\begin{aligned} \rho_t V_t &= D_t - V_{t+1}^N + V_{t+1} - V_t \qquad \Leftrightarrow \\ (1+\rho_t) V_t &= \left[D_t - V_{t+1}^N\right] + V_{t+1}. \end{aligned} \tag{I.1}$$

By iterating (I.1) forward in time we obtain:

$$\begin{aligned} V_t &= \frac{D_t - V_{t+1}^N}{1+\rho_t} + \frac{V_{t+1}}{1+\rho_t} \\ &= \frac{D_t - V_{t+1}^N}{1+\rho_t} + \frac{1}{1+\rho_t}\left[\frac{D_{t+1} - V_{t+2}^N}{1+\rho_{t+1}} + \frac{V_{t+2}}{1+\rho_{t+1}}\right] \\ &= \sum_{s=t}^{\infty}\left[\prod_{z=t}^{s}\frac{1}{1+\rho_z}\right]\left[D_s - V_{s+1}^N\right] + \lim_{T\to\infty}\Xi_V, \end{aligned} \tag{I.2}$$

where Ξ_V is:

$$\Xi_V \equiv \prod_{z=t}^{T}\frac{1}{1+\rho_z}V_{T+1}. \tag{I.3}$$

The so-called no-Ponzi game (NPG) condition rules out "chain letters" and amounts in this context to the requirement:

$$\lim_{T\to\infty}\Xi_V = 0. \tag{I.4}$$

By incorporating (I.4) into (I.2) we obtain (5.57).

The derivation of equations (5.58) and (5.59) proceeds as follows. First we use (I.1) and add appropriate terms to both sides to obtain:

$$\begin{aligned} (1+\rho_t) V_t + B_t\left[1+r_t(1-t_K)\right] &= V_{t+1} + D_t - V_{t+1}^N + B_{t+1} - B_{t+1} \\ &\quad + B_t\left[1+r_t(1-t_K)\right]. \end{aligned} \tag{I.5}$$

Next we note the definition of D_t from (5.47) and derive from (I.5):

$$(1+\rho_t) V_t + B_t\left[1+r_t(1-t_K)\right] = V_{t+1} + x_{t+1} + B_{t+1}. \tag{I.6}$$

The left-hand side of (I.6) can be rewritten as:

$$
\begin{aligned}
LHS &\equiv (B_t + V_t)\left((1+\rho_t)\frac{V_t}{B_t+V_t} + \frac{B_t}{B_t+V_t}\left[1+r_t\left(1-t_K\right)\right]\right) \\
&= (B_t + V_t)\left((1+\rho_t)(1-\beta_t) + \beta_t\left[1+r_t\left(1-t_K\right)\right]\right) \\
&= (B_t + V_t)\left(1+\rho_t(1-\beta_t) + \beta_t r_t\left(1-t_K\right)\right) \\
&= (1+\theta_t)(B_t + V_t)\,,
\end{aligned}
\tag{I.7}
$$

where in the last step use is made of (5.59). Hence, equation (I.6) can be written as:

$$
B_t + V_t = \frac{B_{t+1} + V_{t+1} + x_{t+1}}{1+\theta_t}. \tag{I.8}
$$

In view of (5.56) and (I.8) we obtain the difference equation for Ω_t:

$$
\Omega_t - x_t = \frac{\Omega_{t+1}}{1+\theta_t}. \tag{I.9}
$$

Iterating (I.9) forward in time we find:

$$
\begin{aligned}
\Omega_t &= x_t + \frac{\Omega_{t+1}}{1+\theta_t} \\
&= x_t + \frac{1}{1+\theta_t}\left[x_{t+1} + \frac{\Omega_{t+2}}{1+\theta_{t+1}}\right] \\
&= x_t + \sum_{s=t+1}^{\infty}\left[\prod_{z=t}^{s-1}\frac{1}{1+\theta_z}\right]x_s + \lim_{T\to\infty}\Xi_W,
\end{aligned}
\tag{I.10}
$$

where $\lim_{T\to\infty}\Xi_W$ is again a term that goes to zero in the limit (provided the relevant NPG condition is imposed):

$$
\lim_{T\to\infty}\Xi_W \equiv \prod_{z=t}^{T}\frac{1}{1+\theta_z}\Omega_{T+1} = 0. \tag{I.11}
$$

Substituting (I.11) in (I.10) yields equation (5.58).

****

5.2.3 Wealth maximization: positive personal taxes

We now reinstate the personal tax system by allowing for non-zero tax rates on both capital gains ($t_G \geq 0$) and on dividend income ($t_D > 0$). In most countries capital gains are taxed at a lower rate than other income is, i.e. the case that we focus on assumes that $t_G < t_D$. Recall that in the presence of personal taxation, the expression for stockholders wealth is given by equation (5.54) above. It is clear from that expression that, if $t_D \neq t_G$, the firm should not strive to maximize the sum of current cash flow and the market value of securities ($V_t + B_t + x_t$). The "standard result" mentioned below equation (5.56) no longer holds in the presence of differential taxation of dividends and capital gains.[10] As was pointed out by Edwards and Keen (1984, p. 212), however, both with and without personal taxes, maximization of Ω_t^O is equivalent to maximization of the opening market value of equity, V_{t-1}. To see why this is the case, we use equation (5.51), lagged once, and note (5.52) to derive:

$$\Omega_t^O = (1 + \rho_{t-1}) V_{t-1}. \tag{5.60}$$

Since ρ_{t-1} is exogenous, and $\rho_{t-1} V_{t-1}$ is the post-tax return that share holders obtain over period $t-1$, maximizing V_{t-1} is the same as maximizing Ω_t^O.

By taking the steps leading to (5.57), it is possible to derive the following expression for the ex-dividend value of the firm's equity:

$$V_t = \sum_{s=t}^{\infty} \left[\prod_{z=t}^{s} \frac{1}{1 + \frac{\rho_z}{1-t_G}} \right] \left[\frac{1-t_D}{1-t_G} \left(D_s - V_{s+1}^N \right) - \frac{t_D - t_G}{1-t_G} V_{s+1}^N \right]. \tag{5.61}$$

This expression generalizes (5.57) to the case of personal taxation. If $t_D = t_G$ then the only thing that differs between the two expressions (5.57) and (5.61) is the discount rate (which is higher in the taxation case, i.e. $\rho_z / (1-t_G) > \rho_z$). In the more relevant case, however, with $t_G < t_D$, the differential tax treatment introduces an asymmetric effect of dividends and new share issues in the valuation formula. Indeed, given I_X and I_B, a decrease in V_s^N increases V_t (and hence Ω_t). It is never optimal for the firm to issue new shares *and* pay dividends at the same time, i.e. V_t may be increased by an equal reduction in D_s and V_{s+1}^N.[11]

Given the favourable tax treatment of capital gains, the firm may wish to repurchase its own shares (provided dividends remain non-negative) and thus increase its market value. Although this argument is theoretically correct, share repurchases do not occur very often in reality. The reason is clear: repurchases are very difficult in the United Kingdom (requires a court order) and are forbidden or made very unattractive (by taxing capital gains as dividends) in the United States.

The *tax capitalization view*, developed by *inter alia* King (1974), Auerbach (1979), and Bradford (1981), is based on the assumption that mature firms have sufficiently high after-tax profits and therefore use retained earnings (withheld dividends) as the marginal source of investment funds. According to this view, the firm has exhausted all its low-tax opportunities to channel income to its shareholders and thus

[10] Obviously, if $t_D = t_G$ then maximization of $V_t + B_t + x_t$ is still called for. This case represents the integration of corporate and personal income taxes, i.e. the same tax rate is applied to retained and distributed corporate income.

[11] As is pointed out by Poterba and Summers, however, there exists a *dividend puzzle* in that in reality some firms pay dividends despite the fact that they have "unused opportunities to repurchase shares or engage in equivalent transactions which would effectively transmit tax-free income to shareholders" (1985, p. 234). According to the theory proposed here, this constitutes suboptimal behaviour on the part of the firm.

faces the following binding constraint:

$$V_s^N = 0 \text{ for all } s. \tag{5.62}$$

Dividends are strictly positive and equation (5.61) simplifies to:

$$V_t = \sum_{s=t}^{\infty} \left[\prod_{z=t}^{s} \frac{1}{1 + \frac{\rho_z}{1-t_G}} \right] \frac{1-t_D}{1-t_G} D_s. \tag{5.63}$$

Following the steps leading to (5.58) we find that shareholder wealth in the presence of personal taxation can be written as follows:

$$\Omega_t = (1 - t_D) \left(x_t + \sum_{s=t+1}^{\infty} \left[\prod_{z=t}^{s-1} \frac{1}{1+\theta_z} \right] x_s \right), \tag{5.64}$$

where θ_z is the redefined cost of capital in period z:

$$\theta_z \equiv \frac{\beta_z r_z (1-t_K)(1-t_D) + (1-\beta_z)\rho_z}{(1-t_G) - (t_D - t_G)\beta_z}. \tag{5.65}$$

In equation (5.64), the term in round brackets on the right-hand side is the present value of after-corporate-tax cash flows using the discount rate determined by the firm's optimal financial policy. Shareholder wealth is the after-dividend-tax equivalent of this present value term.

It is clear from equation (5.65) that the cost of capital in period t depends on the leverage parameter, β_t, and on the tax rates t_K and t_G, but *not* on the dividend payout rate or the tax on dividends, t_D. In order to understand this result it is useful to consider some special cases. First, a solely debt-financed firm sets $\beta_t = 1$ so that the cost of capital reduces to the net of tax rate of interest, $\theta_t = r_t(1-t_K)$. Second, a solely equity-financed firm sets $\beta_t = 0$ so that $\theta_t = \frac{\rho_t}{1-t_G}$, again no effect effect of the dividend tax rate on the cost of capital.

As before, the firm's optimal financial policy consists of choosing the appropriate leverage, β_t. By using (5.65) (for period t) and differentiating with respect to β_t we find:

$$\frac{d\theta_t}{d\beta_t} = \frac{(1-t_D)\left[r_t(1-t_K)(1-t_G) - \rho_t\right]}{\left[(1-t_G) - (t_D - t_G)\beta_t\right]^2}. \tag{5.66}$$

It follows from (5.66) that there are three cases to be considered:

- If $r_t(1-t_K)(1-t_G) > \rho_t$ then $d\theta_t/d\beta_t > 0$ and it is optimal to choose $\beta_t = 0$: finance by equity alone;
- If $r_t(1-t_K)(1-t_G) = \rho_t$ then $d\theta_t/d\beta_t = 0$ and the firm is indifferent about its choice of β_t; and
- If $r_t(1-t_K)(1-t_G) < \rho_t$ then $d\theta_t/d\beta_t < 0$ so it is optimal to choose $\beta_t = 1$: finance by bonds alone.

Imposing these three possible outcomes in (5.65), we reach the conclusion that the optimized cost of capital is independent of the dividend tax altogether, i.e. θ_t becomes:

$$\theta_t \equiv \min\left[\frac{\rho_t}{1-t_G}, r_t\left(1-t_K\right)\right]. \tag{5.67}$$

Whilst the dividend tax does not influence the cost of capital, it does affect the value of shareholder wealth, Ω_t. Indeed, as is clear from the expression in (5.64), the dividend tax acts as a lump-sum levy on wealth in the corporate sector. According to the tax capitalization view, current equity is *trapped* in the corporate sector and as a result bears the burden of the dividend tax (Poterba and Summer, 1985, p. 239).

Will corporate debt and equity ever be held simultaneously in the market? Clearly, this is only possible if both firms and household-investors are indifferent between the two instruments. It follows from (5.66) that firms are indifferent if and only if $r_t\left(1-t_K\right)\left(1-t_G\right)=\rho_t$. Household-investors receive ρ_t from holding equity and $r_t\left(1-t_R\right)$ when holding corporate debt, where t_R is the rate at which interest income is taxed at the personal level (this rate may or may not equal the dividend tax rate). In the absence of risk, household-investors are thus indifferent between debt and equity if and only if $\rho_t = r_t\left(1-t_R\right)$. It follows from the two indifference relationships that simultaneous debt *and* equity holding can be an equilibrium phenomenon if and only if $\left(1-t_K\right)\left(1-t_G\right)=1-t_R$. If $\left(1-t_K\right)\left(1-t_G\right)>1-t_R$ then there will only be equity whilst if $\left(1-t_K\right)\left(1-t_G\right)<1-t_R$ there will be only debt.[12]

The key findings of this subsection are as follows. First, the tax system generally affects the validity of the Modigliani-Miller Theorem. Second, the corporation tax favours bond financing because it reduces the interest rate on corporate bonds. Third, the cost of capital is independent of the dividend tax and of the dividend payout ratio. Fourth, the firm engages in a kind of "sequential" decision making; via its financial policy the firm determines the cost of capital, after which it decides on output and capital accumulation plans.

5.3 Taxation and firm investment

In the previous section it has been demonstrated that the cost of capital to a firm depends very much on the details of the firm's financial policy, the corporate tax system, *and* the personal tax system. It follows that the effect of taxes on the firm's investment and output decisions is also critically dependent on these details. We also noted in the previous section that in reality firms do tend to pay out dividends despite their unfavorable tax treatment vis-a-vis capital gains (the dividend puzzle).

The objective of this section is to study the effects of the different personal and corporate taxes on the firm's output and capital accumulation decisions. The argument makes use of a (simplified version of the) model of firm investment due to Turnovsky (1990). The model is based on explicitly dynamic

[12]The proof proceeds as follows. Assume that investors are indifferent between debt and equity so that $\rho_t = r_t\left(1-t_R\right)$. If $\left(1-t_K\right)\left(1-t_G\right)>1-t_R$ it follows that $\left(1-t_K\right)\left(1-t_G\right)>\rho_t/r_t$ or $\left(1-t_K\right)\left(1-t_G\right)r_t>\rho_t$ so firms will want equity.

maximizing behaviour of household-investors and firms. The optimization programs yield explicit expressions for the arbitrage equations and the cost of capital under different dividend policies. In contrast with Turnovsky (1990), our analysis restricts attention to the *partial equilibrium* effects of tax policy on the representative firm.[13] For reasons of analytical convenience and in order to prepare for the analysis in Chapter 7 below, we cast the model in continuous (rather than discrete) time.

5.3.1 Representative household

There are many infinitely-lived identical household-investors who are blessed with perfect foresight. For notational convenience, we normalize the number of households to unity and argue on the basis of a single representative household. The lifetime utility function of the representative household is:

$$\Lambda(0) \equiv \int_0^\infty U(C(t))\, e^{-\rho t} dt, \tag{5.68}$$

where $U(\cdot)$ is the *felicity (or instantaneous utility) function*, featuring positive but diminishing marginal instantaneous utility ($U'(\cdot) > 0 > U''(\cdot)$), $C(t)$ is the flow of consumption, ρ is the pure rate of time preference ($\rho > 0$), and $\Lambda(0)$ is an indicator for *lifetime utility* from the perspective of the planning time, $t = 0$.

The household can save by investing in *shares* or in *government bonds* (by assumption there are no corporate bonds). In this deterministic setting, there is no risk so bonds and shares are perfect substitutes in the household's portfolio. The household budget identity is given by:

$$\begin{aligned}\dot{B}(t) + P_E(t)\dot{E}(t) + C(t) = (1 - t_Y)\left[W(t)\bar{L} + D(t)\right] + (1 - t_R)\, r(t)\, B(t) \\ - t_G \dot{P}_E(t)\, E(t) + Z(t),\end{aligned} \tag{5.69}$$

where $B(t)$ is the stock of government debt, $P_E(t)$ is the price of shares, $E(t)$ is the outstanding stock of equities, t_Y is the common tax on wage income and dividend income, $W(t)$ is the wage rate, $\bar{L}$ is (exogenous) labour supply, $D(t)$ is dividends received from firms, t_R is the tax on interest income, $r(t)$ is the interest rate on government bonds, t_G is the capital gains tax, and $Z(t)$ is the lump-sum transfer received from the government. As usual a variable with a dot is that variable's time rate of change, e.g. $\dot{B}(t) \equiv dB(t)/dt$.

Normalizing the planning period by $t = 0$, the household faces the initial conditions: $E(0) = E_0$ and $B(0) = B_0$, i.e. the initial stocks of bonds and equity are predetermined at time $t = 0$. The household chooses paths for $C(t)$, $B(t)$, and $E(t)$ in order to maximize (5.68) subject to (5.69) and two transversality conditions. The dividend payout ratio, $r_D(t)$, is defined as follows:

$$r_D(t) \equiv \frac{D(t)}{P_E(t)\, E(t)}. \tag{5.70}$$

[13] We return to a general equilibrium version of the model in Chapter 7.

This ratio is determined by the firm and taken parametrically by the household. The household optimization program constitutes a non-standard optimal control problem which can, however, be solved by transforming it (see Intermezzo 5.4 below).[14]

The first-order necessary conditions characterizing the household's interior optimum are:

$$U'(C(t)) = \lambda(t), \tag{5.71}$$

$$(1-t_R)\,r(t) = (1-t_G)\frac{\dot{P}_E(t)}{P_E(t)} + r_D(t)(1-t_Y), \tag{5.72}$$

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = \rho - (1-t_R)\,r(t), \tag{5.73}$$

where $\lambda(t)$ is the co-state variable associated with aggregate financial wealth. Intuitively, $\lambda(t)$ measures the increase in lifetime utility that the household would experience if a Martian visitor to earth were to provide it with a little bit of additional wealth in period t.[15] Equation (5.71) is the (implicit) *Frisch demand* for consumption, (5.72) is the no-arbitrage equation between government bonds and equities, and (5.73) describes the optimal time path for the marginal utility of wealth, $\lambda(t)$.

In the partial equilibrium interpretation of the model, we assume that $\dot{\lambda}(t) = 0$ so that $\dot{C}(t) = 0$ is constant and $r(t) = \rho/(1-t_R)$. Both consumption and the interest rate are time-invariant. The arbitrage equation (5.72) simplifies to:

$$(1-t_Y)\frac{D(t)}{P_E(t)\,E(t)} + (1-t_G)\frac{\dot{P}_E(t)}{P_E(t)} = \rho, \tag{5.74}$$

where we have also used (5.70). Intuitively, (5.74) says that in the interior portfolio equilibrium, the after-tax rate of return on shares (left-hand side) must equal the pure rate of time preference (right-hand side). As usual, the rate of return on shares equals after-tax dividend plus capital gain per share divided by the share price.

Intermezzo 5.4

Solving the household problem. All assets are perfect substitutes in the household portfolio so that we can define total assets, $A(t)$, as follows:

$$A(t) = B(t) + P_E(t)\,E(t). \tag{I.1}$$

By differentiating both sides of (I.1) with respect to time we obtain:

$$\dot{A}(t) = \dot{B}(t) + P_E(t)\,\dot{E}(t) + \dot{P}_E(t)\,E(t). \tag{I.2}$$

[14] The basic tools of optimal control theory are summarized in the brief Technical Appendix to this chapter.
[15] The interested reader is referred to Léonard and Long (1992, pp. 153-155) for a formal proof of this result.

Rewriting (5.69) we get (by adding $\dot{P}_E(t)\,E(t)$ to both sides):

$$\dot{B}(t) + P_E(t)\,\dot{E}(t) + \dot{P}_E(t)\,E(t) = (1-t_Y)\left[W(t)\,\bar{L} + D(t)\right] + Z(t) - C(t) \\ + (1-t_R)\,r(t)\,B(t) + (1-t_G)\,\dot{P}_E(t)\,E(t)\,, \tag{I.3}$$

where it is noted that the left-hand side of (I.3) is equal to $\dot{A}(t)$. By adding and deducting $(1-t_R)\,r(t)\,P_E(t)\,E(t)$ to the right-hand side of (I.3) we obtain:

$$\dot{A}(t) = (1-t_Y)\left[W(t)\,\bar{L} + D(t)\right] + (1-t_R)\,r(t)\,A(t) + Z(t) - C(t) \\ + \left[(1-t_G)\,\dot{P}_E(t) - (1-t_R)\,r(t)\,P_E(t)\right]E(t)\,. \tag{I.4}$$

Noting the relationship $D(t) = r_D(t)\,P_E(t)\,E(t)$ we can rewrite (I.4) as:

$$\dot{A}(t) = (1-t_R)\,r(t)\,A(t) + (1-t_Y)\,W(t)\,\bar{L} + Z(t) - C(t) \\ + \left[(1-t_G)\,\frac{\dot{P}_E(t)}{P_E(t)} + r_D\,(1-t_Y) - (1-t_R)\,r(t)\right]P_E(t)\,E(t)\,. \tag{I.5}$$

We now have a single aggregate state variable $(A(t))$ whose dynamic evolution must be determined.

Dropping time subscripts where no confusion can arise, the current-value Hamiltonian can be written as:

$$\mathcal{H} \equiv U(C) + \lambda\Bigg((1-t_R)\,rA + (1-t_Y)\,W\bar{L} + Z - C \\ + \left[(1-t_G)\,\frac{\dot{P}_E}{P_E} + r_D\,(1-t_Y) - (1-t_R)\,r\right]P_E E\Bigg) \\ + \mu\left[A - B - P_E E\right].$$

The control variables are C, E, and B, the state variable is A, the co-state variable is λ, and the Lagrange multiplier is μ. The key first-order (Kuhn-Tucker) conditions are:

$$\frac{\partial\mathcal{H}}{\partial C} = U'(C) - \lambda \leq 0, \quad C \geq 0, \quad C\frac{\partial\mathcal{H}}{\partial C} = 0, \tag{I.6}$$

$$\frac{\partial\mathcal{H}}{\partial E} = \lambda P_E\left[(1-t_G)\,\frac{\dot{P}_E}{P_E} + r_D\,(1-t_Y) - (1-t_R)\,r\right] - \mu P_E \leq 0, \\ E \geq 0, \quad E\frac{\partial\mathcal{H}}{\partial E} = 0, \tag{I.7}$$

$$\frac{\partial\mathcal{H}}{\partial B} = -\mu \leq 0, \quad B \geq 0, \quad B\frac{\partial\mathcal{H}}{\partial B} = 0, \tag{I.8}$$

$$\dot{\lambda} - \rho\lambda = -\frac{\partial\mathcal{H}}{\partial A} = -\lambda\,(1-t_R)\,r - \mu. \tag{I.9}$$

Assuming that consumption is *essential* (i.e. the marginal felicity of the first infinitesimal amount of consumption is infinite, $\lim_{C\to 0} U'(C) = \infty$) it follows that consumption will always be strictly positive ($C > 0$). Hence, the first expression in (I.6) holds with equality, i.e. $U'(C) = \lambda$. If some government bonds are held by the household ($B > 0$), then it follows from (I.8) that $\mu = 0$. If some shares are also held ($E > 0$), then it follows from (I.72) that $(1 - t_G) \frac{\dot{P}_E}{P_E} + r_D (1 - t_Y) - (1 - t_R) r = 0$. Hence, in the *interior solution*, with both $B > 0$ and $E > 0$, the expressions (I.6)-(I.9) simplify to (5.71)-(5.73).

****

5.3.2 Representative firm

There are many, perfectly competitive firms, using a constant returns to scale technology to produce a single homogeneous good, $Y(t)$. Just as with the households, we normalize the number of firms to unity and argue on the basis of the representative firm. To keep the model as simple as possible, we abstract from corporate debt (so that financing is by retained earnings or by new equities) and assume that there is no depreciation of capital. To ensure that the firm has a well-defined capital accumulation decision we postulate the existence of firm-level adjustment costs.[16]

Gross operating profit of the firm is denoted by $\Pi(t)$ and defined as:

$$\Pi(t) \equiv F(K(t), L(t)) - W(t) L(t), \tag{5.75}$$

where $K(t)$ is the physical capital stock (machines, buildings, cars, PCs, etcetera), $L(t)$ is labour demand, and $F(\cdot)$ is a constant returns to scale production function with the properties stated in equation (5.2) above. Output is used as the numeraire commodity so the price of output has been set equal to unity ($P(t) = 1$).

Corporate profit is taxed at rate t_K and after-corporate-tax profit is either paid out to household-investors in the form of dividends, $D(t)$, or kept in the form of retained earnings, $RE(t)$:

$$(1 - t_K) \Pi(t) = D(t) + RE(t). \tag{5.76}$$

The capital accumulation identity abstracts from physical depreciation and is given by:

$$\dot{K}(t) = I(t), \tag{5.77}$$

[16] In the absence of adjustment costs, the firm would (unrealistically) be able to freely vary its stock of capital at any time. With adjustment costs, the capital stock can only be changed *gradually* over time and a well-defined investment policy can be formulated. Readers in need of a brief introduction to Tobin's q theory of firm investment are referred to Heijdra and van der Ploeg (2002, ch. 4). See also Summers (1981b) and Hayashi (1982).

where $I(t)$ is investment and $\dot{K}(t) \equiv dK(t)/dt$. The cost of investment including convex adjustment costs equals:

$$\phi\left(\frac{I(t)}{K(t)}\right)K(t), \tag{5.78}$$

where $\phi'(\cdot) > 0$ and $\phi''(\cdot) > 0$. The intuition underlying (5.78) is as follows. First, for a given installed capital stock, $K(t)$, the costs of investment rise more than proportionally with investment, i.e. one large investment is more costly to the firm than a sequence of small investments leading to the same change in the capital stock. Second, for a given level of investment, the larger the firm is (in terms of its installed capital stock), the less disruption cost per unit of capital it experiences. By a suitable choice of units we can ensure that $\phi(0) = 0$ and $\phi'(0) = 1$.[17]

In the absence of corporate bonds, the firm can finance its investment plans by retained earnings, or by selling new equity, or both. The financing constraint of the firm is thus given by:

$$RE(t) + P_E(t)\dot{E}(t) = \phi\left(\frac{I(t)}{K(t)}\right)K(t). \tag{5.79}$$

By combining (5.76) and (5.79) and assuming that $RE(t) > 0$ we obtain the following expression for dividends:

$$D(t) = (1 - t_K)\Pi(t) - \phi\left(\frac{I(t)}{K(t)}\right)K(t) + P_E(t)\dot{E}(t). \tag{5.80}$$

The market value for outstanding shares is:

$$V(t) = P_E(t)E(t). \tag{5.81}$$

Finally, by using (5.74), (5.80), and (5.81) we can derive the fundamental differential equation for the value of shares:[18]

$$\dot{V}(t) = \frac{\rho}{1 - t_G}V(t) - \left[(1 - t_K)\Pi(t) - \phi\left(\frac{I(t)}{K(t)}\right)K(t)\right] + \frac{t_Y - t_G}{1 - t_G}D(t). \tag{5.82}$$

Several things are worth noting about this expression. First, provided the dividend income tax differs

[17] An example of a $\phi(\cdot)$ function for which these results hold is:

$$\phi(\cdot) = \frac{I}{K}\left(1 + b\frac{I}{K}\right),$$

with $b > 0$.

[18] The derivation proceeds as follows. Differentiating (5.81) we get:

$$\dot{V}(t) = \dot{P}_E(t)E(t) + P_E(t)\dot{E}(t).$$

The household no-arbitrage equation (5.74) can be rewritten as:

$$\rho V(t) = (1 - t_G)\dot{P}_E(t)E(t) + (1 - t_Y)D(t).$$

By combining these expressions with (5.80) we obtain (5.82).

from the capital gains tax ($t_Y \neq t_G$) dividends matter to the determination of the value of the firm. If dividends are in any way related to the value of the firm they will end up affecting the cost of capital (see below). Second, as we saw in the previous section, provided $t_Y > t_G$ (the standard case) it is not optimal for the firm to pay dividends. However, in reality firms do pay dividends, and for this reason Turnovsky (1990, p. 497) formulates three alternative assumptions regarding dividend policy of the firm.

Under *Rule 1*, the firm offers its stock holders a fixed dividend yield by maintaining a constant dividend payout ratio, r_D. Dividend payments are thus given by:

$$D(t) = r_D V(t), \tag{5.83}$$

where $r_D \geq 0$. Under *Rule 2* it is assumed that the marginal source of financing is new equities only, i.e. $RE(t) = 0$ in equation (5.79) above. Any remaining after-corporate-tax profits are distributed in the form of dividends:

$$D(t) = (1 - t_K)\,\Pi(t). \tag{5.84}$$

Finally, under *Rule 3* the marginal source of financing consists of retained earnings only, i.e. $P_E(t)\,\dot{E}(t) = 0$ in equation (5.79) above, $RE(t) = \phi(I(t)/K(t))\,K(t)$, and dividends are determined residually:

$$D(t) = (1 - t_K)\,\Pi(t) - \phi\left(\frac{I(t)}{K(t)}\right)K(t) \tag{5.85}$$

By using the different dividend rules in (5.82) we obtain three alternative expressions for the fundamental differential equation of $V(t)$. In the interest of brevity we restrict attention to rule 1. (The reader is invited to investigate the consequences of rules 2 and 3.) For dividend rule 1 we obtain by substituting (5.83) into (5.82):

$$\dot{V}(t) = \frac{\rho + (t_Y - t_G)\,r_D}{1 - t_G}V(t) - \left[(1 - t_K)\,\Pi(t) - \phi\left(\frac{I(t)}{K(t)}\right)K(t)\right]. \tag{5.86}$$

Clearly, since the coefficient for $V(t)$ on the right-hand side is positive, equation (5.86) is an unstable differential equation in $V(t)$. The only economically sensible (no-bubble) solution is obtained by solving this differential equation forward in time and imposing the following terminal condition:

$$\lim_{t\to\infty} V(t)\exp\left[-\int_0^t \theta(\tau)\,d\tau\right] = 0, \tag{5.87}$$

where the cost of capital, $\theta(\tau)$, is defined as:

$$\theta(\tau) \equiv \frac{\rho + (t_Y - t_G)\,r_D}{1 - t_G}. \tag{5.88}$$

The terminal condition (5.87) ensures that the value of the firm remains finite. We thus focus on the *fundamental value* of the firm. In (5.88), the cost of capital depends on income and capital gains taxes and on the payout yield r_D. Note that if the firm pays no dividends, and sets $r_D = 0$, then this expression is the same as in (5.67) (simplified for the absence of corporate debt and recast in continuous time). Provided r_D, t_Y, and t_G remain constant over the indefinite future, the cost of capital will be constant also. For future purposes we denote this initial cost of capital level by θ^*.

The value of the firm at time $t = 0$ (the planning period) is given by:

$$V(0) = \int_0^\infty \left[(1 - t_K) \Pi(t) - \phi\left(\frac{I(t)}{K(t)} \right) K(t) \right] \exp\left[-\int_0^t \theta(\tau)\, d\tau \right]. \tag{5.89}$$

Equation (5.89) is the firm's objective function *under dividend rule 1*. It is maximized by the appropriate choice of $K(t)$, $I(t)$, and $L(t)$ subject to the accumulation identity (5.77), the definition of gross operating profit (5.75), and taking as given the initial capital stock, $K(0)$.

The key first-order necessary conditions for an interior solution can be written as follows:[19]

$$W(t) = F_L(K(t), L(t)), \tag{5.90}$$

$$q(t) = \phi'\left(\frac{I(t)}{K(t)} \right), \tag{5.91}$$

$$\dot{q}(t) = \theta(t) q(t) - (1 - t_K) F_K(K(t), L(t)) + \phi\left(\frac{I(t)}{K(t)} \right) - \frac{q(t) I(t)}{K(t)}, \tag{5.92}$$

where $q(t)$ is Tobin's q, representing the replacement value of installed capital. Equation (5.90) is a standard labour demand equation (taking the same form as (5.4) above), (5.91) relates optimal investment demand to Tobin's q and the installed stock of capital at time t, and (5.92) is the dynamic expression for Tobin's q.

Two further simplifications are incorporated. First, under the assumption of labour market clearing, labour demand equals the exogenous supply, i.e. $L(t) = \bar{L}$, and the marginal product of capital becomes $F_K(K(t), \bar{L})$ which is a downward sloping function of $K(t)$ only (since $F_{KK}(K(t), \bar{L}) < 0$). (Of course, once $K(t)$ is determined, the market clearing wage rate is determined residually through equation (5.90).)

The second simplification is obtained by assuming a specific functional form for adjustment costs, namely a quadratic specification, i.e. $\phi(z) = z(1 + bz/2)$ is chosen so that (5.91) and the optimized

[19] The optimization problem is solved with the method of optimal control (see the Technical Appendix to this chapter or Heijdra and van der Ploeg (2002, pp. 700-702)). The *current-value Hamiltonian* is:

$$\mathcal{H} \equiv (1 - t_K)[F(K(t), L(t)) - W(t) L(t)] - \phi\left(\frac{I(t)}{K(t)} \right) K(t) + q(t) I(t),$$

where $K(t)$ is the state variable, $q(t)$ is the co-state variable, and $L(t)$ and $I(t)$ are the control variables. The first-order conditions are $\partial\mathcal{H}/\partial L(t) = \partial\mathcal{H}/\partial I(t) = 0$, $\dot{q}(t) - \theta(t) q(t) = -\partial\mathcal{H}/\partial K(t)$, and $\dot{K}(t) = \partial\mathcal{H}/\partial q(t)$. Upon simplifying we obtain (5.77) and (5.90)-(5.92).

value for $\phi(\cdot)$ can be written as:

$$\frac{I(t)}{K(t)} = \frac{q(t)-1}{b}, \tag{5.93}$$

$$\phi\left(\frac{I(t)}{K(t)}\right) = \frac{(q(t)-1)(q(t)+1)}{2b}. \tag{5.94}$$

Imposing all simplifications, the model can be reduced to two dynamic equations in $K(t)$ and $q(t)$:

$$\dot{K}(t) = \left(\frac{q(t)-1}{b}\right) K(t), \tag{5.95}$$

$$\dot{q}(t) = \theta(t) q(t) - (1-t_K) F_K(K(t), \bar{L}) - \frac{(q(t)-1)^2}{2b}. \tag{5.96}$$

The phase diagram of the model is shown in Figure 5.3. In that figure, the $\dot{K}=0$ line implies a unique value for Tobin's q, denoted by $q^*=1$ (horizontal line). The dynamics of the capital stock is obtained from (5.95):

$$\frac{\partial \dot{K}}{\partial q} = \frac{K}{b} > 0, \tag{5.97}$$

i.e. for points above (below) the $\dot{K}=0$ line, the capital stock increases (decreases) over time. This has been indicated with horizontal arrows in Figure 5.3. The $\dot{q}=0$ line is downward sloping in or near the steady state (where $q^* \approx 1$):

$$\left(\frac{dq}{dK}\right)_{\dot{q}=0} = \frac{b(1-t_K)(F_{KK})^*}{b\theta^* - (q^*-1)} < 0, \tag{5.98}$$

where asterisks denote steady-state values. The dynamics of Tobin's q is obtained from (5.96):

$$\frac{\partial \dot{q}}{\partial K} = -(1-t_K) F_{KK} > 0. \tag{5.99}$$

To the right (left) of the $\dot{q}=0$ line, the capital stock is too high (low), the marginal product of capital is too low (high), so part of the return on capital consists of capital gains ($\dot{q}>0$) (capital losses, $\dot{q}<0$) on installed capital. This has been indicated with vertical arrows in the figure. Given the arrow configuration it is clear that there is a unique saddle-point stable equilibrium at point E_0, where the steady-state equilibrium values for q and K are equal to:

$$q^* = 1, \tag{5.100}$$

$$\theta^* = (1-t_K) F_K(K^*, \bar{L}). \tag{5.101}$$

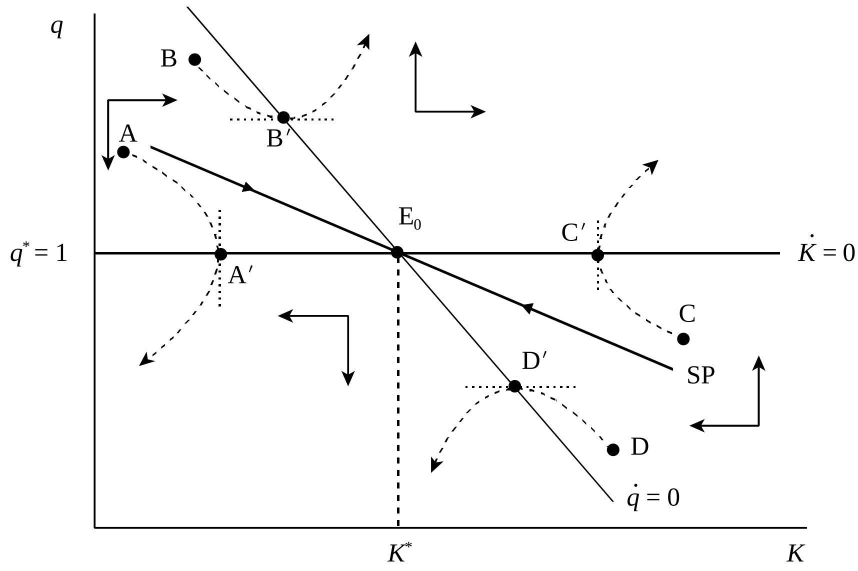

Figure 5.3: Phase diagram investment model

5.3.3 Tax policy

In this subsection it is demonstrated how the model can be used to conduct tax policy analysis at the firm level. Various tax rates affect the investment system (5.95)-(5.96). It is clear from equation (5.88) that the cost of capital under dividend rule 1 depends on the capital gains tax, t_G, for sure, and also on the dividend income tax, t_Y, provided the firm pays out some dividends (by setting $r_D > 0$). In addition, the corporate income tax, t_K, affects the size of the after-tax marginal product of capital. In principle, therefore, all these taxes will have an effect on the firm's investment and output plans.

A *quantitative* local tax policy can be based on a linearized version of the model. Here we will, however, focus on a *qualitative analysis* by deducing the effects by graphical means. We restrict attention to the corporate tax rate. (The reader is invited to study the effects of the other tax rates under the various dividend rules.) In a perfect foresight model the timing of the shock is of crucial importance. Three time points are important to characterize a shock:

- *Announcement* time (T_A): at what time does the shock become known to the relevant agent?
- *Implementation* time (T_I): at what time will the shock actually occur?
- *Ending* time (T_E): at what time will the shock be ended?

If the announcement time coincides with the implementation time ($T_A = T_I$) then we call the shock *unanticipated*: the agent is taken by surprise and the shock occurs immediately. In contrast, if the announcement time predates the implementation time ($T_A < T_I$) then we call the shock *anticipated*: the agent receives the news about a future shock but can brace himself partially for its effects. Finally, if the ending time is infinite ($T_E \to \infty$), then we call the shock *permanent*, whereas the shock is called *temporary* if the ending time is finite ($T_E \ll \infty$).

Using this terminology, it is possible to consider four different cases, namely a shock which is (i) permanent and unanticipated, (ii) permanent and anticipated, (iii) temporary and unanticipated, and (iv) temporary and anticipated. In a perfect foresight world these shocks will have drastically different effects.

Consider a *permanent and unanticipated* increase in the corporate tax rate (t_K) that is announced and implemented at time $T_A = T_I = 0$ and lasts forever ($T_E \to \infty$). In terms of Figure 5.4 the increase in t_K results in a downward shift of the $\dot{q} = 0$ line, say from $(\dot{q} = 0)_0$ to $(\dot{q} = 0)_1$. Formally, we obtain from (5.96):

$$\left(\frac{\partial q}{\partial t_K}\right)_{\dot{q}=0} = -\frac{(F_K)^*}{b\theta^* - (q^* - 1)} < 0. \tag{5.102}$$

Since the cost of capital is unaffected by the corporate tax rate (see (5.88) above), it follows from (5.100)-(5.101) that the long-run effects on q and K are given by:

$$\frac{dq^*}{dt_K} = 0, \tag{5.103}$$

$$\frac{dK^*}{dt_K} = \frac{(F_K)^*}{(1 - t_K)(F_{KK})^*} < 0. \tag{5.104}$$

The long-run equilibrium shifts from E_0 to E_1 and the corporate tax reduces the capital stock in the long run. Assuming that the firm is initially at the steady-state equilibrium point E_0, the transitional dynamics is as follows. At impact ($t = 0$) the capital stock is predetermined and, immediately following the shock, the only stable trajectory to the new equilibrium is the saddle path, SP. Tobin's q jumps down at impact, from point E_0 to point A directly below it. Intuitively, the firm realizes that the path of the after-corporate-tax marginal product of capital, $(1 - t_K) F_K(K(t), \bar{L})$, is temporarily lower because of the tax change. Since $q(0)$ is the capitalized value of these after-tax marginal products from the perspective of the impact period, it falls immediately at impact. At point A, Tobin's q is lower than unity so it follows from (5.93) that investment falls (becomes negative, in fact). As a result, the capital stock gradually falls over time ($\dot{K}(t) < 0$). During transition the after-tax marginal product of capital, $(1 - t_K) F_K(K(t), \bar{L})$, is gradually restored to its initial level.

A *temporary and unanticipated* increase in the corporate tax rate (t_K) has vastly different effects. The shock is announced and implemented at time $T_A = T_I = 0$ and will last until $T_E > 0$. In order to deduce the effects in a graphical setting we employ the following heuristic solution concept. We postulate that the solution must satisfy the following criteria:

- If it occurs at all, a *discrete* adjustment in q must occur at the time the news becomes available (i.e. at time T_A), and there cannot be a further discrete adjustment in q after T_A. Intuitively, an *anticipated* jump in q would imply an infinite (shadow) capital gain or loss (since there would be a finite change in q in an infinitesimal amount of time). Hence, the solution principle amounts

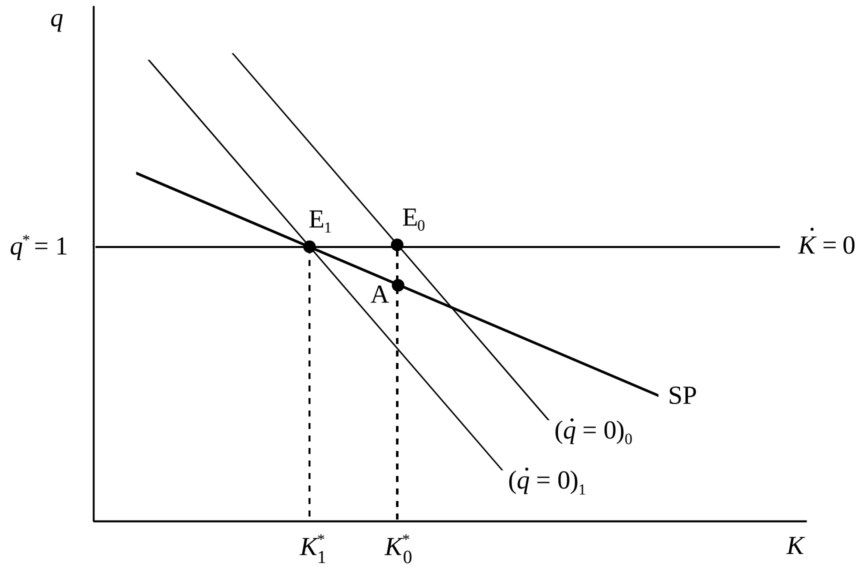

Figure 5.4: Unanticipated and permanent incease in the corporate tax

to requiring that all jumps occur when something truly unexpected occurs (which is at time T_A). Obviously, at T_A there is an infinite capital loss, but it is unanticipated.

- At time $T_A = 0$ the capital stock is predetermined.
- While the shock is relevant (for $T_A \leq t \leq T_E$), the dynamics associated with the "new" situation determines the motion of $q(t)$ and $K(t)$.
- At time $t = T_E$, the (q, K) combination must be exactly on the saddle path leading back to the initial equilibrium. Recall from the first requirement that q is not allowed to jump onto this saddle path. It must arrive smoothly at exactly the right time.

Using the heuristic solution principle we can deduce the impact and transitional effects of the temporary tax change. In terms of Figure 5.5, the shock leads to shift in the $\dot{q} = 0$ line from $(\dot{q} = 0)_0$ to $(\dot{q} = 0)_1$ which lasts until time T_E. The impact effect consists of a jump from E_0 to point B. For $T_A < t < T_E$ the firm follows the trajectory from B to C and then from C to D. The relevant dynamic effects are those implied by the "equilibrium" at point E′ and are indicated by the arrows. At $t = T_E$, the shock is ended, the $\dot{q} = 0$ line is shifted back from $(\dot{q} = 0)_1$ to $(\dot{q} = 0)_0$, and the firm is at point D which lies on the then relevant saddle path, SP. For $t > T_E$, the firm gradually moves along SP to the ultimate equilibrium at E_0. The temporary tax increase exerts a temporary effect on the capital stock.

Using the heuristic solution concept it is also possible to characterize the effects of anticipated permanent or temporary policies. This is left as an exercise for the reader. The key thing to remember about this subsection is the crucial role of timing and expectations in determining the effects of policy changes.

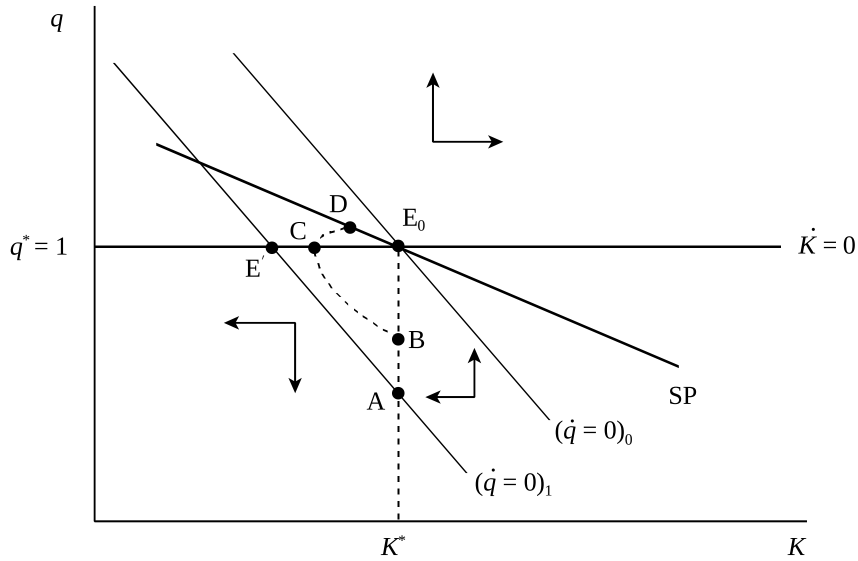

Figure 5.5: Unanticipated and temporary increase in the corporate tax

5.4 Empirical evidence

In an influential paper, Poterba and Summers (1985) use data for the United Kingdom to study the effects of dividend taxation on security returns, dividend payout rates, and corporate investment. The UK data are especially useful because there were various major dividend tax reforms during the sample period, 1950-1983. They postulate three competing hypotheses. First, the *Traditional View* (or *Old View*), according to which the dividend tax acts as an additional tax on corporate profits (over and above the corporate tax). The early proponents of this "double taxation" view, such as Harberger (1962) and McLure (1979), were criticized because they ignored the role of corporate financial policy. In contrast, Poterba and Summers (1985, pp. 241-244) formulate an optimizing model in which it is advantageous for the firm to pay dividends. In particular, they assume that the discount rate that is applied to the firm's income stream (e.g. ρ_z in equation (5.61) above) depends negatively on the dividend payout ratio.

The second hypothesis is the *Tax Capitalisation View* (or *New View*) that was developed by King (1974), Auerbach (1979b), and Bradford (1981). According to this view, the dividend tax does not affect the cost of capital and thus does not affect anything real, including the marginal incentive to invest. The dividend tax is essentially a lump-sum tax on initial holders of corporate capital. See also the discussion surrounding equations (5.64)-(5.65) above.

Finally, the third hypothesis is the *Tax Irrelevance View* pioneered by Miller and Scholes (1978, 1982). According to this view, dividend paying firms are not penalized in the market. This is because the marginal investor does not effectively pay any taxes on dividends or capital gains (e.g. because they are institutional investors, or because they use sophisticated tax strategies). As a result, the dividend tax has no effect on the value of the firm or on its real investment decisions.

Poterba and Summers (1985, p. 244) show that each of the three hypotheses produces a set of testable predictions regarding the cost of capital, the equilibrium value of Tobin's (marginal) q, and the responsiveness of investment and the dividend payout ratio to a permanent change in the dividend tax. These predictions are then confronted with the data, which leads Poterba and Summers (1985) to the conclusion that the traditional view is most consistent with British postwar data on security returns. As is pointed out by Sinn (1991, p. 29) and Sørenson (1994, p. 436), this conclusion is somewhat troublesome because the old view is based on the counterfactual assumption that new share issues are the marginal source of equity finance. Indeed, Sinn cites evidence on US non-financial corporations suggesting that internal financing accounts for about 68% of gross investment, debt financing for 31%, and new share issues for only 1.2%.

5.5 Punchlines

In this chapter we study the effects of taxation on firm behaviour. We start out by formulating a basic static model of firm behaviour under perfect competition. The competitive firm faces a constant returns to scale technology, is a price taker on the markets for both its output and all its inputs, and produces a homogeneous commodity by hiring factors of production in a cost minimizing fashion. Any tax which affects the relative price of inputs will cause a *factor substitution effect*. When all production factors can be varied at will, marginal cost depends only on tax-inclusive factor prices, the market supply curve is horizontal, and any cost change is passed on to consumers on a one-for-one basis. This is the *price shifting* phenomenon.

Next we assume that the firm operates under conditions of imperfect competition and investigate whether *price overshifting* is possible, i.e. whether a cost change is shifted more than one-for-one to consumers. In the standard monopoly case with a constant elasticity of demand, price overshifting must occur! However, following an increase in marginal cost, the maximized profit level of the monopolist must decrease, i.e. *profit overshifting* cannot occur in this case. The monopolist could have imposed the higher cost on himself but did not choose to do so.

Matters are different in an oligopolistic setting. Using a homogeneous-good quantity-setting oligopoly model with conjectural variations, it is shown that price overshifting occurs in the symmetric equilibrium provided the elasticity of the slope of the inverse demand function (i.e., Seade's $\bar{E}$) is sufficiently high ($\bar{E} > 1$). In contrast to the monopoly case, profit overshifting is also a distinct possibility in this case—it occurs provided $\bar{E} > 2$. If the industry demand features a constant elasticity, the requirements for price overshifting and profit overshifting are both fulfilled if the demand curve is inelastic. Intuitively, the cost increase acts as an implicit collusion device prompting all firms to restrict output, thus increasing their profit.

In the second section of this chapter we study the optimal financial decisions of a representative perfectly competitive firm. An infinite-period dynamic model of a *mature* firm is formulated in discrete time. This firm operates under conditions of perfect foresight and is able to finance its real investment plans by means of profit retentions, corporate debt issues, or share emissions. In the absence of personal taxes on dividends and accrued capital gains, the Modigliani-Miller result is obtained which shows that the particular source of equity funds used to finance a given investment policy has no impact on the firm's valuation. The cost of capital to the firm is the minimum of the after-corporate-tax interest rate on corporate bonds and the discount rate of equity owners.

In reality, both dividend income and capital gains on shares are taxed. Typically, the tax on dividends exceeds that on capital gains. In such a setting, it is never optimal for the firm to issue new shares and pay dividends at the same time. The relatively mild tax treatment of capital gains makes it attractive for the firm to repurchase its own shares, thus transmitting lightly taxed income to its shareholders. According to the *tax capitalization view*, a firm which has exhausted all such low-tax income transmission opportunities finances its investment plans by means of profit retentions. The cost of capital is then the minimum of the after-corporate-tax interest rate on corporate bonds and the equity owners' discount factor corrected for the capital gains tax. Neither the dividend payout ratio nor the dividend income tax affects the cost of capital. The dividend tax acts as a lump-sum tax on wealth in the corporate sector (*trapped equity*).

In the third section of this chapter we study the interaction between the firm's real and financial decisions in a continuous-time perfect foresight model of the firm. The firm faces adjustment costs of investment so that the value of installed capital may deviate from its replacement value in the market (Tobin's q theory). Depending on the assumptions regarding the firm's dividend policy, expressions for the cost of capital are obtained. A qualitative comparative dynamics exercise shows how tax changes affect the firm's investment plans, both at impact, during transition, and in the long run. Because the economic agents are forward looking and blessed with perfect foresight, the timing of any tax change critically determines its economic effects.

The chapter concludes with a brief discussion of the empirical evidence regarding the effects of dividend taxation on security returns, dividend payout rates, and corporate investment. The evidence seems to suggest that the tax capitalization view misses out on important aspects of reality. Although there are many partial insights into how the theory of the firm's financial decision making could be brought into closer accordance with reality, no consensus model exists at this stage.

Further reading

Basic theory. Atkinson and Stiglitz (1980, lecture 5) cover much of the same topics as we do. In addition, they also discuss additional issues such as the tax treatment of capital depreciation allowances and the effects of inflation on the cost of capital. Good surveys on capital taxation, the firm's financial policy, and the cost of capital are Auerbach (1983, 2002), Sinn (1987, 1991), and Sørenson (1994).

Oligopoly. Key references to the oligopoly model are Seade (1980a, 1980b, 1985), Katz and Rosen

(1985), Kurz (1985), Dixit (1986), Stern (1987b), and Hamilton (1999). Anderson et al. (2001) study tax incidence in an oligopoly model with differentiated (rather than homogeneous) products. Good surveys of the modern literature are Vives (1999) and Fullerton and Metcalf (2002, pp. 1823-1832). There is a large empirical literature on price overshifting and/or profit overshifting. See, for example, Sumner (1981), Bulow and Pfleiderer (1983), Sullivan (1985), Poterba (1996), Besley and Rosen (1999), and Feuerstein (2002).

Tobin's q theory. The interaction between the real and financial policies of the firm is typically studied in a setting with adjustment costs of investment. Prominent contributions are Auerbach (1979a, 1984), Poterba and Summers (1983), Hayashi (1985), Goulder and Summers (1989), Brock and Turnovsky (1981), and Turnovsky (1990). Howitt and Sinn (1989) study gradual capital income tax reform in a closed-economy general equilibrium model. See Hassett and Hubbard (2002) for a survey of the recent theoretical and empirical literature on taxation and firm investment.

Firm behaviour under uncertainty. There is a large literature on firm behaviour under conditions of risk. Prominent contributions are Sandmo (1971), Leland (1972), Ishii (1977), Batra and Ullah (1974), Eeckhoudt and Hansen (1980), Eeckhoudt and Gollier (1995, ch. 11), Eldor and Zilcha (1990), and Zilcha and Eldor (2004).

Duality and the cost function. A very accessible discussion of the cost function is presented by Jehle and Reny (2001, ch. 3) and Varian (1992, p. 72). The restricted cost function, in which one or more of the factors of production cannot be varied in the short-run, is explained by Chamber (1988, ch. 93). A very advanced treatment of duality methods is McFadden (1978).

Technical appendix on optimal control theory

In the text of this chapter we make use of the technique of optimal control. While a full treatment of this technique is beyond the scope of the book, it is nevertheless useful to briefly discuss some of the key results here. Our treatment is based on material taken from Heijdra and van der Ploeg (2002, pp. 700-702).

The proto-typical optimal control problem encountered in economics takes the following form. The objective function is defined as:

$$y(0) = \int_0^\infty F\left(x(t), u(t), t\right) e^{-\rho t} dt, \tag{A5.1}$$

where $x(t)$ is the *state variable*, $u(t)$ is the control variable, $e^{-\rho t}$ is the discount factor, and t is time. The state and control variable are related according to the following *state equation*:

$$\dot{x}(t) = f\left(x(t), u(t), t\right). \tag{A5.2}$$

The state equation thus describes the law of motion for the state variable. The *initial condition* for the state variable is given by:

$$x(0) = x_0, \tag{A5.3}$$

where x_0 is a given constant (e.g. the accumulated stock of some resource).

The objective is to find a time path for the control variable, $u(t)$ for $t \in [0, \infty)$, such that the objective function (A5.1) is maximized given the state equation (A5.2) and the initial condition (A5.3). To solve this problem one formulates a so-called *Hamiltonian* which takes the following form:

$$\mathcal{H} \equiv F\left(x(t), u(t), t\right) e^{-\rho t} + \lambda(t) f\left(x(t), u(t), t\right), \tag{A5.4}$$

where $\lambda(t)$ is the *co-state variable* which plays a role similar to the Lagrange multiplier encountered in static optimization problems. The *Maximum Principle* furnishes the following conditions (for $t \in [0, \infty)$):

$$\frac{\partial \mathcal{H}}{\partial u(t)} = 0, \tag{A5.5}$$

$$\dot{x}(t) = \frac{\partial \mathcal{H}}{\partial \lambda(t)}, \tag{A5.6}$$

$$\dot{\lambda}(t) = -\frac{\partial \mathcal{H}}{\partial x(t)}. \tag{A5.7}$$

Equation (A5.5) says that the control variable should be chosen such that the Hamiltonian is maximized, (A5.6) gives the equation of motion for the state variable, and (A5.7) gives the equation of motion for the co-state variable.

Current-value Hamiltonian

An equivalent way of solving the same problem is to work with the *current-value Hamiltonian*, which is defined as follows:

$$\mathcal{H}_C \left[\equiv \mathcal{H}e^{\rho t}\right] = F\left(x(t), u(t), t\right) + \mu(t) f\left(x(t), u(t), t\right), \tag{A5.8}$$

where $\mu(t) \equiv \lambda(t)e^{\rho t}$ is the redefined co-state variable. The first-order conditions expressed in terms of the current-value Hamiltonian are:

$$\frac{\partial \mathcal{H}_C}{\partial u(t)} = 0, \tag{A5.9}$$

$$\dot{x}(t) = \frac{\partial \mathcal{H}_C}{\partial \mu(t)}, \tag{A5.10}$$

$$\dot{\mu}(t) - \rho\mu(t) = -\frac{\partial \mathcal{H}_C}{\partial x(t)}. \tag{A5.11}$$

If there are n state variables and m controls then the same methods carry over except, of course, that $x(t) \equiv [x_1(t), ..., x_n(t)]$ and $u(t) \equiv [u_1(t), ..., u_m(t)]$ must be interpreted as vectors and the set of conditions is suitable expanded:

$$\frac{\partial \mathcal{H}_C}{\partial u_j(t)} = 0, \tag{A5.12}$$

$$\dot{x}_i(t) = \frac{\partial \mathcal{H}_C}{\partial \mu_i(t)}, \tag{A5.13}$$

$$\dot{\mu}_i(t) - \rho\mu_i(t) = -\frac{\partial \mathcal{H}_C}{\partial x(t)}, \tag{A5.14}$$

where $\lambda_i(t)$ is the co-state variable corresponding to the state variable $x_i(t)$, $j = 1, ..., m$, and $i = 1, ..., n$.

(In)equality constraints

Recall the original problem (A5.1)-(A5.3). Suppose that there is an additional constraint in the form of:

$$g\left(x(t), u(t), t\right) \leq c, \tag{A5.15}$$

where c is some constant. Suppose furthermore that there is a non-negativity constraint on the control variable, i.e. $u(t) \geq 0$ is required. The way to deal with these inequalities is to form the following *current-value Lagrangian*:

$$\mathcal{L}_C = F\left(x(t), u(t), t\right) + \mu(t) f\left(x(t), u(t), t\right) + \theta(t)\left[c - g\left(x(t), u(t), t\right)\right], \tag{A5.16}$$

where $\theta(t)$ is the Lagrange multiplier associated with the inequality constraint (A5.15).

The first-order conditions are now:

$$\frac{\partial \mathcal{L}_C}{\partial u(t)} \leq 0, \quad u(t) \geq 0, \quad u(t)\frac{\partial \mathcal{L}_C}{\partial u(t)} = 0, \tag{A5.17}$$

$$\frac{\partial \mathcal{L}_C}{\partial \theta(t)} \geq 0, \quad \theta(t) \geq 0, \quad \theta(t)\frac{\partial \mathcal{L}_C}{\partial \theta(t)} = 0, \tag{A5.18}$$

$$\dot{x}(t) = \frac{\partial \mathcal{L}_C}{\partial \mu(t)}, \tag{A5.19}$$

$$\dot{\mu}(t) - \rho\mu(t) = -\frac{\partial \mathcal{L}_C}{\partial x(t)}. \tag{A5.20}$$

Equation (A5.17) gives the *Kuhn-Tucker conditions* taking care of the non-negativity constraint on the control variable, (A5.18) gives the Kuhn-Tucker conditions for the inequality constraint (A5.15), and (A5.19) and (A5.20) give the laws of motion of, respectively, the state variable and the co-state variable.

Chapter 6

Tax incidence in general equilibrium

The purpose of this chapter is to discuss the following topics:

- Why is a tax levied on agents situated on one side of the market in some cases ultimately borne (entirely or partially) by agents on the other side of the market?
- How can we model tax incidence in a partial equilibrium setting and what are the main limitations of this Marshallian approach?
- How does tax shifting occur in the prototypical two-factor-two-commodity general equilibrium model, and what role is played by output effects and factor substitution effects?
- What are Applied General Equilibrium models and how can they be used for the analysis of drastic tax policy changes?
- How robust is the standard two-factor-two-commodity model to deviations from its basic assumptions? In particular, what is the role of imperfectly competitive firm behaviour in the goods market and of market frictions in the labour market?

6.1 Introduction

In the previous chapters we have studied the effects of various taxes on the behaviour of the different market participants in isolation. For example, in Chapter 2 we studied how taxes levied on labour income affect the labour supply decision of households whilst in Chapter 3 we discussed how the various tax rates affect consumption and saving. In each case attention was restricted to the behavioural response by the agent upon whom the particular tax was actually levied, i.e. any market-induced repercussions were ignored.[1] In this chapter we move beyond the partial equilibrium approach and study the topic of *tax incidence* in general equilibrium.

[1] In Chapter 5 interactions between households and firms have already been introduced. The analysis there remained partial equilibrium in nature.

The basic insight of tax incidence analysis is the notion that the agent upon whom the tax is levied may not actually fully bear the tax, i.e. the tax may be *shifted* to other agents. Consider, for example, the labour market. A labour income tax is levied on the household whose labour supply will generally be affected as a result. The change in labour supply, following a change in the tax, may give rise to a change in the wage rate which adversely affects the demanders of labour (i.e. the firms). Part of the labour income tax may thus be borne by producers, even though they are not the ones upon whom the tax is levied by the tax authorities.

As is pointed out by Atkinson and Stiglitz (1980, pp. 160-161) there are different types of tax incidence analysis. Following a change in one or more taxes one could look at the effects on various economic quantities. First, one could look at the incomes of producers, consumers, and the suppliers of the production factors. Second, one could look at the functional income distribution by investigating the effects on the payments to the main production factors (e.g. capital and labour). Third, one could look at the personal income distribution. Fourth, one could take a spatial perspective and look at how different regions in an economy are affected. Finally, one could investigate how different generations are affected.

In this chapter we restrict attention to the first two types of tax incidence analysis. We study tax effects in a *neutral fashion*, i.e. only one tax is changed at a time and tax revenue is recycled in a lump-sum fashion to households. The reason for doing so is that it allows us to study the effect of that particular tax change in isolation. An alternative approach would be to consider budgetarily neutral *tax reform* packages, i.e. a change in one tax offset by a change in another tax in such a way that the government budget constraint is satisfied. In such a tax reform experiment two taxes are changed at the same time so that it is difficult to attribute the resulting effects on quantities and factor prices to each individual tax in isolation. We briefly return to the issue of tax reform in Chapter 9 below.

6.2 Tax incidence in partial equilibrium

Although the classical economists were keen to argue on the basis of general equilibrium principles, tax incidence analysis during the first half of the 20th century was dominated by the partial equilibrium methods of Alfred Marshall (1920). In the Marshallian type of incidence analysis, markets are studied in isolation, i.e. the interaction between demanders and suppliers on one market is taken into account but the spill-overs that may exist with other markets are not.

To illustrate the strengths and weaknesses of the Marshallian partial equilibrium approach, consider the following example taken from Atkinson and Stiglitz (1980, p. 162). There is a crop of some agricultural product (say "grapes") which is produced with the production factors land, K, and labour, L. Land can only be used for growing this particular crop and its quantity is fixed, i.e. $K = \bar{K}$. The supply of land is thus perfectly inelastic. In contrast, the supply of labour is perfectly elastic at the given wage rate W. The linearly homogeneous production function is defined as:

$$Y = F(L, \bar{K}), \tag{A5.1}$$

where Y is output and $F(\cdot)$ features a positive but diminishing marginal product of labour, i.e. $F_L(L, \bar{K}) > 0 > F_{LL}(L, \bar{K})$.[2] The competitive demand for labour, L^D, is obtained by equating the marginal product of labour to the real wage rate, i.e. $F_L\left(L^D, \bar{K}\right) = W/P$ or:

$$L^D = \bar{K} l(W/P), \tag{A5.2}$$

where P is the *producer price* of the good and we have used the fact that F_L is homogeneous of degree zero in L and $\bar{K}$. Obviously, the $l(\cdot)$ function features a negative derivative ($l'(\cdot) < 0$). By substituting (A5.2) into (A5.1) we obtain the expression for the competitive supply curve:[3]

$$Y^S = \bar{K} S(W/P), \tag{A5.3}$$

where $S_P \equiv \partial S/\partial P = -F_L(\cdot)^2 \bar{K} l'(\cdot)/P > 0$.

The demand curve on the market for grapes is assumed to take the following form:

$$Y^D = D(P_D, Z), \tag{A5.4}$$

where P_D is the *consumer price* of the good and Z is a vector of other variables influencing demand (such as the prices of other goods, income, etcetera). We assume that demand is downward sloping ($D_P \equiv \partial D(\cdot)/\partial P_D < 0$).

In Figure 6.1 the situation on the grape market is illustrated. Initially there is no tax on grape consumption so the producer price equals the consumer price, $P_D = P$. The initial equilibrium is at E_0, the equilibrium price is P_0, and the equilibrium quantity is Y_0. It is easy to show that the rents received by the land owners (Π_0) is represented by the area BE_0P_0:

$$\begin{aligned} \Pi_0 &\equiv P_0 Y_0 - \int_0^{Y_0} MVC(Y, W, \bar{K})\, dY \\ &= P_0 Y_0 - \left[TVC(Y_0, W, \bar{K}) - TVC(0, W, \bar{K})\right] \\ &= P_0 Y_0 - TVC(Y_0, W, \bar{K}), \end{aligned} \tag{A5.5}$$

[2] The production function is strictly quasi-concave and features the properties stated in equation (5.2) above.

[3] Consider, for example, the following Cobb-Douglas production function: $F(\cdot) \equiv A L^{\varepsilon} \bar{K}^{1-\varepsilon}$ (with $0 < \varepsilon < 1$). For this case, $l(\cdot)$ and $S(\cdot)$ are given, respectively, by $l(W/P) \equiv \left(\frac{W}{\varepsilon A P}\right)^{1/(\varepsilon-1)}$ and $S(W/P) = A\left(\frac{W}{\varepsilon A P}\right)^{\varepsilon/(\varepsilon-1)}$.

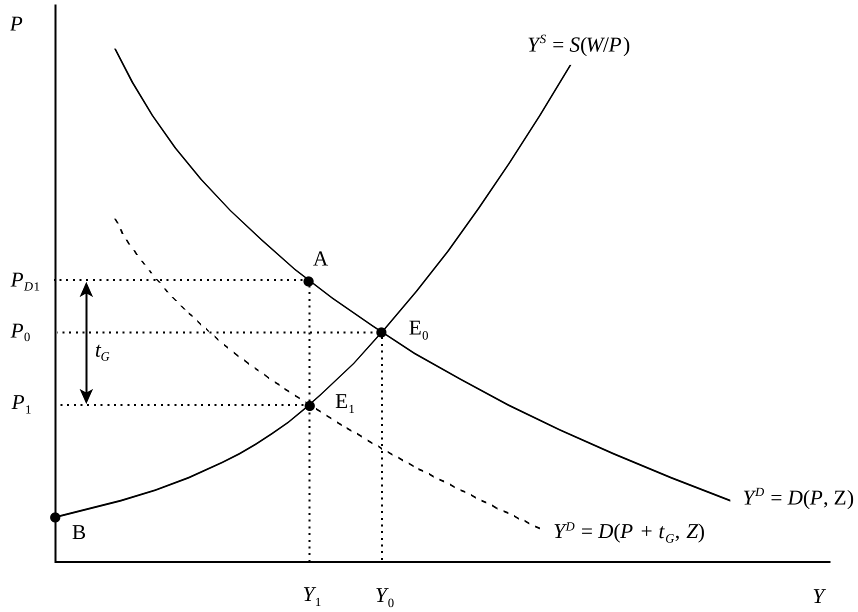

Figure 6.1: Tax incidence in partial equilibrium

where $TVC(\cdot)$ and $MVC(\cdot)$ are, respectively, total and marginal variable cost of production:

$$TVC(Y, W, \bar{K}) \equiv W\bar{F}(Y, \bar{K}), \tag{A5.6}$$

$$MVC(Y, W, \bar{K}) \equiv \frac{\partial W\bar{F}(Y, \bar{K})}{\partial Y}, \tag{A5.7}$$

and where $L = \bar{F}(Y, \bar{K})$ represents the minimum amount of labour that is needed to produce output Y, taking as given the available amount of land $\bar{K}$. Hence, L is implicitly determined via the production technology, $Y = F(L, \bar{K})$.[4]

The introduction of a tax on the consumption of grapes (t_G) increases the demand price to $P_D = P + t_G$, shifts the demand curve downwards (to the dashed line), and shifts the market equilibrium point to E_1. As a result of the tax, the producer price falls from P_0 to P_1, the demand price rises from P_0 to P_{D1}, the wage remains unchanged by assumption, and landowner rents decline from BE_0P_0 to BE_1P_1. It follows that the tax is borne by both consumers and landowners. The households used to pay P_0 for grapes and must now pay P_{D1}. The producers used to get P_0 for their grapes and now only receive P_1.

[4]With only one variable production factor, the short-run variable cost function is thus obtained by inverting the short-run production function. In the general case, there are n variable production factors, L_i, with associated rental rates, W_t. The definition for total variable cost is given by:

$$TVC(Y, W_1, \cdots W_n, \bar{K}) \equiv \min_{\{L_i\}} \sum_{i=1}^{n} W_i L_i \quad \text{subject to} \quad Y = F(\bar{K}, L_1, \cdots, L_n).$$

Of course, $TVC(Y, W_1, \cdots W_n, \bar{K})$ is increasing in Y. Furthermore, the constrained conditional input demand functions are obtained by Shephard's Lemma:

$$L_i = \frac{\partial TVC(Y, W_1, \cdots W_n, \bar{K})}{\partial W_i}.$$

It is clear from the diagram that the division of the burden depends on the elasticities of demand and supply. For example, if grapes were produced with land only, then supply would be perfectly inelastic (vertical), and the tax would be borne by landowners only. Conversely, if grapes were produced with labour only, then the supply curve would be perfectly elastic (horizontal), and the tax would be borne by consumers only.

There are several limitations associated with the partial equilibrium approach. As the grape example shows, the method is only valid under extreme assumption regarding the *supply side*, i.e. factor supplies are either totally elastic (labour) or totally inelastic (land). In the more general case, both factors of production may be used in other sectors of the economy so that the factor supply curves facing the grape industry are a lot more complicated. Similarly, via the *demand side* a change in the demand for grapes may affect the household's demands for goods produced in other sectors, and this may affect factor demands. In the partial equilibrium model, all factors other than the grape price are subsumed in the term Z, which is treated as an exogenous variable. As we shall argue in the next section, the general equilibrium approach to tax incidence analysis is much better suited to address these interactions across markets and across supply and demand.

6.3 A simple general equilibrium model

In a classic paper, Harberger (1962) formulated a simple analytical general equilibrium model of an economy with two production sectors and two factors of production. He used this "two-by-two" model to study the effects of the corporate tax on sectoral outputs, factor employments, and factor rewards. Subsequent major analytical contribution to this literature were made by Jones (1965, 1971a, 1971b). The Harberger-Jones approach is important for at least two reasons. First, it has been used successfully not only in the area of public economics but also in many other fields of economics, such as the pure theory of international trade and two-sector growth theory (to mention just a few). Second, following the advent of low-cost computing power, the Harberger-Jones approach has stimulated the construction of multi-sector, multi-factor applied general equilibrium (AGE) models of the economy.

The objective of this section is to construct a basic two-sector-two-factor general equilibrium model (without taxes) and to study its key properties. In order to facilitate the interpretation of the results, a simple geometric illustration of the model is developed. In Section 6.4 we introduce taxes into the model and study their effects.

6.3.1 A two-sector model

The basic static general equilibrium model has the following features. There are two sectors in the economy. In each sector, technology features constant returns to scale, and firms are perfect competitors in input and output markets. Outputs in the two sectors are denoted by X and Y, respectively, whilst prices are denoted by P_X and P_Y, respectively. Both sectors use the two factors of production, capital and labour. The amounts of capital and labour used in sector i are denoted by, respectively, K_i and L_i (where $i = X, Y$). The total supplies of capital and labour are fixed (at $\bar{K}$ and $\bar{L}$, respectively) and there is full employment of both factors. In addition, there is perfect intersectoral mobility of factors implying common rental rates, W and R, on labour and capital, respectively.

Following Atkinson and Stiglitz (1980, pp. 165-167) we develop the production side of the model by making use of one of the modern tools of duality theory, namely the *cost function* (see Intermezzo 5.1 in the previous chapter). Recall that the cost function represents the minimum level of factor costs needed to produce a given level of output when faced with given rental rates. Since technology features constant returns to scale, and both factors can be adjusted freely, the cost functions in the two sectors are linear in the respective outputs:

$$C^x \equiv c^x (W, R) X, \tag{A5.8}$$

$$C^y \equiv c^y (W, R) Y, \tag{A5.9}$$

where C^i and c^i denote, respectively, total cost and unit cost in sector i. Since both factors are perfectly mobile across sectors the same rental rates for capital and labour feature in (A5.8) and (A5.9).

Using Shephard's Lemma, the conditional factor demands are obtained from (A5.8)-(A5.9):

$$L_X = \frac{\partial c^x (W, R)}{\partial W} X \equiv c_W^x X, \tag{A5.10}$$

$$K_X = \frac{\partial c^x (W, R)}{\partial R} X \equiv c_R^x X, \tag{A5.11}$$

$$L_Y = \frac{\partial c^y (W, R)}{\partial W} Y \equiv c_W^y Y, \tag{A5.12}$$

$$K_Y = \frac{\partial c^y (W, R)}{\partial R} Y \equiv c_R^y Y, \tag{A5.13}$$

where c_W^x, c_R^x, c_W^y, and c_R^y are *unit input coefficients* depending in general on W and R (see below).

Full employment in the two factor markets implies:

$$c_W^x X + c_W^y Y = \bar{L}, \tag{A5.14}$$

$$c_R^x X + c_R^y Y = \bar{K}, \tag{A5.15}$$

where $\bar{L}$ is the total supply of labour and $\bar{K}$ is the total supply of capital. In equation (A5.14), $c_W^x X$ and $c_W^y Y$ represent labour absorbed in, respectively, the X-sector and the Y-sector. Hence, the left-hand side of (A5.14) is the total demand for labour and similarly, the left-hand side of (A5.15) represents the total demand for capital.

Perfectly competitive firms equate price to marginal cost in each sector:

$$P_X = c^x (W, R), \tag{A5.16}$$

$$P_Y = c^y (W, R), \tag{A5.17}$$

where c^i is unit-cost in sector i (see above).

The demand side of the model is as follows. There is a single representative agent, whose utility function depends on the consumption of both goods:

$$U = U (X, Y). \tag{A5.18}$$

The utility function is strictly quasi-concave, i.e. it features positive but diminishing marginal utility for both goods, and possesses indifference curves which bulge toward the origin, $U_X > 0$, $U_Y > 0$, $U_{XX} < 0$, $U_{YY} < 0$, and $U_{XX}U_{YY} - U_{XY}^2 > 0$. The household's budget restriction is given by:

$$P_X X + P_Y Y = M, \tag{A5.19}$$

where M is total income. The household chooses its consumption of X and Y in order to maximize (A5.18) subject to (A5.19). The first-order conditions are (A5.19) and:

$$\frac{U_X (X, Y)}{U_Y (X, Y)} = \frac{P_X}{P_Y}. \tag{A5.20}$$

According to (A5.20), the household equates the marginal rate of substitution between the two goods (left-hand side) to their relative market price (right-hand side). To keep things as simple as possible, we assume that the household's preferences are *homothetic* so that the Engel curves are linear in income and the Marshallian demands can be written as follows:

$$X = d^x (P_X, P_Y) M, \tag{A5.21}$$

$$Y = d^y (P_X, P_Y) M, \tag{A5.22}$$

where $d^x (\cdot)$ and $d^y (\cdot)$ are homogeneous of degree minus one in P_X and P_Y.[5] Obviously, it follows from (A5.19) and (A5.21)-(A5.22) that $P_X d^x (P_X, P_Y) + P_Y d^y (P_X, P_Y) = 1$. Finally, since the household is ultimately the owner of both factors of production, aggregate income is equal to:

$$M = W\bar{L} + R\bar{K}. \tag{A5.23}$$

[5] See Section 1.2 in Chapter 2 for an extended discussion of homothetic preferences. Recall that Marshallian demand functions are homogeneous of degree zero in P_X, P_Y, and M. It follows that for homothetic preferences d^x and d^y are homogeneous of degree minus one in P_X and P_Y.

In summary, the basic general equilibrium model can be written in compact form as follows.

$$\bar{L} = c_W^x(W,R)X + c_W^y(W,R)Y, \tag{A5.24}$$

$$\bar{K} = c_R^x(W,R)X + c_R^y(W,R)Y, \tag{A5.25}$$

$$P_X = c^x(W,R), \tag{A5.26}$$

$$P_Y = c^y(W,R), \tag{A5.27}$$

$$X = d^x(P_X,P_Y)\left[W\bar{L} + R\bar{K}\right], \tag{A5.28}$$

$$Y = d^y(P_X,P_Y)\left[W\bar{L} + R\bar{K}\right]. \tag{A5.29}$$

Equations (A5.24)-(A5.25) are obtained by combining (A5.10)-(A5.15), (A5.26)-(A5.27) are the same as (A5.16)-(A5.17), and (A5.28)-(A5.29) are obtained by substituting (A5.23) into (A5.21)-(A5.22). The endogenous variables of the model are the outputs (X and Y), the goods prices (P_X and P_Y), and the factor prices (W and R). The exogenous variables are the factor supplies ($\bar{K}$ and $\bar{L}$). Since the number of endogenous variables equals the number of equations of the model, one would be tempted to conclude that all variables are determinate. This is not the case, however, because not all six equations are independent. Indeed, according to the *Law of Walras*, if all but one markets are in equilibrium then so is the last market, i.e. one equation out of (A5.24)-(A5.29) is redundant and we only have five independent equations to determine six endogenous variables. As a result the model can only determine *relative* prices.

The validity of Walras' Law can be demonstrated as follows in the context of the basic model. Clearly, by definition, equations (A5.28) and (A5.29) imply that spending on goods equals total factor payments:

$$P_X X + P_Y Y = W\bar{L} + R\bar{K}. \tag{A5.30}$$

By substituting the factor market clearing conditions (A5.24) and (A5.25) into (A5.30) we obtain:

$$\begin{aligned} P_X X + P_Y Y &= W\left[c_W^x X + c_W^y Y\right] + R\left[c_R^x X + c_R^y Y\right] \\ &= \left[Wc_W^x + Rc_R^x\right]X + \left[Wc_W^y + Rc_R^y\right]Y. \end{aligned} \tag{A5.31}$$

By the linear homogeneity property of the unit-cost functions we have $c^x = Wc_W^x + Rc_R^x$ and $c^y = Wc_W^y + Rc_R^y$ so that (A5.31) can be rewritten as:

$$P_X X + P_Y Y = c^x X + c^y Y \quad \Leftrightarrow \quad (P_X - c^x)X = -(P_Y - c^y)Y. \tag{A5.32}$$

We have now deduced a dependency between two equations featuring in the model description. Indeed, if equation (A5.26) holds then so does (A5.27) and vice versa. One of the equations can be dropped.

6.3.2 Loglinearizing the model

The basic model is given in (A5.24)-(A5.29) above. In order to discover some of its key properties we adopt the usual strategy of loglinearizing it around an initial equilibrium. Since these loglinearizations are far from trivial, we show some of the details of the derivation.[6]

6.3.2.1 Loglinearized demand equations

The goods demand equations can be loglinearized as follows.[7] By using (A5.28)-(A5.29) we find that $X/Y = d^x(P_X, P_Y)/d^y(P_X, P_Y)$. Since d^x and d^y are both homogeneous of degree minus one in P_X and P_Y, it follows that d^x/d^y will depend only on the *relative* price of good X, i.e. on P_X/P_Y. Hence, it is clear that X/Y will also depend on that relative price only. In order to deduce the properties of d^x/d^y, we recall the key first-order condition for utility maximization (A5.20) and note that for homothetic preferences the elasticity of substitution, σ_D, is defined as follows:

$$\sigma_D \equiv \frac{d\ln(Y/X)}{d\ln(U_X/U_Y)} > 0. \tag{A5.33}$$

It follows from (A5.20) and (A5.33) that:

$$d\ln\left(\frac{U_X}{U_Y}\right) = d\ln\left(\frac{P_X}{P_Y}\right) = \frac{1}{\sigma_D} d\ln\left(\frac{Y}{X}\right), \tag{A5.34}$$

or (using the final two expressions):

$$\tilde{X} - \tilde{Y} = -\sigma_D\left[\tilde{P}_X - \tilde{P}_Y\right], \tag{A5.35}$$

where $\tilde{X} \equiv dX/X$, $\tilde{Y} \equiv dY/Y$, $\tilde{P}_X \equiv dP_X/P_X$, and $\tilde{P}_Y \equiv dP_Y/P_Y$. We have thus established that $\tilde{d}^x - \tilde{d}^y = -\sigma_D[\tilde{P}_X - \tilde{P}_Y]$, i.e. the relative demand for good X depends negatively on the relative price of good X. This negative effect is stronger, the larger is the elasticity of substitution, σ_D.

6.3.2.2 Loglinearized price equations

The pricing equations (A5.26)-(A5.27) can be loglinearized as follows. By totally differentiating equation (A5.26) we obtain:

$$\begin{aligned} dP_X &= c^x_W dW + c^x_R dR \qquad \Rightarrow \\ \frac{dP_X}{P_X} &= \frac{Wc^x_W}{c^x}\frac{dW}{W} + \frac{Rc^x_R}{c^x}\frac{dR}{R} \qquad \Rightarrow \\ \tilde{P}_X &= \theta_{LX}\tilde{W} + \theta_{KX}\tilde{R}, \end{aligned} \tag{A5.36}$$

[6] Readers familiar with the loglinearization techniques may skim or skip this subsection and proceed to the next.
[7] The more general non-homothetic case is discussed by Atkinson and Stiglitz (1980, p. 168).

where $\tilde{W} \equiv dW/W$, $\tilde{R} \equiv dR/R$, and where θ_{LX} and θ_{KX} represent the factor shares of, respectively, labour and capital in the X-sector:

$$\theta_{LX} \equiv \frac{Wc_W^x}{c^x}, \tag{A5.37}$$

$$\theta_{KX} \equiv \frac{Rc_R^x}{c^x}. \tag{A5.38}$$

Since the unit-cost function, $c^x(W,R)$, is linear homogeneous in W and R, it follows that the factor shares add to unity:

$$\begin{aligned} c^x &= c_W^x W + c_R^x R \qquad \Leftrightarrow \\ 1 &= \frac{c_W^x W}{c^x} + \frac{c_R^x R}{c^x} \qquad \Leftrightarrow \\ 1 &= \theta_{LX} + \theta_{KX}. \end{aligned} \tag{A5.39}$$

Similarly, by totally differentiating equation (A5.27) we obtain:

$$\tilde{P}_Y = \theta_{LY}\tilde{W} + \theta_{KY}\tilde{R}, \tag{A5.40}$$

$$\theta_{LY} \equiv \frac{Wc_W^y}{c^y}, \tag{A5.41}$$

$$\theta_{KY} \equiv \frac{Rc_R^y}{c^y} = 1 - \theta_{LY}, \tag{A5.42}$$

where θ_{LY} and θ_{KY} represent the factor shares of, respectively, labour and capital in the Y-sector.

Finally, by deducting (A5.40) from (A5.36) we obtain an expression linking the (change in the) relative price of good X to the (change in the) relative rental rate on labour:

$$\begin{aligned} \tilde{P}_X - \tilde{P}_Y &= \theta_{LX}\tilde{W} + \theta_{KX}\tilde{R} - \left[\theta_{LY}\tilde{W} + \theta_{KY}\tilde{R}\right] \\ &= \theta_{LX}\tilde{W} + (1-\theta_{LX})\,\tilde{R} - \left[\theta_{LY}\tilde{W} + (1-\theta_{LY})\,\tilde{R}\right] \\ &= \theta^*\left[\tilde{W} - \tilde{R}\right], \end{aligned} \tag{A5.43}$$

where θ^* is defined as follows:

$$\theta^* \equiv \theta_{LX} - \theta_{LY}, \tag{A5.44}$$

$$= \theta_{KY} - \theta_{KX}. \tag{A5.45}$$

Note that θ^* measures the relative factor intensity in the two sectors by means of the *factor income shares* in the two sectors. (Obviously, (A5.45) follows directly from (A5.44) by noting (A5.39) and (A5.42).) According to (A5.43), if the X-sector is relatively labour intensive ($\theta_{LX} > \theta_{LY}$ so that $\theta^* > 0$) then an increase in the relative price of labour (W/R) results in an increase in the relative price of good X

(P_X/P_Y).

6.3.2.3 Loglinearized factor market clearing equations

The factor market clearing conditions (A5.24)-(A5.25) can be loglinearized as follows. First, we totally differentiate the labour market clearing condition (A5.24), taking into account that c_W^x and c_W^y (may) depend on W/R and allowing for exogenous changes in labour supply:

$$\begin{aligned} d\bar{L} &= c_W^x dX + X dc_W^x + c_W^y dY + Y dc_W^y \qquad \Rightarrow \\ \frac{d\bar{L}}{\bar{L}} &= \frac{Xc_W^x}{\bar{L}}\frac{dX}{X} + \frac{c_W^x X}{\bar{L}}\frac{dc_W^x}{c_W^x} + \frac{Yc_W^y}{\bar{L}}\frac{dY}{Y} + \frac{c_W^y Y}{\bar{L}}\frac{dc_W^y}{c_W^y} \qquad \Rightarrow \\ \tilde{\bar{L}} &= \lambda_{LX}\left(\tilde{X} + \tilde{c}_W^x\right) + \lambda_{LY}\left(\tilde{Y} + \tilde{c}_W^y\right), \end{aligned} \tag{A5.46}$$

where $\tilde{c}_W^x \equiv dc_W^x/c_W^x$, $\tilde{c}_W^y \equiv dc_W^y/c_W^y$, $\tilde{\bar{L}} \equiv d\bar{L}/\bar{L}$, and where λ_{LX} and λ_{LY} are the shares of the labour force employed in, respectively, the X and the Y sector.

$$\lambda_{LX} \equiv \frac{L_X}{\bar{L}} = \frac{Xc_W^x}{\bar{L}}, \qquad \lambda_{LY} \equiv \frac{L_Y}{\bar{L}} = \frac{Yc_W^y}{\bar{L}} = 1 - \lambda_{LX}. \tag{A5.47}$$

In a similar fashion, total differentiation of the capital market equilibrium condition (A5.25) yields:

$$\tilde{\bar{K}} = \lambda_{KX}\left(\tilde{X} + \tilde{c}_R^x\right) + \lambda_{KY}\left(\tilde{Y} + \tilde{c}_R^y\right), \tag{A5.48}$$

where $\tilde{c}_R^x \equiv dc_R^x/c_R^x$, $\tilde{c}_R^y \equiv dc_R^y/c_R^y$, $\tilde{\bar{K}} \equiv d\bar{K}/\bar{K}$, and where λ_{KX} and λ_{KY} are the shares of the capital stock used in, respectively, the X and the Y sector.

$$\lambda_{KX} \equiv \frac{K_X}{\bar{K}} = \frac{Xc_R^x}{\bar{K}}, \qquad \lambda_{KY} \equiv \frac{K_Y}{\bar{K}} = \frac{Yc_R^y}{\bar{K}} = 1 - \lambda_{KX}. \tag{A5.49}$$

Before going on we note that for the special case of *Leontief technologies* (zero substitutability in production) all c_j^i coefficients are fixed. The loglinearized model is then given by (A5.35), (A5.43), (A5.46), and (A5.48) with $\tilde{c}_W^x = \tilde{c}_W^y = \tilde{c}_R^x = \tilde{c}_R^y = 0$ imposed in the latter two equations. We occasionally look at the Leontief case to build intuition. In the general case, however, a change in the relative price of labour will induce producers to change the input coefficients.

6.3.2.4 Loglinearized production coefficients

The production coefficients can be loglinearized as follows. First we totally differentiate $c_W^x(W, R)$:

$$\begin{aligned} dc_W^x &= c_{WW}^x dW + c_{WR}^x dR \qquad \Rightarrow \\ \frac{dc_W^x}{c_W^x} &= \frac{Wc_{WW}^x}{c_W^x}\frac{dW}{W} + \frac{Rc_{WR}^x}{c_W^x}\frac{dR}{R}. \end{aligned} \tag{A5.50}$$

Next we recall that the production coefficients are homogeneous of degree zero in W and R so that, by *Euler's Theorem*,[8] we have:

$$0 \times c_W^x = W c_{WW}^x + R c_{WR}^x. \tag{A5.51}$$

By using this result in (A5.50) we find:

$$\tilde{c}_W^x = -\frac{R c_{WR}^x}{c_W^x} \left[\tilde{W} - \tilde{R}\right]. \tag{A5.52}$$

The Allen-Uzawa substitution elasticity between capital and labour in the technology of the X-sector is defined, via the cost function, as follows:[9]

$$\sigma_X \equiv \frac{C^x C_{WR}^x}{C_R^x C_W^x} = \frac{c^x c_{WR}^x}{c_R^x c_W^x} \geq 0, \tag{A5.53}$$

where the second equality follows from the fact that the cost function is linear in output (see (A5.8) above). By using (A5.53) in (A5.52) and gathering terms we get a rather convenient expression after some straightforward steps:

$$\begin{aligned}
\tilde{c}_W^x &= -\frac{c_R^x}{c^x}\frac{c^x}{c_R^x}\frac{R c_{WR}^x}{c_W^x}\left[\tilde{W} - \tilde{R}\right] \\
&= -\frac{c_R^x R}{c^x}\frac{c^x c_{WR}^x}{c_R^x c_W^x}\left[\tilde{W} - \tilde{R}\right] \\
&= -\sigma_X \theta_{KX}\left[\tilde{W} - \tilde{R}\right],
\end{aligned} \tag{A5.54}$$

where θ_{KX} is the factor share of capital in the X-sector (see (A5.38) above). According to (A5.54), provided there are technological substitution possibilities ($\sigma_X > 0$), an increase in the relative rental price of labour leads to a decrease in the unit-input coefficient for labour in the X-sector. Producers substitute capital for labour along a given isoquant—the factor *substitution effect* studied in Chapter 5. Not surprisingly, loglinearization of $c_R^x(W, R)$ confirms this result for the unit-input coefficient for capital:

$$\tilde{c}_R^x = \sigma_X \theta_{LX}\left[\tilde{W} - \tilde{R}\right]. \tag{A5.55}$$

By using the same approach we find that the unit-input coefficients for the Y-sector can be loglinearized as follows:

$$\tilde{c}_W^y = -\sigma_Y \theta_{KY}\left[\tilde{W} - \tilde{R}\right], \tag{A5.56}$$

[8] Euler's Theorem (for two variables): if $F(X_1, X_2)$ is homogeneous of degree k, then it can be written as:

$$kF(X_1, X_2) = F_1 X_1 + F_2 X_2,$$

where $F_i \equiv \partial F / \partial X_i$. See Sydsæter, Strøm, and Berck (2000, p. 28).

[9] See, for example, Sydsæter, Strøm, and Berck (2000, p. 155).

$$\tilde{c}_R^y = \sigma_Y \theta_{LY} \left[\tilde{W} - \tilde{R}\right], \tag{A5.57}$$

where σ_Y is the substitution elasticity in the Y-sector:

$$\sigma_Y \equiv \frac{c^y c_{WR}^y}{c_R^y c_W^y} \geq 0. \tag{A5.58}$$

6.3.2.5 Factor market clearing equations again

By substituting (A5.54)-(A5.57) into the relevant places in (A5.46) and (A5.48) we obtain the most general expressions for the factor market equilibrium loci:

$$\tilde{\bar{L}} = \lambda_{LX}\tilde{X} + \lambda_{LY}\tilde{Y} - [\lambda_{LX}\theta_{KX}\sigma_X + \lambda_{LY}\theta_{KY}\sigma_Y]\left[\tilde{W} - \tilde{R}\right], \tag{A5.59}$$

$$\tilde{\bar{K}} = \lambda_{KX}\tilde{X} + \lambda_{KY}\tilde{Y} + [\lambda_{KX}\theta_{LX}\sigma_X + \lambda_{KY}\theta_{LY}\sigma_Y]\left[\tilde{W} - \tilde{R}\right]. \tag{A5.60}$$

In the final step, we deduct (A5.60) from (A5.59) to get:

$$\begin{aligned}\tilde{\bar{L}} - \tilde{\bar{K}} = &- [\lambda_{LX}\theta_{KX}\sigma_X + \lambda_{LY}\theta_{KY}\sigma_Y + \lambda_{KX}\theta_{LX}\sigma_X + \lambda_{KY}\theta_{LY}\sigma_Y]\left[\tilde{W} - \tilde{R}\right] \\ &+ (\lambda_{LX} - \lambda_{KX})\tilde{X} + (\lambda_{LY} - \lambda_{KY})\tilde{Y},\end{aligned} \tag{A5.61}$$

or (after simplification, using $\lambda_{LY} = 1 - \lambda_{LX}$, $\lambda_{KY} = 1 - \lambda_{KX}$, and gathering terms):

$$\lambda^*\left[\tilde{X} - \tilde{Y}\right] = \left[\tilde{\bar{L}} - \tilde{\bar{K}}\right] + [a_X\sigma_X + a_Y\sigma_Y]\left[\tilde{W} - \tilde{R}\right], \tag{A5.62}$$

where λ^*, a_X and a_Y are defined as follows:

$$\lambda^* \equiv \lambda_{LX} - \lambda_{KX}, \tag{A5.63}$$

$$a_X \equiv \lambda_{LX}\theta_{KX} + \lambda_{KX}\theta_{LX} > 0, \tag{A5.64}$$

$$a_Y \equiv \lambda_{LY}\theta_{KY} + \lambda_{KY}\theta_{LY} > 0. \tag{A5.65}$$

Like θ^*, defined in (A5.44)-(A5.45) above, λ^* is a measure for the relative factor intensity. The difference between the two measures lies in the fact that λ^* is in terms of physical units whereas θ^* is in terms of factor shares. According to (A5.62), if X is relatively labour intensive in the physical sense ($\lambda^* > 0$), then an increase in X/Y is associated with a rise in W/R.

We now have two measures for the relative factor intensity which may or may not give the same ranking. In the absence of distortions and/or taxes, however, λ^* and θ^* always have the same sign:

$$\lambda^* \overset{>}{\underset{<}{=}} 0 \quad \Leftrightarrow \quad \theta^* \overset{>}{\underset{<}{=}} 0. \tag{A5.66}$$

The proof of this important result proceeds as follows (see also Jones (1965, p. 559) or Atkinson and Stiglitz (1980, pp. 169-170)). Using $\bar{L} = L_X + L_Y$ and $\bar{K} = K_X + K_Y$ in (A5.63), λ^* can be rewritten as:

$$\begin{aligned}\lambda^* &\equiv \frac{L_X}{\bar{L}} - \frac{K_X}{\bar{K}} \\ &= \frac{L_X\left(K_X + K_Y\right) - \left(L_X + L_Y\right) K_X}{\bar{L}\bar{K}} \\ &= \frac{L_X K_Y - L_Y K_X}{\bar{L}\bar{K}}.\end{aligned} \tag{A5.67}$$

Similarly, by using (A5.8)-(A5.10), (A5.12), and noting that $C^i = WL_i + RK_i$ (for $i \in (X, Y)$), θ^* can be rewritten as:

$$\begin{aligned}\theta^* &\equiv \frac{W c_W^x}{c^x} - \frac{W c_W^y}{c^y} \\ &= \frac{W L_X}{c^x X} - \frac{W L_Y}{c^y Y} \\ &= W\left[\frac{L_X\left(WL_Y + RK_Y\right) - L_Y\left(WL_X + RK_X\right)}{C^x C^y}\right] \\ &= \frac{WR\left[L_X K_Y - L_Y K_X\right]}{C^x C^y} = \frac{W\bar{L}R\bar{K}}{C^x C^y}\lambda^*,\end{aligned} \tag{A5.68}$$

where we have used (A5.67) to arrive at the final expression. It follows from (A5.68) that λ^* and θ^* have the same sign.

6.3.3 Qualitative analysis of the model

We now possess all the ingredients needed to characterize the key properties of the simple general equilibrium model. The key equations of the model are restated here for convenience:

$$\tilde{X} - \tilde{Y} = -\sigma_D\left[\tilde{P}_X - \tilde{P}_Y\right], \tag{A5.69}$$

$$\tilde{P}_X - \tilde{P}_Y = \theta^*\left[\tilde{W} - \tilde{R}\right], \tag{A5.70}$$

$$\lambda^*\left[\tilde{X} - \tilde{Y}\right] = \left[\tilde{\bar{L}} - \tilde{\bar{K}}\right] + \left[a_X\sigma_X + a_Y\sigma_Y\right]\left[\tilde{W} - \tilde{R}\right]. \tag{A5.71}$$

Equation (A5.69) is the relative demand equation relating X/Y to P_X/P_Y, (A5.70) is the competitive (relative) pricing relationship establishing the link between relative prices of goods and factors (i.e. of P_X/P_Y to W/R), and (A5.71) represents the factor market equilibrium conditions relating W/R to X/Y

and $\bar{L}/\bar{K}$. The endogenous variables of the model are X/Y, P_X/P_Y, and W/R. The exogenous variable is $\bar{L}/\bar{K}$.

In Figure 6.2 we illustrate the determination of the general equilibrium under the assumption that the X-sector is relatively labour intensive (so that $\lambda^* > 0$ and $\theta^* > 0$). In the top right-hand panel, the D curve is the relative demand equation (A5.69). It slopes down because $\sigma_D > 0$. In the top left-hand panel, the FME curve represents the factor markets equilibrium locus (A5.71), holding constant the economy-wide labour-capital ratio, $\bar{L}/\bar{K}$. In view of the assumption that the X-sector is labour intensive ($\lambda^* > 0$), FME is an upward sloping line.[10] In the bottom left-hand panel, the 45° line projects the W/R ratio onto the vertical axis. Finally, in the bottom right-hand panel the CPR curve represents the competitive pricing relationship (A5.70). The intensity assumption (that $\theta^* > 0$) implies that this curve is upward sloping.

Taken in combination, CPR and FME characterize the *supply side* of the model. The supply curve, S, in the top right-hand panel is constructed graphically by "completing the boxes" for different relative price levels (see, for example, the boxes abcd and efgh). The thus constructed supply curve is upward sloping. Of course, the mathematical expression for the supply curve can be obtained directly by combining (A5.70) and (A5.71) and eliminating $\tilde{W} - \tilde{R}$:

$$\tilde{X} - \tilde{Y} = \frac{1}{\lambda^*}\left[\tilde{\bar{L}} - \tilde{\bar{K}}\right] + \frac{a_X\sigma_X + a_Y\sigma_Y}{\lambda^*\theta^*}\left[\tilde{P}_X - \tilde{P}_Y\right]. \tag{A5.72}$$

Regardless of the factor intensity assumption, the S curve is upward sloping (because (A5.66) implies that $\lambda^*\theta^* > 0$).

The general equilibrium occurs at the intersection of the S curve with the D curve, i.e. at point E_0 in Figure 6.2. Using (A5.69)-(A5.71) we find the mathematical representation of the general equilibrium solutions for X/Y, W/R, and P_X/P_Y:

$$\tilde{P}_X - \tilde{P}_Y = \frac{\theta^*[\tilde{\bar{K}} - \tilde{\bar{L}}]}{\lambda^*\theta^*\sigma_D + a_X\sigma_X + a_Y\sigma_Y}, \tag{A5.73}$$

$$\tilde{X} - \tilde{Y} = -\frac{\sigma_D\theta^*[\tilde{\bar{K}} - \tilde{\bar{L}}]}{\lambda^*\theta^*\sigma_D + a_X\sigma_X + a_Y\sigma_Y}, \tag{A5.74}$$

$$\tilde{W} - \tilde{R} = \frac{\tilde{\bar{K}} - \tilde{\bar{L}}}{\lambda^*\theta^*\sigma_D + a_X\sigma_X + a_Y\sigma_Y}. \tag{A5.75}$$

Provided there is some substitutability in the economy, and at least one of σ_D, σ_X, and σ_Y is strictly positive, the denominator of these expressions is unambiguously positive because $\lambda^*\theta^* > 0$ and $a_i > 0$.

Consider the effects of an increase in the relative abundance of labour (an increase in the $\bar{L}/\bar{K}$ ratio). In Figure 6.3, the only curve that is affected is the FME curve (which shifts up, from FME_0 to FME_1). As a result of this shift, the supply curve also shifts up (from S_0 to S_1) and the equilibrium shifts from E_0 to

[10]Note that all axes in Figure 6.2 are positively defined. Though FME looks (at first view) like a downward sloping function it is actually an upward sloping relationship between X/Y and W/R.

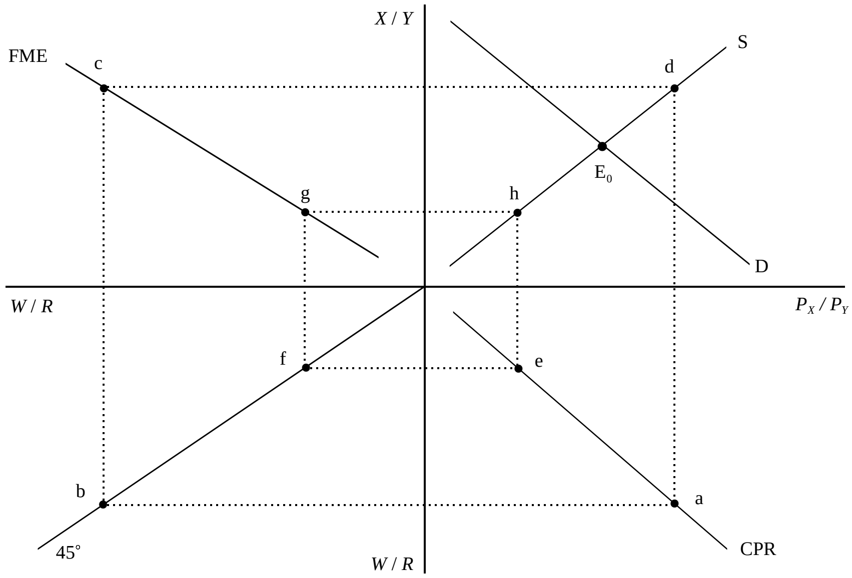

Figure 6.2: Two-sector general equilibrium (X labour-intensive)

E_1. In the new equilibrium, P_X/P_Y and W/R are lower whilst X/Y is higher than before the shock. In mathematical terms these comparative static results are verified by (A5.73)-(A5.75).

In order to demonstrate the role of factor substitutability on the production side, consider the case with Leontief technology in both sectors (i.e. $\sigma_X = \sigma_Y = 0$). This case has been illustrated in Figure 6.4. Since $\sigma_X = \sigma_Y = 0$ it follows from (A5.71) that the FME curve is independent of the W/R ratio, i.e. it is horizontal (see FME_0 in Figure 6.4). As a result, the supply curve is independent of P_X/P_Y, i.e. it is also horizontal–see S_0, or alternatively equation (A5.72). The initial equilibrium is at point E_0. Equality between demand, D, and supply, S_0, determines the equilibrium relative goods price (P_X/P_Y) which, via the pricing relationship CPR, determines the equilibrium relative price of labour (W/R). An increase in $\bar{L}/\bar{K}$ leads to an upward shift of the FME curve (from FME_0 to FME_1), an upward shift in supply (from S_0 to S_1), a decrease in P_X/P_Y, and a decrease in W/R.

6.4 Tax incidence in general equilibrium

In this section we perform tax policy analysis using the general equilibrium model developed above. Following Atkinson and Stiglitz (1980, pp. 173-183), we consider a range of *ad valorem* taxes, namely (i) taxes on factor prices in both sectors, denoted by t_{KX}, t_{KY}, t_{LX}, and t_{LY}, and (ii) excise taxes on the consumption of the two goods, denoted by t_X and t_Y. All tax revenues are recycled to the representative household in the form of lump-sum transfers. To keep matters simple, we continue to restrict attention to the homothetic case.

Rather than redoing all derivations, we simply note where the changes in the model occur. First, in view of the fact that both factors of production are taxed, firms react to tax-inclusive factor prices, i.e.

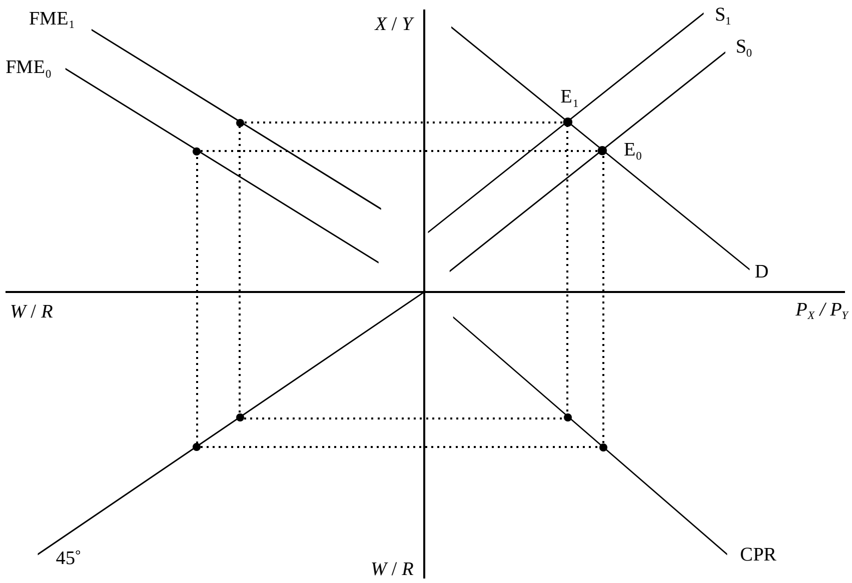

Figure 6.3: Increase in labour endowment (*X* labour-intensive)

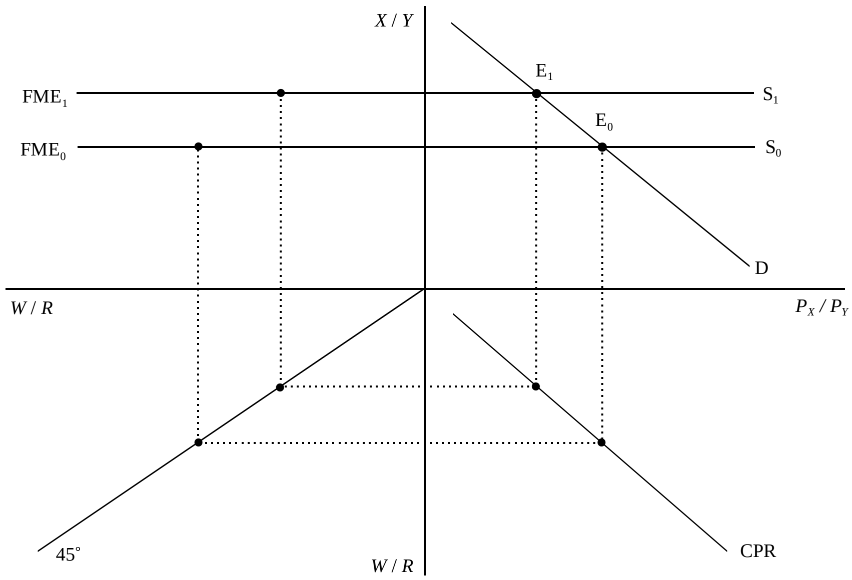

Figure 6.4: Leontief technology (*X* labour-intensive)

to $W(1+t_{Li})$ for labour and to $R(1+t_{Ki})$ for capital, where $i = X, Y$. The factor market equilibrium conditions (A5.14)-(A5.15) are thus changed to:

$$\bar{L} = c_W^x \left(W(1+t_{LX}), R(1+t_{KX})\right) X + c_W^y \left(W(1+t_{LY}), R(1+t_{KY})\right) Y, \tag{A5.76}$$

$$\bar{K} = c_R^x \left(W(1+t_{LX}), R(1+t_{KX})\right) X + c_R^y \left(W(1+t_{LY}), R(1+t_{KY})\right) Y. \tag{A5.77}$$

and the (producer) price equations (A5.16)-(A5.17) are changed to:

$$P_X = c^x \left(W(1+t_{LX}), R(1+t_{KX})\right), \tag{A5.78}$$

$$P_Y = c^y \left(W(1+t_{LY}), R(1+t_{KY})\right). \tag{A5.79}$$

In the presence of excise taxes, households react to consumer prices, $P_i(1+t_i)$, so that the demand equations (A5.21)-(A5.22) are modified to:

$$X = d^x \left(P_X(1+t_X), P_Y(1+t_Y)\right) M, \tag{A5.80}$$

$$Y = d^y \left(P_X(1+t_X), P_Y(1+t_Y)\right) M, \tag{A5.81}$$

where M is after-tax household income:

$$M \equiv W(1-t_L)\bar{L} + R(1-t_K)\bar{K} + T, \tag{A5.82}$$

and T is the total tax revenue:

$$T = W\left[t_L\bar{L} + t_{LX}L_X + t_{LY}L_Y\right] + R\left[t_K\bar{K} + t_{KX}K_X + t_{KY}K_Y\right] + t_X P_X X + t_Y P_Y Y. \tag{A5.83}$$

In equation (A5.82), t_L and t_K are the taxes levied on, respectively, labour income and capital income of the household. The model consists of equations (A5.76)-(A5.83), of which one is redundant. The endogenous variables are X, Y, P_X, P_Y, W, R, and T, whilst the exogenous variables are the factor endowments and the tax rates $\bar{K}$, $\bar{L}$, t_{LX}, t_{LY}, t_{KX}, t_{KY}, t_X, t_Y, t_L, and t_K. Even though tax revenue is determined endogenously within the general equilibrium model, in formulating its optimal consumption decision the household treats it parametrically (i.e., as a lump-sum transfer).

As was shown by Musgrave (1959, ch. 15) and McLure (1975, p. 137), several tax equivalency results can now be demonstrated within the context of the model. These equivalencies have been summarized in Table 6.1. First, an equal tax on both factors in the same industry (e.g. $t_{LX} = t_{KX}$) has no factor substitution effect and is equivalent to an excise tax on the good produced by that industry. This equivalency is reported in the first *row* of Table 6.1. Formally, since $t_{LX} = t_{KX}$ it follows from the linear homogeneity of the unit cost function (A5.78) that $P_X = c^x\left(W(1+t_{LX}), R(1+t_{KX})\right) = (1+t_{LX})\, c^x(W,R)$, i.e. the

$$\begin{array}{ccccc} t_{KX} & + & t_{LX} & = & t_X \\ + & & + & & + \\ t_{KY} & + & t_{LY} & = & t_Y \\ \shortparallel & & \shortparallel & & \shortparallel \\ t_K & + & t_L & = & t \end{array}$$

Table 6.1: Tax equivalencies

common factor tax raises the price of good X just as an excise tax t_X does. By analogy, the second row of Table 6.1 says that there is an equivalency between $t_{LY} = t_{KY}$ and t_Y.

Second, an equal tax on the production factor capital ($t_{KX} = t_{KY}$) is equivalent to a tax on capital income received by households (t_K). This equivalency result is reported in the first *column* of Table 6.1. Formally, since $t_{KX} = t_{KY}$, it follows from (A5.76)-(A5.79) that the tax-inclusive rental rate on capital faced by producers is $R' \equiv (1 + t_{KX}) R$. Similarly, since $t_K = 0$ it follows from (A5.82)-(A5.83) that the rental rate entering household income also equals R'. If instead capital is untaxed at firm level ($t_{KX} = t_{KY} = 0$) but capital income of households is taxed ($t_K > 0$), then the only rental rate appearing in the model is R. The systems of capital taxation therefore must yield exactly the same allocation. Again, by analogy, the second column in Table 6.1 shows that the systems $(t_L = 0, t_{LX} = t_{LY} > 0)$ and $(t_L > 0, t_{LX} = t_{LY} = 0)$ are equivalent.

Third, a tax on capital income and labour income at the same rate (i.e., a general income tax, $t = t_K = t_L$) has the same effect as a tax on both products at the same rate (i.e., a general consumption tax, $t = t_X = t_Y$). In both cases there is no effect on relative prices or relative factor returns. This equivalency result has been reported in the third row and the third column in Table 6.1.

An implication of these equivalency results is that we only need to study four (out of the nine) taxes set out in the model, provided they are independent, i.e. provided they do not all come from the same row or column (McLure, 1975, p. 138). For example, if we know the incidence of t_{KX}, t_K, t_X, and t, then we can deduce the incidence of the remaining taxes with the aid of Table 6.1. Of course, the incidence of the general income tax, t, is particularly easy to deduce for the present case with fixed factor supplies. Such a tax has no effect on relative product prices or relative factor prices.

Following exactly the same steps as before we find the loglinearized model (see subsection 6.3.2 above for details on loglinearization):

$$\tilde{X} - \tilde{Y} = -\sigma_D \left[\tilde{P}_X - \tilde{P}_Y + \tilde{t}_X - \tilde{t}_Y \right], \tag{A5.84}$$

$$\tilde{P}_X - \tilde{P}_Y = \theta^* \left[\tilde{W} - \tilde{R} \right] + \theta_{LX}\tilde{t}_{LX} - \theta_{LY}\tilde{t}_{LY} + \theta_{KX}\tilde{t}_{KX} - \theta_{KY}\tilde{t}_{KY}, \tag{A5.85}$$

$$\begin{aligned} \lambda^* \left[\tilde{X} - \tilde{Y} \right] &= \left[a_X\sigma_X + a_Y\sigma_Y \right] \left[\tilde{W} - \tilde{R} \right] + a_X\sigma_X \left[\tilde{t}_{LX} - \tilde{t}_{KX} \right] \\ &\quad + a_Y\sigma_Y \left[\tilde{t}_{LY} - \tilde{t}_{KY} \right], \end{aligned} \tag{A5.86}$$

where $\tilde{t}_i \equiv dt_i / (1 + t_i)$ and $\tilde{t}_{ji} \equiv dt_{ji} / \left(1 + t_{ji}\right)$ for $j \in (K, L)$ and $i \in (X, Y)$. Equation (A5.84) is the relative demand equation relating X/Y to $P_X (1 + t_X) / [P_Y (1 + t_Y)]$. Equation (A5.85) is the competitive (relative) pricing relationship relating P_X/P_Y to W/R and the various factor tax rates. Finally, equation (A5.86) represents the factor market equilibrium conditions relating W/R to X/Y and the different factor tax rates.

As is pointed out by Atkinson and Stiglitz (1980, pp. 180-181), a complication arises because in the presence of taxes (or other non-tax distortions), the intensity ranking according to physical factor intensity (λ^*) may not be the same as the intensity ranking according to factor shares (θ^*). The reason for this complication is that the θ_{ji}-shares appearing in the loglinearized model are tax-inclusive, i.e. instead of (A5.37)-(A5.38) and (A5.41)-(A5.42) we now have:

$$\theta_{LX} \equiv \frac{W(1+t_{LX})c_W^x}{c^x}, \qquad \theta_{KX} \equiv \frac{R(1+t_{KX})c_R^x}{c^x} = 1-\theta_{LX}, \tag{A5.87}$$

$$\theta_{LY} \equiv \frac{W(1+t_{LY})c_W^y}{c^y}, \qquad \theta_{KY} \equiv \frac{R(1+t_{KY})c_R^y}{c^y} = 1-\theta_{LY}. \tag{A5.88}$$

By using these expressions, $\theta^* \equiv \theta_{LX} - \theta_{LY}$ can be rewriting as follows:

$$\begin{aligned}\theta^* &\equiv \frac{W(1+t_{LX})c_W^x}{c^x} - \frac{W(1+t_{LY})c_W^y}{c^y} \\ &= \frac{W(1+t_{LX})L_X}{c^x X} - \frac{W(1+t_{LY})L_Y}{c^y Y} \\ &= \frac{WR\left[(1+t_{LX})(1+t_{KY})L_X K_Y - (1+t_{LY})(1+t_{KX})L_Y K_X\right]}{C^x C^y},\end{aligned} \tag{A5.89}$$

where we have used the fact that $C^i = W(1+t_{Li})L_i + R(1+t_{Ki})K_i$ (for $i = X, Y$) in going from the second to the third line. Comparing this expression to the one for λ^* (given in (A5.67) above) reveals why the ranking may differ. In the special case, with $(1+t_{LX})(1+t_{KY})$ equal to $(1+t_{LY})(1+t_{KX})$, the signs of λ^* and θ^* are guaranteed to be the same. In the general case they may not be the same. For the remainder of this chapter, however, we simply assume that the intensity rankings remain the same.

6.4.1 General equilibrium tax effects

In this subsection we conduct a number of comparative static experiments. We assume throughout that X is relatively labour-intensive according to both intensity measures, i.e. $\theta^* > 0$ and $\lambda^* > 0$.

The effect on a marginal change in the excise tax on good X is illustrated in Figure 6.5. In terms of the model (A5.84)-(A5.86), the shock is such that $\tilde{t}_X > 0$ whilst all other taxes are kept unchanged ($\tilde{t}_Y = \tilde{t}_{LX} = \tilde{t}_{LY} = \tilde{t}_{KX} = \tilde{t}_{KX} = 0$). The model reduces to:

$$\tilde{X} - \tilde{Y} = -\sigma_D\left[\tilde{P}_X - \tilde{P}_Y + \tilde{t}_X\right], \tag{A5.90}$$

$$\tilde{P}_X - \tilde{P}_Y = \theta^*\left[\tilde{W} - \tilde{R}\right], \tag{A5.91}$$

$$\lambda^*\left[\tilde{X} - \tilde{Y}\right] = \left[a_X\sigma_X + a_Y\sigma_Y\right]\left[\tilde{W} - \tilde{R}\right]. \tag{A5.92}$$

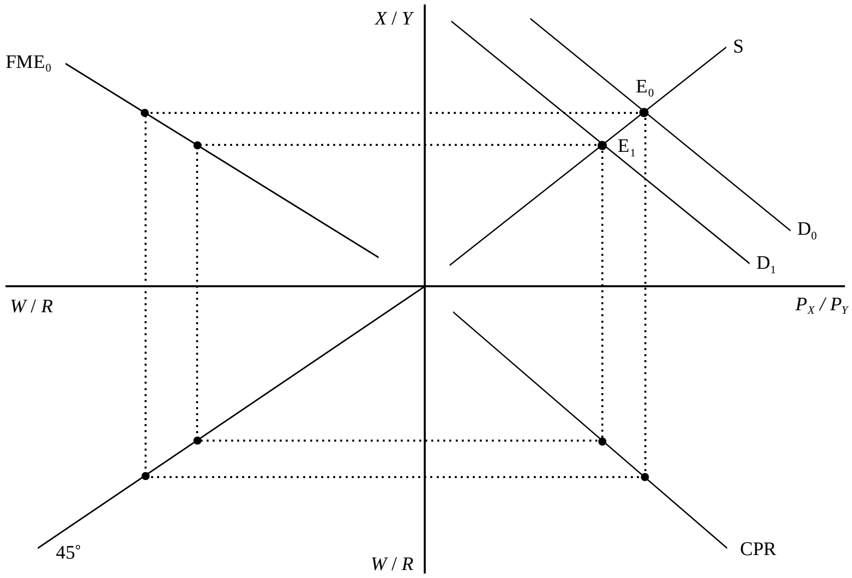

Figure 6.5: Increase in the output tax t_X (X labour-intensive)

The increase in t_X shifts the demand curve down, say from D_0 to D_1. As a result, the equilibrium shifts from E_0 to E_1, and P_X/P_Y, X/Y, and W/R all fall. In this case there is only an *output* (or *volume*) *effect*: as a result of the shock, the pattern of demand changes in favour of good Y. Since X/Y falls and since X is relatively labour intensive, the wage-rental ratio falls. This effect holds regardless of the magnitude of the substitution elasticity in the X-sector. The mathematically expressions for the comparative static effects can be obtained from (A5.90)-(A5.92):

$$\tilde{X}-\tilde{Y}=-\frac{(a_X\sigma_X+a_Y\sigma_Y)\,\sigma_D}{\lambda^*\theta^*\sigma_D+a_X\sigma_X+a_Y\sigma_Y}\tilde{t}_X<0, \tag{A5.93}$$

$$\tilde{P}_X-\tilde{P}_Y=-\frac{\lambda^*\theta^*\sigma_D}{\lambda^*\theta^*\sigma_D+a_X\sigma_X+a_Y\sigma_Y}\tilde{t}_X<0, \tag{A5.94}$$

$$\tilde{W}-\tilde{R}=-\frac{\lambda^*\sigma_D}{\lambda^*\theta^*\sigma_D+a_X\sigma_X+a_Y\sigma_Y}\tilde{t}_X<0. \tag{A5.95}$$

The second tax policy experiment consists of a marginal change in the capital tax in sector X is illustrated, i.e. $\tilde{t}_{KX}>0$ and all other taxes are kept constant ($\tilde{t}_X=\tilde{t}_Y=\tilde{t}_{LX}=\tilde{t}_{LY}=\tilde{t}_{KY}=0$). The model (A5.84)-(A5.86) simplifies to:

$$\tilde{X}-\tilde{Y}=-\sigma_D\left[\tilde{P}_X-\tilde{P}_Y\right], \tag{A5.96}$$

$$\tilde{P}_X-\tilde{P}_Y=\theta^*\left[\tilde{W}-\tilde{R}\right]+\theta_{KX}\tilde{t}_{KX}, \tag{A5.97}$$

$$\lambda^*\left[\tilde{X}-\tilde{Y}\right]=\left[a_X\sigma_X+a_Y\sigma_Y\right]\left[\tilde{W}-\tilde{R}\right]-a_X\sigma_X\tilde{t}_{KX}. \tag{A5.98}$$

In the general case (with $\sigma_X>0$), both the FME and CPR curves are affected, i.e. there are both *output effects* and *factor substitution effects*. To disentangle these effects we first consider the output effect in isolation. Figure 6.6 is based on the Leontief assumption that the substitution elasticity in the X-sector

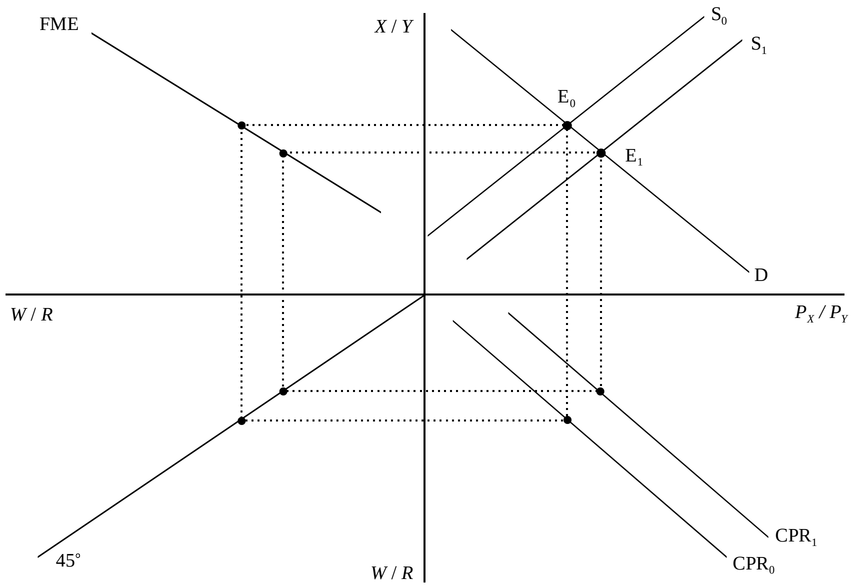

Figure 6.6: Increase in the corporate tax t_{KX} (X labour-intensive, $\sigma_X = 0$)

is zero ($\sigma_X = 0$). It follows from (A5.98) that the FME is not affected by the tax change in that case. The CPR curve shifts to the right, from CPR_0 to CPR_1, as does the supply curve, which shifts from S_0 to S_1. The equilibrium shifts from E_0 to E_1, P_X/P_Y rises, X/Y falls, and W/R falls. Despite the fact that capital is taxed, the rental rate on capital rises relative to wages. There is only an *output* (or *volume*) *effect*: P_X/P_Y falls and demand shifts toward good Y (X/Y falls). Since X is labour-intensive, labour demand drops off whilst capital demand is boosted. In equilibrium W/R has to fall.

In Figure 6.7 we consider the general case (i.e. $\sigma_X > 0$). As in the Leontief case, the CPR curve shifts to the right, say from CPR_0 to CPR_1. Because factor substitution is now possible, the FME curve shifts down, say from FME_0 to FME_1. Combining these results we find that the supply curve shifts to the right, from S_0 to S_1. In the top right-hand panel the equilibrium shifts from E_0 to E_1, P_X/P_Y rises and X/Y falls. The resulting effect on W/R is ambiguous, although the figure is drawn under the assumption that the rental rate on capital *falls* relative to wages. In order to determine the conditions under which this result obtains, we solve the model (A5.96)-(A5.98):

$$\tilde{X} - \tilde{Y} = -\frac{\sigma_D \left[a_X \sigma_X \theta_{KY} + a_Y \sigma_Y \theta_{KX}\right]}{\lambda^* \theta^* \sigma_D + a_X \sigma_X + a_Y \sigma_Y} \tilde{t}_{KX} < 0, \tag{A5.99}$$

$$\tilde{P}_X - \tilde{P}_Y = \frac{a_X \sigma_X \theta_{KY} + a_Y \sigma_Y \theta_{KX}}{\lambda^* \theta^* \sigma_D + a_X \sigma_X + a_Y \sigma_Y} \tilde{t}_{KX} > 0, \tag{A5.100}$$

$$\tilde{W} - \tilde{R} = \frac{a_X \sigma_X - \lambda^* \theta_{KX} \sigma_D}{\lambda^* \theta^* \sigma_D + a_X \sigma_X + a_Y \sigma_Y} \tilde{t}_{KX} \gtreqless 0. \tag{A5.101}$$

Equation (A5.101) confirms that the effect on the relative price of labour, W/R, is ambiguous. Since $\lambda^* > 0$, the sign of the effect is determined by the numerator which has two terms, namely (i) a positive factor substitution effect (represented by $a_X \sigma_X$) and (ii) a negative output effect (represented by $-\lambda^* \theta_{KX} \sigma_D$).

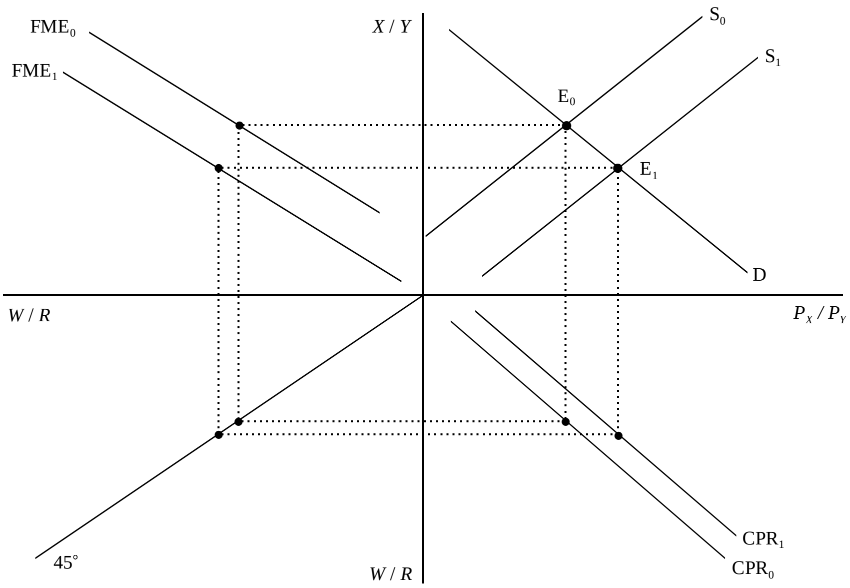

Figure 6.7: Increase in the corporate tax t_{KX} (X labour-intensive)

Figure 6.7 is thus drawn under the assumption that the output effect dominates.

6.4.2 Applied general equilibrium model

Up to this point attention has been restricted to the analysis of *differential tax incidence*, i.e. only relative output and price effects are considered. Of course, once a numeraire has been chosen, the general equilibrium model given in equations (A5.76)-(A5.83) above can also be solved in level terms. Two approaches are possible. First, for small tax changes it suffices to loglinearize the model in order to find the comparative static effects. Second, for large tax changes the loglinearization approach is generally invalid but comparative static effects can be computed numerically. In this subsection we demonstrate both approaches. To keep things simple, we restrict attention to the case of a capital tax in the X sector. This tax is assumed to be zero initially, but it is raised in the comparative static exercise. The output of the Y sector is chosen as the numeraire commodity so that $P_Y = 1$. The model considered in this subsection is given in general terms as:

$$\bar{L} = c_W^x \left(W, R\left(1+t_{KX}\right)\right) X + c_W^y \left(W, R\right) Y, \tag{A5.102}$$

$$\bar{K} = c_R^x \left(W, R\left(1+t_{KX}\right)\right) X + c_R^y \left(W, R\right) Y, \tag{A5.103}$$

$$P_X = c^x \left(W, R\left(1+t_{KX}\right)\right), \tag{A5.104}$$

$$1 = c^y \left(W, R\right), \tag{A5.105}$$

$$X = d^x \left(P_X, 1\right) M, \tag{A5.106}$$

$$M = W\bar{L} + R\bar{K} + T, \tag{A5.107}$$

$$T = t_{KX} R c_R^x \left(W, R\left(1+t_{KX}\right)\right) X. \tag{A5.108}$$

Note that the demand equation for good Y is dropped because, by the *Law of Walras*, it is implied by the remaining equations of the model (see also the discussion below equations (A5.24)-(A5.29) above).

6.4.2.1 Small changes

The comparative static effects of the change in t_{KX} on relative output, X/Y, the relative goods price P_X/P_Y, and the relative wage rate, W/R, are given in equations (A5.99)-(A5.101) above. In order to obtain expressions for the output effects in absolute terms, it suffices to loglinearize equations (A5.106)-(A5.108), taking into account that $t_{KX} = 0$ initially, and holding constant the exogenous factor supplied ($\tilde{\bar{L}} = \tilde{\bar{K}} = 0$). After some straightforward manipulations we find:

$$\tilde{X} = -\left[\omega_X + \sigma_D\left(1-\omega_X\right)\right]\tilde{P}_X + \tilde{M}, \tag{A5.109}$$

$$\tilde{M} = \omega_L \tilde{W} + \left(1-\omega_L\right)\tilde{R} + \tilde{T}, \tag{A5.110}$$

$$\tilde{T} = \omega_X \theta_{KX} \tilde{t}_{KX}, \tag{A5.111}$$

where $\omega_X \equiv P_X X/M$ is share of household income spent on good X, $\omega_L \equiv W\bar{L}/\left(W\bar{L}+R\bar{K}\right)$ is the income share of labour, and $1-\omega_L \equiv R\bar{K}/\left(W\bar{L}+R\bar{K}\right)$ is the income share of capital.[11] In equations (A5.110) and (A5.111), $\tilde{T}$ is defined as $\tilde{T} \equiv dT/M$.

The expression for the demand equation in equation (A5.109) can be simplified substantially. First, $\tilde{M}$ and $\tilde{T}$ can be eliminated by substituting (A5.110)-(A5.111) into (A5.109). Second, it follows from the loglinearized version of equation (A5.105) that $\tilde{W}$ and $\tilde{R}$ are related according to $\theta_{LY}\tilde{W} + \theta_{KY}\tilde{R} = 0$, or:

$$\tilde{R} = -\frac{\theta_{LY}}{\theta_{KY}}\tilde{W}, \tag{A5.112}$$

$$\tilde{W} - \tilde{R} = \frac{1}{\theta_{KY}}\tilde{W}, \tag{A5.113}$$

where we have used the fact that $\theta_{KY} + \theta_{LY} = 1$. Third, the share parameters ω_L and ω_X can be expressed in terms of the "fundamental parameters" appearing elsewhere in the general equilibrium model (i.e.

[11] Equation (A5.109) is obtained as follows. First, we loglinearize the household budget constraint, $P_X X + P_Y Y = M$, to obtain:

$$\omega_X[\tilde{X} + \tilde{P}_X] + (1-\omega_X)[\tilde{Y} + \tilde{P}_Y] = \tilde{M}.$$

Next we use this expression in combination with (A5.35) to solve for $\tilde{X}$ and $\tilde{Y}$ in terms of $\tilde{P}_X$, $\tilde{P}_Y$, and $\tilde{M}$.

$$\tilde{X} = -\left[\omega_X + \sigma_D\left(1-\omega_X\right)\right]\tilde{P}_X + \left(\sigma_D - 1\right)\left(1-\omega_X\right)\tilde{P}_Y + \tilde{M},$$
$$\tilde{Y} = \left(\sigma_D - 1\right)\omega_X \tilde{P}_X - \left[1-\omega_X + \sigma_D\omega_X\right]\tilde{P}_Y + \tilde{M}.$$

Finally, by setting $\tilde{P}_Y = 0$ in the first equation we obtain (A5.109).

the intensity coefficients):[12]

$$\omega_X = \frac{\theta_{LY}\lambda_{LX}}{\theta_{LX}\lambda_{LY} + \theta_{LY}\lambda_{LX}}, \tag{A5.114}$$

$$\omega_L = \frac{\theta_{LY}\theta_{LX}}{\theta_{LX}\lambda_{LY} + \theta_{LY}\lambda_{LX}}. \tag{A5.115}$$

By imposing all these simplifications, the loglinearized demand equation (A5.109) can be written as:

$$\begin{aligned}\tilde{X} = &-\frac{\sigma_D\theta_{LX}\lambda_{LY} + \theta_{LY}\lambda_{LX}}{\theta_{LX}\lambda_{LY} + \theta_{LY}\lambda_{LX}}\tilde{P}_X + \frac{\theta_{LY}\lambda_{LX}}{\theta_{LX}\lambda_{LY} + \theta_{LY}\lambda_{LX}}\theta^* \left[\tilde{W} - \tilde{R}\right] \\ &+ \frac{\theta_{LY}\theta_{KX}\lambda_{LX}}{\theta_{LX}\lambda_{LY} + \theta_{LY}\lambda_{LX}}\tilde{t}_{KX},\end{aligned} \tag{A5.116}$$

where the relative price effect is given in (A5.100) (with $\tilde{P}_Y = 0$ imposed), and the relative wage effect is stated in (A5.101). Equation (A5.116) identifies the three mechanisms explaining the general equilibrium effect on output in the X sector. The first term on the right-hand side is negative because the capital tax drives up the price. The second term on the right-hand side is ambiguous due to offsetting output and factor substitution effects (as was explained above). The third term on the right-hand side is positive because the lump-sum transfer received from the tax authority boosts demand.

6.4.2.2 Large changes

As we have shown above, it is in principle possible to compute the level effects of small changes in tax rates. Even for the simple two-by-two model under consideration, however, general analytical conclusions are typically hard to come by. Indeed, as is clear from our discussion surrounding equation (A5.116) above, the effects on outputs and the wage-rental rate of an isolated change in the capital tax in the X sector are ambiguous in general. Unless drastic simplifications are incorporated in the model, this ambiguity will remain.

In order to circumvent the fundamental limitations of small analytical models, researchers have redirected their attention to so-called *applied general equilibrium* (AGE) models over the last three decades. AGE models are typically based on a realistically detailed multi-sectoral description of the economy and the effects of complex policy measures are investigated with the aid of computational experiments. With the advent of cheap computing power, the simulation of large non-linear models is feasible even on relatively cheap personal computers. In the remainder of this subsection we will formulate a simple AGE model based on Shoven and Whalley (1984). This toy model is then used to simulate the effects of large tax changes.

In contrast to Shoven and Whalley, and in order to facilitate the comparison with the theoretical

[12] Here are two pieces of advice to students who want to work with small analytical models. First, in order to deduce the maximum amount of interpretable analytical results from such a model, it is imperative to state the model in terms of its most fundamental parameters only. Inclusion of non-fundamental (composite) parameters inevitably hide patterns that exist in the result. Second, write down the definitions of all parameters and their inter-relationships on a big piece of paper and display it prominently on your desk so that you can refer to it at all times.

model presented above, our example AGE model is formulated in its dual format. There are two productions sectors, X and Y, and the unit cost functions in the two sectors are given by:

$$c^x = \frac{1}{\phi_X}\left[\varepsilon_X^{\sigma_X} W^{1-\sigma_X} + (1-\varepsilon_X)^{\sigma_X}\left[R\left(1+t_{KX}\right)\right]^{1-\sigma_X}\right]^{1/(1-\sigma_X)}, \tag{A5.117}$$

$$c^y = \frac{1}{\phi_Y}\left[\varepsilon_Y^{\sigma_Y} W^{1-\sigma_Y} + (1-\varepsilon_Y)^{\sigma_Y} R^{1-\sigma_Y}\right]^{1/(1-\sigma_X)}, \tag{A5.118}$$

where ϕ_i is an exogenous scale parameter, ε_i is the efficiency parameter of labour in the production function, and σ_i is the substitution elasticity between capital and labour (for $i = X, Y$).[13] The input coefficients appearing in (A5.102)-(A5.103) and (A5.108) are thus given by:

$$(c_W^x \equiv)\ \frac{\partial c^x}{\partial W} = \frac{1}{\phi_X}\left[\varepsilon_X + (1-\varepsilon_X)^{\sigma_X}\left(\frac{\varepsilon_X R_X}{W}\right)^{1-\sigma_X}\right]^{\sigma_X/(1-\sigma_X)}, \tag{A5.119}$$

$$(c_R^x \equiv)\ \frac{\partial c^x}{\partial R_X} = \frac{1}{\phi_X}\left[1-\varepsilon_X + \varepsilon_X^{\sigma_X}\left(\frac{(1-\varepsilon_X)W}{R_X}\right)^{1-\sigma_X}\right]^{\sigma_X/(1-\sigma_X)}, \tag{A5.120}$$

$$\left(c_W^y \equiv\right)\ \frac{\partial c^y}{\partial W} = \frac{1}{\phi_Y}\left[\varepsilon_Y + (1-\varepsilon_Y)^{\sigma_Y}\left(\frac{\varepsilon_Y R}{W}\right)^{1-\sigma_Y}\right]^{\sigma_Y/(1-\sigma_Y)}, \tag{A5.121}$$

$$\left(c_R^y \equiv\right)\ \frac{\partial c^y}{\partial R} = \frac{1}{\phi_Y}\left[1-\varepsilon_Y + \varepsilon_Y^{\sigma_Y}\left(\frac{(1-\varepsilon_Y)W}{R}\right)^{1-\sigma_Y}\right]^{\sigma_Y/(1-\sigma_Y)}, \tag{A5.122}$$

where $R_X \equiv R\left(1+t_{KX}\right)$.

The representative household features the following indirect utility function:

$$V\left(P_X, P_Y, M\right) = \frac{M}{P_V}, \tag{A5.123}$$

$$P_V \equiv \left[\alpha_X P_X^{1-\sigma_D} + (1-\alpha_X) P_Y^{1-\sigma_D}\right]^{1/(1-\sigma_D)}, \tag{A5.124}$$

where P_V is the true price index (linking income and utility), α_X and $1-\alpha_X$ are, respectively, the share parameter for goods X and Y, and σ_D is the substitution elasticity between the two goods.[14] The Marshallian demands are obtained from (A5.123) by applying Roy's Identity:

$$P_X X = \frac{\alpha_X P_X^{1-\sigma_D} M}{\alpha_X P_X^{1-\sigma_D} + (1-\alpha_X) P_Y^{1-\sigma_D}}, \tag{A5.125}$$

[13]The underlying production functions take the following form:

$$X = \phi_X\left[\varepsilon_X L_X^{(\sigma_X-1)/\sigma_X} + (1-\varepsilon_X) K_X^{(\sigma_X-1)/\sigma_X}\right]^{\sigma_X/(\sigma_X-1)},$$

$$Y = \phi_Y\left[\varepsilon_Y L_Y^{(\sigma_Y-1)/\sigma_Y} + (1-\varepsilon_Y) K_Y^{(\sigma_Y-1)/\sigma_Y}\right]^{\sigma_Y/(\sigma_Y-1)}$$

[14]The underlying direct utility function takes the following form:

$$U = \left[\alpha_X^{1/\sigma_D} X^{(\sigma_D-1)/\sigma_D} + (1-\alpha_X)^{1/\sigma_D} Y^{(\sigma_D-1)/\sigma_D}\right]^{\sigma_D/(\sigma_D-1)}.$$

Technology Parameters			
Sector	ϕ_i	ε_i	σ_i
X	1.5	0.6	2.0
Y	2.0	0.7	0.5
Taste Parameters			
	α_X	σ_D	
	0.5	1.5	
Endowments			
$\bar{K}$	20		
$\bar{L}$	60		

Table 6.2: Parameters and endowments

$$P_Y Y = \frac{(1-\alpha_X) P_Y^{1-\sigma_D} M}{\alpha_X P_X^{1-\sigma_D} + (1-\alpha_X) P_Y^{1-\sigma_D}}, \tag{A5.126}$$

where household income is defined in (A5.107) above.

The full model consists of equations (A5.102)-(A5.108), with the functional forms as given in (A5.117)-(A5.122) and (A5.125). In order to simulate the model, numerical values must be assigned to the parameters (ε_X, σ_X, ϕ_X, ε_Y, σ_Y, ϕ_Y, α_X, and σ_D) and the factor supplies ($\bar{L}$ and $\bar{K}$). In real-world applications of the AGE methodology there are two main sources for these parameters. First, there may be *econometric* evidence regarding some of the elasticities appearing in the model, e.g. household demand studies may yield estimates for σ_D and α_X. Second, some parameters (such as ϕ_X and ϕ_Y) may be used in order to *calibrate* the model to some base-year situation. Basic data for the economy are gathered (e.g. from the national income accounts) and the calibration parameters are set in such a way that the AGE model best captures the initial situation.

Since the objective here is to present a simple toy version of an AGE model, we follow Shoven and Whalley by simply postulating the parameters of the model. The values chosen are reported in Table 6.2. Substitution between capital and labour is relatively easy in the X-sector ($\sigma_X = 2$) and relatively difficult in the Y-sector ($\sigma_Y = 0.5$). The substitution elasticity between the two goods in the household's utility function is relatively high ($\sigma_D = 1.5$) ensuring that the two goods are *gross substitutes* (Varian, 1992, p. 395). Finally, note that the economy-wide capital labour ratio is equal to $\bar{K}/\bar{L} = 1/3$.

In the initial equilibrium, there is no tax on capital use in the X-sector, i.e. $t_{KX} = 0$. The general equilibrium results for all variables have been reported in the columns marked (a) in Table 6.3. The interested student should verify that this is indeed the solution, satisfying the zero profit conditions, marginal cost pricing, and goods and labour market equilibrium conditions. Next we introduce a huge shock, in that a large tax is introduced on the use of capital in the X-sector, i.e. t_{KX} is changed from $t_{KX} = 0$ to $t_{KX} = 0.5$. The new general equilibrium is reported in the columns marked (b) in Table 6.3. As a result of the tax increase, P_X, W, and Y increase, whereas R and X decrease. Obviously, therefore, the model predicts that P_X/P_Y and W/R both increase and X/Y decreases. The results in

Equilibrium Prices	(a)	(b)	*Simulation Features*
P_X	1.297	1.488	Parameters: see Table 6.2
P_Y	1.000	1.000	Column (a): zero tax on capital in X-sector ($t_{KX} = 0$)
W	0.944	1.039	Column (b): 50% tax on capital in X-sector ($t_{KX} = 0.5$)
R	1.205	1.050	

Production Accounts	(a)	(b)		(a)	(b)	*Household Accounts*	(a)	(b)
X	31.285	27.767	Y	46.203	50.410			
$P_X X$	40.572	41.323	$P_Y Y$	46.203	50.410	$P_X X + P_Y Y$	86.775	91.733
L_X	31.870	30.746	L_Y	28.130	29.254			
WL_X	30.092	31.953	WL_Y	26.561	30.403	$W\bar{L}$	56.653	62.356
K_X	8.698	5.948	K_Y	16.302	19.052			
RK_X	10.480	6.246	RK_Y	19.643	20.007	$R\bar{K}$	30.123	26.254
$t_{KX} RK_X$	0	3.123				T	0	3.123
C^x	40.572	41.323	C^y	46.203	50.410			

Table 6.3: A large tax change

Table 6.3 imply that the X-sector is relatively labour intensive according to both intensity measures, i.e. the relative results implied by the simulation are consistent with the situation depicted in Figure 6.7 above. Given the parameter values and endowments reported in Table 6.2, the factor substitution effect dominates the negative output effect in equation (A5.101), so that W/R rises unambiguously.

As a result of the tax increase, household income increases from 86.775 to 91.733. Labour income increases (from 56.653 to 62.356), capital income decreases (from 30.123 to 26.254), but the latter decrease is more than compensated by the increase in the lump-sum transfer received from the government (from 0 to 3.123). It would be tempting (but very wrong) to conclude that the household is better off as a result of the tax increase. Income has increased but so has the (relative) price of goods from the X-sector. The welfare comparison can only be made by computing the change in indirect utility (A5.123) resulting from the tax shock. Using the results in Table 6.3, it is not difficult to compute that V changes from $V_{old} = 76.521$ "utils" to $V_{new} = 75.941$ "utils". Even though income increases, the true cost of living index increases by even more, namely from $P_V^{old} = 1.134$ to $P_V^{new} = 1.208$.

But how bad is the loss of 0.58 "utils" for the household? As is pointed out by Shoven and Whalley (1984, p. 1014), there are two widely employed measures to evaluate the welfare effect in economically intuitive terms, namely the *Compensating Variation* (CV) measure and the *Equivalent Variation* (EV) measure. The former takes the new values for prices and income and computes how much income must be taken away (or added) in order to restore the household to its *initial* indifference curve. In contrast, the EV measure takes the old values for prices and income, and computes how much income must be added (or taken away) in order to bring the household to its *post-shock* indifference curve. For a welfare increase (decrease), CV and EV are both positive (negative). Using obvious notation, we find that the

Measure	*Value*	*Units of measurement*
CV	-0.701	units of good Y
EV	-0.658	units of good Y
$(CV/M_{new}) \times 100$	-0.764%	percentage of national income
$(EV/M_{new}) \times 100$	-0.717%	percentage of national income
$(CV/T_{new}) \times 100$	-22.435%	percentage of ultimate tax revenue
$(EV/T_{new}) \times 100$	-21.062%	percentage of ultimate tax revenue

Table 6.4: Welfare effect of capital taxation

two measures are given by:

$$\frac{M_{new} - CV}{P_V^{new}} \equiv \frac{M_{old}}{P_V^{old}} \quad \Leftrightarrow \quad CV \equiv M_{new} - P_V^{new} \frac{M_{old}}{P_V^{old}}, \tag{A5.127}$$

$$\frac{M_{new}}{P_V^{new}} \equiv \frac{M_{old} + EV}{P_V^{old}} \quad \Leftrightarrow \quad EV \equiv P_V^{old} \frac{M_{new}}{P_V^{new}} - M_{old}. \tag{A5.128}$$

The computed welfare losses according to the two measures are reported in Table 6.4. In physical terms, the welfare cost of capital taxation lies in the range of 0.658-0.701 units of good Y. Given that over 50 units of good Y are produced, the welfare cost appears to be small. The second block of figures in Table 6.4 shows that the welfare cost amounts to about 0.7 percent of national income. Although this appears to be a small amount, the third block of entries in the table confirms that the welfare cost is actually quite substantial when expressed in terms of the ultimate tax revenue that the tax gives rise to. Indeed, the welfare cost amounts to more than 20 percent of tax revenue, suggesting that the capital tax is a costly instrument to raise revenue for the government.

6.4.2.3 Approximation error

In this section we have used a simple two-by-two general equilibrium model to illustrate the two major approaches to tax incidence analysis. If one is primarily interested in qualitative results, then the analysis based on the loglinearized model is appropriate. On the other hand, if one is interested also in the size of the effects, then the simulation approach is the appropriate tool. It remains to determine the relationship between the two approaches. To what extent does the loglinear approach understate or overstate the results of a change in the capital tax? In Table 6.5 we present the results from such a comparison. The entries in the table refer to elasticities, e.g., according to column (a), a 1 percent increase in t_{KX} leads to a reduction in X of 0.292 percent. Column (a) presents the effects predicted by the loglinear model. Because the model is loglinear, the *size* of the shock that is administered does not matter, i.e. the predicted elasticities are the same for a shock of $\Delta t_{KX} = 0.01$ and one of $\Delta t_{KX} = 0.50$.

Column (b) presents the results predicted by the nonlinear simulation model for a small shock, i.e.

	(a) *linearized*	(b) *exact*	(c) *exact*	(d) *exact*	(e) *exact*
$\Delta t_{KX} =$		0.01	0.10	0.25	0.50
Variable					
$\tilde{X} - \tilde{Y}$	-0.548	-0.543	-0.501	-0.443	-0.373
		(0.9)	(9.4)	(23.5)	(46.8)
$\tilde{X}$	-0.292	-0.290	-0.275	-0.254	-0.225
		(0.6)	(6.0)	(15.0)	(29.7)
$\tilde{Y}$	0.256	0.254	0.237	0.213	0.182
		(0.8)	(7.9)	(20.0)	(40.6)
$\tilde{P}_X$	0.365	0.363	0.348	0.326	0.295
		(0.5)	(4.9)	(12.0)	(23.9)
$\tilde{W} - \tilde{R}$	0.640	0.638	0.615	0.579	0.526
		(0.4)	(4.2)	(10.7)	(21.8)
$\tilde{W}$	0.272	0.270	0.255	0.232	0.201
		(0.7)	(7.0)	(17.5)	(35.2)
$\tilde{R}$	-0.368	-0.365	-0.339	-0.303	-0.257
		(0.9)	(8.6)	(21.5)	(43.3)
$\tilde{T}$	0.121	0.119	0.109	0.092	0.072
		(1.2)	(12.2)	(31.8)	(67.8)
$\tilde{M}$	0.171	0.169	0.156	0.138	0.114
		(0.9)	(9.4)	(24.0)	(49.4)

Table 6.5: Approximate and exact elasticities

the introduction of a 1 percent capital tax ($\Delta t_{KX} = 0.01$). The comparison of columns (a) and (b) reveals what is to be expected: the loglinear model gives virtually identical results as the nonlinear model does. For small changes the loglinear model is fine. The figures in round brackets underneath the elasticities quantify the size of the approximation error. For variable z, the approximation error is defined as follows:

$$err\,(z) \equiv \left| \frac{z_E - z_A}{z_E} \right| \times 100\%, \tag{A5.129}$$

where z_E and z_A are, respectively, the exact and approximate solutions for z. As column (b) reveals, for a 1 percent shock, the approximation error of the loglinear model amounts to 0.6 percent for X, 0.7 percent for W, and 1.2 percent for T.

Columns (c)-(e) present the results for the nonlinear model for larger shocks. For example, column (e) shows the results of an introduction of a huge capital tax ($\Delta t_{KX} = 0.50$). The comparison between columns (e) and (a) reveals that the results are qualitatively the same (signs are identical) but quantitatively rather different. Indeed, the approximation error is 21.8 percent for W/R, and a whopping 67.8 percent for the tax revenue T! Comparing across columns, one observes that, *for the particular model under consideration*, the approximation error is roughly linear in the shock, i.e. $err\,(z) \approx \gamma_z \Delta t_{KX}$.

6.5 Monopolistic competition

Up to this point attention has been restricted to the standard two-by-two model first formulated by Harberger (1962). In the remainder of this chapter we investigate how robust the standard approach is to deviations from its basic assumptions. Two of the most crucial assumptions are, first, that the market structure is one of perfect competition throughout the economy and, second, that all markets are in equilibrium (clearing markets). In this section we challenge the first of these assumptions by including an imperfectly competitive production sector in the model. In the next section, the second crucial assumption is challenged by incorporating an imperfection in the labour market which gives rise to a positive unemployment rate.

In this section we study a relatively simple general equilibrium model, due to Atkinson and Stiglitz (1980, pp. 208-217), in which there is *monopolistic competition* in one sector of the goods market. The modelling approach is based on the classic paper by Dixit and Stiglitz (1977).[15] As in the basic Harberger model, the economy consists of two sectors of production. Whilst the Y-sector is characterized by perfect competition, there is, however, monopolistic competition in the X-sector, i.e. the small firms populating that sector each possess a little bit of market power. In the X-sector a *differentiated product* is manufactured. Each "slightly different" variety of the good is sold by a single firm with some (but not much) market power under conditions of increasing returns to scale at firm level. Exit/entry of firms drives excess profits to zero. In the Y-sector a homogeneous good is produced by a representative price-taking firm facing a constant returns to scale technology (just as in the basic Harberger-Jones model studied above).

6.5.1 Households

The representative household has the following utility function:

$$U = Z^{\alpha_X} Y^{1-\alpha_X}, \qquad 0 < \alpha_X < 1, \tag{A5.130}$$

where U is utility, Y is consumption of the homogeneous good, and Z is the consumption of a *composite* differentiated good. To keep the model as simple as possible, the utility function is Cobb-Douglas, i.e. the substitution elasticity between Z and Y is equal to unity (in the notation used above, $\sigma_D = 1$).

The composite differentiated good consists of a bundle of closely related product "varieties" which are close but imperfect substitutes for each other:

$$Z \equiv \left[\sum_{i=1}^{N} X_i^{(\sigma_V - 1)/\sigma_V} \right]^{\sigma_V/(\sigma_V - 1)}, \qquad 1 < \sigma_V \ll \infty, \tag{A5.131}$$

[15] For technical background on the Dixit-Stiglitz model, see Heijdra and van der Ploeg (2002, pp. 360-367). See also Brakman and Heijdra (2004) for an extensive discussion of this influential model.

where N is the existing number of different varieties (N is assumed to be large), X_i is consumption of variety i, σ_V is the Allen-Uzawa cross-partial elasticity of substitution between any X_i and any other X_j. Intuitively, the higher is σ_V, the better substitutes the varieties are for each other, i.e. the less differentiated the good is.

The household budget constraint is:

$$\sum_{i=1}^{N} P_i X_i + P_Y Y = M, \tag{A5.132}$$

where P_i is the price of variety i, P_Y is the price of good Y, and M is household income (see below). The household chooses Y and X_i (for $i = 1, \cdots, N$) in order to maximize (A5.130), subject to the definition (A5.131) and the budget constraint (A5.132), taking as given the goods prices and its income. Two-stage budgeting yields the following solutions:[16]

$$Z = \alpha_X \frac{M}{P_Z}, \tag{A5.133}$$

$$Y = (1 - \alpha_X) \frac{M}{P_Y}, \tag{A5.134}$$

$$X_i = \alpha_X \left(\frac{P_i}{P_Z}\right)^{-\sigma_V} \frac{M}{P_Z}, \qquad (i = 1, \cdots, N), \tag{A5.135}$$

where P_Z is the *true price index* of the composite consumption good Z:

$$P_Z \equiv \left[\sum_{i=1}^{N} P_i^{1-\sigma_V}\right]^{1/(1-\sigma_V)}. \tag{A5.136}$$

Intuitively, P_Z represents the price of one unit of Z given that the quantities of all varieties are chosen in an optimal (utility-maximizing) fashion by the household. Notice that (A5.133)-(A5.134) show that income spending shares of Z and Y are constant for the Cobb-Douglas utility function. Finally, (A5.135) is the demand curve facing the producer of variety i. It features a constant price elasticity (in absolute value) which is equal to σ_V–see also equation (A5.142) below.

6.5.2 Firms

The representative firm in the Y-sector is the same as in the standard model. In the absence of taxes, the firm's technology is represented by the following cost function:

$$C^y \equiv c^y (W, R) Y, \tag{A5.137}$$

[16]See Intermezzo 3.1 above for a detailed discussion of the method of two-stage budgeting.

where C^y is total cost, c^y is unit-cost, W is the wage rate, R is the rental price on capital, and Y is output. The conditional factor demands are thus given by:

$$L_Y = c_W^y (W, R) Y, \tag{A5.138}$$

$$K_Y = c_R^y (W, R) Y, \tag{A5.139}$$

where $c_W^y \equiv \partial c^y / \partial W$ and $c_R^y \equiv \partial c^y / \partial R$ are the unit input coefficients. The firm operates under perfect competition so its pricing decision is given by:

$$P_Y = c^y (W, R). \tag{A5.140}$$

The price equals marginal (and average) cost: there are no fixed costs and pure profits are zero.

In the X-sector each producer of variety i faces a price-elastic demand for its own good given in (A5.135) above. The inverse demand function (relating P_i to X_i) can be determined by using (A5.133) and (A5.135):

$$P_i = \frac{\alpha_X M Z^{(\sigma_V - 1)/\sigma_V}}{X_i^{1/\sigma_V}}. \tag{A5.141}$$

Ceteris paribus household income, M, and composite consumption, Z, this demand curve features a constant price elasticity denoted by ε_i:

$$\varepsilon_i \equiv -\frac{\partial \ln X_i}{\partial \ln P_i} = \sigma_V. \tag{A5.142}$$

In order to operate at all, the firm must incur *fixed costs*, i.e. costs not related to scale of production. For example, it must have an administrative office which utilizes labour and/or capital to provide services to the firm. The cost function associated with the maintenance of the office is:

$$C^o = c^o (W, R), \tag{A5.143}$$

where $c^o (W, R)$ is the unit cost of an office (and we note that only one office is needed per firm). Firm i's derived input demands arising from fixed costs are thus equal to:

$$L_{Oi} = c_W^o (W, R), \tag{A5.144}$$

$$K_{Oi} = c_R^o (W, R). \tag{A5.145}$$

In order to manufacture its output, the firm must also incur *variable production costs*. The variable cost function is:

$$C^x = c^x(W,R)\,X_i, \tag{A5.146}$$

and the derived factor demands are given by:

$$L_{Xi} = c^x_W(W,R)\,X_i, \tag{A5.147}$$

$$K_{Xi} = c^x_R(W,R)\,X_i. \tag{A5.148}$$

Profit of firm i is denoted by Π_i and equals revenue minus total costs (i.e. the sum of variable cost and fixed cost):

$$\begin{aligned}\Pi_i &\equiv P_i(X_i)\,X_i - C^x - C^o \\ &= P_i(X_i)\,X_i - c^x(W,R)\,X_i - c^o(W,R)\,,\end{aligned} \tag{A5.149}$$

where $P_i(X_i)$ is the inverse demand curve for variety X_i (see (A5.141) above). Firm i chooses its output in order to maximize profit (A5.149), taking as given the input prices (W and R), and operating under the *Cournot-Nash* assumption that other producers in the X-sector do not change their production (so that Z in (A5.141) can be considered constant).[17] The first-order condition for profit maximization is:

$$\begin{aligned}\frac{d\Pi_i}{dX_i} = P_i + X_i\frac{\partial P_i}{\partial X_i} - c^x(W,R) = 0 \quad &\Leftrightarrow \\ c^x(W,R) = P_i\left[1 + \frac{X_i}{P_i}\frac{\partial P_i}{\partial X_i}\right].&\end{aligned} \tag{A5.150}$$

By noting (A5.142) we can simplify (A5.150) and obtain the familiar markup pricing rule:

$$P_i = \mu c^x(W,R)\,, \tag{A5.151}$$

$$\mu \equiv \frac{\sigma_V}{\sigma_V - 1}, \tag{A5.152}$$

where μ (> 1) is the *gross markup* of price over marginal cost.

The X-sector is characterized by *Chamberlinian* monopolistic competition and free entry/exit of firms. Since all firms are the same, entry/exit occurs until all *active* firms exactly break even and make zero profit:

$$\Pi_i = 0, \qquad \text{for } i = 1, 2, \cdots, N. \tag{A5.153}$$

By using the pricing rule (A5.151) and the profit function (A5.149) in (A5.153) we obtain the condition

[17] In addition, the firm also ignores the effect its own production level (X_i) has on the magnitude of Z. This assumption is justified by noting that the number of firms (N) is assumed to be large so that X_i is only a very small part of Z.

determining individual equilibrium firm size:

$$\begin{aligned}\Pi_i &= P_i\left(X_i\right)X_i - c^x\left(W,R\right)X_i - c^o\left(W,R\right)\\ &= X_i\left[\mu c^x\left(W,R\right) - c^x\left(W,R\right)\right] - c^o\left(W,R\right)\\ &= X_i\left(\mu-1\right)c^x\left(W,R\right) - c^o\left(W,R\right) = 0,\end{aligned} \tag{A5.154}$$

or:

$$X_i = \frac{c^o\left(W,R\right)}{\left(\mu-1\right)c^x\left(W,R\right)}. \tag{A5.155}$$

According to (A5.155), optimal equilibrium firm size is the same for all active firms (because c^o, c^x, and μ are the same), increasing in fixed cost (c^o), decreasing in unit variable cost (c^x), and decreasing in μ (i.e. increasing in σ_V).

The model is completely symmetric, i.e. for $i = 1, \cdots, N$ we have:

$$\begin{aligned}P_i &= P_X, \quad L_{Xi} = \bar{L}_X, \quad L_{Oi} = \bar{L}_O,\\ X_i &= \bar{X}, \quad K_{Xi} = \bar{K}_X, \quad K_{Oi} = \bar{K}_O,\end{aligned} \tag{A5.156}$$

and the model can be studied in aggregate terms. Finally, we need to tie up some loose ends. First, in the symmetric equilibrium the factor market clearing conditions are given by:

$$\begin{aligned}\bar{L} &= N\left(\bar{L}_X + \bar{L}_O\right) + L_Y,\\ &= c_W^x\left(W,R\right)N\bar{X} + Nc_W^o\left(W,R\right) + c_W^y\left(W,R\right)Y,\end{aligned} \tag{A5.157}$$

$$\begin{aligned}\bar{K} &= N\left(\bar{K}_X + \bar{K}_O\right) + K_Y,\\ &= c_R^x\left(W,R\right)N\bar{X} + Nc_R^o\left(W,R\right) + c_R^y\left(W,R\right)Y.\end{aligned} \tag{A5.158}$$

Second, since there are no excess profits and no taxes, household income is simply equal to factor payments:

$$M = W\bar{L} + R\bar{K}. \tag{A5.159}$$

The model consists of the demand equations (A5.133)-(A5.134), the pricing equations (A5.140) and (A5.151), the expression for the individual firm size (A5.155), the factor market clearing conditions (A5.157)-(A5.158), and the income definition (A5.159). The endogenous variables are $\bar{X}$, Y, P_X, P_Y, W, R, and N. The exogenous variables are $\bar{K}$ and $\bar{L}$, and μ is a structural parameter. As in the basic Harberger-Jones model, *Walras' Law* implies that one equation is redundant and that only the relative price (P_X/P_Y) is determined within the model.

6.5.3 Tax incidence in the monopolistic competition model

In this subsection we demonstrate some of the properties of the monopolistic competition model by performing a small-scale tax policy analysis. In the interest of brevity we only consider two tax instruments, both impacting on the firms in the differentiated sector. First, we assume that the use of capital for production purposes is taxed at rate t_{KX}. Second, we assume that the policy maker wishes to subsidize fixed costs in the X-sector, say to stimulate small-scale "entrepreneurship" in the economy. It does so by subsidizing the use of overhead capital at rate s_{KO}. No other taxes or subsidies exist.

The modified general equilibrium model consists of the following equations:

$$\bar{L} = c_W^x\left(W, R\left(1+t_{KX}\right)\right)X + Nc_W^o\left(W, R\left(1-s_{KO}\right)\right) + c_W^y\left(W, R\right)Y, \tag{A5.160}$$

$$\bar{K} = c_R^x\left(W, R\left(1+t_{KX}\right)\right)X + Nc_R^o\left(W, R\left(1-s_{KO}\right)\right) + c_R^y\left(W, R\right)Y, \tag{A5.161}$$

$$P_X = \mu c^x\left(W, R\left(1+t_{KX}\right)\right), \tag{A5.162}$$

$$P_Y = c^y\left(W, R\right), \tag{A5.163}$$

$$P_X X = \alpha_X\left[W\bar{L} + R\bar{K} + T\right], \tag{A5.164}$$

$$P_Y Y = \left(1-\alpha_X\right)\left[W\bar{L} + R\bar{K} + T\right], \tag{A5.165}$$

$$\bar{X} = \frac{X}{N} = \frac{c^o\left(W, R\left(1-s_{KO}\right)\right)}{\left(\mu-1\right)c^x\left(W, R\left(1+t_{KX}\right)\right)}, \tag{A5.166}$$

$$T \equiv t_{KX}Rc_R^x\left(W, R\left(1+t_{KX}\right)\right)X - s_{KO}RNc_R^o\left(W, R\left(1-s_{KO}\right)\right), \tag{A5.167}$$

where $X \equiv N\bar{X}$ is "aggregate" output in the X-sector and T is the tax revenue which is recycled in a lump-sum fashion to households. The capital tax, t_{KX}, affects (i) factor demands for production purposes in equations (A5.160) and (A5.161), (ii) marginal production cost in equations (A5.162) and (A5.166), and (iii) tax revenue in (A5.167). The subsidy, s_{KO}, affects (i) factor demands for overhead purposes in equations (A5.160) and (A5.161), (ii) unit fixed costs in equation (A5.166), and (iii) tax revenue in (A5.167).

As before, we restrict attention to the *relative* effects of the capital tax and the fixed cost subsidy. Following similar steps as before the model can be loglinearized. By dividing (A5.164) by (A5.165) and loglinearizing we find the expression for relative demand:

$$\tilde{X} - \tilde{Y} = -\left[\tilde{P}_X - \tilde{P}_Y\right]. \tag{A5.168}$$

The pricing equations (A5.162)-(A5.163) are loglinearized as follows:

$$\tilde{P}_X = \theta_{LX}\tilde{W} + \theta_{KX}\left[\tilde{R} + \tilde{t}_{KX}\right], \tag{A5.169}$$

$$\tilde{P}_Y = \theta_{LY}\tilde{W} + \theta_{KY}\tilde{R}, \tag{A5.170}$$

where we have incorporated the fact that the gross markup is constant (so that $\tilde{\mu} = 0$), $\tilde{t}_{KX} \equiv dt_{KX}/\left(1 + t_{KX}\right)$, and the θ_{ji} coefficients are defined as follows:

$$\theta_{LX} \equiv \frac{Wc_W^x}{c^x}, \quad \theta_{KX} \equiv \frac{R\left(1 + t_{KX}\right)c_R^x}{c^x} = 1 - \theta_{LX}, \tag{A5.171}$$

$$\theta_{LY} \equiv \frac{Wc_W^y}{c^y}, \quad \theta_{KY} \equiv \frac{Rc_R^y}{c^y} = 1 - \theta_{LY}. \tag{A5.172}$$

By deducting (A5.172) from (A5.171) we obtain the expression for the relative goods price:

$$\tilde{P}_X - \tilde{P}_Y = \theta^* \left[\tilde{W} - \tilde{R}\right] + \theta_{KX}\tilde{t}_{KX}, \tag{A5.173}$$

where θ^* is defined as:

$$\theta^* \equiv \theta_{LX} - \theta_{LY} = \theta_{KY} - \theta_{KX} > 0. \tag{A5.174}$$

In (A5.174) we have incorporated the assumption (also made in the standard Harberger-Jones model) that *production* activities in the X-sector are relatively labour intensive.

By loglinearizing the expression for optimal equilibrium firm size, equation (A5.166), we obtain:

$$\tilde{\bar{X}} = \tilde{c}^o - \tilde{c}^x, \tag{A5.175}$$

where $\tilde{c}^x = \tilde{P}_X$ (by (A5.162) and the constancy of the gross markup) and $\tilde{c}^o$ is defined as follows:

$$\tilde{c}^o = \theta_{LO}\tilde{W} + \theta_{KO}\left[\tilde{R} - \tilde{s}_{KO}\right], \tag{A5.176}$$

where $\tilde{s}_{KO} \equiv ds_{KO}/\left(1 - s_{KO}\right)$, and where θ_{LO} and θ_{KO} are defined as follows:

$$\theta_{LO} \equiv \frac{Wc_W^o}{c^o}, \quad \theta_{KO} \equiv \frac{R\left(1 - s_{KO}\right)c_R^o}{c^o} = 1 - \theta_{LO}. \tag{A5.177}$$

In equation (A5.176), θ_{LO} measure the labour intensity of fixed cost in terms of factor shares. By substituting equation (A5.176) into (A5.175) and noting $\tilde{c}^x = \tilde{P}_X$ and (A5.169) we find:

$$\tilde{\bar{X}} = -\eta^* \left[\tilde{W} - \tilde{R}\right] - \theta_{KX}\tilde{t}_{KX} - \theta_{KO}\tilde{s}_{KO}, \tag{A5.178}$$

where η^* is defined as follows:

$$\eta^* \equiv \theta_{LX} - \theta_{LO} = \theta_{KO} - \theta_{KX} > 0. \tag{A5.179}$$

In equation (A5.179) we have incorporated the assumption that variable (production) cost is relatively labour intensive compared to fixed (overhead) cost in terms of factor shares. It follows from (A5.178) that an increase in W/R leads to a decrease in individual firm size. Furthermore, ceteris paribus W/R, an increase in the capital tax increases marginal production cost and reduces equilibrium firm size. Similarly, an increase in the fixed cost subsidy leads to a decrease in fixed cost and a reduction in equilibrium firm size.

It remains to derive the loglinearized factor market equilibrium conditions. After some manipulation we obtain from (A5.160) and (A5.161):

$$\tilde{\bar{L}} = \lambda_{LX}\left[\tilde{X} + \tilde{c}_W^x\right] + \lambda_{LY}\left[\tilde{Y} + \tilde{c}_W^y\right] + \lambda_{LO}\left[\tilde{N} + \tilde{c}_W^o\right], \tag{A5.180}$$

$$\tilde{\bar{K}} = \lambda_{KX}\left[\tilde{X} + \tilde{c}_R^x\right] + \lambda_{KY}\left[\tilde{Y} + \tilde{c}_R^y\right] + \lambda_{KO}\left[\tilde{N} + \tilde{c}_R^o\right], \tag{A5.181}$$

where the λ-coefficients are defined as:

$$\lambda_{LX} \equiv \frac{N\bar{L}_X}{\bar{L}}, \quad \lambda_{LY} \equiv \frac{L_Y}{\bar{L}}, \quad \lambda_{LO} \equiv \frac{N\bar{L}_O}{\bar{L}}, \tag{A5.182}$$

$$\lambda_{KX} \equiv \frac{N\bar{K}_X}{\bar{K}}, \quad \lambda_{KY} \equiv \frac{K_Y}{\bar{K}}, \quad \lambda_{KO} \equiv \frac{N\bar{K}_O}{\bar{K}}, \tag{A5.183}$$

$$1 = \lambda_{LX} + \lambda_{LY} + \lambda_{LO} = \lambda_{KX} + \lambda_{KY} + \lambda_{KO}, \tag{A5.184}$$

and where the loglinearized unit input coefficients are given by:

$$\tilde{c}_W^x = -\theta_{KX}\sigma_X\left[\tilde{W} - \tilde{R} - \tilde{t}_{KX}\right], \tag{A5.185}$$

$$\tilde{c}_W^y = -\theta_{KY}\sigma_Y\left[\tilde{W} - \tilde{R}\right], \tag{A5.186}$$

$$\tilde{c}_R^x = \theta_{LX}\sigma_X\left[\tilde{W} - \tilde{R} - \tilde{t}_{KX}\right], \tag{A5.187}$$

$$\tilde{c}_R^y = \theta_{LY}\sigma_Y\left[\tilde{W} - \tilde{R}\right], \tag{A5.188}$$

$$\tilde{c}_W^o = -\theta_{KO}\sigma_O\left[\tilde{W} - \tilde{R} + \tilde{s}_{KO}\right], \tag{A5.189}$$

$$\tilde{c}_R^o = \theta_{LO}\sigma_O\left[\tilde{W} - \tilde{R} + \tilde{s}_{KO}\right]. \tag{A5.190}$$

In equations (A5.189)-(A5.190), the parameter σ_O denotes the substitution elasticity between capital and labour in fixed cost. For a given wage-rental ratio, W/R, an increase in the fixed cost subsidy induces firms in the differentiated sector to substitute capital for labour in the production of overhead services. By substituting (A5.185)-(A5.190) in the relevant places in (A5.180)-(A5.181) we obtain:

$$\begin{aligned}\tilde{\bar{L}} &= \lambda_{LX}\tilde{X} + \lambda_{LY}\tilde{Y} + \lambda_{LO}\tilde{N} + \lambda_{LX}\theta_{KX}\sigma_X\tilde{t}_{KX} - \lambda_{LO}\theta_{KO}\sigma_O\tilde{s}_{KO} \\ &\quad - \left[\lambda_{LX}\theta_{KX}\sigma_X + \lambda_{LY}\theta_{KY}\sigma_Y + \lambda_{LO}\theta_{KO}\sigma_O\right]\left[\tilde{W} - \tilde{R}\right],\end{aligned} \tag{A5.191}$$

and:

$$\tilde{\bar{K}} = \lambda_{KX}\tilde{X} + \lambda_{KY}\tilde{Y} + \lambda_{KO}\tilde{N} - \lambda_{KX}\theta_{LX}\sigma_X\tilde{t}_{KX} + \lambda_{KO}\theta_{LO}\sigma_O\tilde{s}_{KO}$$

$$+\left[\lambda_{KX}\theta_{LX}\sigma_X+\lambda_{KY}\theta_{LY}\sigma_Y+\lambda_{KO}\theta_{LO}\sigma_O\right]\left[\tilde{W}-\tilde{R}\right]. \tag{A5.192}$$

Finally, by deducting (A5.192) from (A5.191) we get:

$$\tilde{\tilde{L}}-\tilde{\tilde{K}}=\lambda^*\left[\tilde{X}-\tilde{Y}\right]+\mu^*\tilde{\tilde{X}}-\sigma^*\left[\tilde{W}-\tilde{R}\right]+a_X\sigma_X\tilde{t}_{KX}-a_O\sigma_O\tilde{s}_{KO}, \tag{A5.193}$$

where we have used the fact that $\tilde{N}=\tilde{X}-\tilde{\tilde{X}}$, $\lambda_{LX}+\lambda_{LO}=1-\lambda_{LY}$, and $\lambda_{KX}+\lambda_{KO}=1-\lambda_{KY}$. The composite parameters, λ^*, μ^*, σ^*, a_X, a_Y, and a_O are defined as follows:

$$\lambda^*\equiv\lambda_{KY}-\lambda_{LY}=(\lambda_{LX}+\lambda_{LO})-(\lambda_{KX}+\lambda_{KO})>0, \tag{A5.194}$$

$$\mu^*\equiv\lambda_{KO}-\lambda_{LO}>0, \tag{A5.195}$$

$$\sigma^*\equiv a_X\sigma_X+a_Y\sigma_Y+a_O\sigma_O>0, \tag{A5.196}$$

$$a_X\equiv\lambda_{LX}\theta_{KX}+\lambda_{KX}\theta_{LX}>0, \tag{A5.197}$$

$$a_Y\equiv\lambda_{LY}\theta_{KY}+\lambda_{KY}\theta_{LY}>0, \tag{A5.198}$$

$$a_O\equiv\lambda_{LO}\theta_{KO}+\lambda_{KO}\theta_{LO}>0. \tag{A5.199}$$

In equation (A5.194) we have already incorporated the assumption that the Y-sector is relatively capital intensive, i.e. that the X-sector as a whole (including production and overhead factors) is relatively labour intensive (both measured in physical units). Furthermore, in (A5.195) it is assumed that fixed cost are relatively capital intensive (in terms of physical units).

For future purposes we restate the key equations of the loglinearized model in its most convenient format:

$$\tilde{X}-\tilde{Y}=-\theta^*\left[\tilde{W}-\tilde{R}\right]-\theta_{KX}\tilde{t}_{KX}, \tag{A5.200}$$

$$\lambda^*\left[\tilde{X}-\tilde{Y}\right]=\left[\tilde{\tilde{L}}-\tilde{\tilde{K}}\right]+\left[\eta^*\mu^*+\sigma^*\right]\left[\tilde{W}-\tilde{R}\right]$$

$$+\left[\mu^*\theta_{KX}-a_X\sigma_X\right]\tilde{t}_{KX}+\left[\mu^*\theta_{KO}+a_O\sigma_O\right]\tilde{s}_{KO}. \tag{A5.201}$$

Equation (A5.200) is obtained by combining the relative demand equation (A5.168) with the relative pricing rule equation (A5.173). It thus expresses relative demand in terms of relative factor prices (W/R) rather than the relative goods price (P_X/P_Y). In view of the assumptions made ($\theta^*>0$, $\lambda^*>0$, $\eta^*>0$, and $\mu^*>0$), it follows that the relative demand curve slopes down–see the DPR curves in Figures 6.8-6.10. Equation (A5.201) is obtained by substituting (A5.178) into the relative factor market equilibrium condition (A5.193). This equation has been labeled FME in Figures 6.8-6.10.

6.5.3.1 Fixed cost subsidy

In Figure 6.8 the effect of an increase in the fixed cost subsidy is illustrated, i.e. $\tilde{s}_{KO} > 0$ and $\tilde{t}_{KX} = 0$. In that figure, the original relative demand curve (A5.168) is plotted in the top right-hand panel (the D curve) and the expression for equilibrium firm size (A5.178) is plotted in the bottom left-hand panel (the FS curve). As a result of the shock, the DPR curve (A5.200) stays put and the FME curve (A5.201) shifts up, say from FME_0 to FME_1. The equilibrium shifts from E_0 to E_1, the relative demand for X rises, and both W/R and P_X/P_Y fall. It follows from (A5.178) that, for a given value of W/R, an increase in the fixed cost subsidy reduces the equilibrium firm size, i.e. in terms of Figure 6.8 the FS curve shifts up, say from FS_0 to FS_1. There are clearly two offsetting effects on equilibrium firm size. The direct effect of the subsidy is to decrease firm size (because it reduces fixed cost at a given value of W/R). This is the move from point A to point B in the bottom left-hand panel. As a result of general equilibrium interactions, however, the W/R ratio falls which leads to an increase in equilibrium firm size because fixed cost is capital intensive. This is the move from point B to point C. The figure has been drawn under the assumption that the direct effect dominates the general equilibrium effect, so that equilibrium firm size declines. But is that the case? To find out we solve the model mathematically. By using (A5.178) and (A5.200)-(A5.201) we obtain:

$$\tilde{X} - \tilde{Y} = -\theta^* \left[\tilde{W} - \tilde{R}\right] = \frac{\theta^* \left[\mu^* \theta_{KO} + a_O \sigma_O\right]}{\lambda^* \theta^* + \eta^* \mu^* + \sigma^*} \tilde{s}_{KO} > 0, \tag{A5.202}$$

$$\tilde{\bar{X}} = \left[\frac{\eta^* \left[\mu^* \theta_{KO} + a_O \sigma_O\right]}{\lambda^* \theta^* + \eta^* \mu^* + \sigma^*} - \theta_{KO}\right] \tilde{s}_{KO}. \tag{A5.203}$$

The first term within square brackets on the right-hand side of (A5.203) is the general equilibrium effect whilst the second term is the direct effect. Using the definition of η^* from (A5.179) we find that (A5.203) can be further simplified to:

$$\tilde{\bar{X}} = -\left[\frac{\theta_{KO} \left(\lambda^* \theta^* + a_X \sigma_X + a_Y \sigma_Y\right) + \theta_{KX} a_O \sigma_O}{\lambda^* \theta^* + \eta^* \mu^* + \sigma^*}\right] \tilde{s}_{KO} < 0, \tag{A5.204}$$

i.e. firm size declines unambiguously.

6.5.3.2 Capital tax

The effects of an increase in the corporate tax ($\tilde{t}_{KX} > 0$) are illustrated in Figure 6.9. By setting $\tilde{s}_{KO} = \tilde{\bar{L}} = \tilde{\bar{K}} = 0$ we find that (A5.201) is simplified to:

$$\lambda^* \left[\tilde{X} - \tilde{Y}\right] = \left[\eta^* \mu^* + \sigma^*\right] \left[\tilde{W} - \tilde{R}\right] + \left[\mu^* \theta_{KX} - a_X \sigma_X\right] \tilde{t}_{KX}. \tag{A5.205}$$

In the top right-hand panel, the tax shock shifts the DPR curve (A5.200) down, say from DPR_0 to DPR_1. The effect on the FME curve (A5.205) is ambiguous as it is determined by the interplay between an

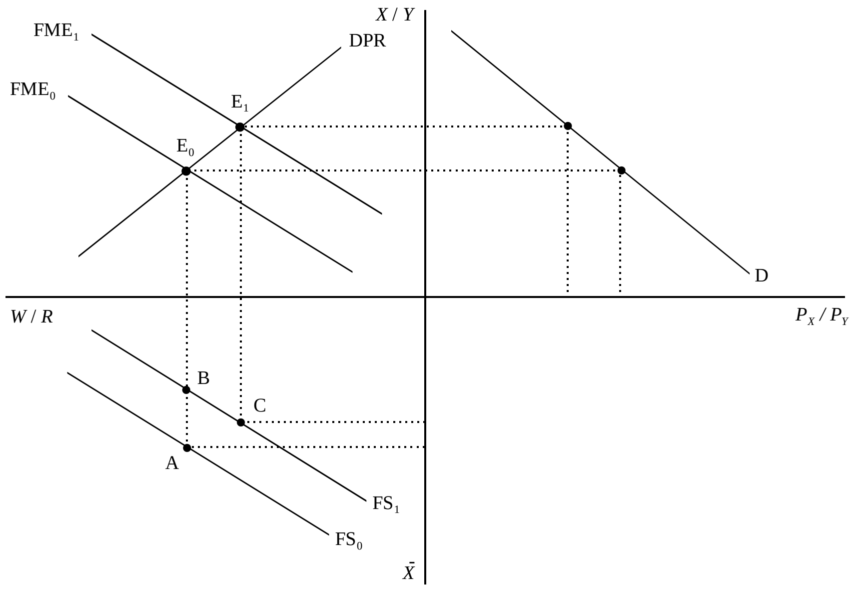

Figure 6.8: Increase in the fixed cost subsidy s_{KO} under monopolistic competition

output effect ($\mu^*\theta_{KX}$) and a factor substitution effect ($-a_X\sigma_X$). In the diagram it is assumed that the latter effect dominates the former effect ($a_X\sigma_X > \mu^*\theta_{KX}$) so that the FME curve shifts down, say from FME_0 to FME_1. The equilibrium shifts from E_0 to E_1 and (for the case drawn) X/Y falls whilst W/R rises. Mathematically, the comparative static effects obtained from equations (A5.200) and (A5.205) are:

$$\tilde{X} - \tilde{Y} = -\left[\frac{\theta_{KX}\left[\eta^*\mu^* + \sigma^*\right] + \theta^*\left[a_X\sigma_X - \mu^*\theta_{KX}\right]}{\lambda^*\theta^* + \eta^*\mu^* + \sigma^*}\right]\tilde{t}_{KX} \gtreqless 0, \tag{A5.206}$$

$$\tilde{W} - \tilde{R} = \left[\frac{-\lambda^*\theta_{KX} + \left[a_X\sigma_X - \mu^*\theta_{KX}\right]}{\lambda^*\theta^* + \eta^*\mu^* + \sigma^*}\right]\tilde{t}_{KX} \gtreqless 0. \tag{A5.207}$$

In the bottom left-hand panel of Figure 6.9, the FS curve (A5.178) shifts up, say from FS_0 to FS_1. Not surprisingly, in view of the fact that the effect on W/R cannot be signed unambiguously, the effect on the equilibrium firm size is also ambiguous. Indeed, by using (A5.207) in (A5.178) we find:

$$\tilde{\bar{X}} = \left[\frac{\lambda^*\eta^*\theta_{KX} - \eta^*\left[a_X\sigma_X - \mu^*\theta_{KX}\right]}{\lambda^*\theta^* + \eta^*\mu^* + \sigma^*}\right]\tilde{t}_{KX} - \theta_{KX}\tilde{t}_{KX} \gtreqless 0. \tag{A5.208}$$

In the bottom left-hand panel of Figure 6.9, the direct effect is the move from point A to point B and the induced general equilibrium effect is the move from point B to point C. In the case drawn, the direct effect dominates and equilibrium firm size falls as a result of the capital tax.

6.5.3.3 Combined shock

Up to this point we have implicitly assumed that the tax authority can distinguish between the firm's use of capital for production purposes ($\bar{K}_X$) and for overhead purposes ($\bar{K}_O$). This is clearly a heroic assumption. Following Atkinson and Stiglitz (1980, pp. 213-216), we now consider the case in which the

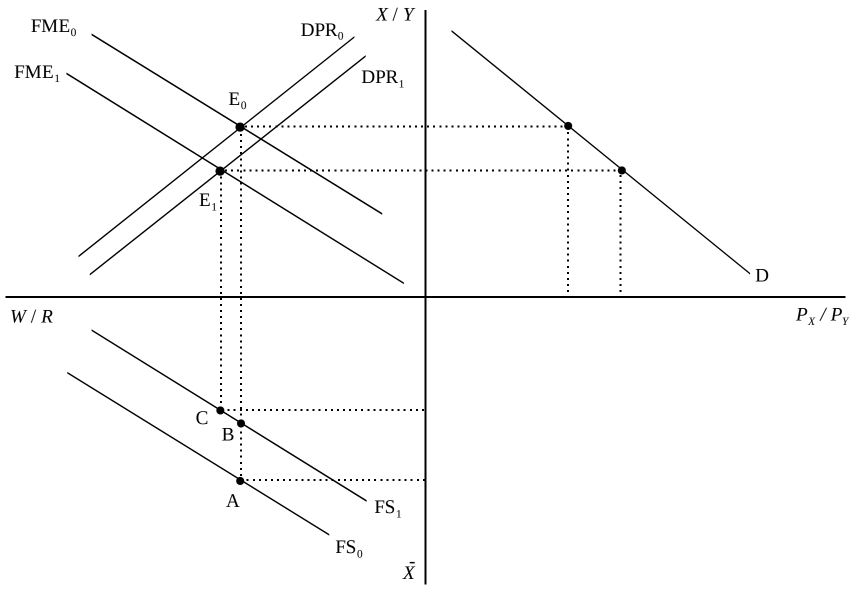

Figure 6.9: Increase in the corporate tax t_{KX} under monopolistic competition

capital tax applies to all capital used in the X-sector. In our model this amounts to setting $s_{KO} = -t_{KX}$ so that $\tilde{s}_{KO} = -\tilde{t}_{KX}$. By using this result in (A5.200)-(A5.201) and (A5.178) we find that the general equilibrium effects of a marginal change in the capital tax rate are given by:

$$\tilde{X} - \tilde{Y} = -\frac{\theta_{KX}\left[\eta^*\mu^* + \sigma^*\right] + \theta^*\left[\eta^*\mu^* + a_X\sigma_X + a_O\sigma_O\right]}{\lambda^*\theta^* + \eta^*\mu^* + \sigma^*}\tilde{t}_{KX} < 0, \tag{A5.209}$$

$$\tilde{W} - \tilde{R} = \frac{-\lambda^*\theta_{KX} + \eta^*\mu^* + a_X\sigma_X + a_O\sigma_O}{\lambda^*\theta^* + \eta^*\mu^* + \sigma^*}\tilde{t}_{KX} \gtreqless 0, \tag{A5.210}$$

$$\tilde{\bar{X}} = -\frac{\lambda^*\theta_{KY} + a_Y\sigma_Y}{\lambda^*\theta^* + \eta^*\mu^* + \sigma^*}\tilde{t}_{KX} < 0. \tag{A5.211}$$

In terms of Figure 6.10, an increase in the capital tax ($\tilde{t}_{KX} > 0$) shifts both the FME and DPR curves (from, respectively, FME$_0$ to FME$_1$ and from DPR$_0$ to DPR$_1$). The effect on X/Y is unambiguously negative (see (A5.209)) but the effect on the W/R ratio is ambiguous (see (A5.210)). Despite this ambiguity, the equilibrium firm size rises as a result of the tax increase (see (A5.211)). As Atkinson and Stiglitz (1980, p. 216) point out, the increase in firm scale can be interpreted as an increase in industrial concentration.

6.5.4 Concluding remarks

This subsection has demonstrated that it is quite feasible to construct a fully tractable general equilibrium model with monopolistic competition in the goods market. The tractability is attributable to the analytical simplifications built into the Dixit-Stiglitz framework employed in the model. The key simplifications are (i) the very specific functional form for preferences (in (A5.131)) leading to the simple variety demand functions (A5.135), and (ii) the Cournot-Nash assumption with respect to producer behaviour in the X-sector. Since the model is based on such very special assumptions, it cannot be easily

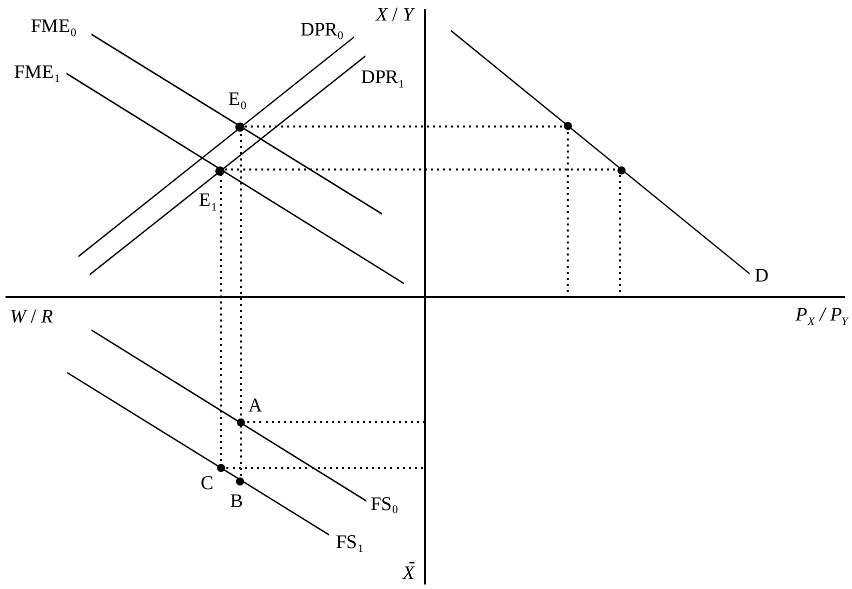

Figure 6.10: Taxing all capital in the *X*-sector

generalized. Clearly, the monopolistic competition model does capture some aspects of imperfect competition rather well (such as the existence of firm-level scale economies and price setting behaviour). It cannot, however, capture the small-group oligopolistic interactions that may be important in some sectors in the economy. A general equilibrium version of the conjectural variations oligopoly model (discussed in Chapter 5 above) may be better suited in that case.

6.6 Labour market frictions and unemployment

In this subsection the emphasis shifts from the goods market to the labour market. Recall that the original Harberger-Jones model assumes perfectly competitive behaviour on all market as well as flexibility of all prices of goods and production factors. In this subsection we study the implications of labour market frictions and unemployment in the context of a two-sector Harberger-Jones model.

Although there are many different ways of modeling labour market frictions, here we wish to focus on the idea of dual labour markets and efficiency wages.[18] The basic *dual labour market* idea was formulated by Doeringer and Piore (1971). There are two sectors in the economy, namely a high-wage *primary sector* and a low-wage *secondary* sector. In the primary sector workers occupy the attractive jobs: employment is stable, there are good provisions for training, jobs involve skilled work and workers are allowed to carry responsibility. In contrast, in the secondary sector the unattractive jobs are located: there is only casual attachment between firm and worker, work is mostly unskilled, and there are little training or promotion opportunities.

[18]See Heijdra and van der Ploeg (chs. 7-9) for an extensive survey of the different labour market approaches. The efficiency wage model is discussed in detail in Weiss (1991).

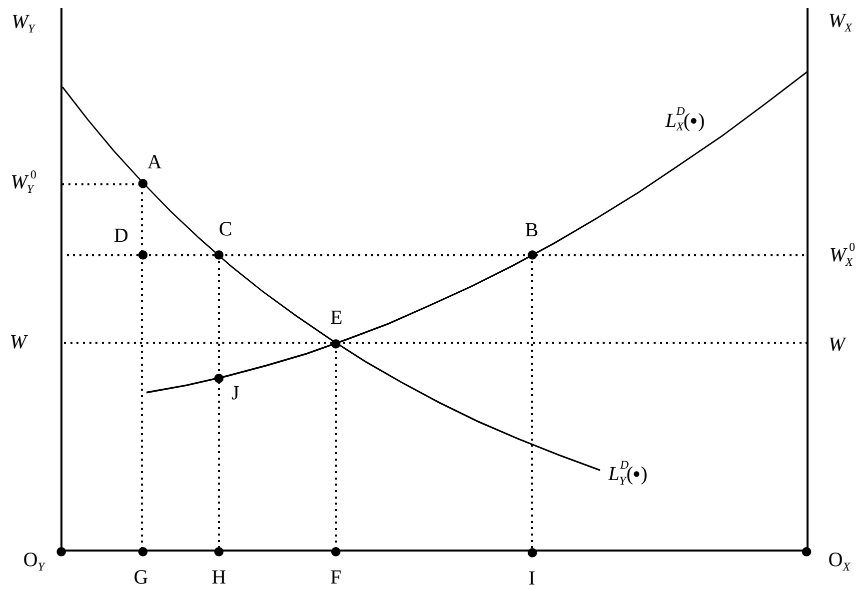

Figure 6.11: Wage differential in a dual labour market

The dual labour market notion can be explained further with the aid of Figure 6.11. In that figure, the given supply of labour ($\bar{L}$) is equal to the horizontal distance O_YO_X. Labour demand in the (primary) Y-sector is measured relative to O_Y on the horizontal axis from left to right. Labour demand in the (secondary) X-sector is measured with respect to O_X, i.e. from right to left on the horizontal axis. The respective labour demand curves are denoted by $L_Y^D(\cdot)$ and $L_X^D(\cdot)$ and the wage rates are measured on the right-hand axis for W_X and on the left-hand axis for W_Y. Note that the figure is only a partial equilibrium representation in that goods prices and the capital stocks in the two sectors are held constant.[19]

With full mobility of labour and a flexible wage rate, equilibrium is attained at point E, where the wage rate in the two sectors is equal (at $W = W_Y = W_X$) and there is full employment of labour. If the wage is too high *and inflexible*, say at the level $W = W_X^0$ and labour is fully mobile, then the equilibrium is at point C for the Y-sector and at point B for the X-sector. There is unemployment represented by the horizontal distance CB. This still does not capture the notion of a dual labour market, however, because the wage rate is the same in both sectors. What is needed is some other kind of labour market friction.[20]

Following Bulow and Summers (1986) and Shapiro and Stiglitz (1984), we explain the joint occurrence of unemployment and a persistent wage differential (between identical workers) by appealing to the notion of *efficiency wages*. In particular, we assume that the cost of supervising workers differs between the two sectors. In the primary sector supervision is costly and firms pay a wage premium to

[19]Indeed, the labour demand functions appearing in Figure 6.11 are defined implicitly by $W_i = P_i \partial F^i(L_i, K_i)/\partial L_i$, where $F^i(L_i, K_i)$ is the production function in sector i (for $i = X, Y$).

[20]McDonald and Solow (1985) introduce a quantity restriction (fixed union membership) and assume that the Y-sector is unionized. In terms of Figure 6.11, union membership is equal to O_YH, the union picks point A and sets the wage at W_Y^0. Unemployment in the Y-sector equals GH. In the secondary sector labour supply equals O_XH, the wage is flexible and the equilibrium is at point J. There is both a wage differential and unemployment.

induce effort (the so-called *shirking model*). In contrast, in the secondary sector supervision is easy and wages are competitive.

The basic model used in this subsection is due to Atkinson (1994). The following key assumptions are made. First, the goods markets are competitive and capital is perfectly mobile across sectors so that its rental rate is the same in both sectors. Second, workers are identical and risk-neutral, i.e. we continue to assume homothetic preferences. Third, in order to keep the model as simple as possible, technology is assumed to feature fixed coefficients, i.e. the substitution elasticities in the two sectors are both zero ($\sigma_X = \sigma_Y = 0$). Fourth, the model is static and the labour market reconvenes at the start of each period.

From an *ex ante* perspective, labour is fully mobile across sectors. At the beginning of each period, workers must choose between two options. Option 1 is to accept a job in the secondary sector at the going wage rate W_X. Option 2 is to enter the primary sector and join a "queue of workers" waiting for a job in that sector. Only a fraction of the workers in the queue obtain a job (at wage rate W_Y which exceeds W_X) while the rest of them remain unemployed (they cannot turn around and join the secondary sector by assumption!). Since the workers are risk neutral, the decision between the two options is such that the *expected* wage obtained by joining the queue equals the certain wage in the secondary sector. In terms of Figure 6.11, O_XI is employment in the secondary sector, O_YI is the queue in the primary sector, O_YG primary sector workers are employed and GI workers are unemployed. The model determines endogenously both the location of points I and G and wages in the two sectors.

6.6.1 The intersectoral wage premium

In the primary sector there are two possible effort levels by the worker, namely $E = 0$ (no effort, the worker is "shirking" on the job) and $E = \bar{E}$ (the worker is expending full effort). In the secondary sector there is no cost of supervision and worker effort always equals $\bar{E}$. The cost of supervision in the primary sector is high and firms pay a wage premium (over and above the wage in the secondary sector) in order to induce worker effort $\bar{E}$. Only intermittently the firm engages in monitoring of its workers. Any worker caught shirking is fired and takes a job in the secondary sector during the current period.

The cost of exerting effort $\bar{E}$ (in monetary terms) equals P_E for the worker. Any worker employed in the primary sector faces a positive (and exogenous) probability, q, of being monitored by the firm (with $0 < q < 1$). The worker in the primary sector thus faces the following options: (a) exert effort $\bar{E}$ and obtain a net-of-effort wage $W_Y - P_E$ for sure, or (b) exert no effort ($E = 0$) and obtain W_Y with probability $1 - q$ and $W_X - P_E$ with probability q (shirking is impossible in the secondary sector so that the net-of-effort wage, $W_X - P_E$, is relevant). A primary sector worker exerts effort $\bar{E}$ provided option (a) is preferred to option (b), or:

$$\begin{aligned} W_Y - P_E \geq & \quad (1-q)\, W_Y + q\,(W_X - P_E) \quad \Leftrightarrow \\ W_Y \geq & \quad W_X + P_E \frac{1-q}{q}. \end{aligned} \tag{A5.212}$$

This is the so-called *no-shirking condition*. Firms in the primary sector know the effort-inducing condition (A5.212) and set W_Y such that the high-effort level, $\bar{E}$, is just guaranteed:

$$W_Y = W_X + P_E \frac{1-q}{q}. \tag{A5.213}$$

According to (A5.213), the effort-inducing wage premium is an increasing function of the pecuniary cost of effort, P_E, and a decreasing function of the monitoring probability, q.

Workers queue for jobs in the primary sector. In that sector there are V vacancies at the beginning of each period and there are U unemployed job seekers so that the probability of finding a job in the primary sector is equal to V/U. We assume that $U > V > 0$ so that $0 < V/U < 1$, i.e. the probability of finding a job is positive but less than unity. Workers are indifferent between certain secondary employment and joining the queue in the primary sector provided the expected net-of-effort reward is the same:

$$\begin{aligned} W_X - P_E &= \frac{V}{U}(W_Y - P_E) + \left[1 - \frac{V}{U}\right] \times 0 \quad \Leftrightarrow \\ W_X &= \frac{V}{U} W_Y + \left[1 - \frac{V}{U}\right] P_E. \end{aligned} \tag{A5.214}$$

The expression in (A5.214) governs the intersectoral allocation of labour.[21]

The process of job destruction and vacancy creation is as follows. At the start of each period a random fraction, sL_Y, of all primary sector jobs is destroyed and the same number of vacancies is created. This implies, of course, that as far as job destruction is concerned L_Y stays the same and that new vacancies are given by:

$$V = sL_Y, \tag{A5.215}$$

where s is the exogenous *job destruction rate*. The labour market "equilibrium" condition is given by:

$$\bar{L} = U + L_X + L_Y, \tag{A5.216}$$

where $\bar{L}$ is the total labour force, L_X and L_Y represent employment in, respectively the X-sector and the Y-sector, and U is unemployment. By using (A5.215) and (A5.216) we find:

$$\frac{V}{U} = \frac{sL_Y}{\bar{L} - (L_X + L_Y)}, \tag{A5.217}$$

where V/U is an index of labour market tightness. The higher is V/U, the easier it is for workers to find a job in the primary sector.

[21] Note that we abstract from unemployment benefits altogether, i.e. an unemployed worker receives nothing.

6.6.2 Remainder of the model

As was pointed out above, the model abstracts from factor substitutability, i.e. there are fixed input coefficients in both sectors. The Y-sector is relatively capital intensive and the cost functions in the two sectors are given by:

$$C^x = c^x (W_X, R) X, \tag{A5.218}$$
$$C^y = c^y (W_Y, R) Y, \tag{A5.219}$$

where W_X and W_Y are the wage rate in, respectively, the X-sector and the Y-sector, and R is the common rental rate on capital. The conditional factor demands are thus:

$$L_X = \frac{\partial c^x (W_X, R)}{\partial W_X} X = c^x_W X, \tag{A5.220}$$
$$K_X = \frac{\partial c^x (W_X, R)}{\partial R} X = c^x_R X, \tag{A5.221}$$
$$L_Y = \frac{\partial c^y (W_Y, R)}{\partial W_Y} Y = c^y_W Y, \tag{A5.222}$$
$$K_Y = \frac{\partial c^y (W_Y, R)}{\partial R} Y = c^y_R Y, \tag{A5.223}$$

where the c^i_j coefficients are fixed. Firms in both sectors operate under perfect competition and the pricing equations are thus based on marginal costs:

$$P_X \left[= c^x (W_X, R)\right] = c^x_W W_X + c^x_R R, \tag{A5.224}$$
$$P_Y \left[= c^y (W_Y, R)\right] = c^y_W W_Y + c^y_R R, \tag{A5.225}$$

where we have used the linear homogeneity of the unit cost functions to get from the first to the second equality.

The factor market clearing conditions are given by:

$$\bar{K} = c^x_R X + c^y_R Y, \tag{A5.226}$$
$$\bar{L} - U = c^x_W X + c^y_W Y, \tag{A5.227}$$

where (A5.226) is the capital market equilibrium condition and (A5.227) is the labour market equilibrium condition in the presence of non-zero unemployment. By using (A5.217) and (A5.222) we can relate unemployment to output in the primary sector and the labour market tightness variable:

$$\begin{aligned} U &= \bar{L} - (L_X + L_Y), \\ &= s \frac{U}{V} L_Y, \end{aligned}$$

$$= s\frac{U}{V}c_W^y Y, \tag{A5.228}$$

so that (A5.227) can be expressed as:

$$\bar{L} = c_W^x X + \left[1 + s\frac{U}{V}\right] c_W^y Y, \tag{A5.229}$$

If unemployment were zero, this expression would be a standard labour market equilibrium condition (like equation (6.14) above).

Households have homothetic preferences and the aggregate demands for the two goods are as given in the standard:

$$X = d^x (P_X, P_Y) M, \tag{A5.230}$$
$$Y = d^y (P_X, P_Y) M, \tag{A5.231}$$

where $d^x(\cdot)$ and $d^y(\cdot)$ are homogeneous of degree minus one in P_X and P_Y and M is aggregate income:

$$M = W_X L_X + W_Y L_Y + R\bar{K}. \tag{A5.232}$$

6.6.3 Payroll tax

In this subsection we study some of the properties of the model by looking at the general equilibrium effects of a payroll tax in the primary sector. This tax on the use of labour is denoted by t_{LY}. As usual, the revenue from the tax is recycled in a lump-sum fashion to households. The model consists of equations (A5.213)-(A5.214), (A5.224), (A5.226), and (A5.229)-(A5.231). The pricing equation in the Y-sector, (A5.225), is modified to:

$$P_Y = c_W^y W_Y (1 + t_{LY}) + c_R^y R. \tag{A5.233}$$

The payroll tax increases the cost of labour in the primary sector and thus increases the goods price in that sector. (It does not induce a capital-labour substitution effect because $\sigma_Y = 0$ by assumption.)

The household income definition (A5.232) becomes:

$$M = W_X c_W^x X + W_Y c_W^y Y + R\bar{K} + T, \tag{A5.234}$$

where T is lump-sum transfers from the government:

$$T \equiv t_{LY} W_Y c_W^y Y. \tag{A5.235}$$

The endogenous variables of the model X, Y, P_X, P_Y, W_X, W_Y, R, M, T and V/U. Exogenous are $\bar{K}$ and

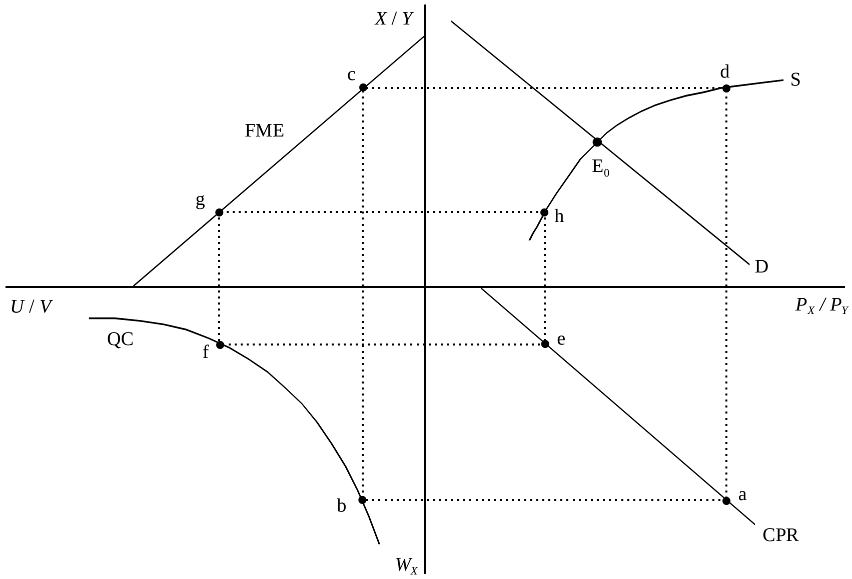

Figure 6.12: Harberger model with efficiency wages and unemployment

$\bar{L}$ and the constant parameters are c_W^x, c_W^y, c_R^x, c_R^y, P_E, q, and s. By the Law of Walras, one equation is redundant and we can only determine relative prices.

The model can be represented graphically by means of Figure 6.12. In the top right-hand panel, the D curve represents the relative demand for good X. It is obtained by dividing (A5.230) by (A5.231), and is downward sloping under the assumption that the substitution elasticity between X and Y in household utility is strictly positive ($\sigma_D > 0$). The FME curve in the top left-hand panel represents factor markets equilibrium in $(X/Y, U/V)$ space. Formally it is derived as follows. First, by solving (A5.226) and (A5.229) for X and Y we get:

$$X = \frac{c_R^y \bar{L} - \left(1 + s\frac{U}{V}\right) c_W^y \bar{K}}{c_R^y c_W^x - c_R^x \left(1 + s\frac{U}{V}\right) c_W^y}, \tag{A5.236}$$

$$Y = \frac{-c_R^x \bar{L} + c_W^x \bar{K}}{c_R^y c_W^x - c_R^x \left(1 + s\frac{U}{V}\right) c_W^y}, \tag{A5.237}$$

where the denominator is positive by virtue of the assumption regarding the relative capital intensity of the Y-sector. It follows from (A5.236)-(A5.237) that X/Y can be written as:

$$\frac{X}{Y} = \frac{c_R^y \bar{L} - \left(1 + s\frac{U}{V}\right) c_W^y \bar{K}}{c_W^x \bar{K} - c_R^x \bar{L}}, \tag{A5.238}$$

which is downward sloping and linear in U/V.[22]

The QC curve in the bottom left-hand panel of Figure 6.12 represents the combination of the no-shirking condition (A5.213) and the expression determining the intersectoral allocation of labour (A5.214)

[22] We obviously restrict attention to the case where both outputs are positive. It follows from (A5.236)-(A5.237) that the economy-

in $(U/V, W_X)$ space. Formally the QC curve is derived by substituting (A5.213) into (A5.214) and solving for W_X:

$$W_X = P_E\left[1+\frac{1-q}{q}\frac{1}{\frac{U}{V}-1}\right]. \tag{A5.239}$$

Equation (A5.239) shows that W_X is a hyperbolic downward sloping function of U/V (recall that $U/V > 1$ by assumption).

Finally, the CPR curve in the bottom right-hand panel of Figure 6.12 represents the combined competitive pricing equations (A5.224) and (A5.233) in $(P_X/P_Y, W_X)$ space, using the capital stock as the numeraire (i.e. $R = 1$). The formal expression of the CPR curve is obtained by combining (A5.213), (A5.224), and (A5.233) and setting $R = 1$:

$$\frac{P_X}{P_Y} = \frac{c_W^x W_X + c_R^x}{c_W^y\left(1+t_{LY}\right)\left[W_X + P_E\frac{1-q}{q}\right] + c_R^y}. \tag{A5.240}$$

The slope of the CPR curve is positive:

$$\frac{\partial\left(P_X/P_Y\right)}{\partial W_X} = \frac{P_Y c_W^x - P_X c_W^y\left(1+t_{LY}\right)}{P_Y^2} > 0, \tag{A5.241}$$

where the inequality follows from the assumption regarding the relative capital intensity of the Y-sector.[23] Starting at a given X/Y ratio and completing the boxes we derive the supply curve, S, in the top right-hand panel of Figure 6.12. The initial equilibrium is at point E_0.

The effects of an increase in the payroll tax are illustrated in Figure 6.13. It follows from (A5.240) that $\partial\left(P_X/P_Y\right)/\partial t_{LY} < 0$, i.e. an increase in the payroll tax shifts the CPR curve to the left, say from CPR_0 to CPR_1. As a result, the supply curve shifts up, from S_0 to S_1, and the equilibrium shifts from E_0 to E_1. As a result of the shock, X/Y rises, and P_X/P_Y, W_X, and U/V fall.

6.7 Punchlines

In this chapter we study the static theory of tax incidence. We start out by demonstrating the incidence of a consumption tax in the context of a partial equilibrium model. Depending on the slopes of the demand

wide capital-labour ratio must lie in the following interval:

$$\frac{c_R^x}{c_W^x} < \frac{\bar{K}}{\bar{L}} < \frac{c_R^y}{\left(1+s\frac{U}{V}\right)c_W^y}.$$

See also Atkinson (1994, p. 284).

[23] Indeed, the intensity condition is given by $\theta_{LX} \equiv W_X c_W^x/c^x > W_Y\left(1+t_{LY}\right)c_W^y/c^y \equiv \theta_{LY}$. By noting that $P_X = c^x$ and $P_Y = c^y$, we find:

$$P_Y c_W^x > P_X c_W^y\left(1+t_{LY}\right)\frac{W_Y}{W_X} > P_X c_W^y\left(1+t_{LY}\right),$$

where the last inequality follows from the fact that $W_Y > W_X$ (see equation (A5.213) above).

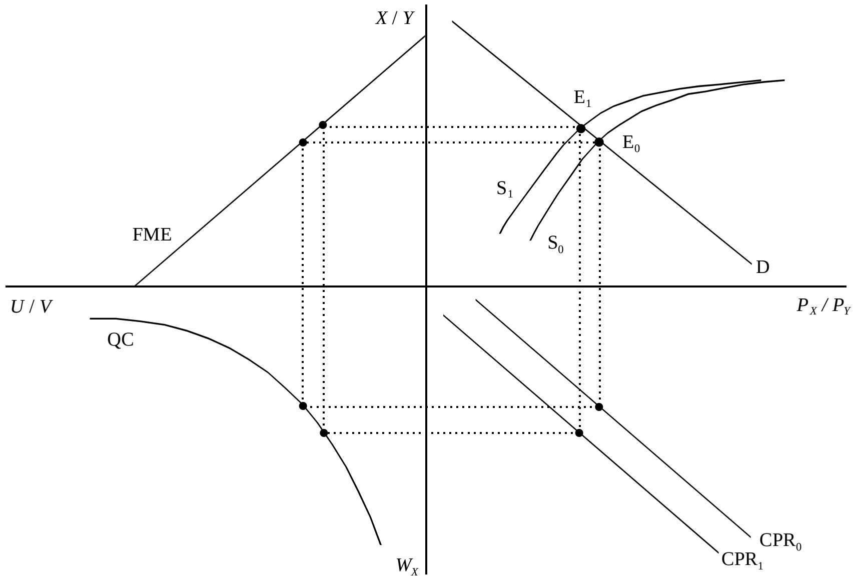

Figure 6.13: The payroll tax in an efficiency wage model

and supply curves, part of the consumption tax is shifted to the owners of the fixed factor (landowners in the example) and part of it is borne by consumers. The conditions under which the Marshallian approach to tax incidence is valid are very restrictive. In the two-factor case under consideration, for example, one factor must be totally elastic in supply and the supply curve of the other factor must be totally inelastic.

Next we turn to the classic model of Harberger to study tax incidence in a general equilibrium context. The most basic general equilibrium model one can consider features two commodities (goods X and Y) and two production factors (labour and capital). Technology features constant returns to scale, factors are perfectly mobile across sectors, prices are flexible, all markets attain equilibrium, and firms act as perfect competitors on input and output markets. In this "two-by-two" model, outputs, household income, factor demands, and the relative prices of goods and factors are all determined in general equilibrium.

In the third section we show how the comparative static effects of small changes in the exogenous variables can be computed by loglinearizing the two-by-two model around an initial equilibrium. By focussing on the case of homothetic household preferences, the model can be solved entirely in relative terms, i.e. in terms of relative output, X/Y, the relative goods price, P_X/P_Y, and the relative wage-rental rate, W/R. The advantage of this *differential approach* is that it allows the use of a very simple graphical apparatus to illustrate the effects.

In the fourth section we introduce a number of different taxes into the two-by-two model and demonstrate some famous equivalency results between them. Next, we study the general equilibrium tax effects of small changes in output and factor taxes. In general, a tax change produces both a *factor substitution effect* and an *output effect*. The former refers to a movement along a given isoquant, whereas the latter refers to a shift of the isoquant. For large tax changes, the solutions implied by the loglinearized model are potentially misleading. In such a setting it is more appropriate to base the tax incidence analysis directly on the nonlinear model. By choosing specific functional forms for preferences and technology, it is possible to construct a non-linear general equilibrium model which can be *calibrated* for an actual economy and subsequently simulated with the aid of a computer. In the calibration phase, the parameters of the Applied General Equilibrium (AGE) models are typically set at such levels that the theoretical model mimics the base case for an actual economy. We present a simple example of an AGE model and show that the approximation error of the loglinearized model gets larger, the larger is the shock that is administered.

In the last two sections of this chapter we demonstrate how the basic two-by-two model can be changed in a number of important aspects. In the fifth section it is assumed that one sector in the economy is characterized by *Chamberlinian monopolistic competition*. In that sector, many small firms produce "slightly unique" varieties of a differentiated product and consequently enjoy a small amount of market power. Technology features increasing returns to scale at the level of the firm, and under free entry/exit, the equilibrium size of each firm is determined endogenously. We use the monopolistic competition model to demonstrate the effects of capital taxes and subsidies on relative output, the wage-rental rate, and equilibrium firm size.

Finally, in the sixth section of this chapter, we shift attention from the goods market to the labour market. Using a *dual labour market* model, in which one sector pays *efficiency wages* to induce worker effort, we are able to model frictional unemployment as a general equilibrium phenomenon. The key properties of the efficiency wage model are demonstrated by computing the relative effects of a change in the payroll tax. The models discussed in the final two sections of this chapter demonstrate the feasibility of tax incidence analysis under non-standard assumptions.

Further reading

- Atkinson and Stiglitz (1980, lectures 6-7), Jha (1998, chs. 11-12), and Myles (1995, ch. 8) on theory. Surveys: Atkinson (1994), Fullerton and Metcalf (2002) on recent theory and empirics on tax incidence. Kotlikoff and Summers (1987) on theory and empirics on tax incidence.

- Classics: Marshall (1920), Pigou (1947), Harberger (1962), Jones (1965, 1971a, 1971b) Mieszkowski (1969), McLure (1975), McLure and Thirsk (1975), McLure et al. (1975), Diamond and McFadden (1974),

- Applied GE models: Shoven and Whalley (1972, 1984, 1992), Ballard, Shoven, Whalley (1985),

- Items from reading list Poterba: Bradford (1978), and Cutler (1988).

- Labour market: Pissarides (1998), Lockwood and Manning (1993), Heijdra and Van der Ploeg

(2002, chs. 7-9), Summers et al. (1993), Bovenberg and Van der Ploeg (1994), Hoel (1990), Koskela and Vilmunen (1996) Boone and Bovenberg (2002, 2004).

- Goods market: Atkinson (1994) and Myles (1995, ch. 11) on theory
- Corporate tax and sector definition: non-corporate firms produce same good as corporate firms, see Gravelle and Kotlikoff (1989).
- Ad hoc stability conditions: Neary (1978)
- Non-traded goods: Jones (1974).
- Tax analysis and oligopoly: Katz and Rosen (1985).
- Two-sector models (without taxes): Doeringer and Piore (1971), McDonald and Solow (1985), Bulow and Summers (1986).

Old stuff from old Section 6.5: The basic two-by-two model can easily be extended.

- factor supplies can be made endogenous
  - [static] endogenous labour supply [add leisure to household utility function]
  - [dynamic] labour supply, saving, and capital accumulation [studied in Chapter 8]
- intersectoral mobility assumption can be augmented:
  - [static] Mussa-Neary: labour mobile but capital sector-specific
  - [static] McLure: capital mobile but labour sector-specific
  - [dynamic] adjustment costs on capital and/or labour
- representative agent model can be replaced by heterogeneous agent model
  - now we can also study distribution issues [how do taxes affect different households etc.]
- open economy version of the model
- allow for imperfections on the goods and / or labour market
- using computers we can formulate, calibrate, and run simulations with highly detailed/complex *computable general equilibrium models*
  - sky is the limit
  - information on key elasticities shaky
  - scenario analyses

Chapter 7

Taxation and economic growth

The purpose of this chapter is to discuss the following topics:

- What is the effect of taxation on macroeconomic growth in general equilibrium?
- Exogenous growth models.
  - Solow-Swan model: ad hoc savings function.
  - Ramsey model: dynamic optimization under perfect foresight.
  - Extended Ramsey model: endogenous labour supply.
  - Brief aside on two-sector exogenous growth models.
- Endogenous growth models.
  - Capital fundamentalist models.
  - Human capital and growth.
  - R&D and growth.

7.1 Introduction

In this chapter we extend our discussion of general equilibrium tax incidence theory to a dynamic setting. Whereas the previous chapter restricted attention to a static setting with exogenous factor supplies, in this chapter both these limitations will be relaxed. The emphasis in this chapter will be on endogenizing the supply of physical capital by modelling the dynamic saving decisions by households. We thus enter the huge field of economic growth theory.[1]

[1] Readers in need of a more extensive introduction to the theories of economic growth discussed in this chapter are referred to Heijdra and van der Ploeg (2002, ch. 14). More advanced sources are Burmeister and Dobell (1970), Barro and Sala-i-Martin (1995), and Aghion and Howitt (1998).

Broadly put there have been two waves of interest in economic growth. The first wave, which occurred during the period 1955-1970, views capital accumulation as the so-called *engine of growth*. In Section 7.2 we present a selective overview of *exogenous growth* models with a particular emphasis on the effects of taxation. The term "exogenous growth model" was coined not long ago to distinguish these types of models from more recent "endogenous growth models" that were developed from the late 1980s onward. In exogenous growth models the *long-run* rate of economic growth is exogenously determined, i.e. it depends only on parameters such as population growth and technological change which are not themselves determined within the model (and thus cannot be affected by government policy). Of course, economic growth during transition to the long-run equilibrium may depend on policy parameters such as tax rates and the like.

The second wave of interest in growth theory has produced a huge body of literature on *endogenous growth*. This literature accepts the importance of physical capital to the growth process, but it identifies additional sources of growth, namely human capital accumulation, knowledge creation and transfer, and the development of new products and/or production processes. The second wave of theorizing typically gives rise to endogenous growth, i.e. the long-run growth rate depends on government policy variables. Section 7.3 presents an overview of tax effects in some of the key endogenous growth models.

7.2 Exogenous growth models

In this section we discuss tax effects in the context of exogenous growth models. In the first subsection we study the classic *Solow-Swan model*. This model combines a neoclassical technology with a Keynesian savings function (featuring a constant rate of saving out of current income) and derives the growth properties. In the second subsection we remove the Keynesian ad hoc savings function and model the dynamic optimization problem of the representative household under perfect foresight and with exogenous labour supply. The growth properties of this *Ramsey model* are investigated. Finally, in the third subsection we study the *extended Ramsey model* in which labour supply is also endogenous. This model is used to study the macroeconomic effects of the corporate tax.

7.2.1 Solow-Swan model

The modern theory of economic growth was initiated by Solow (1956) and Swan (1956) who reacted to an earlier "Keynesian" literature on growth theory by Harrod (1939, 1948) and Domar (1946, 1947). The key distinguishing feature between the Harrod-Domar approach and the Solow-Swan approach is the explicit recognition by the latter of the possibility of capital-labour substitutability in the aggregate production function.

In this section we first provide a brief overview of the basic Solow-Swan model. Aggregate output,

$Y(t)$, is produced according to the aggregate production function:

$$Y(t) = F(K(t), A(t) L(t)), \tag{A5.1}$$

where $K(t)$ is the aggregate stock of capital (machines, buildings, PCs, and the like), $L(t)$ is employment (measured in number of workers), $A(t)$ is the exogenous index of *labour-augmenting productivity* (workers get more productive as time evolves), and $N(t) \equiv A(t) L(t)$ is employment measured in *efficiency units*.

The following technological assumptions are made regarding the production function:

(P1) Constant returns to scale (CRTS), i.e. $F(\lambda K(t), \lambda N(t)) = \lambda F(K(t), N(t))$ for $\lambda > 0$, and $F = F_K K + F_N N$ (Euler's Theorem), where $F_K \equiv \partial F/\partial K$ and $F_N \equiv \partial F/\partial N$;

(P2) Positive but diminishing marginal products (F_K, $F_N > 0$ and F_{KK}, $F_{NN} < 0$), cooperative factors ($F_{KN} > 0$), and strict quasi-concavity ($F_{KK}F_{NN} - F_{KN}^2 > 0$), where $F_{KK} \equiv \partial^2 F/\partial K^2$, $F_{NN} \equiv \partial^2 F/\partial N^2$, and $F_{KN} \equiv \partial^2 F/\partial K \partial N = \partial^2 F/\partial N \partial K \equiv F_{NK}$;

(P3) *Inada conditions*: convenient curvature properties around the origin (with K or N close to zero) and in the limit (with K or N approaching infinity):

$$\lim_{K \to 0} F_K = \lim_{N \to 0} F_N = +\infty, \tag{A5.2}$$

$$\lim_{K \to \infty} F_K = \lim_{N \to \infty} F_N = 0. \tag{A5.3}$$

Assumptions (P1)-(P2) we also routinely adopted in Chapter 5 (see the expressions in (5.2) above), but assumption (P3) is new. It deals with the behaviour of marginal products in extreme cases. As is shown in subsection 7.3.1 below, assumption (P3) is not generally satisfied for all standard production functions. It therefore does not constitute an innocuous assumption at all.

Aggregate household saving is described by the following Keynesian savings function:

$$S(t) = sY(t), \qquad 0 < s < 1, \tag{A5.4}$$

where s is the constant (and exogenous) propensity to save, and $S(t)$ is total saving. In a closed economy and in the absence of government consumption and taxation, the following two identities hold for aggregate output:

$$Y(t) = C(t) + I(t), \tag{A5.5}$$

$$= C(t) + S(t), \tag{A5.6}$$

where $C(t)$ is household consumption and $I(t)$ is aggregate gross investment:

$$I(t) = \delta K(t) + \dot{K}(t). \tag{A5.7}$$

In this expression, $\delta K(t)$ is *replacement investment* and δ is the constant depreciation rate ($\delta > 0$). As usual, a dotted variable denotes that variable's time rate of change, i.e. $\dot{K}(t) \equiv dK(t)/dt$ denotes *net investment*.

It remains to specify the forcing equations for the exogenous variables in the model, i.e. for the number of workers ($L(t)$) and for the index of technological advance ($A(t)$). In the typical representation of the Solow-Swan model, both are assumed to grow at some constant exponential rate. In the absence of unemployment, labour supply equals the population which grows according to:

$$\frac{\dot{L}(t)}{L(t)} = n_L, \tag{A5.8}$$

where n_L is the constant population growth rate. Solving (A5.8) subject to the initial condition $L(0) = L_0$ we obtain the time path for labour supply:

$$L(t) = L_0 e^{n_L t}, \tag{A5.9}$$

where L_0 is the base-year population size. Similarly, labour-augmenting technology is exogenous and grows at the constant exponential rate n_A:

$$\frac{\dot{A}(t)}{A(t)} = n_A, \tag{A5.10}$$

$$A(t) = A_0 e^{n_A t}, \tag{A5.11}$$

where A_0 is the base-year technology level. By combining (A5.8) and (A5.11) we find both the path for labour in efficiency units and its growth rate:

$$N(t) = N_0 e^{(n_L + n_A)t}, \tag{A5.12}$$

$$\frac{\dot{N}(t)}{N(t)} \equiv \frac{\dot{A}(t)}{A(t)} + \frac{\dot{L}(t)}{L(t)} = n_A + n_L, \tag{A5.13}$$

where $N_0 \equiv A_0 L_0$.

In summary, the most basic version of the growth model is described by the following equations:

$$I(t) = S(t), \tag{A5.14}$$

$$S(t) = sY(t), \tag{A5.15}$$

$$I(t) = \delta K(t) + \dot{K}(t), \tag{A5.16}$$

$$Y(t) = F(K(t), N(t)), \tag{A5.17}$$

where (A5.14) follows from (A5.5) and (A5.6). The endogenous variables of the model are $I(t)$, $S(t)$, $K(t)$, and $Y(t)$, the exogenous variable is $N(t)$, and the parameters are s and δ. Unless both n_A and n_L are zero, it follows from (A5.12) that $N(t)$ will grow over time so that the system (A5.14)-(A5.17) will not possess a steady state in level terms. By measuring all variables relative to the path of efficiency units of labour, however, it is more likely (and indeed guaranteed, given the assumption made thus far) that a steady state will exist in transformed variables.

The *fundamental differential equation* for capital per efficiency unit of labour is given by:

$$\dot{k}(t) = sf(k(t)) - (\delta + n_L + n_A)k(t), \tag{A5.18}$$

where $k(t) \equiv K(t)/N(t)$, $y(t) \equiv Y(t)/N(t)$, and $f(k(t))$ is the intensive-form production function:

$$y(t) = f(k(t)) \equiv F(K(t)/N(t), 1). \tag{A5.19}$$

It is not difficult to show that the intensive-form production function has the following properties:

$$f'(k(t)) \equiv F_K(k(t), 1), \tag{A5.20}$$

$$f(k(t)) - k(t) f'(k(t)) = F_N(k(t), 1), \tag{A5.21}$$

$$f''(k(t)) \equiv N(t) F_{KK}(K(t), N(t)) = F_{KK}(k(t), 1). \tag{A5.22}$$

Intermezzo 7.1

Deriving equations (A5.18) and (A5.20)-(A5.22). To derive (A5.18) we substitute (A5.15)-(A5.17) into (A5.14) to obtain:

$$\delta K(t) + \dot{K}(t) = sF(K(t), N(t)). \tag{I.1}$$

Since $F(\cdot)$ features constant returns to scale we can write $F(K(t), N(t)) = N(t) f(K(t)/N(t), 1)$. Hence, (I.1) can be simplified to:

$$\frac{\dot{K}(t)}{N(t)} = sf(k(t)) - \delta k(t), \tag{I.2}$$

where we have used (A5.19) and the definition $k(t) \equiv K(t)/N(t)$. Clearly, the definition for $k(t)$ implies:

$$\dot{k}(t) = \frac{\dot{K}(t)}{N(t)} - k(t)\frac{\dot{N}(t)}{N(t)}. \tag{I.3}$$

By substituting (I.3) into (I.2) and noting (A5.13) we obtain (A5.18).

To derive (A5.20)-(A5.22), we first write $Y = NF(K/N, 1)$ and differentiate with respect to K:

$$(F_K \equiv)\,\frac{\partial Y}{\partial K} = NF_K\left(\frac{K}{N}, 1\right)\frac{1}{N} = F_K\left(\frac{K}{N}, 1\right). \tag{I.4}$$

Similarly, since $Y = Nf(k)$, we also have that:

$$\frac{\partial Y}{\partial K} = Nf'(k)\frac{1}{N} = f'(k). \tag{I.5}$$

Combining (I.4) and (I.5) we find $f'(k) = F_K\left(\frac{K}{N}, 1\right)$. We also find from (I.5) that:

$$(F_{KK} \equiv)\,\frac{\partial^2 Y}{\partial K^2} = f''(k)\frac{1}{N}. \tag{I.6}$$

For labour we find:

$$(F_N \equiv)\,\frac{\partial Y}{\partial N} = F\left(\frac{K}{N}, 1\right) + NF_K(.)\frac{-K}{N^2} = f(k) - f'(k)k. \tag{I.7}$$

To get the final equality in (A5.22) we note that $F_{KK}(\cdot)$ is homogeneous of degree minus one, i.e. by Euler's Theorem it can be written as:

$$NF_{KK}(K, N) = F_{KK}(k, 1). \tag{I.8}$$

****

In Figure 7.1 we illustrate the phase diagram for $k(t)$. In that figure, the straight line $(\delta + n_L + n_A)k(t)$ represents the amount of investment required to replace worn-out capital *and* to endow each efficiency unit of labour with the same amount of capital (recall from (A5.13) that the stock of efficiency units of labour grows at rate $n_L + n_A$). For a constant savings rate, s, the per capita saving curve $sf(k(t))$, has the same shape as the intensive-form production function. The Inada conditions say that $f(k(t))$ is vertical at the origin, is concave, and flattens out as more and more capital per efficiency unit of labour is accumulated. It follows that there is a unique (non-trivial) stable equilibrium at E_0. At point A saving exceeds required investment, $sf(k(t)) > (\delta + n_L + n_A)k(t)$, so that net investment is positive,

$\dot{k}(t) > 0$. The opposite holds at point B where $\dot{k}(t) < 0$.

In the steady state, the capital-effective-labour ratio is constant and equal to k^*. The following properties of the so-called *balanced growth path* associated with this steady state can be deduced. First, in the steady state we have $\dot{k}(t) = 0$ so that it follows from, respectively, (A5.18) and (A5.19) that:

$$k^* = \frac{sf(k^*)}{\delta + n_L + n_A}, \tag{A5.23}$$

$$y^* = f(k^*). \tag{A5.24}$$

Since $k^* \equiv (K(t)/N(t))^*$ and $y^* \equiv (Y(t)/N(t))^*$ are constant whilst $N(t)$ grows exponentially at rate $n_L + n_A$, it follows that $K(t)$ and $Y(t)$ must grow at the same rate as $N(t)$ along the balanced growth path:

$$\left(\frac{\dot{K}(t)}{K(t)}\right)^* = \left(\frac{\dot{Y}(t)}{Y(t)}\right)^* = \left(\frac{\dot{N}(t)}{N(t)}\right)^* = n_L + n_A. \tag{A5.25}$$

In the second step we note that $S(t) = I(t) = sY(t)$ and that s is constant, so that along the balanced growth path $S(t)$ and $I(t)$ must grow at the same rate as $Y(t)$:

$$\left(\frac{\dot{S}(t)}{S(t)}\right)^* = \left(\frac{\dot{I}(t)}{I(t)}\right)^* = \left(\frac{\dot{Y}(t)}{Y(t)}\right)^* = n_L + n_A. \tag{A5.26}$$

Finally, output per worker (i.e. labour productivity) is defined as $Y(t)/L(t)$ and it grows at the following rate along the balanced growth path:

$$\left(\frac{\dot{Y}(t)}{Y(t)}\right)^* - \left(\frac{\dot{L}(t)}{L(t)}\right)^* = n_A. \tag{A5.27}$$

7.2.1.1 Corporate tax and growth

It is clear from our discussion up to this point that in the Solow-Swan model the *long-run growth rate* in the economy is fully explained by exogenous factors, viz. the rate of population growth and the rate of labour-augmenting technological change. The most likely place for taxes to have any effect in the model at all is via the aggregate savings rate s. Up to this point it was simply asserted that s was exogenous and constant.[2] Suppose now that s depends negatively on the capital tax t_K, i.e. we write $s = s(t_K)$ and assume that $\partial s/\partial t_K < 0$. An increase in the capital tax, t_K, will then have the effects as illustrated in Figure 7.2. In that figure the economy is initially in the steady state at point E_0. An increase in the capital tax reduces the savings propensity from s_0 to s_1 and rotates the savings curve in a clockwise fashion from $s_0 f(k(t))$ to $s_1 f(k(t))$. At impact, both $K(t)$ and $N(t)$ and thus $k(t)$ are predetermined and the economy jumps from point E_0 to A. At that point, actual investment falls short of required

[2]In Chapter 3 intertemporal consumption and saving theories were discussed. It was shown that, in the absence of borrowing constraints, saving generally depends of lifetime wealth rather than on current income.

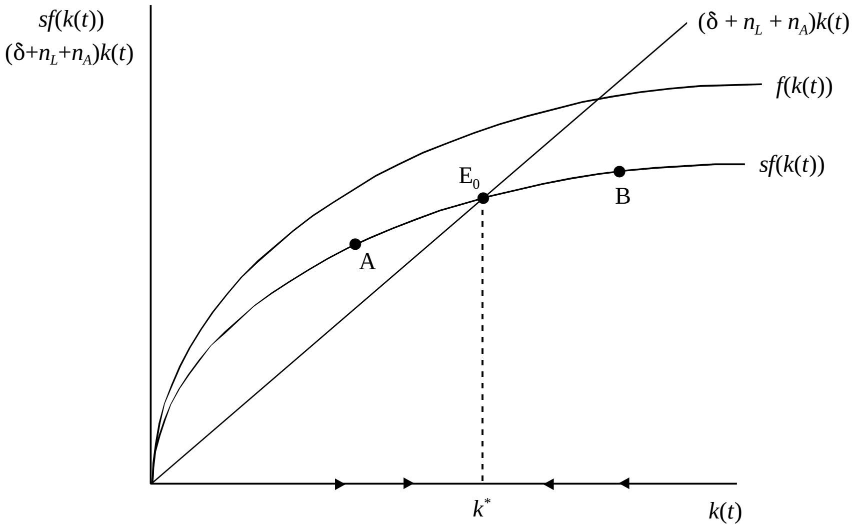

Figure 7.1: The Solow-Swan model

investment so net investment is negative, i.e. $\dot{k}(t) < 0$. Gradually over time the economy moves from point A to the new steady state at E_1.

During the transition period, economic growth is less than it is in the balanced growth path, i.e. the capital tax negatively affects economic growth during that time.[3] Not surprisingly, in view of (A5.25) and (A5.26), the long-run growth rate in the economy is not affected by the tax. In the Solow-Swan model this long-run growth rate is independent of the savings rate. Of course the long-run *levels* of y and k are affected by the capital tax as Figure 7.3 shows. At the time of the shock ($t = 0$) the economy is at point E_0 which lies on the balanced growth path $[\ln K^*(t)]_0$. This growth path is parallel to the path for $\ln N(t)$, the slope of which is $n_A + n_L$. The vertical difference between $[\ln K^*(t)]_0$ and $\ln N(t)$ is equal to $\ln k_0^*$, where k_0^* is illustrated in Figure 7.2. During transition, the economy moves from point E_0 in the direction of the new balanced growth path $[\ln K^*(t)]_1$ which is again parallel to the path for $\ln N(t)$.

7.2.2 Ramsey model

The Solow-Swan model is somewhat ill-equipped for studying tax incidence because one of the key variables in the model, viz. the savings rate, is exogenous. The objective of this subsection is to extend the Solow-Swan model by providing a theoretical foundation for intertemporal consumption and saving behaviour.[4] In doing so we develop the so-called *Ramsey growth model*, which derives its name from the mathematician-economist Frank Ramsey, who formulated a forward-looking theory of saving almost eighty years ago–see Ramsey (1928).

[3]Since $k(t)$ falls during transition, it follows from (A5.19) that $y(t)$ also falls during that time, i.e. $\dot{y}(t)/y(t) < 0$. Since $y(t) \equiv Y(t)/N(t)$ it is easy to show that $\dot{Y}(t)/Y(t) < n_A + n_L$ during transition.

[4]See Chapter 3 for the two-period consumption saving model and Section 4.1 of Chapter 5 for its multiperiod generalization.

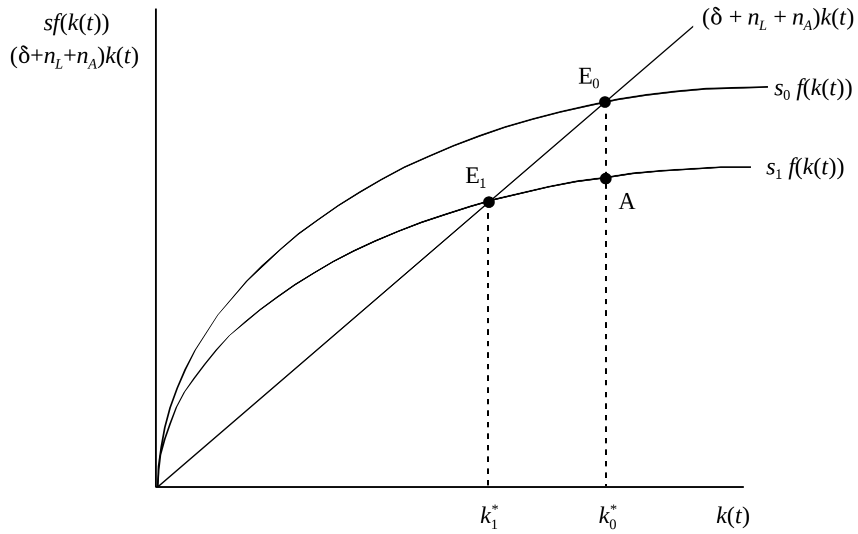

Figure 7.2: Fall in the Ssavings rate in the Solow-Swan model

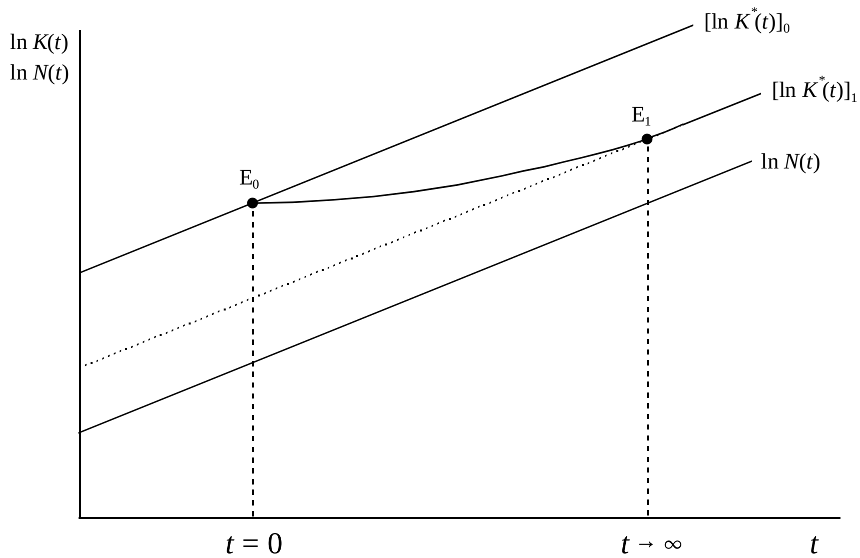

Figure 7.3: Transitional dynamics

7.2.2.1 Representative household

To keep the model as simple as possible, we abstract from exogenous sources of economic growth by assuming that there is no population growth ($n_L = 0$) and no labour-augmenting technological progress ($n_A = 0$). As in Chapter 5, we develop the model by postulating the existence of a representative household who is infinitely lived, enjoys perfect foresight, and supplies an exogenous amount of labour. The household's objective function is:

$$\Lambda(0) \equiv \int_0^\infty U(C(t))\, e^{-\rho t} dt, \tag{A5.28}$$

where $\Lambda(0)$ is life-time utility, $U(\cdot)$ is the *felicity function* (exhibiting $U'(\cdot) > 0 > U''(\cdot)$), $C(t)$ is (the flow of) consumption of the household, and ρ is the pure rate of time preference ($\rho > 0$).

Portfolio investment opportunities of the household consist of purchasing shares in existing firms or buying (short-period) government bonds. There is no risk so that shares and bonds are perfect substitutes in the portfolio. The budget identity is given by:

$$\dot{B}(t) + P_E(t)\,\dot{E}(t) + C(t) = (1 - t_L)\,W(t)\,\bar{L} + (1 - t_R)\,r(t)\,B(t) + Z(t)\,, \tag{A5.29}$$

where $B(t)$ is the stock of government debt, $P_E(t)$ is the market price of shares, $E(t)$ is the outstanding stock of equities, t_L is the tax on wage income, $W(t)$ is the wage rate, $\bar{L}$ is labour supply, t_R is the tax on interest income, $r(t)$ is the interest rate on government bonds, and $Z(t)$ is the lump-sum transfer received from the government. As usual we have $\dot{B}(t) \equiv dB(t)/dt$, and $\dot{E}(t) \equiv dE(t)/dt$. Compared to the model discussed in Chapter 5, we abstract from capital gains taxation ($t_G = 0$) and the firm pays no dividends ($D(t) = 0$).

The household chooses paths for $C(t)$, $B(t)$, and $E(t)$ in order to maximize (A5.28) subject to (A5.29) and some transversality conditions. In addition the household faces some initial conditions, i.e. $E(0)$ and $B(0)$ are predetermined at time $t = 0$. The first-order conditions for the household's optimization problem are:[5]

$$U'(C(t)) = \lambda(t)\,, \tag{A5.30}$$

$$(1 - t_R)\,r(t) = \frac{\dot{P}_E(t)}{P_E(t)}, \tag{A5.31}$$

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = \rho - (1 - t_R)\,r(t)\,, \tag{A5.32}$$

where $\lambda(t)$ is the co-state variable associated with aggregate financial wealth. Equation (A5.30) is the (implicit) Frisch demand for consumption (relating $C(t)$ to the marginal utility of wealth $\lambda(t)$), (A5.31) is the no-arbitrage equation between government bonds and equities (with the returns on equity con-

[5]The details of the derivation are explained in Intermezzo 5.4 above.

sisting solely of capital gains), and (A5.32) describes the optimal time path for $\lambda(t)$.

7.2.2.2 Representative firm

Our description of the representative firm is very similar to the one presented in Chapter 5. For convenience we quickly re-derive firm behaviour. We continue to abstract from corporate debt and simplify the model further by assuming that the firm pays no dividends and that there are no adjustment costs of firm investment. Gross profit, $\Pi(t)$, is defined as:

$$\Pi(t) \equiv F(K(t), L(t)) - W(t) L(t), \tag{A5.33}$$

where $K(t)$ is the capital stock and $L(t)$ is labour demand. Gross profit forms the tax base for the corporate income tax and retained earnings are equal to after-tax profit:

$$(1 - t_K) \Pi(t) = RE(t), \tag{A5.34}$$

where t_K is the corporate tax, $RE(t)$ is retained earnings, and we have incorporated $D(t) = 0$ (no dividend payments).

There are no adjustment costs of investment so the capital accumulation identity is as given in (A5.7). The financing constraint of the firm is thus:

$$RE(t) + P_E(t) \dot{E}(t) = I(t), \tag{A5.35}$$

where $\dot{E}(t)$ is the sale of new equities and $I(t)$ is gross investment. By combining (A5.34) and (A5.35) (under the assumption that $RE(t) > 0$) we obtain:

$$P_E(t) \dot{E}(t) = I(t) - (1 - t_K) \Pi(t). \tag{A5.36}$$

According to (A5.36), if investment exceeds after-tax profit, then the firm issues new shares ($\dot{E}(t) > 0$). In the opposite case the firm buys back its own shares ($\dot{E}(t) < 0$).

The market value of outstanding shares is:

$$V(t) = P_E(t) E(t). \tag{A5.37}$$

Using (A5.37) and the household no-arbitrage equation (A5.31) we find the differential equation for $V(t)$:[6]

$$\dot{V}(t) = (1 - t_R) r(t) V(t) - [(1 - t_K) \Pi(t) - I(t)]. \tag{A5.38}$$

[6] Differentiating (A5.37) we get:

$$\dot{V}(t) = \dot{P}_E(t) E(t) + P_E(t) \dot{E}(t).$$

The only economically sensible (no-bubble) solution for $V(0)$ is obtained in the usual manner by solving (A5.38) forward in time and imposing a terminal condition:[7]

$$V(0) = \int_0^\infty \left[(1-t_K)\,\Pi(t) - I(t)\right] \exp\left[-\int_0^t \theta(\tau)\,d\tau\right] dt, \tag{A5.39}$$

where $\theta(\tau)$ is the cost of capital to the firm:

$$\theta(\tau) \equiv r(\tau)\,(1 - t_R(\tau)). \tag{A5.40}$$

Equation (A5.39) shows that the fundamental value of the firm is equal to the present value of after-corporate-tax cash flows, using the cost of capital for discounting purposes. In the absence of dividend payments, dividend taxes, and capital gains taxes, the cost of capital is equal to the after–tax interest rate on government bonds–see (A5.40).

The firm maximizes its stockmarket value (A5.39) subject to the capital accumulation constraint (A5.7). Because there are no adjustment costs on firm investment, it follows that the firm can vary its desired capital stock at will. Indeed, by substituting (A5.7) into (A5.39) and integrating we find that the objective function for the firm can be written as:

$$\begin{aligned} V(0) &= \int_0^\infty \left[(1-t_K)\,\Pi(t) - \delta K(t) - \dot{K}(t)\right] \exp\left[-\int_0^t \theta(\tau)\,d\tau\right] dt, \\ &= \int_0^\infty \Big[(1-t_K)\,\Pi(t) - R(t)\,K(t) - \dot{K}(t) + (R(t)-\delta)\,K(t)\Big] \exp\left[-\int_0^t \theta(\tau)\,d\tau\right] dt \\ &= K(0) + \int_0^\infty \Big[(1-t_K)\,\Pi(t) - R(t)\,K(t)\Big] \exp\left[-\int_0^t \theta(\tau)\,d\tau\right] dt, \end{aligned} \tag{A5.41}$$

where $K(0)$ is the initial capital stock[8] and $R(t)$ is the *rental rate on capital*:

$$R(t) \equiv r(t)\,(1 - t_R(t)) + \delta. \tag{A5.42}$$

Equation (A5.41) has two important implications. First, it shows that the firm's decision about factor

---

Using (A5.37), the household no-arbitrage equation (A5.31) can be written as:

$$(1-t_R)\,r(t)\,V(t) = \dot{P}_E(t)\,E(t).$$

By combining these expressions with (A5.36) we obtain (A5.38).

[7] As in Chapter 5, this terminal condition is:

$$\lim_{t\to\infty} V(t) \exp\left[-\int_0^t \theta(\tau)\,d\tau\right] = 0.$$

[8] In going from the second to the third line in (A5.41) we use the fact that:

$$\begin{aligned} &\int_0^\infty \Big[\dot{K}(t) - r(t)\,(1-t_R(t))\,K(t)\Big] \exp\left[-\int_0^t \theta(\tau)\,d\tau\right] dt \\ &= \int_0^\infty d\left[K(t) \exp\left[-\int_0^t \theta(\tau)\,d\tau\right]\right] = -K(0), \end{aligned}$$

where we have used the fact that $\lim_{K(t)\to\infty} K(t) \exp\left[-\int_0^t \theta(\tau)\,d\tau\right] = 0$ in the final step.

inputs is essentially a static one, i.e. maximization of $V(0)$ by choice of $L(t)$ and $K(t)$ yields the familiar marginal productivity conditions for labour and capital:

$$W(t) = F_L(K(t), L(t)), \tag{A5.43}$$

$$R(t) = (1 - t_K) F_K(K(t), L(t)). \tag{A5.44}$$

Equation (A5.43) is the standard labour demand equation. Since the wage bill is exempt from the corporate tax, labour demand is not directly affected by it. Equation (A5.44) is the demand for physical capital by the firm. The decision rule calls for an equalization of the rental rate on capital and the after-corporate-tax marginal product of labour.

The second implication of (A5.41) is that the maximized value of $V(0)$ equals the firm's capital stock, i.e. $V(0) = K(0)$. This result follows from the fact that the production function exhibits constant returns to scale, so that factor payments exhaust output, i.e. the integral appearing on the right-hand side of (A5.41) is zero.

Two further expressions suffice to close the model. First, in the absence of government consumption the goods market clearing condition can be written as follows:

$$Y(t) = C(t) + \delta K(t) + \dot{K}(t). \tag{A5.45}$$

Formally, (A5.45) is obtained by substituting (A5.7) into (A5.5). Second, the labour market equilibrium condition is:

$$L(t) = \bar{L}, \tag{A5.46}$$

where $L(t)$ is labour demand and $\bar{L}$ is (exogenous) labour supply.

7.2.2.3 Summary of the model

By gathering the various expressions, the full Ramsey growth model can be summarized in the following compact format:

$$U'(C(t)) = \lambda(t), \tag{A5.47}$$

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = \rho - [(1 - t_K) F_K(K(t), \bar{L}) - \delta], \tag{A5.48}$$

$$\dot{K}(t) = F(K(t), \bar{L}) - C(t) - \delta K(t). \tag{A5.49}$$

Equation (A5.47) is simply the Frisch demand for consumption restated (see also (A5.30) above), (A5.48) is the household Euler equation expressed in terms of $\lambda(t)$ (instead of $C(t)$). It is obtained by using (A5.32), (A5.42), and (A5.44) and imposing labour market equilibrium (A5.46). The term in square brackets on the right-hand side is the after-tax net marginal product of capital. Finally, equation (A5.49) is the goods market clearing condition in the absence of government consumption. It is obtained by using (A5.45), substituting the production function, and imposing labour market equilibrium (A5.46). The endogenous variables in the model are C, λ, and K, the exogenous variable is the corporate tax rate, t_K, and the parameters are ρ and δ.

In Figure 7.4 we present the phase diagram for the Ramsey model in $(\lambda(t), K(t))$ space. The $\dot{\lambda}(t) = 0$ line is obtained from (A5.48) and defines a unique capital-labour ratio, $K^{KR}/\bar{L}$, where the superscript KR stands for "Keynes-Ramsey":

$$\dot{\lambda}(t) = 0 \quad \Leftrightarrow \quad F_K\left(\frac{K^{KR}}{\bar{L}}, 1\right) = \frac{\rho + \delta}{1 - t_K}. \tag{A5.50}$$

For points to the left (right) of the $\dot{\lambda}(t) = 0$ line, $K(t)$ is too low (too high), $F_K(K(t), \bar{L})$ is too high (too low), and $\lambda(t)$ decreases (increases) over time, i.e. $\dot{\lambda}(t) < 0$ (> 0). This is indicated with the vertical arrows in Figure 7.4.

The $\dot{K}(t) = 0$ line is obtained by setting $\dot{K}(t) = 0$ in (A5.49) and substituting the resulting expression for $C(t)$ into (A5.47):

$$\dot{K}(t) = 0 \quad \Leftrightarrow \quad U'(F(K(t), \bar{L}) - \delta K(t)) = \lambda(t). \tag{A5.51}$$

It follows from (A5.51) that the slope of the $K(t) = 0$ line is equal to:

$$\left(\frac{d\lambda(t)}{dK(t)}\right)_{\dot{K}(t)=0} = U''(\cdot)\left[F_K\left(\frac{K(t)}{\bar{L}}, 1\right) - \delta\right]. \tag{A5.52}$$

In view of the assumption that $U''(\cdot) < 0$, we derive from (A5.52) that:

$$\left(\frac{d\lambda(t)}{dK(t)}\right)_{\dot{K}(t)=0} \begin{array}{c}<\\=\\>\end{array} 0 \quad \Leftrightarrow \quad F_K\left(\frac{K(t)}{\bar{L}}, 1\right) \begin{array}{c}>\\=\\<\end{array} \delta. \tag{A5.53}$$

We define the *golden-rule* capital-labour ratio by that point along the $\dot{K}(t) = 0$ line for which consumption is at its maximum. We find in a straightforward fashion from (A5.49) that this golden rule capital labour ratio is defined implicitly by:

$$F_K\left(\frac{K^{GR}}{\bar{L}}, 1\right) = \delta, \tag{A5.54}$$

where the superscript GR stands for "golden rule". Comparing (A5.50) and (A5.54), it is clear that K^{GR} exceeds K^{KR} because $(\rho + \delta) / (1 - t_K) > \delta$ and the marginal product of capital declines with the amount of capital ($F_{KK} < 0$, so that a high value for K implies a low value for F_K). Given the definition of

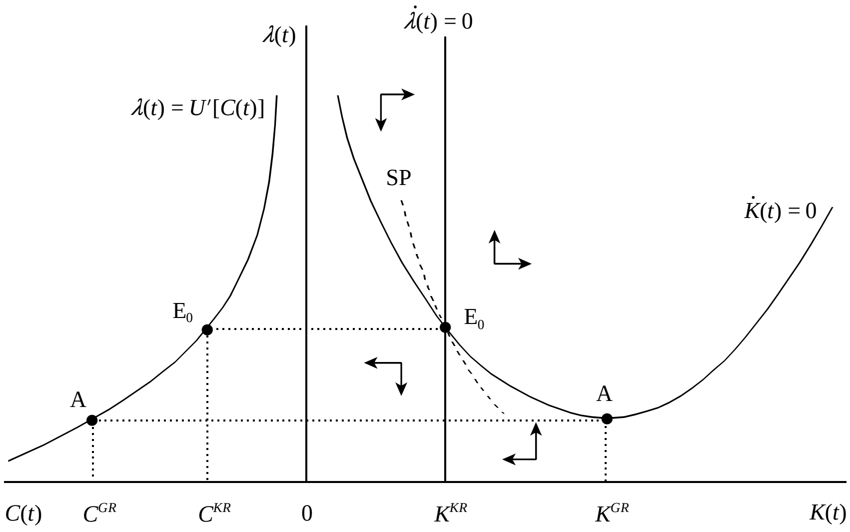

Figure 7.4: The Ramsey growth model

the golden-rule capital-labour ratio, it follows from (A5.53) that the $\dot{K}(t) = 0$ line is downward (upward) sloping for values of $K(t)$ less than (greater than) than K^{GR}.

The dynamic adjustment in the capital stock can be deduced from (A5.49). By differentiating this expression with respect to $\lambda(t)$ we find:

$$\frac{\partial \dot{K}(t)}{\partial \lambda(t)} = -\frac{dC(t)}{d\lambda(t)} = -\frac{1}{U''(\cdot)} > 0, \tag{A5.55}$$

i.e. for points above (below) the $\dot{K}(t) = 0$ line, $\lambda(t)$ is too high (too low), $C(t)$ is too low (too high), and the capital stock increases (decreases) over time, i.e. $\dot{K}(t) > 0$ (< 0). See the horizontal arrows in Figure A5.4.

Given the configuration of horizontal and vertical arrows, it is clear that there is a downward sloping saddle path, SP, which passes through the unique equilibrium at point E_0 in the right-hand panel of Figure 7.4. The downward sloping relationship between $\lambda(t)$ and $C(t)$, defined by the Frisch demand (A5.47), is plotted in the left-hand panel of Figure 7.4.

The growth properties of the Ramsey model are as follows. In the steady state, the capital-labour ratio is constant and equal to $K^{KR}/\bar{L}$. Since labour supply is constant, the capital stock itself also settles at a constant value in the steady state. With both factors constant, it follows that all other variables are also constant in the steady state. The Ramsey model is thus an example of an exogenous growth model because all variables grow at the same exogenous rate along the balanced growth path (here this rate is zero because we have set $n_L = n_A = 0$).

7.2.2.4 Corporate tax again

We can now re-examine the effects of the capital tax t_K. Figure 7.5 illustrates the effects of an unanticipated and permanent increase in the corporate tax t_K, at impact, during the transition phase, and in the long-run. It is assumed that the economy is initially in the steady-state equilibrium at E_0. By totally differentiating (A5.50) with respect to the golden-rule capital stock and the corporate tax we find:

$$\begin{aligned} F_{KK}\left(\frac{K^{KR}}{\bar{L}},1\right)\frac{dK^{KR}}{\bar{L}} &= \frac{\rho+\delta}{(1-t_K)^2}dt_K \qquad \Rightarrow \\ \frac{dK^{KR}}{dt_K} &= \frac{(\rho+\delta)\,\bar{L}}{(1-t_K)^2\,F_{KK}\left(\frac{K^{KR}}{\bar{L}},1\right)} < 0, \end{aligned} \tag{A5.56}$$

where the sign follows from the fact that $F_{KK}(\cdot) < 0$. The increase in the corporate tax, reduces the Keynes-Ramsey capital stock and thus shifts the $\dot{\lambda}(t) = 0$ line to the left. Clearly, since t_K does not feature in (A5.51) there is no effect on the $\dot{K}(t) = 0$ line. In Figure 7.5, the long-run equilibrium shifts from point E_0 to E_1, the capital stock and consumption both decline and the marginal utility of wealth increases.

The transitional dynamics is as follows. In the right-hand panel of Figure 7.5, the immediate effect consists of a jump from E_0 to A. At impact the capital stock is predetermined and the economy jumps onto the only stable trajectory leading to the new steady-state equilibrium. The impact reduction in the marginal utility of wealth is associated with an increase in consumption as is illustrated in the left-hand panel of the figure.

At point A in the right-hand panel, actual investment falls short of required investment so $\dot{K}(t) < 0$ and the after-tax net marginal product of capital falls short of the rate of time preference, i.e. $(1-t_K)\,F_K(K(t),\bar{L}) - \delta < \rho$ so that $\dot{\lambda}(t) > 0$. There is a gradual move in north-westerly direction (along the saddle path) to the new steady-state at E_1. Growth is less than in the balanced growth path during transition, as capital and output both fall during that time (negative growth). As in the Solow-Swan model, the long-run growth rate is unaffected by the capital tax, i.e. it remains zero.

An interesting implication of this dynamic model is that the incidence of the corporate tax differs in the short-run and the long-run. By imposing labour market equilibrium (A5.46) in (A5.43)-(A5.44) we find the following expressions for factor prices:

$$W(t) = F_L(K(t),\bar{L}), \tag{A5.57}$$

$$R(t) = (1-t_K)\,F_K(K(t),\bar{L}). \tag{A5.58}$$

In the short run, the capital stock is fixed so it follows from (A5.57) that the wage rate is unchanged and

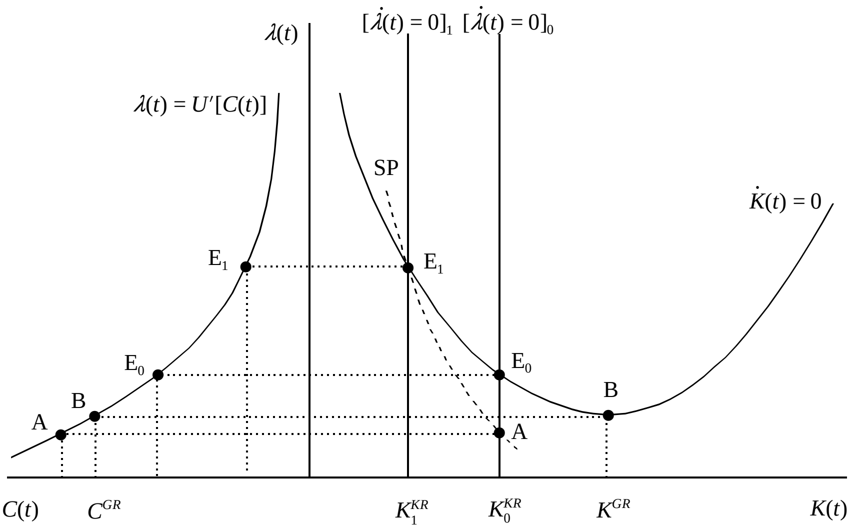

Figure 7.5: A rise in the corporate tax in the Ramsey model

from (A5.58) that the rental rate on capital falls:

$$\frac{dW(0)}{dt_K} = 0, \qquad \frac{dR(0)}{dt_K} = -F_K\left(K^{KR}, \bar{L}\right) < 0. \tag{A5.59}$$

Hence, in the impact period capital owners bear the full burden of the corporate tax.

In the long run, however, the capital stock is crowded out, and the after-tax reward to capital owners is restored, i.e. $R(\infty) = \rho + \delta$. Since capital and labour are cooperative production factors ($F_{LK} > 0$), it follows from (A5.57) that the wage rate falls in the long run:

$$\frac{dW(\infty)}{dt_K} = F_{LK}\left(K^{KR}, \bar{L}\right)\frac{dK^{KR}(\infty)}{dt_K} < 0, \qquad \frac{dR(t)}{dt_K} = 0. \tag{A5.60}$$

Hence, labour bears the full burden of the corporate tax in the long run!

7.2.3 Extended Ramsey model

One of the shortcomings of the standard Ramsey model (at least as a tool for tax policy analysis) is the lack of substitution possibilities it incorporates. With exogenous labour supply and a single-good economy, the only substitutability that exists in the model concerns intertemporal substitution of consumption. This limits the kinds of distortions that can be distinguished with the model. There are at least two ways to make the model more suitable for tax policy analysis, one of which is pursued in this subsection. First, one might forge the link with the Harberger-Jones model studied in Chapter 6 by postulating the existence of two goods produced in separate sectors of the economy. In the context of the growth model, one sector could be producing the investment good and the other sector the consumption good. This approach was taken by Uzawa (1961, 1963, 1964). A modern application of the Uzawa approach is studied in Section 7.3.1 below.

Here we pursue the second way to make the model more suitable for tax policy analysis: we retain the single-good assumption but endogenize the labour supply decision by households.[9] The objective function of the representative household is modified to:

$$\Lambda(0) \equiv \int_0^\infty U(C(t), 1 - L(t))\, e^{-\rho t} dt, \tag{A5.61}$$

where $L(t)$ is labour supply and $1 - L(t)$ is leisure (the household's time endowment is unity). Compared to (A5.28), the only thing that has changed is the inclusion of leisure into the felicity function, $U(\cdot)$. We assume that $U(\cdot)$ is a strictly quasi-concave function in C and $1 - L$, i.e. $U_C > 0$, $U_{1-L} > 0$, $U_{CC} < 0$, $U_{1-L,1-L} < 0$, and $U_{CC}U_{1-L,1-L} - (U_{C,1-L})^2 > 0$ (see Silberberg and Suen, 2001, pp. 140, 260). This implies that indifference curves bulge toward the origin. [REFER TO CHAPTER 2]

The budget identity of the representative household is modified to recognize the endogeneity of labour supply:

$$\dot{B}(t) + s(t)\dot{E}(t) + C(t) = (1 - t_L) W(t) L(t) + (1 - t_R) r(t) B(t) + Z(t), \tag{A5.62}$$

where wage income is now $W(t) L(t)$ (rather than $W(t)\bar{L}$, as in (A5.29) above) and t_L is the labour income tax.

The household chooses paths for $C(t)$ and $L(t)$ (the real decisions) and for $B(t)$ and $E(t)$ (the portfolio decisions) in order to maximize (A5.61) subject to (A5.62) and some transversality conditions, and taking as given the initial conditions regarding $E(0)$ and $B(0)$ (which are predetermined at time $t = 0$). The key first-order conditions for this optimization problem are given by (A5.31)-(A5.32) and:

$$U_C(C(t), 1 - L(t)) = \lambda(t), \tag{A5.63}$$

$$U_{1-L}(C(t), 1 - L(t)) = \lambda(t)(1 - t_L) W(t), \tag{A5.64}$$

where $\lambda(t)$ is the co-state variable associated with aggregate financial wealth, i.e. the marginal utility of wealth. Equations (A5.63)-(A5.64) implicitly define the Frisch demands for consumption and leisure. Compared to the standard Ramsey model, (A5.63) differs from (A5.30) in that the marginal utility of consumption may depend on leisure (if $U_{C,1-L} \neq 0$ this is indeed the case). Equation (A5.64) is new compared to the standard Ramsey model.

Note that the labour supply model gives rise to the same first-order condition as was obtained in Chapter 2 for the static model. Indeed, by dividing (A5.64) by (A5.63) we can eliminate $\lambda(t)$ and obtain:

$$\frac{U_{1-L}(C(t), 1 - L(t))}{U_C(C(t), 1 - L(t))} = (1 - t_L) W(t), \tag{A5.65}$$

[9] For background on the extended Ramsey model, see Heijdra and van der Ploeg (2002, pp. 478-483). Judd (1987) is the classic paper dealing with tax policy analysis in the extended Ramsey model.

which says that at each instant in time the household equates the marginal rate of substitution between consumption and leisure to the after-tax wage rate. This is exactly the same condition we found for the static model, see equation (2.7) above.

The rest of the model is unchanged compared to the standard Ramsey model. From the firm side, factor demands are still given by (A5.43)-(A5.44) and the capital accumulation constraint is still given by (A5.7). In the absence of government consumption, aggregate output is still the sum of consumption and gross investment (as in (A5.5) above) and tax revenues are equal to:

$$Z(t) = t_K \left[F(K(t), L(t)) - W(t) L(t)\right] + t_L W(t) L(t). \tag{A5.66}$$

7.2.3.1 Compact summary of the model

The extended Ramsey model can be written in a very compact format as follows:

$$C(t) = c(\lambda(t), K(t), t_L), \tag{A5.67}$$

$$L(t) = l(\lambda(t), K(t), t_L), \tag{A5.68}$$

$$\dot{K}(t) = L(t) f\left(\frac{K(t)}{L(t)}\right) - C(t) - \delta K(t), \tag{A5.69}$$

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = \rho + \delta - (1 - t_K) f'\left(\frac{K(t)}{L(t)}\right). \tag{A5.70}$$

Equation (A5.67) and (A5.68) are the expressions for, respectively, consumption and labour supply, conditional on the state variables ($\lambda(t)$ and $K(t)$) and the labour income tax rate (t_L). They are obtained by noting that (A5.63), (A5.64) and (A5.43) define implicit functions ($c(\cdot)$ and $l(\cdot)$) for these variables (see the Intermezzo). The capital stock enters these functions because the wage rate depends on it. Equation (A5.69) is the capital accumulation expression obtained by combining (A5.5) and (A5.7) and noting that $Y(t) = L(t) f(K(t)/L(t))$ where $f(K(t)/L(t)) \equiv F(K(t)/L(t), 1)$. Finally, (A5.70) is the usual expression for the marginal utility of wealth which is obtained by combining (A5.32), (A5.42) and (A5.44) and noting that $F_K(\cdot) = f'(\cdot)$. The endogenous variables of the model are $C(t)$, $L(t)$, $\lambda(t)$, and $K(t)$, the exogenous variables are t_K and t_L, and the parameters are ρ and δ.

Several things are worth noting about the extended Ramsey model. First, its growth properties are as follows. In the steady state, we have $\dot{\lambda}(t) = \dot{K}(t) = 0$, and (A5.70) defines a unique capital-labour ratio ($k^{KR} \equiv (K(t)/L(t))^{KR}$), where the superscript KR again stands for "Keynes-Ramsey". Using this value for k^{KR}, equations (A5.67)-(A5.69) can then be used to obtain the steady-state solutions for λ, C, and L. There is no long-run growth in the model (because the population is constant) and taxes therefore do not affect long-run growth either. In that sense the extended Ramsey model does not differ significantly from the standard Ramsey model.

The second noteworthy feature of the model concerns its implications for tax incidence. Judd (1987) uses the most general version of the model for tax policy analysis. Such an analysis is rather complex because of the plethora of different elasticities affecting the implicit functions $c(\cdot)$ and $l(\cdot)$. Instead of working with the general model, here we study a "Mickey Mouse" version of it in which all important substitution elasticities are set equal to unity.

Intermezzo 7.2

Frisch consumption demand and labour supply. The Frisch consumption demand (A5.67) and labour supply (A5.68) are derived as follows. We drop the time index for convenience. By using (A5.43), (A5.63), and (A5.64) we find an implicit system of equations relating C and $1-L$ to λ, t_L, and K:

$$U_C(C, 1-L) = \lambda, \tag{A}$$

$$U_{1-L}(C, 1-L) = \lambda(1-t_L) F_L(K, L). \tag{B}$$

We postulate the existence of implicit functions $c(\cdot)$ and $l(\cdot)$ and wish to determine their partial derivatives with respect to their arguments. By totally differentiating (A)-(B) we find:

$$\begin{aligned} \Delta \begin{bmatrix} dC \\ dL \end{bmatrix} &= \begin{bmatrix} 1 \\ -(1-t_L) F_L \end{bmatrix} d\lambda + \begin{bmatrix} 0 \\ \lambda F_L \end{bmatrix} dt_L \\ &\quad - \begin{bmatrix} 0 \\ \lambda(1-t_L) F_{KL} \end{bmatrix} dK, \end{aligned} \tag{C}$$

where Δ is defined as follows:

$$\Delta \equiv \begin{bmatrix} U_{CC} & -U_{C,1-L} \\ -U_{C,1-L} & U_{1-L,1-L} + \lambda(1-t_L) F_{LL} \end{bmatrix}, \tag{D}$$

and where we have used the fact that $d(1-L) = -dL$. The determinant of Δ is positive because felicity is quasi-concave and there are diminishing returns to the labour input, i.e. $|\Delta| > 0$. By inverting Δ we find from (C):

$$\begin{aligned} |\Delta|\, dC &= [U_{1-L,1-L} + \lambda(1-t_L) F_{LL} - (1-t_L) F_L U_{C,1-L}]\, d\lambda \\ &\quad + \lambda F_L U_{C,1-L} dt_L - \lambda(1-t_L) F_{KL} U_{C,1-L} dK, \end{aligned} \tag{E}$$

$$\begin{aligned} |\Delta|\, dL &= [U_{C,1-L} - (1-t_L) F_L U_{CC}]\, d\lambda + \lambda F_L U_{CC} dt_L \\ &\quad - \lambda(1-t_L) F_{KL} U_{CC} dK. \end{aligned} \tag{F}$$

By changing one of λ, K, and t_L at a time we obtain the desired partial derivatives of the implicit functions. The partial derivatives for λ are thus:

$$c_\lambda \equiv \frac{\partial C}{\partial \lambda} = \frac{U_{1-L,1-L} + \lambda\left(1-t_L\right)F_{LL} - \left(1-t_L\right)F_L U_{C,1-L}}{|\Delta|} \lesseqgtr 0, \tag{G}$$

$$l_\lambda \equiv \frac{\partial L}{\partial \lambda} = \frac{U_{C,1-L} - \left(1-t_L\right)F_L U_{CC}}{|\Delta|} \lesseqgtr 0, \tag{H}$$

where the ambiguity arises from the fact that we have not yet made an assumption regarding the cross partial derivative, $U_{C,1-L}$. Clearly, if $U_{C,1-L} \geq 0$ then $c_\lambda < 0$ and $l_\lambda > 0$.

The partial derivatives for K are:

$$c_K \equiv \frac{\partial C}{\partial K} = -\frac{\lambda\left(1-t_L\right)F_{KL}U_{C,1-L}}{|\Delta|} \lesseqgtr 0, \tag{I}$$

$$l_K \equiv \frac{\partial L}{\partial K} = -\frac{\lambda\left(1-t_L\right)F_{KL}U_{CC}}{|\Delta|} > 0. \tag{J}$$

The labour supply effect is unambiguously positive but the consumption effect is fully determined by the sign of $U_{C,1-L}$.

Finally, the partial derivatives for t_L are:

$$c_{t_L} \equiv \frac{\partial C}{\partial t_L} = \frac{\lambda F_L U_{C,1-L} dt_L}{|\Delta|} \lesseqgtr 0, \tag{K}$$

$$l_{t_L} \equiv \frac{\partial L}{\partial t_L} = \frac{\lambda F_L U_{CC} dt_L}{|\Delta|} < 0. \tag{L}$$

Not surprisingly (since both K and $1-t_L$ enter via the wage rate), the labour supply effect is unambiguously negative but the consumption effect is fully determined by the sign of $U_{C,1-L}$.

Note that for the loglinear felicity function, $U_{C,1-L} = 0$ and $c_K = c_{t_L} = 0$ (see the text).

****

7.2.3.2 Tax incidence in the unit-elastic model

In the so-called *unit-elastic model* all elasticities appearing in the extended Ramsey model are set equal to unity. Specifically, on the household side we assume that the felicity function is loglinear:

$$U\left(C, 1-L\right) = \ln\left(C^\alpha \left(1-L\right)^{1-\alpha}\right), \qquad 0 < \alpha < 1. \tag{A5.71}$$

This expression implies, of course, that the *inter*temporal substitution elasticity for consumption is equal to unity (logarithmic felicity) and that the *intra*temporal substitution elasticity between consumption and leisure is also unity (Cobb-Douglas sub-felicity function). On the production side we use the following Cobb-Douglas production function:

$$F(K,L) \equiv K^{\varepsilon}L^{1-\varepsilon}, \qquad 0<\varepsilon<1, \tag{A5.72}$$

which implies that the intratemporal substitution elasticity between capital and labour is unity.

By using (A5.71) and (A5.72) we find that the model (A5.67)-(A5.70) is simplified quite a lot:

$$C(t) = \frac{\alpha}{\lambda(t)}, \tag{A5.73}$$

$$\frac{1-\alpha}{1-L(t)} = \lambda(t)(1-t_L)(1-\varepsilon)\left(\frac{K(t)}{L(t)}\right)^{\varepsilon}, \tag{A5.74}$$

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = \rho+\delta-(1-t_K)\,\varepsilon\left(\frac{K(t)}{L(t)}\right)^{\varepsilon-1}, \tag{A5.75}$$

$$\dot{K}(t) = L(t)\left(\frac{K(t)}{L(t)}\right)^{\varepsilon} - C(t) - \delta K(t). \tag{A5.76}$$

Equations (A5.73) and (A5.74) are the Frisch demands for, respectively, consumption and leisure. Since $C(t)$ only depends on $\lambda(t)$, we can condense the model even more by eliminating $\lambda(t)$ from (A5.73)-(A5.75):

$$\frac{(1-\alpha)\,C(t)}{\alpha\,[1-L(t)]} = (1-t_L)(1-\varepsilon)\left(\frac{K(t)}{L(t)}\right)^{\varepsilon}, \tag{A5.77}$$

$$\frac{\dot{C}(t)}{C(t)} = (1-t_K)\,\varepsilon\left(\frac{K(t)}{L(t)}\right)^{\varepsilon-1} - (\rho+\delta). \tag{A5.78}$$

The model consists of the capital accumulation expression (A5.76), the labour market equilibrium condition (A5.77), expressing combinations of C, L, and K for which labour supply (left-hand side) equals labour demand (right-hand side), and the household Euler equation (A5.78).

The phase diagram of the unit-elastic model is presented in Figure 7.6. The construction of the phase diagram is considerably more complicated than for the standard Ramsey model because labour supply depends on the state variables of the model, viz. consumption and the capital stock. Indeed, the labour market equilibrium (LME) condition (A5.77) defines an implicit function relating equilibrium employment to consumption and the capital stock:

$$L(t) = l\left(\frac{C(t)\,K(t)^{-\varepsilon}}{1-t_L}\right), \tag{A5.79}$$

with $l'(\cdot)<0$ (see Intermezzo). We find the following partial effects:

$$\frac{\partial L(t)}{\partial C(t)} = l'(\cdot)\,\frac{K(t)^{-\varepsilon}}{1-t_L} < 0, \tag{A5.80}$$

$$\frac{\partial L(t)}{\partial K(t)} = -\varepsilon l'(\cdot)\,\frac{C(t)\,K(t)^{-(1+\varepsilon)}}{1-t_L} > 0, \tag{A5.81}$$

$$\frac{\partial L(t)}{\partial t_L} = l'(\cdot)\frac{K(t)^{-\varepsilon}}{(1-t_L)^2} < 0. \tag{A5.82}$$

Intermezzo 7.3

Derivation of (A5.79). Dropping the time index, we write (A5.77) as:

$$[\phi(L) \equiv] \quad [1-L] L^{-\varepsilon} = \omega_0 \frac{CK^{-\varepsilon}}{1-t_L}, \tag{A}$$

where $\omega_0 \equiv (1-\alpha) / [\alpha(1-\varepsilon)] > 0$ is a constant. We find that $\phi(L)$ is positive and downward sloping in the feasible range $L \in [0,1]$ and can thus be inverted, i.e. we can write:

$$L = l(\cdot) \equiv \phi^{-1}\left(\omega_0 \frac{CK^{-\varepsilon}}{1-t_L}\right). \tag{B}$$

The derivative of $l(\cdot)$ with respect to its argument is obtained from (A) by using the Implicit Function Theorem:

$$\phi'(L)\, dL = \omega_0 d\left(\frac{CK^{-\varepsilon}}{1-t_L}\right) \quad \Rightarrow \quad l'(\cdot) \equiv \frac{dL}{d\left(\frac{CK^{-\varepsilon}}{1-t_L}\right)} = \frac{\omega_0}{\phi'(L)} < 0. \tag{C}$$

****

The $\dot{C}(t) = 0$ line is the combination of $C(t)$ and $K(t)$ such that the capital-labour ratio is constant. The equilibrium capital-labour ratio is obtained from (A5.78):

$$k^{KR} \equiv \left(\frac{K}{L}\right)^{KR} = \left[\frac{\rho+\delta}{(1-t_K)\,\varepsilon}\right]^{1/(\varepsilon-1)}. \tag{A5.83}$$

Using (A5.83) in (A5.77) we immediately find that $C(t) / (1-L(t))$ is constant along the $\dot{C}(t) = 0$ line. Since $K(t) / L(t)$ is also constant along that line, it follows that the $\dot{C}(t) = 0$ line is a straight downward sloping line–see Figure 7.6.

Consumption dynamics can be found by noting that (by using (A5.79) and (A5.83)) equation (A5.78) can be written as:

$$\frac{\dot{C}(t)}{C(t)} = (1-t_K)\,\varepsilon \left[\left[\frac{K(t)}{l\left(\frac{C(t)K(t)^{-\varepsilon}}{1-t_L}\right)}\right]^{\varepsilon-1} - \left(k^{KR}\right)^{\varepsilon-1}\right]. \tag{A5.84}$$

It follows from (A5.84) that $\partial\left[\dot{C}(t) / C(t)\right] / \partial C(t) < 0$ (since $\partial L / \partial C < 0$ by (A5.80)), i.e. consumption rises (falls) over time for points below (above) the $\dot{C}(t) = 0$ line. This has been indicated with vertical arrows in Figure 7.6.

The $\dot{K}(t) = 0$ line is the combination of $C(t)$ and $K(t)$ such that the capital stock is constant over time. By using (A5.76) and (A5.79) and setting $\dot{K}(t) = 0$ we find:

$$C(t) = \left[l\left(\frac{C(t)\, K(t)^{-\varepsilon}}{1 - t_L} \right) \right]^{1-\varepsilon} K(t)^{\varepsilon} - \delta K(t). \tag{A5.85}$$

It is straightforward (though a little tedious) to show that the $\dot{K}(t) = 0$ line is as drawn in Figure 7.6.[10] The $\dot{K}(t) = 0$ lines has two roots on the horizontal axis and attains a maximum where the capital *stock* is equal to its golden-rule level K^{GR}. Note that, since labour supply is endogenous, this golden-rule capital stock is different from the one attained in the standard Ramsey model, i.e. K^{GR} in Figures 7.4 and 7.6 are not the same.

In order to derive the dynamic behaviour of the capital stock, we first substitute (A5.79) into (A5.76):

$$\dot{K}(t) = \left[l\left(\frac{C(t)\, K(t)^{-\varepsilon}}{1 - t_L} \right) \right]^{1-\varepsilon} K(t)^{\varepsilon} - C(t) - \delta K(t). \tag{A5.86}$$

By partially differentiating this expression with respect to consumption we find:

$$\frac{\partial \dot{K}(t)}{\partial C(t)} = (1-\varepsilon)\, l(\cdot)^{-\varepsilon} K(t)^{\varepsilon} \frac{\partial L(t)}{\partial C(t)} - 1 < 0, \tag{A5.87}$$

where the sign follows from the fact that $\partial L / \partial C < 0$ (see (A5.80) above). Above (below) the $\dot{K}(t) = 0$ line, consumption is too high (low) and labour supply is too low (high), so net investment is negative (positive). This is indicated by the horizontal arrows in Figure 7.6.

There is a unique equilibrium at point E_0 where the $\dot{K}(t) = 0$ line intersects the $\dot{C}(t) = 0$ line. Given the configuration of vertical and horizontal arrows, it is clear that this equilibrium is a saddle point (just as in the standard Ramsey model).

In Figure 7.7 we illustrate the effects of an increase in the corporate tax rate, t_K. Since this tax rate does not feature in (A5.86), it follows that the $\dot{K}(t) = 0$ line is not affected by the shock. It follows from (A5.84) that the $\dot{C}(t) = 0$ line rotates in a counter-clockwise fashion around point K_C (at that limiting point, $L(t) = 1$ and $C(t) = 0$), say from $[\dot{C}(t) = 0]_0$ to $[\dot{C}(t) = 0]_1$. The long-run equilibrium shifts from E_0 to point E_1. At impact, the capital stock is predetermined and the economy jumps from point E_0 to point A on the saddle path SP_1. At that point net investment is negative ($\dot{K}(t) < 0$) and the consumption time profile is downward sloping ($\dot{C}(t) < 0$). Over time, the economy therefore moves gradually from A to E_1. Just as in the standard Ramsey model, both consumption and the capital stock are reduced in the long run as a result of the tax increase. So as far as the macroeconomic effects are concerned the two models do not yield vastly (qualitatively) different conclusions.

The tax incidence effects on factor prices are as follows. At impact, the increase in consumption

[10] The interested reader is referred to Heijdra and van der Ploeg (2002, pp. 530-533) for a detailed derivation of this curve.

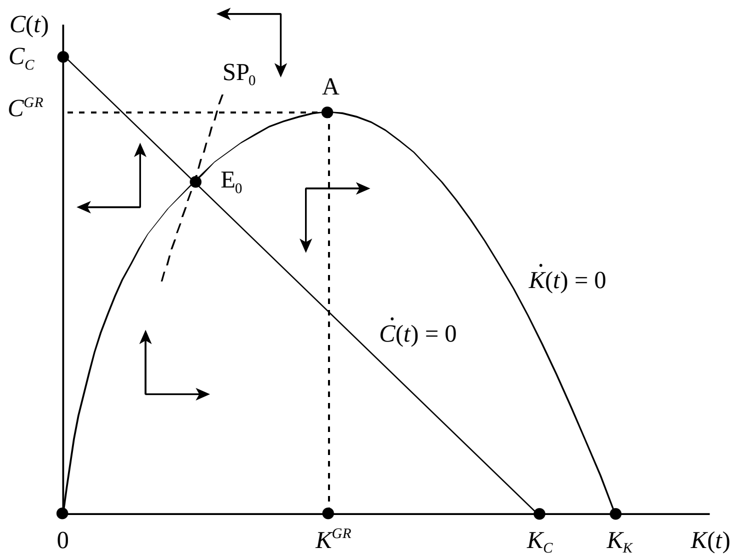

Figure 7.6: Phase diagram of the extended Ramsey model

($dC(0)/dt_K > 0$) is associated with a decrease in labour supply and equilibrium employment ($dL(0)/dt_K < 0$). Since the capital stock is fixed at impact, the capital-labour ratio rises ($dk(0)/dt_K > 0$), the marginal product of capital falls and the marginal product of labour rises. Hence, at impact the wage rate increases whilst the rental rate on capital decreases:

$$\frac{dW(0)}{dt_K} = \frac{dF_L}{dk(0)} \frac{dk(0)}{dt_K} > 0, \tag{A5.88}$$

$$\frac{dR(0)}{dt_K} = -F_K\left(k^{KR}, 1\right) + (1 - t_K) \frac{dF_K}{dk(0)} \frac{dk(0)}{dt_K} < 0. \tag{A5.89}$$

In the standard Ramsey model the wage is unchanged and the rental rate on capital declines (see (A5.59) above). In contrast, in the extended Ramsey model, the impact reduction in labour supply pushes up the wage rate and increases the relative abundance of capital (which reduces the marginal product of capital). In equation (A5.89) there is both a direct effect of the tax (first term on the right-hand) and an indirect or induced effect (second term).

It is not difficult to show that in the long run, labour continues to bear the full burden of the tax. In the long run the rental rate is pinned down by the pure rate of time preference ($R(\infty) = \rho + \delta$) and the reduction in the capital labour ratio leads to a decrease in the wage rate ($dW(\infty)/dt_K < 0$).

We close this subsection with two further remarks on the extended Ramsey model. First, the general model (with non-unitary substitution elasticities) can easily be handled by performing *local policy analysis*, i.e. by adopting the loglinearization approach used elsewhere in this book. Details of this approach are found in Judd (1987). Second, it cannot be overemphasized that the extended Ramsey model is still one of exogenous growth, i.e. *per definition* tax rates cannot affect the long-run growth rate in the economy! It is time to move on to the endogenous growth literature.

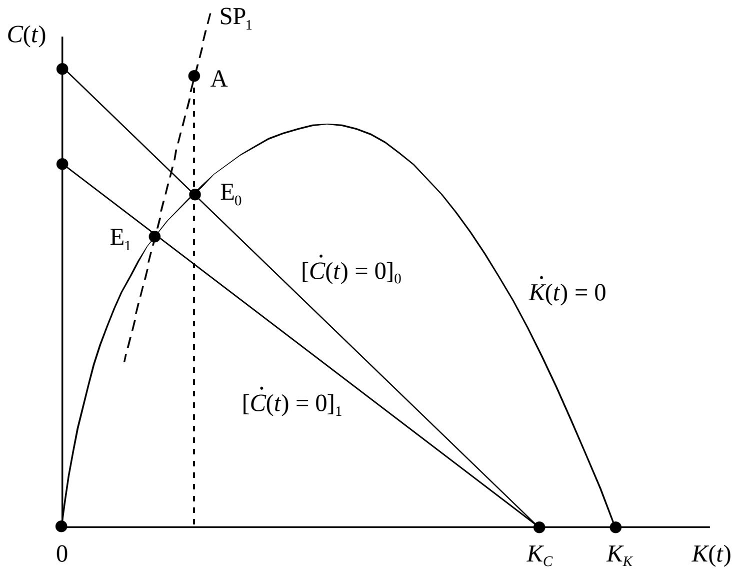

Figure 7.7: A Rise in the corporate tax in the extended Ramsey model

7.3 Endogenous growth models

In this section we discuss the effects of taxation on economic growth in the context of a number of key endogenous growth models. In endogenous growth models the *long-run* growth rate is endogenously determined, i.e. growth during transition *and* in the long run may depend on policy parameters such as tax rates. During the first wave of growth theory (1955-1970) the concept of endogenous growth was known but not taken seriously. In that literature, capital accumulation is seen as the *engine of growth* and the existence of (sufficiently strong) diminishing returns to capital ensures that labour gets scarce and the growth process is choked off in the long run.

In the second wave of growth theory (1985-) there are some authors who deny the existence of (sufficiently strong) diminishing returns to capital in the long run. We call this the *capital-fundamentalist* approach. Others see *human capital* accumulation as an additional engine of growth. Finally, yet others see *endogenous technological change* as the key engine of growth. In this section we present a selective overview of endogenous growth models and the effects of taxation.

7.3.1 Capital-fundamentalism

Within the group of capital-fundamentalist models one can distinguish at least two sub-groups. The first sub-group derives its inspiration from an insight by Solow (1956, pp. 77-78) himself. If there is "easy substitutability" between labour and capital, then labour never becomes an effective constraint to economic growth because firms can substitute capital for labour indefinitely. Suppose that the production function features the following CES form:

$$Y(t) = F[K(t), L(t)] \equiv \left[\varepsilon K(t)^{(\sigma_{KL}-1)/\sigma_{KL}} + (1-\varepsilon)L(t)^{(\sigma_{KL}-1)/\sigma_{KL}}\right]^{\sigma_{KL}/(\sigma_{KL}-1)}, \tag{A5.90}$$

where $0 < \varepsilon < 1$, $K(t)$ is the capital stock, $L(t)$ is employment, and σ_{KL} is the (non-negative) substitution elasticity between capital and labour. Substitution between capital and labour is labeled "difficult" if $0 \leq \sigma_{KL} \leq 1$ and "easy" if $\sigma_{KL} > 1$.

We abstract from labour-augmenting technological change, assume that the labour force grows at a constant exponential rate n_L and that the savings rate, s, is constant (Solow-Swan model). Recall from (A5.18) that the fundamental differential equation for the capital-labour ratio can then be written as:

$$g_k(t) \equiv \frac{\dot{k}(t)}{k(t)} = s\frac{f(k(t))}{k(t)} - (\delta + n_L), \tag{A5.91}$$

where $k(t) \equiv K(t)/L(t)$, $g_k(t)$ is the growth rate in $k(t)$, and $y(t) \equiv Y(t)/L(t) = f(k(t))$:

$$f(k(t)) \equiv \left[1 - \varepsilon + \varepsilon k(t)^{(\sigma_{KL}-1)/\sigma_{KL}}\right]^{\sigma_{KL}/(\sigma_{KL}-1)}. \tag{A5.92}$$

It follows from (A5.91) that the path of average capital productivity, $f(k(t))/k(t)$, determines the path for the growth rate of $k(t)$. We can deduce from (A5.92) that with easy substitutability of capital and labour ($\sigma_{KL} > 1$), average capital productivity satisfies:

$$\lim_{k(t)\to 0} \frac{f(k(t))}{k(t)} = \lim_{k(t)\to 0} \frac{f'(k(t))}{1} = +\infty, \tag{A5.93}$$

$$\lim_{k(t)\to \infty} \frac{f(k(t))}{k(t)} = \varepsilon^{\sigma_{KL}/(\sigma_{KL}-1)} > 0. \tag{A5.94}$$

The average product of capital is very high for low values of $k(t)$, falls as $k(t)$ increases (due to diminishing returns to capital), but bottoms out at a *positive* value given in (A5.94). In terms of Figure 7.8, the curve $sf(k(t))/k(t)$ reaches a horizontal asymptote as more and more capital is accumulated (see the dashed line). Provided the savings ratio is high enough, the asymptote lies above the line labeled $\delta + n_L$, and there is no steady-state for $k(t)$ even in the very long run. If the economy is initially endowed with a capital-labour ratio of $k(0)$ then the growth rate in the capital-labour ratio is equal to the vertical difference between points A and B. In general, we obtain from (A5.91) and (A5.94) that the long-run (or asymptotic) endogenous *growth rate* is given by:

$$g_k(\infty) = \lim_{k(t)\to\infty} \left[s\frac{f(k(t))}{k(t)} - (\delta + n_L)\right] = s\varepsilon^{\sigma_{KL}/(\sigma_{KL}-1)} - (\delta + n_L) > 0. \tag{A5.95}$$

In the long run, the $y(t)/k(t)$ ratio is constant so that $g_y(\infty) \equiv \dot{y}(t)/y(t) = g_k(\infty)$ and the level variables grow asymptotically according to $g_K(\infty) = g_Y(\infty) = g_k(\infty) + n_L$.

We call the asymptotic growth rate, $g_k(\infty)$, "endogenous" because it depends on more things than the growth rate of the population alone. A decrease in the savings rate, prompted for example by an increase in the corporate tax (as in Subsection 7.2.1 above), will have one of two drastically different effects. Case 1: if the shock leaves the horizontal asymptote in Figure 7.8 strictly above the $\delta + n_L$ line then $g_k(\infty)$ stays positive and perpetual growth in $k(t)$ is reduced but not eliminated. Case 2: if the shock pushes the asymptote below the $\delta + n_L$ line, then there will be a steady state in $k(t)$ and $g_k(\infty)$ will be zero as in the standard Solow-Swan model.

One of the problems with this type of endogenous growth is that it runs foul of some well-known stylized facts about the growth process. As the capital-labour ratio expands, labour becomes less and less important and eventually the income share of capital goes to unity and that of labour goes to zero. Indeed, denoting these shares by $\omega_K(t)$ and $1 - \omega_K(t)$ respectively, we find from (A5.92):

$$\omega_K(t) \equiv \frac{k(t) f'(k(t))}{f(k(t))} = \varepsilon \left(\frac{f(k(t))}{k(t)} \right)^{(1-\sigma_{KL})/\sigma_{KL}}. \tag{A5.96}$$

By letting $k(t) \to +\infty$ in this expression and using (A5.94) we find:

$$\begin{aligned} \lim_{k(t)\to\infty} \omega_K(t) &= \varepsilon \lim_{k(t)\to\infty} \left(\frac{f(k(t))}{k(t)} \right)^{(1-\sigma_{KL})/\sigma_{KL}} \\ &= \varepsilon \left(\varepsilon^{\sigma_{KL}/(\sigma_{KL}-1)} \right)^{(1-\sigma_{KL})/\sigma_{KL}} = 1. \end{aligned} \tag{A5.97}$$

As was documented by Kaldor (1961, pp. 178-179), one of the stylized facts of economic growth is that factor shares of capital and labour are both non-zero and fairly constant over time. For that reason we do not pursue this type of capital-fundamentalist model any further.

The second type of capital-fundamentalist models looks for various reasons (other than easy substitutability between factors of production) for there to exist constant returns to scale with respect to capital. Examples of this approach are Barro (1990), who argues that productive government spending can offset diminishing returns to private capital and Arrow (1962) and Romer (1986) who model a learning-by-doing effect with spillovers across firms.

Here we study the approach suggested by Rebelo (1991). The key idea is that there is a "core" of capital goods that is produced under constant returns to scale (CRTS) using only accumulable factors of production. In the remainder of this subsection we show a simple example of the basic Rebelo (1991) model. Just as in the Uzawa model, there are two production sectors. The *capital good* sector produces investment goods using only existing capital under CRTS. The production function is thus:

$$I(t) = A_I K_I(t), \tag{A5.98}$$

where $I(t)$ is output in the capital good sector ($I(t) \geq 0$), A_I is the general technology index, and $K_I(t)$

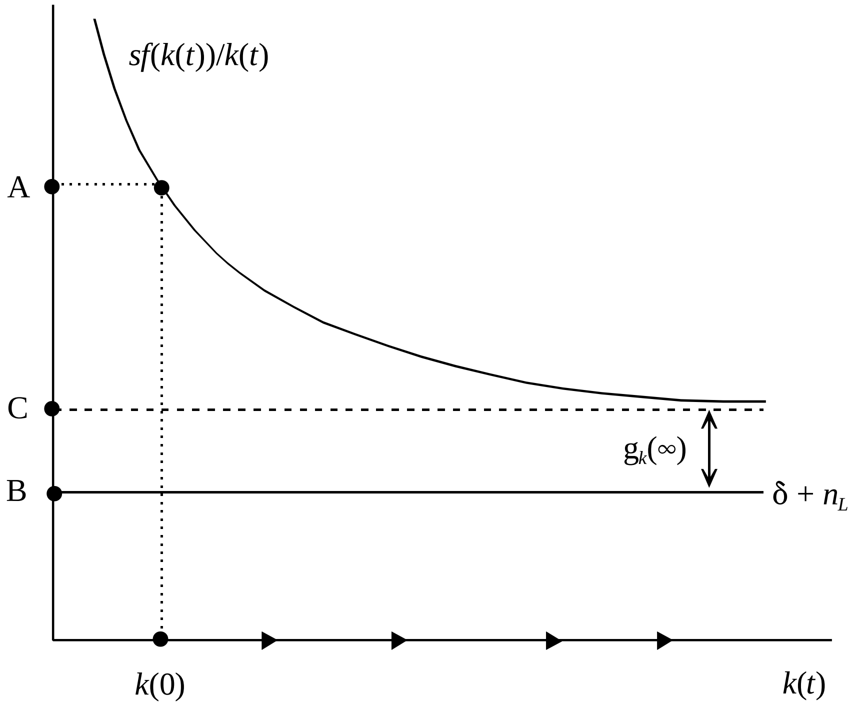

Figure 7.8: The capital-fundamentalist model

is capital use in the sector. The *consumption good* sector produces a good which is used for consumption purposes, using existing capital and land under CRTS. Technology is assumed to be Cobb-Douglas:

$$C(t) = A_C [K_C(t)]^{\varepsilon} T^{1-\varepsilon}, \tag{A5.99}$$

where $0 < \varepsilon < 1$, $C(t)$ is output in the consumption good sector, A_C is the general technology index, $K_C(t)$ is capital use in the sector, and T is land (a fixed factor).

Conceptually there are two types of capital in the model. *Reproducible capital* is capital which can be accumulated over time and used in the two sectors:

$$K(t) = K_I(t) + K_C(t), \tag{A5.100}$$

$$\dot{K}(t) = I(t) - \delta K(t), \tag{A5.101}$$

where $K(t)$ is the aggregate stock of reproducible capital and δ is its depreciation rate.[11] The second type of capital is *non-reproducible capital*, which is available in same quantity in each period, e.g. land T.

The institutional setting in the Rebelo model is standard. The market structure is competitive throughout. There are competitive firms in both sectors and we argue on the basis of a representative firm per sector. Firms rent production factors from the representative household, and there is perfect mobility of capital between sectors.

[11] As Rebelo points out, K can also be interpreted as a composite of physical and human capital (e.g. skilled labour).

The representative household is infinitely lived and is blessed with perfect foresight. The household's life-time utility function is:

$$\Lambda(0) \equiv \int_0^{\infty} \frac{C(t)^{1-1/\sigma} - 1}{1 - 1/\sigma} e^{-\rho t} dt, \tag{A5.102}$$

where $C(t)$ is (the flow of) consumption, σ is the intertemporal substitution elasticity ($\sigma > 0$), ρ is the pure rate of time preference ($\rho > 0$), and $\Lambda(0)$ is an index for life-time utility.

The household directly decides on the capital accumulation decision, i.e. it chooses $I(t)$ in an optimal fashion. The budget identity is given by:

$$C(t) + p_I(t) I(t) = R^T(t) T + R^K(t) K(t) + Z(t), \tag{A5.103}$$

where $R^T(t)$ is the rental rate on land, $R^K(t)$ is the rental rate on capital, $Z(t)$ is the lump-sum transfer received from the government, and $p_I(t)$ is the *relative* price of the investment good ($p_I(t) \equiv P_I(t)/P_C(t)$ and we use the consumption good as the numeraire commodity and set $P_C(t) = 1$).

The household chooses paths for $C(t)$, $I(t)$, and $K(t)$ in order to maximize (A5.102) subject to (A5.103) and (A5.101), some transversality conditions, and taking as given its initial stock of capital (i.e. $K(0)$ is predetermined at time $t = 0$). The first-order conditions for the household's optimum are:[12]

$$C(t)^{-1/\sigma} = \frac{\lambda(t)}{p_I(t)}, \tag{A5.104}$$

$$\dot{\lambda}(t) = \left[\rho + \delta - \frac{R^K(t)}{p_I(t)}\right] \lambda(t), \tag{A5.105}$$

where $\lambda(t)$ is the marginal utility of wealth (i.e. the co-state variable associated with the capital accumulation constraint (A5.101)). Equation (A5.104) is the Frisch demand for consumption, and (A5.105) characterizes the optimal path for the marginal utility of wealth.

Firm behaviour is quite straightforward. The representative firm in the capital good sector has the following profit function (expressed in terms of the consumption good):

$$\Pi_I(t) \equiv (1 - t_I) p_I A_I K_I(t) - R^K(t) K_I(t), \tag{A5.106}$$

where t_I is an output tax. Competitive behaviour leads to $d\Pi_I/dK_I = 0$ or:

$$R^K(t) = (1 - t_I) p_I(t) A_I, \tag{A5.107}$$

[12]Provided $I(t) > 0$, the current-value Hamiltonian is:

$$\mathcal{H} \equiv \frac{C(t)^{1-1/\sigma} - 1}{1 - 1/\sigma} + \lambda(t) \left[\frac{R^T(t) T + R^K(t) K(t) + Z(t) - C(t)}{p_I(t)} - \delta K(t)\right],$$

where $C(t)$ is the control variable, $K(t)$ is the state variable, and $\lambda(t)$ the co-state variable. The first-order conditions are $\partial\mathcal{H}/\partial C(t) = 0$ and $-\partial\mathcal{H}/\partial K(t) = \dot{\lambda}(t) - \rho\lambda(t)$. If the constraint $I(t) \geq 0$ becomes binding the household simply chooses to set $I(t) = 0$ and to consume its net income.

and as a result there are no excess profits ($\Pi_I(t) = 0$). Similarly, the representative firm in the consumption good sector has the following profit function:

$$\Pi_C(t) \equiv (1 - t_C) A_C K_C(t)^{\varepsilon} T^{1-\varepsilon} - R^K(t) K_C(t) - R^T(t) T, \tag{A5.108}$$

where t_C is an output tax. Competitive behaviour leads to $\partial\Pi_C/\partial K_C = \partial\Pi_C/\partial T = 0$ or:

$$R^K(t) = \varepsilon (1 - t_C) A_C \left(\frac{K_C(t)}{T}\right)^{\varepsilon-1}, \tag{A5.109}$$

$$R^T(t) = (1 - \varepsilon)(1 - t_C) A_C \left(\frac{K_C(t)}{T}\right)^{\varepsilon}, \tag{A5.110}$$

and $\Pi_C(t) = 0$.

Drawing things together we find that the following expressions make up (this version of) the Rebelo model:

$$\frac{\dot{C}(t)}{C(t)} = \sigma \left[\frac{\dot{p}_I(t)}{p_I(t)} - \frac{\dot{\lambda}(t)}{\lambda(t)}\right], \tag{A5.111}$$

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = \rho + \delta - (1 - t_I) A_I, \tag{A5.112}$$

$$\frac{\dot{p}_I(t)}{p_I(t)} = (\varepsilon - 1)\left[\frac{\dot{\phi}(t)}{\phi(t)} + \frac{\dot{K}(t)}{K(t)}\right], \tag{A5.113}$$

$$\frac{\dot{C}(t)}{C(t)} = \varepsilon \left[\frac{\dot{\phi}(t)}{\phi(t)} + \frac{\dot{K}(t)}{K(t)}\right], \tag{A5.114}$$

where $\phi(t) \equiv K_C(t)/K(t)$ is the fraction of the capital stock which is employed in the consumption good sector. Equation (A5.111) is obtained by differentiating (A5.104) with respect to time, and (A5.112) is obtained by substituting (A5.107) into (A5.105). The derivation of (A5.113) is a little more complex. First, we combine (A5.107) and (A5.109) to obtain the following expression:

$$p_I(t) = \varepsilon \frac{1 - t_C}{1 - t_I} \frac{A_C}{A_I} \left(\frac{K_C(t)}{T}\right)^{\varepsilon-1}. \tag{A5.115}$$

Next we differentiate (A5.115) with respect to time (under the assumption of time-invariant taxes, i.e. $\dot{t}_C(t) = \dot{t}_I(t) = 0$) and obtain:

$$\frac{\dot{p}_I(t)}{p_I(t)} = (\varepsilon - 1)\frac{\dot{K}_C(t)}{K_C(t)}. \tag{A5.116}$$

Finally, since $\phi(t) \equiv K_C(t)/K(t)$ we find by definition:

$$\frac{\dot{K}_C(t)}{K_C(t)} = \frac{\dot{\phi}(t)}{\phi(t)} + \frac{\dot{K}(t)}{K(t)}. \tag{A5.117}$$

By substituting (A5.117) into (A5.116) we obtain (A5.113). Finally, equation (A5.114) is obtained by differentiating the production function (A5.99) with respect to time and noting (A5.117).

It is straightforward to show that the model defined in (A5.111)-(A5.114) exhibits endogenous growth. We start by computing the growth rate in consumption. By combining (A5.113) and (A5.114) we find:

$$\frac{\dot{p}_I(t)}{p_I(t)} = \frac{\varepsilon - 1}{\varepsilon}\frac{\dot{C}(t)}{C(t)}. \tag{A5.118}$$

By using (A5.118) and (A5.112) into (A5.111) we find a linear equation in the growth rate in consumption:

$$\frac{\dot{C}(t)}{C(t)} = \sigma\left[\frac{\varepsilon - 1}{\varepsilon}\frac{\dot{C}(t)}{C(t)} - (\rho + \delta) + (1 - t_I)A_I\right]. \tag{A5.119}$$

Solving for the growth rate we find:

$$g_C \equiv \frac{\dot{C}(t)}{C(t)} = \frac{\varepsilon\sigma\left[(1 - t_I)A_I - (\rho + \delta)\right]}{\sigma + \varepsilon(1 - \sigma)}. \tag{A5.120}$$

Several things are worth noting about (A5.120). First, the growth rate of consumption, g_C, is time invariant, i.e. there is no transitional dynamics at all! Second, the tax rate on the investment good sector decreases the rate of growth. This result stands in stark contrast with the predictions obtained on the basis of exogenous growth models. Intuitively, increasing t_I is equivalent to decreasing A_I and this directly affects the engine of growth in this economy. Third, the tax rate on the consumption good sector does not affect the rate of growth (just as in the exogenous growth literature). Increasing t_C is equivalent to decreasing A_C and this only affects the level of consumption (but not its rate of growth) in this economy. Intuitively, t_C is like a lump-sum tax.

The next task is to determine the allocation of capital over the two sectors (i.e. $\phi(t)$). By using (A5.98) and (A5.101) we find:

$$\frac{\dot{K}(t)}{K(t)} = A_I\left[1 - \phi(t)\right] - \delta. \tag{A5.121}$$

By combining (A5.121) with (A5.114) and noting that $g_C \equiv \dot{C}(t)/C(t)$ we derive the differential equation in $\phi(t)$:

$$\frac{\dot{\phi}(t)}{\phi(t)} = \frac{g_C}{\varepsilon} + \delta + A_I\left[\phi(t) - 1\right]. \tag{A5.122}$$

Clearly, this is an unstable differential equation for which the only economically sensible solution is that $\phi(t)$ adjusts at all times to ensure that $\dot{\phi}(t) = 0$, i.e. the equilibrium level is given by:

$$1 - \phi^* = \frac{g_C + \varepsilon\delta}{\varepsilon A_I}. \tag{A5.123}$$

In terms of Figure 7.9, the differential equation (A5.122) is plotted as the straight line CAL_0. The unique equilibrium is at E_0 and any other values for $\phi(t)$ lead to economically nonsensical outcomes. An increase in the tax on the investment sector, t_I, leads to a decrease in g_C and a downward shift in the capital allocation line, say from CAL_0 to CAL_1 in Figure 7.9. The economy jumps from E_0 to E_1 and ϕ^* increases, i.e. a larger proportion of capital is used in the consumption good sector as a result of the shock.

Finally, the growth rate of net aggregate output can be determined as follows. We define net aggregate output (in terms of the consumption good) as the sum of consumption and net investment:

$$Y(t) = C(t) + p_I(t)\left[I(t) - \delta K(t)\right]. \tag{A5.124}$$

By differentiating this expression with respect to time we obtain:

$$\begin{aligned}\frac{\dot{Y}(t)}{Y(t)} &= \frac{C(t)}{Y(t)}\frac{\dot{C}(t)}{C(t)} + \frac{p_I(t)\,I(t)}{Y(t)}\left[\frac{\dot{I}(t)}{I(t)} + \frac{\dot{p}_I(t)}{p_I(t)}\right] \\ &\quad - \frac{\delta p_I(t)\,K(t)}{Y(t)}\left[\frac{\dot{K}(t)}{K(t)} + \frac{\dot{p}_I(t)}{p_I(t)}\right].\end{aligned} \tag{A5.125}$$

Since $\phi(t) = \phi^*$ is constant we know that $I(t)/K(t)$ is constant. By also using (A5.113) and (A5.114) we derive:

$$\begin{aligned}g_Y \equiv \frac{\dot{Y}(t)}{Y(t)} &= \frac{C(t)}{Y(t)}\varepsilon\frac{\dot{K}(t)}{K(t)} + \frac{p_I(t)\,I(t)}{Y(t)}\left[\frac{\dot{K}(t)}{K(t)} + (\varepsilon-1)\frac{\dot{K}(t)}{K(t)}\right] \\ &\quad - \frac{\delta p_I(t)\,K(t)}{Y(t)}\left[\frac{\dot{K}(t)}{K(t)} + (\varepsilon-1)\frac{\dot{K}(t)}{K(t)}\right] \\ &= \varepsilon\frac{\dot{K}(t)}{K(t)} = g_C.\end{aligned} \tag{A5.126}$$

Hence, net output grows at the same rate as consumption. The savings rate is constant.

We close this subsection with a number of remarks on the capital fundamentalist model. First, it is abundantly clear what is the engine of growth in the model described above. The key model assumption is the one underlying (A5.98): output in the investment sector is a linear function of accumulable capital. Endogenous growth is thus almost a direct assumption (rather than a result) of the model. Second, Rebelo (1991) has shown that the model can be generalized by disaggregating capital into physical and human capital. Provided there is a core of capital goods that is produced under constant returns to scale using no non-reproducible factors (directly or indirectly), the story remains valid (see also below).

Third, although there is no transitional dynamics in the present version of the Rebelo model, in some variations of the capital fundamentalist model there is non-trivial transitional dynamics (thus adding realism to the model).

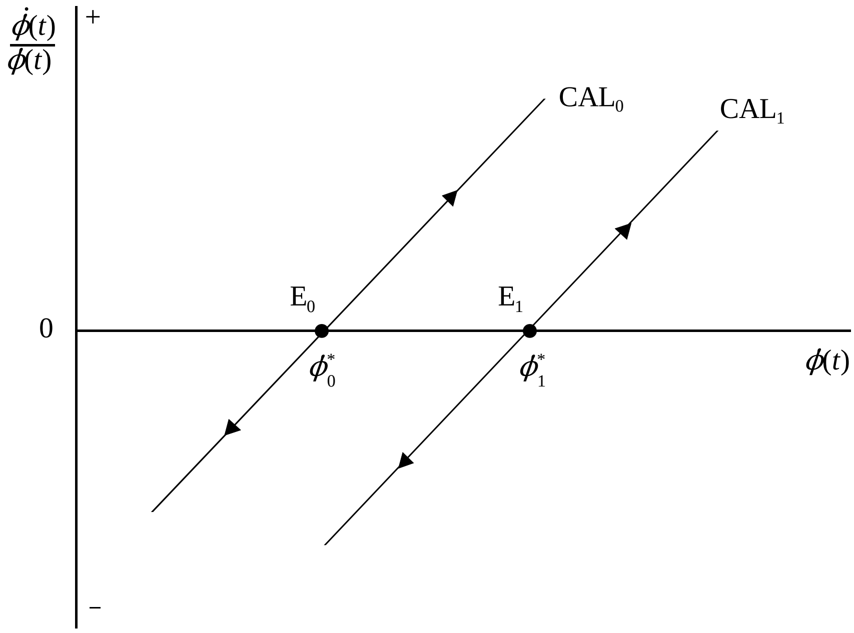

Figure 7.9: Intersectoral allocation of capital

7.3.2 Human capital

Four decades ago, Uzawa (1965) argued that (labour-augmenting) technological progress should not be seen as some kind of "manna from heaven" but instead should be regarded as the outcome of the intentional actions by economic agents employing scarce resources in order to advance the state of technological knowledge. Uzawa (1965) formalized his notions by assuming that all technological knowledge is embodied in labour and proposed a theory which endogenizes labour-augmenting productivity ($A(t)$ in (A5.1) above). Uzawa postulates the existence of a broadly defined educational sector which uses resources in order to augment the state of knowledge in the economy. Somewhat surprisingly, Uzawa's ideas lay dormant for almost a quarter century until Lucas (1988) once again placed human capital at the forefront of the economic growth process.

In this subsection we study a simplified version of the Lucas model due to Rebelo (1991, pp. 507-511). Lifetime utility of the representative household is still as in (A5.102) above. There are two sectors of production, namely a *goods* sector and an *education* sector. The goods sector produces consumption and investment goods using the following Cobb-Douglas technology:

$$Y(t) = A_Y [K_Y(t)]^{1-\gamma} [L_Y(t) H(t)]^{\gamma}, \qquad \text{(A5.127)}$$

where $0 < \gamma < 1$, $Y(t)$ is aggregate output in the goods sector, $K_Y(t)$ is the physical capital use, $L_Y(t)$ is the raw labour input, $H(t)$ is human capital, and A_Y is an index of technology. Each unit of "raw" labour has productivity level $H(t)$ so $L_Y(t) H(t)$ is the labour input in efficiency units. We define the following ratio:

$$\phi(t) \equiv \frac{K_Y(t)}{K(t)}, \tag{A5.128}$$

and rewrite (A5.127) as:

$$Y(t) = A_Y [\phi(t) K(t)]^{1-\gamma} [L_Y(t) H(t)]^{\gamma}. \tag{A5.129}$$

In the *education* sector, capital and efficiency units of labour are used by the household to produce new human capital. Human capital is embodied in the worker and appreciates at rate δ. The accumulation of human capital proceeds according to the following technology:

$$\dot{H}(t) = A_H [(1-\phi(t)) K(t)]^{1-\beta} [(\bar{L} - L_Y(t)) H(t)]^{\beta} - \delta H(t), \tag{A5.130}$$

where $0 < \beta < 1$, $\dot{H}(t)$ is *net* investment in human capital, A_H is an efficiency index, $(1-\phi(t)) K(t)$ is the amount of capital used for human capital accumulation, $\bar{L}$ is the (fixed) labour supply so $(\bar{L} - L_Y(t)) H(t)$ is the labour input (in efficiency units) devoted to the creation of new human capital, and $\delta H(t)$ is the depreciation on existing human capital.

The household directly decides on (i) the physical capital accumulation decision (choice of $I(t)$), (ii) the human capital accumulation decision (choice of $\dot{H}(t)$), and (iii) the consumption decision (choice of $C(t)$). We interpret the education sector as an "in-house" activity, so the household rents out $\phi(t) K(t)$ units of physical capital and $L_Y(t) H(t)$ units of human capital to firms in the goods sector at respective rental rates $R^K(t)$ and $R^H(t)$. The household budget identity is:

$$C(t) + I(t) = \phi(t) R^K(t) K(t) + R^H(t) L_Y(t) H(t) + Z(t), \tag{A5.131}$$

where $C(t)$ is consumption, $I(t)$ is gross investment in physical capital, and $Z(t)$ is the lump-sum transfer received from the government. Total output in the goods sector equals:

$$Y(t) = C(t) + I(t), \tag{A5.132}$$

and the capital accumulation equation is given by:

$$\dot{K}(t) = I(t) - \delta K(t), \tag{A5.133}$$

where δ is the depreciation rate of physical capital.[13]

The household chooses paths for $C(t)$, $I(t)$, and $\dot{H}(t)$ in order to maximize (A5.102) subject to (A5.131) and (A5.133), and some transversality conditions, and taking as given the initial stocks of the

[13] To keep things simple, human and physical capital are assumed to feature the same depreciation rate δ.

two types of capital (i.e. $K(0)$ and $H(0)$ are both predetermined at time $t = 0$). The current-value Hamiltonian for this problem is (provided $I(t) > 0$):

$$
\begin{aligned}
\mathcal{H} \equiv & \frac{C(t)^{1-1/\sigma} - 1}{1 - 1/\sigma} \\
& + \lambda(t) \left[\phi(t) R^K(t) K(t) + R^H(t) L_Y(t) H(t) + Z(t) - C(t) - \delta K(t)\right] \\
& + \mu(t) \left[A_H \left[(1 - \phi(t)) K(t)\right]^{1-\beta} \left[(\bar{L} - L_Y(t)) H(t)\right]^{\beta} - \delta H(t)\right],
\end{aligned}
\tag{A5.134}
$$

where the control variables are $C(t)$, $L_Y(t)$, and $\phi(t)$, the state variables are $K(t)$ and $H(t)$, and the co-state variable are $\lambda(t)$ and $\mu(t)$. The key first-order conditions for the household's optimum are:

$$
C(t)^{-1/\sigma} = \lambda(t), \tag{A5.135}
$$

$$
R^H(t) = \beta \frac{\mu(t)}{\lambda(t)} A_H \left[k_E(t)\right]^{1-\beta}, \tag{A5.136}
$$

$$
R^K(t) = (1 - \beta) \frac{\mu(t)}{\lambda(t)} A_H \left[k_E(t)\right]^{-\beta}, \tag{A5.137}
$$

$$
\frac{\dot{\lambda}(t)}{\lambda(t)} = \rho + \delta - \phi(t) R^K(t) - (1 - \beta) \frac{\mu(t)}{\lambda(t)} (1 - \phi(t)) A_H \left[k_E(t)\right]^{-\beta}, \tag{A5.138}
$$

$$
\frac{\dot{\mu}(t)}{\mu(t)} = \rho + \delta - R^H(t) L_Y(t) \frac{\lambda(t)}{\mu(t)} - \beta (\bar{L} - L_Y(t)) A_H \left[k_E(t)\right]^{1-\beta}, \tag{A5.139}
$$

where $k_E(t)$ is the capital-labour ratio in the education sector:

$$
k_E(t) \equiv \frac{(1 - \phi(t)) K(t)}{(\bar{L} - L_Y(t)) H(t)}. \tag{A5.140}
$$

The representative firm in the goods sector has the following profit function (in real terms):

$$
\begin{aligned}
\Pi(t) \equiv & (1 - t_Y) A_Y \left[\phi(t) K(t)\right]^{1-\gamma} \left[L_Y(t) H(t)\right]^{\gamma} - R^K(t) \phi(t) K(t) \\
& - R^H(t) L_Y(t) H(t),
\end{aligned}
\tag{A5.141}
$$

where t_Y is an output tax. The firm chooses its inputs in order to maximize profit. Competitive behaviour leads to the usual factor demands:

$$
R^K(t) = (1 - \gamma)(1 - t_Y) A_Y \left[k_Y(t)\right]^{-\gamma}, \tag{A5.142}
$$

$$
R^H(t) = \gamma (1 - t_Y) A_Y \left[k_Y(t)\right]^{1-\gamma}, \tag{A5.143}
$$

where $k_Y(t)$ is the capital-labour ratio in the goods sector:

$$
k_Y(t) \equiv \frac{\phi(t) K(t)}{L_Y(t) H(t)}, \tag{A5.144}
$$

and excess profit is zero ($\Pi(t) = 0$) in the optimum.

We can now combine the various expressions. First, by combining (A5.136) and (A5.143) we obtain the static expression determining the optimal allocation of human capital across the two sectors:

$$\left[R^H(t) =\right] \; \beta \frac{\mu(t)}{\lambda(t)} A_H \left[k_E(t)\right]^{1-\beta} = \gamma (1 - t_Y) A_Y \left[k_Y(t)\right]^{1-\gamma}. \tag{A5.145}$$

Similarly, by combining (A5.137) and (A5.142) we obtain the corresponding expression for the optimal allocation of physical capital across sectors:

$$\left[R^K(t) =\right] \; (1-\beta) \frac{\mu(t)}{\lambda(t)} A_H \left[k_E(t)\right]^{-\beta} = (1-\gamma)(1 - t_Y) A_Y \left[k_Y(t)\right]^{-\gamma}. \tag{A5.146}$$

By combining (A5.145) and (A5.146) we can eliminate $\mu(t)/\lambda(t)$ to get:

$$\frac{\beta}{1-\beta} k_E(t) = \frac{\gamma}{1-\gamma} k_Y(t). \tag{A5.147}$$

It is optimal to maintain a constant ratio of capital intensities in the two sectors.

Second, by using (A5.137) and (A5.138) the dynamic evolution of the shadow price of physical capital can be written as:

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = \rho + \delta - R^K(t). \tag{A5.148}$$

In a similar fashion, equations (A5.136) and (A5.139) can be combined to obtain an expression for the dynamic evolution of the shadow price of human capital:

$$\frac{\dot{\mu}(t)}{\mu(t)} = \rho + \delta - \frac{\lambda(t)}{\mu(t)} \bar{L} R^H(t). \tag{A5.149}$$

In the *steady state*, $k_E(t)$ and $k_Y(t)$ are constant and it follows from (A5.145) and (A5.146) that $\mu(t)/\lambda(t)$ is also constant, i.e.:

$$\frac{\dot{\mu}(t)}{\mu(t)} = \frac{\dot{\lambda}(t)}{\lambda(t)}. \tag{A5.150}$$

By combining (A5.148)-(A5.150) we find :

$$\begin{aligned} \left(R^K\right)^* &= \left(\frac{\lambda(t)}{\mu(t)}\right)^* \bar{L} \left(R^H\right)^* \qquad \Leftrightarrow \\ (1-\gamma)(1-t_Y) A_Y \left[k_Y^*\right]^{-\gamma} &= \bar{L} \beta A_H \left[k_E^*\right]^{1-\beta}, \end{aligned} \tag{A5.151}$$

where the starred variables denote steady-state values, and we have used (A5.145) and (A5.146) to get to the second line.

We now have two expressions for k_E and k_Y, namely equation (A5.147) which holds at all times, and equation (A5.151), which holds in the steady state. By using these two expressions, the steady-state capital-labour ratios in the two activities can be determined. After some manipulation we find:

$$k_Y^* = \left[\frac{(1-\gamma)(1-t_Y) A_Y}{\beta \bar{L} A_H} \left(\frac{\beta (1-\gamma)}{\gamma (1-\beta)} \right)^{1-\beta} \right]^{1/(1-\beta+\gamma)}, \tag{A5.152}$$

$$k_E^* = \left[\frac{(1-\gamma)(1-t_Y) A_Y}{\beta \bar{L} A_H} \left(\frac{\beta (1-\gamma)}{\gamma (1-\beta)} \right)^{-\gamma} \right]^{1/(1-\beta+\gamma)}. \tag{A5.153}$$

In view of (A5.152) and (A5.146), the steady-state rental rate on physical capital can be written as:

$$\begin{aligned} \left(R^K\right)^* &= (1-\gamma)(1-t_Y) A_Y [k_Y^*]^{-\gamma} \\ &= \Psi [(1-t_Y) A_Y]^{\theta} (A_H \bar{L})^{1-\theta}, \end{aligned} \tag{A5.154}$$

where Ψ and θ are defined as follows:

$$\Psi \equiv \left[(1-\gamma)^{(1-\beta)(1-\gamma)} (1-\beta)^{\gamma(1-\beta)} \beta^{\beta\gamma} \gamma^{\gamma(1-\beta)} \right]^{1/(1-\beta+\gamma)} > 0, \tag{A5.155}$$

$$\theta \equiv \frac{1-\beta}{1-\beta+\gamma}, \qquad 0 < \theta < 1. \tag{A5.156}$$

We are now in the position to derive the steady-state growth rate in the economy. By differentiating the Frisch demand for consumption, (A5.135), with respect to time we obtain the Euler equation:

$$\begin{aligned} g_C^* \equiv \left(\frac{\dot{C}(t)}{C(t)} \right)^* &= -\sigma \left(\frac{\dot{\lambda}(t)}{\lambda(t)} \right)^* \\ &= \sigma \left[\left(R^K\right)^* - \delta - \rho \right] \\ &= \sigma \left[\Psi [(1-t_Y) A_Y]^{\theta} (A_H \bar{L})^{1-\theta} - \delta - \rho \right], \end{aligned} \tag{A5.157}$$

where we have used (A5.148) to get from the first to the second line, and (A5.154) to get to the third line. According to (A5.157), consumption grows in the steady-state at an exponential rate g_C^*. It is easy to show that $I(t)$, $K(t)$, and $H(t)$ all grow at the rate g_C^* in the steady state. Similarly, net output, $Y \equiv C + I - \delta K$, grows at that rate, i.e. $g_Y^* = g_C^*$.

Several things are worth noting about the steady-state growth rate given in (A5.157). First, an increase in the tax rate levied on the goods sector (t_Y) leads to a decrease in the steady-state growth rate. The shock reduces the steady-state rental rate on capital as the capital-labour ratio in both sectors falls (see (A5.152)-(A5.153)). Intuitively, the private sector substitutes away from the production factor whose production is taxed more heavily (i.e. away from capital). Second, it is clear that the engine of growth of the human capital model is provided by the assumption of constant returns to scale in the production of goods and new human capital, i.e. the production functions (A5.129)-(A5.130) both feature constant returns to scale in the accumulable factors $K(t)$ and $H(t)$. This is the "core property" mentioned above. Third, a problematic aspect of the human capital model is the presence of a so-called *scale effect*. In equation (A5.157), the steady-state growth rate of consumption depends *inter alia* on the size of the total labour force ($\bar{L}$). This means, of course, that large countries should grow faster than small countries do (since $\bar{L}$ is larger for the former). This prediction of the model is easily falsified empirically. We return to the scale effect below in the context of the R&D model.

In closing this subsection we note that, in accordance with reality, there exists non-trivial transitional dynamics in the model, i.e. during transition to the steady state the growth rates of the different variables are time-dependent. The formal analysis of the transitional dynamics is quite complicated, however, because there are two slow-moving state variables in the model, namely the stocks of physical and human capital. The interested reader is referred to Bond, Wang, and Yip (1996) for a thorough discussion of this issue.

7.3.3 Endogenous technology

In the previous subsection we have shown that the purposeful accumulation of human capital forms the key ingredient of the Uzawa-Lucas theory of economic growth. In this subsection we briefly review a branch of the (huge) literature in which the purposeful conduct of research and development (R&D) activities forms the key source of growth.[14] In order to demonstrate the key mechanism by which R&D affects economic growth we follow Grossman and Helpman (1991, ch. 3) and Bénassy (1998) by abstracting from physical and human capital altogether. In such a setting all saving by households is directed towards the creation of new technology.

There are three production sectors in the economy. The *final goods sector* produces a homogeneous good using varieties of a differentiated intermediate good as productive inputs. Production is subject to constant returns to scale (in these inputs) and perfect competition prevails. The *R&D sector* is also perfectly competitive. In this sector units of labour are used to produce blueprints of new varieties of the differentiated input. Finally, the *intermediate goods sector* is populated by a large number of small firms, each producing a single variety of the differentiated input, who engage in Chamberlinian monopolistic competition (see also Chapter 7 for a detailed account of this market structure).

In the final goods sector the representative firm produces a homogenous good under conditions of perfect competition. The technology is give by the following CES function:

$$Y(t) \equiv N(t)^{\eta} \left[N(t)^{-1} \int_0^{N(t)} X_j(t)^{1/\mu} dj \right]^{\mu}, \tag{A5.158}$$

where $Y(t)$ is output, $X_j(t)$ is intermediate input j, $N(t)$ is the existing number of input varieties, and μ and η are parameters satisfying $\mu > 1$, and $1 \leq \eta \leq 2$. Intuitively, if η is strictly greater than

[14] Key contributions to this literature are Paul Romer (1987, 1990), Aghion and Howitt (1998), and Grossman and Helpman (1991).

unity, then there are so-called *returns to specialization* as in Adam Smith's famous pin factory example. If intermediate inputs are more finely differentiated then firms can use a more "roundabout" production process and reap productivity gains as a result (note that $\eta < 2$ is a mild and reasonable assumption used later). The assumption that μ exceeds unity implies that the intermediate inputs are close but *imperfect* substitutes in the production function.

The cost function associated with the technology (A5.158) is given by:

$$C^y \left[P_j(t), Y(t), N(t)\right] \equiv c^y \left[P_j(t), N(t)\right] Y(t), \tag{A5.159}$$

where the unit cost function is defined as follows:

$$c^y \left[P_j(t), N(t)\right] \equiv N(t)^{-\eta} \left[N(t)^{\mu/(1-\mu)} \int_0^{N(t)} P_j(t)^{1/(1-\mu)} dj\right]^{1-\mu}, \tag{A5.160}$$

and where $P_j(t)$ is the price of input j. The firm's pricing decision amounts to equating the output price, $P_Y(t)$, to marginal cost:

$$P_Y(t) \equiv c^y \left[P_j(t), N(t)\right], \tag{A5.161}$$

Finally, the derived demand for input j is obtained by applying Shephard's Lemma to (A5.159):

$$\frac{X_j(t)}{Y(t)} = \left[\frac{\partial c^y[\cdot]}{\partial P_j(t)} =\right] N(t)^{(\eta-\mu)/(\mu-1)} \left(\frac{P_j(t)}{P_Y(t)}\right)^{\mu/(1-\mu)}, \tag{A5.162}$$

for $j \in [0, N(t)]$. The key thing to note about (A5.162) is that the derived demand for input j is a downward sloping function of the price of input j, with $-\mu/(\mu-1)$ representing the demand elasticity.

In the R&D sector labour is used to create blueprints for new input varieties. The sector is perfectly competitive and technology features constant returns to scale:

$$\dot{N}(t) = (1/k_R)N(t)L_R(t), \tag{A5.163}$$

where $\dot{N}(t)$ is the output of the R&D sector (new varieties), k_R is a productivity index, and $L_R(t)$ is labour employed in the R&D sector. Note that (A5.163) incorporates an external effect in that labour engaged in the R&D sector becomes more productive as more patents already exist. Intuitively, today's engineers "stand on the shoulders of giants." Profit of the representative R&D firm is equal to:

$$\Pi_N(t) \equiv P_N(t)\dot{N}(t) - (1-s_R)W(t)L_R(t), \tag{A5.164}$$

where $P_N(t)$ is the market price of a (new or existing) patent (determined below), s_R is a wage subsidy in the R&D sector, and $W(t)$ is the wage rate. The firm chooses its labour input in order to maximize

(A5.164) subject to (A5.163), taking as given its output price ($P_N(t)$), its input price ($W(t)$), and the existing number of product varieties ($N(t)$). The first-order condition sets price equal to marginal cost:

$$P_N(t) = \frac{k_R(1-s_R)W(t)}{N(t)}. \tag{A5.165}$$

In the intermediate goods sector there are many small monopolistically competitive firms. Each firm holds a patent allowing it to use labour to produce its own slightly unique variety of the intermediate input. Technology is given by:

$$X_j(t) = (1/k_X)L_j(t), \tag{A5.166}$$

where $L_j(t)$ is the labour input and k_X is a technology index. Operating profit of firm j is defined as:

$$\Pi_j(t) \equiv P_j(t)X_j(t) - W(t)L_j(t). \tag{A5.167}$$

The firm chooses its output level, $X_j(t)$, given the elastic demand for its output (A5.162) and the production function (A5.166), and taking the actions of all other producers in the intermediate goods sector as given (the Cournot-Nash assumption). As is familiar from the detailed discussion in Chapter 7 above, the optimal choice of the firm is to set price according to a fixed markup over marginal production cost:

$$P_j(t) = \mu W(t)k_X, \tag{A5.168}$$

where μ is thus the gross monopoly markup. The model is completely symmetric, so all firms charge the same price ($P_j(t) = \bar{P}_j(t)$, produce the same quantity with the same amount of labour ($X_j(t) = \bar{X}(t)$ and $L_j(t) = \bar{L}_j(t)$), and make the same profits ($\Pi_j(t) = \bar{\Pi}(t)$).

The representative household has the lifetime utility function (A5.102) and faces the following budget identity:

$$P_Y(t)C(t) + P_N(t)\dot{N}(t) = W(t)\bar{L} + N(t)\bar{\Pi}(t) + Z(t), \tag{A5.169}$$

where $C(t)$ is consumption, $\bar{L}$ is exogenous labour supply, $N(t)\bar{\Pi}(t)$ is aggregate profit income received from the intermediate goods sector, and $Z(t)$ is lump-sum transfers received from the government. The household saves by accumulating patents ($P_N(t)\dot{N}(t)$). By owning a patent, the household receives the profit derived from it ($\bar{\Pi}(t)$).

The household chooses $C(t)$ and $\dot{N}(t)$ in order to maximize lifetime utility (A5.102) subject to (A5.169), a transversality condition, and taking as given the initial stock of patents (i.e. $N(0)$ is predetermined). Assuming an interior solution (with positive saving, i.e. $\dot{N}(t) > 0$) the first-order conditions are given by:

$$\frac{\dot{C}(t)}{C(t)} = \sigma\left[r(t) - \rho\right], \tag{A5.170}$$

$$r(t) = \frac{\bar{\Pi}(t) + \dot{P}_N(t)}{P_N(t)}, \tag{A5.171}$$

where $r(t)$ is the rate of return on blueprints.

The model is closed by two market clearing conditions. The final goods market clears provided output equals consumption:

$$Y(t) = C(t). \tag{A5.172}$$

The labour market equilibrium condition requires the total supply of labour to equal the sum of labour demand in the intermediate and R&D sectors, i.e. $L_X(t) + L_R(t) = \bar{L}$. Since $L_X(t) = k_X N(t)\bar{X}(t)$ and $L_R(t) = k_R \dot{N}(t)/N(t)$ we can rewrite this labour market equilibrium condition as:

$$\frac{\dot{N}(t)}{N(t)} = \frac{\bar{L} - k_X N(t)\bar{X}(t)}{k_R}, \tag{A5.173}$$

where we assume implicitly that the differentiated sector is not too large and thus does not absorb all available labour (i.e. the numerator on the right-hand side is strictly positive).

The growth rate in the R&D model can be derived as follows. First we note some intermediate results:

$$\frac{\bar{\Pi}(t)}{P_N(t)} = (\mu - 1)\left(\frac{k_X N(t)\bar{X}(t)}{k_R\left(1 - s_R\right)}\right), \tag{A5.174}$$

$$\frac{\dot{P}_N(t)}{P_N(t)} = (\eta - 2)\left(\frac{\dot{N}(t)}{N(t)}\right), \tag{A5.175}$$

$$C(t) = N(t)^{\eta-1} N(t)\bar{X}(t). \tag{A5.176}$$

Assuming a time-invariant subsidy ($\dot{s}_R(t) = 0$) we obtain the following expressions for the various growth rates:

$$\gamma_C(t) = \sigma\left[\left(\frac{\mu - 1}{k_R\left(1 - s_R\right)}\right) L_X(t) + (\eta - 2)\gamma_N(t) - \rho\right], \tag{A5.177}$$

$$\gamma_C(t) = (\eta - 1)\gamma_N(t) + \frac{\dot{L}_X(t)}{L_X(t)}, \tag{A5.178}$$

$$\gamma_N(t) = \frac{\bar{L} - L_X(t)}{k_R}. \tag{A5.179}$$

By substituting (A5.177) and (A5.179) into (A5.178) we obtain a differential equation for $L_X(t)$:

$$\frac{\dot{L}_X(t)}{L_X(t)} = \frac{\sigma(\mu-1)}{k_R(1-s_R)}L_X(t) - \sigma\rho + [\sigma(2-\eta)+(\eta-1)]\left(\frac{L_X(t)-\bar{L}}{k_R}\right). \tag{A5.180}$$

It follows from that:

$$\frac{\partial}{\partial L_X(t)}\left[\frac{\dot{L}_X(t)}{L_X(t)}\right] = \frac{\sigma(\mu-1)}{k_R(1-s_R)} + \frac{[\sigma(2-\eta)+(\eta-1)]}{k_R} > 0, \tag{A5.181}$$

where we have used the fact that $\mu > 1$ and $1 \leq \eta < 2$. Hence, (A5.180) is an unstable differential equation for which the only economically sensible solution is the steady-state solution, for which $\dot{L}_X(t) = 0$ and $L_X(t) = L_X^*$. By imposing the steady state in (A5.180) we obtain an expression for the equilibrium amount of labour employed in the R&D sector and thus (by (A5.179)) for the rate of innovation:

$$\gamma_N = \frac{\bar{L} - L_X^*}{k_R} = \frac{\sigma(\mu-1)(\bar{L}/k_R) - \sigma\rho(1-s_R)}{\sigma(\mu-1) + [\sigma(2-\eta)+\eta-1](1-s_R)} > 0. \tag{A5.182}$$

Finally, by using (A5.178) and (A5.172) we find the growth rates for $C(t)$, and $Y(t)$:

$$\gamma_C = \gamma_Y = (\eta-1)\gamma_N. \tag{A5.183}$$

The innovation rate, γ_N, increases with the monopoly markup (μ), the R&D subsidy (s_R), and the size of the labour force ($\bar{L}$), and decreases with the rate of time preference (ρ). The effect on the innovation rate of the intertemporal substitution elasticity is:

$$\frac{\partial\gamma_N}{\partial\sigma} = \frac{(\eta-1)(1-s_R)\gamma_N}{\sigma[\sigma(\mu-1)+[\sigma(2-\eta)+\eta-1](1-s_R)]}. \tag{A5.184}$$

Provided the returns to specialization are operative (so that $\eta > 1$), an increase in the willingness of the representative household to substitute consumption across time raises the rate of innovation ($\partial\gamma_N/\partial\sigma > 0$). As is evident from (A5.183), the common growth rate of consumption and output also depends critically on whether or not the technology in the final goods sector is characterized by returns from specialization.

We end this subsection with a number of remarks on the R&D model. First, it is clear from the discussion that there is no transitional dynamics in this version of the R&D model. This is not surprising because there is no physical and/or human capital in the model. Second, the engine of growth in the model is the production function in the R&D sector, i.e. equation (A5.163) above. This expression relates a growth rate to the absolute amount of labour employed in the R&D sector (i.e. $\gamma_N(t) = L_R(t)/k_R$). So, just as in the capital-fundamentalist model, endogenous growth is more a direct assumption than a result of the model.

Third, like the human capital model, the R&D model has the *problematic property* that the growth rate depends on the scale of the economy ($\bar{L}$ in this case). Hence, large countries should grow faster than small countries. This is not observed in reality. Jones (1995) therefore removes the scale effect by replacing the R&D technology (A5.163) by the following expression:

$$\dot{N}(t) = (1/k_R)L_R(t)N(t)^{\phi_1}\left[\bar{L}_R(t)\right]^{\phi_2-1}, \tag{A5.185}$$

where $\bar{L}_R$ is *average* R&D labour per firm in the R&D sector. The R&D technology has changed in two aspects. First, we used to have $\phi_1 = 1$ but now we assume $0 < \phi_1 < 1$. Intuitively, there are diminishing return to "giants' shoulders." Second, we used to have $\phi_2 = 1$ but now we assume $0 < \phi_2 \leq 1$. Jones defends this assumption by appealing to a *duplication externality*: individual R&D firms think the production function is linear (in the labour input), but in actuality it features diminishing returns to labour. Using the more general R&D technology, it is possible to derive the following expressions for the rate of innovation and the growth rates in consumption and output: $\gamma_N(t) = \gamma_C(t) = \gamma_Y(t) = 0$. Hence, we reach the striking conclusion that by eliminating the scale effect we are back in the realm of exogenous growth and the Solow model.

Key literature

- Atkinson & Stiglitz (1980, lecture 8) on theory.
- Optimal control and duality: Benveniste and Scheinkman (1982).
- Feldstein (1974a-b), Newbery and Stern (1987) on tax policy in exogenous growth model.
- Judd (1985, 1987a, 1987b) and Bovenberg (1993, 1994) on the (open economy) Ramsey model.
- Endogenous growth models: Barro and Sala-i-Martin (1992), Stokey and Rebelo (1995), and Jones, Manuelli and Rossi (1993, 1997), Lucas (1988), Bond et al. (1996).
- Bovenberg and Heijdra (1998) and Heijdra and Ligthart (2001, 2002) on the OLG model.

Part II

Normative economics

Chapter 8

Introduction to normative public economics

The purpose of this chapter is to discuss the following topics:

- What is normative public economics and how does it differ from positive public economics?
- The two fundamental theorems of welfare economics.
- Social welfare function.
- First-best versus second-best.

8.1 Introduction

Up to this point in the book the focus has been on *positive economics*, i.e. we have studied the economy *as it is* (rather than *as it should be*). In particular, we have discussed a number of key issues in positive economics. For example, in Chapter 2 we studied how a household determines consumption and labour supply in the presence of a tax system, in Chapter 3 we looked at the effects of taxation on household savings decisions, and in Chapter 4 we extended the discussion by incorporating the effects of uncertainty on household consumption and portfolio decisions. In Chapter 5 we studied the effect of the corporate tax system on the typical firm's real and financial decisions, and in Chapters 6 and 7 we looked at the general equilibrium repercussions of taxation under perfect and imperfect competition. Finally, in Chapter 8 the effect of taxation on the economic growth process was studied. In all of these chapters we took the existing tax system for granted and determined how households, firms, and the economy as a whole react to this system.

In the next set of chapters we change tack and look at questions in *normative economics*. What shape should the tax system have? We start answering this question by studying the problem of optimal *indirect taxation* in Chapter 10. The issue there is to determine the socially optimal commodity tax system. Should luxuries be taxed heavily or not? Should necessities be taxed at all, or should they even be subsidized? Those are some of the issues discussed in that chapter.

In Chapter 11 we study the optimal structure of *direct taxation*. The key issue in the determination of the optimal income tax system is the amount of information the policy maker is assumed to possess. Finally, in Chapter 12 we study so-called public goods and externalities. Which goods should be provided by the government, and which by the private sector? What is the appropriate way to provide public goods under various tax systems? How should the government react to external effects?

The common theme in Chapters 10-12 is thus an explicit welfare-theoretic stance. We must be able to somehow provide a social ranking between alternative states the economy could be in. Central in normative economics is the so-called *Social Welfare Function* (SWF) first proposed by Bergson (1938) and Samuelson (1947). A typical example of an *individualistic* SWF is:

$$SW = \Psi\left(U^1, U^2, \dots, U^H\right), \tag{A5.1}$$

where SW is an indicator for social welfare, U^h is the utility level of individual household h (where $h = 1, 2, \dots, H$), and $\Psi(\cdot)$ is some function featuring positive partial derivatives, i.e. $\Psi_h \equiv \partial\Psi/\partial U^h > 0$ for all h (every household has at least some weight in social welfare). Note that we call the SWF in (A5.1) *individualistic* because its only arguments are the individual utility functions. A comparison between alternatives states only involves individual utility levels in those states. The SWF confers social (or ethical) weights on the different individuals and, since $\Psi_h > 0$, social welfare rankings made with the SWF honour the *Pareto Principle*: if one individual's utility level increases (decreases) and all other individual utility levels are held constant, then SW increases (decreases). The SWF goes beyond the Pareto Principle, however, because it assumes that gains and losses can be compared. In that sense it also goes beyond the "New Welfare Economics" of the 1930s which denied the validity of making interpersonal comparisons. In defense of the SWF approach one could argue that in most cases the Pareto Principle simply does not provide sufficient guidance for public policy decisions as it contains an inherent bias to maintain the status quo.

Different types of social welfare functions have been used in the literature. In a so-called *Benthamite SWF*, social welfare is the *sum* of individual utility levels (or a positive linear transformation of this sum):[1]

$$SW = \sum_{h=1}^{H} U^h. \tag{A5.2}$$

Technically, in this type of formulation, every household has the same (constant) weight in social wel-

[1]This form of the social welfare function is named after the classical economist Jeremy Bentham (1748-1832). To him is attributed the idea that "it is the greatest happiness of the greatest number that is the measure of right and wrong" (cited by Harrison (1987, p. 226)).

fare. At the other extreme, in a so-called *Rawlsian SWF*, social welfare is the welfare of the worst-off individual ("maxi-min"):[2]

$$SW = \min_h \left[U^1, U^2, \dots, U^H\right]. \tag{A5.3}$$

In this formulation only the worst-off individual has a positive weight in social welfare. Atkinson and Stiglitz (1980, p. 340) suggest the following generalized iso-elastic SWF:

$$SW = \sum_{h=1}^{H} \frac{\left(U^h\right)^{1-1/\zeta} - 1}{1 - 1/\zeta}, \tag{A5.4}$$

where ζ (≥ 0) measures the degree of substitutability between the U^h's. This formulation is quite convenient because it nests several special case. Indeed, by letting $\zeta \to 0$ in (A5.4) we obtain the Rawlsian SWF (A5.3), for $\zeta = 1$ we obtain a logarithmic SWF, and for $\zeta \to \infty$ we obtain the Benthamite SWF (A5.2). For pragmatic reasons we will often use the iso-elastic formulation (A5.4) below.

8.2 Brief overview of welfare economics

In order to prepare for things to come, this section presents a brief and selective overview of basic welfare economics.[3] We use the general equilibrium model of Samuelson (1947) as presented by Tresch (2002, ch. 2). The mathematical model is kept fairly general. There are H households, F production factors, and G goods. The key elements of the model are individual preferences, firm production technologies, and market clearing.

Individual preferences of household h are defined as follows:

$$U^h = U^h\left(X_1^h, \dots, X_G^h, V_1^h, \dots, V_F^h\right), \tag{A5.5}$$

where U^h is utility of household h ($h = 1, 2, \dots, H$), X_g^h is consumption of good g by household h ($g = 1, 2, \dots, G$), and V_f^h is the supply of production factor f by household h ($f = 1, 2, \dots, F$).

The production technology for good g is assumed to be given by:

$$Y_g = F^g\left(Z_1^g, \dots, Z_F^g\right), \tag{A5.6}$$

where Y_g is the aggregate production of good g, $F^g(\cdot)$ is the production function for good g, and Z_f^g is factor f used in the production of good g.

[2] This type of SWF function is named in honour of the American philosopher, John Rawls (1921-2002). In Rawls (1971) he derives the maxi-min rule as the outcome of decision making under a "veil of ignorance" (about one's own position).

[3] Interested readers are referred to Nath (1969), Boadway and Bruce (1984), Myles (1995), and Mas-Colell, Whinston, and Green (1995) for further details.

The market clearing conditions are as follows. For the goods markets the clearing conditions are:

$$Y_g = \sum_{h=1}^{H} X_g^h, \qquad (\text{for } g = 1, 2, \ldots, G), \tag{A5.7}$$

where the left-hand side is aggregate production and the right-hand side is total demand. Since there are G different goods markets, there are in total G goods market clearing conditions.

The factor market clearing conditions are:

$$\sum_{h=1}^{H} V_f^h = \sum_{g=1}^{G} Z_f^g, \qquad (\text{for } f = 1, 2, \ldots, F), \tag{A5.8}$$

where the left-hand side is the total supply of factor f and the right-hand side is the total demand for that factor. Note, there are F factors and thus F factor market clearing equations.

The general equilibrium model consists of equations (A5.5)-(A5.8).

8.2.1 Efficiency: Pareto optimality

An important concept in welfare economics is that of *Pareto optimality*. A given allocation of resources is Pareto optimal if no one consumer can be made better off by a reallocation of resources without at the same time making at least one other consumer worse off. The locus of Pareto-optimal allocations defines the so-called *Utility Possibility Frontier* (UPF). In Figure 8.1 such a UPF is illustrated for the case with only two households ($H = 2$). Note that the UPF may have a very irregular shape even if individual utility functions are well-shaped. This is because it is an envelope around infinitely many so-called *Point UPF*'s which are each associated with a particular point of the Production Possibility Frontier, i.e. a particular bundle of goods (see, e.g., Nath (1969, p. 20)).

In terms of Figure 8.1, point A is a Pareto optimal point (as are all other points on the UPF by definition). Points within the UPF are not Pareto optimal. At point C, for example, it is possible to keep household 1 equally well of and make household 2 strictly better off (vertical move). Conversely, it is possible to move horizontally and make household 1 strictly better off whilst keeping household 2's utility the same. In fact any move in north-easterly direction makes both households strictly better off. Finally, points which lie beyond the UPF are unattainable (e.g. point B).

In formal terms, the set of all Pareto optimal allocations can be characterized as follows. We focus on an arbitrary household, say household $h = 1$, and hold every other household's utility constant, i.e. $U^h = U_0^h$ for $h = 2, \ldots, H$. Then we maximize household 1's utility subject to the restrictions, i.e. the social planner chooses X_g^h, V_f^h, Z_f^g, Y_g in order to maximize:

$$U^1 = U^1\left(X_1^1, \ldots, X_G^1, V_1^1, \ldots, V_F^1\right), \tag{A5.9}$$

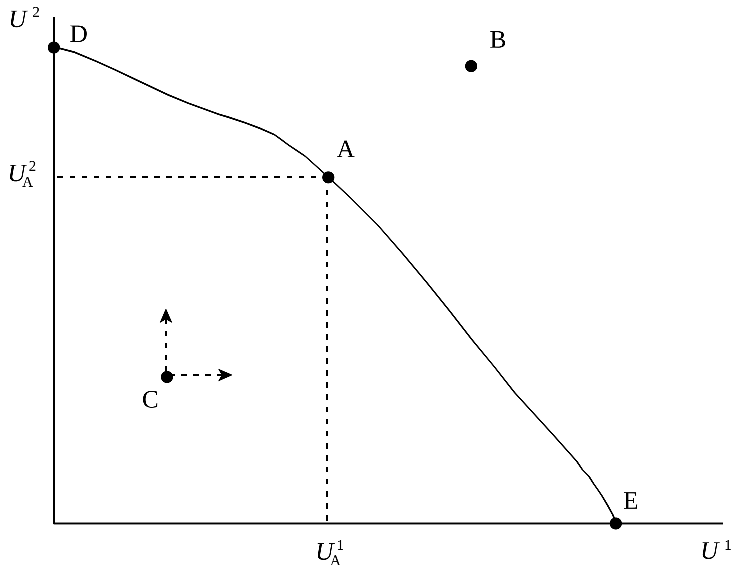

Figure 8.1: The utility possibility frontier

subject to (A5.6)-(A5.8) and:

$$U_0^h = U^h\left(X_1^h, \ldots, X_G^h, V_1^h, \ldots, V_F^h\right), \qquad (\text{for } h = 2, \ldots, H). \tag{A5.10}$$

Equation (A5.10) is the set of constraints that keeps all households other than household 1 equally well off, (A5.6) represents the constraints imposed by existing technology, (A5.7) ensures goods market clearing, and (A5.8) ensures factor market clearing.

The Lagrangian for this social optimization problem is defined as follows:

$$\begin{aligned}\mathcal{L} \equiv\ & \lambda_1 U^1\left(X_1^1, \ldots, X_G^1, V_1^1, \ldots, V_F^1\right) \\ & + \sum_{h=2}^{H} \lambda_h \left[U^h\left(X_1^h, \ldots, X_G^h, V_1^h, \ldots, V_F^h\right) - U_0^h\right] + \sum_{g=1}^{G} \mu_g \left[Y_g - F^g\left(Z_1^g, \ldots, Z_F^g\right)\right] \\ & + \sum_{g=1}^{G} \nu_g \left[\sum_{h=1}^{H} X_g^h - Y_g\right] + \sum_{f=1}^{F} \xi_f \left[\sum_{h=1}^{H} V_f^h - \sum_{g=1}^{G} Z_f^g\right],\end{aligned} \tag{A5.11}$$

where the Lagrange multipliers are λ_h (for $h = 2, \ldots, H$), μ_g (for $g = 1, \ldots, G$), ν_g (for $g = 1, \ldots, G$), and ξ_f (for $f = 1, \ldots, F$). In total equation (A5.11) thus features $(H-1) + 2G + F$ Lagrange multipliers. Note that in the first line of (A5.11) we define an auxiliary variable for household 1, λ_1, which we set equal to unity ($\lambda_1 = 1$). This is done to cut down on notation below.

The first-order necessary conditions (assuming an interior solution) are the constraints and the following. For the goods demands there are $G \times H$ equations:

$$\frac{\partial \mathcal{L}}{\partial X_g^h} = \lambda_h \frac{\partial U^h}{\partial X_g^h} + \nu_g = 0, \tag{A5.12}$$

for the factor supplies there are $H \times F$ equations:

$$\frac{\partial \mathcal{L}}{\partial V_f^h} = \lambda_h \frac{\partial U^h}{\partial V_f^h} + \xi_f = 0, \tag{A5.13}$$

for the output decisions there are G equations:

$$\frac{\partial \mathcal{L}}{\partial Y_g} = \mu_g - \nu_g = 0, \tag{A5.14}$$

and for the factor demands there are $F \times G$ equations:

$$\frac{\partial \mathcal{L}}{\partial Z_f^g} = -\mu_g \frac{\partial F^g}{\partial Z_f^g} - \xi_f = 0. \tag{A5.15}$$

Although these conditions look rather forbidding, we can eliminate the various Lagrange multipliers and derive the condensed statement of the necessary conditions for the Paretian optimum. We look at the following pairings. First, for any given arbitrary household h we derive:

$$\frac{\nu_{g_1}}{\nu_{g_2}} = \frac{\partial U^h / \partial X_{g_1}^h}{\partial U^h / \partial X_{g_2}^h}, \tag{A5.16}$$

$$\frac{\xi_{f_1}}{\xi_{f_2}} = \frac{\partial U^h / \partial V_{f_1}^h}{\partial U^h / \partial V_{f_2}^h}, \tag{A5.17}$$

$$\frac{\xi_f}{\nu_g} = \frac{\partial U^h / \partial V_f^h}{\partial U^h / \partial X_g^h}. \tag{A5.18}$$

Equation (A5.16) is obtained by using (A5.12) for any two goods g_1 and g_2. It says that the marginal rate of substitution (MRS) between any two consumption goods g_1 and g_2 (the right-hand side) is the same for all households h (since the left-hand side does not feature an h-index). Graphically this condition can be illustrated for the case with two goods ($G = 2$) and two households ($H = 2$) with the aid of Figure 8.2. In that figure the total availability of the two goods is held constant. The origin for household 1 is O_1 and that for household 2 is O_2. The slope of each indifference curve represents the MRS between the two goods at that point. Points A, B, C, and D are such that the indifference curves have the same slope for the two households, i.e. they are all points of efficient exchange. Point E, in contrast, is inefficient in exchange because the MRS is not the same for both households. At point A, household 2 is equally well of as at E but household 1 is strictly better off. The dashed line connecting O_1 and O_2 is the contract curve, i.e. the locus of points that are efficient in exchange.

Equation (A5.17) is obtained by using (A5.13) for any two factors f_1 and f_2. It says that the MRS between any two supplied factors f_1 and f_2 (right-hand side) is the same for all households h. Finally, equation (A5.18) is obtained from (A5.12) and (A5.13). It shows that the MRS between any supplied factor f and any consumption good g is the same for all households h. In principle one could draw

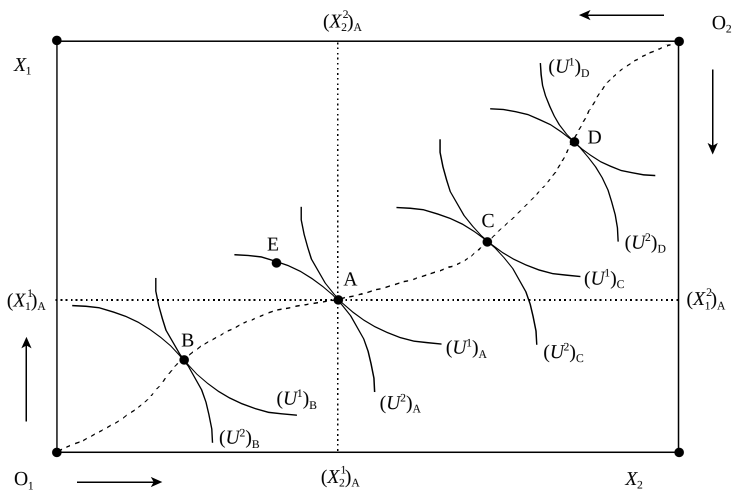

Figure 8.2: Efficient exchange

diagrams similar to Figure 8.2 to visualize equations (A5.17) and (A5.18) but this is left as an exercise to the reader.

Second, on the production side we derive:

$$1 = \frac{\mu_{g_1}}{\mu_{g_2}} \frac{\partial F^{g_1}/\partial Z_f^{g_1}}{\partial F^{g_2}/\partial Z_f^{g_2}} = \frac{\nu_{g_1}}{\nu_{g_2}} \frac{\partial F^{g_1}/\partial Z_f^{g_1}}{\partial F^{g_2}/\partial Z_f^{g_2}} \tag{A5.19}$$

$$\frac{\xi_{f_1}}{\xi_{f_2}} = \frac{\partial F^g/\partial Z_{f_1}^g}{\partial F^g/\partial Z_{f_2}^g}. \tag{A5.20}$$

Equation (A5.19) is obtained by using (A5.14) and (A5.15) for any two goods g_1 and g_2. It requires the social value of the marginal product of factor f to be the same for any two goods g_1 and g_2. Equation (A5.20) is derived by using (A5.15) for any two factors f_1 and f_2. It requires the *marginal rate of technical substitution* (MRTS) between any two factors f_1 and f_2 to be the same for all goods g. Graphically condition (A5.20) can be illustrated for the case with two factors ($F = 2$) and two goods ($G = 2$) with the aid of Figure 8.3. In that figure total supplies of the two factors is held constant. The origin for good 1 is O_1 and that for good 2 is O_2, and the factors are labour (L) and capital (K). The slope of each isoquant represents the MRTS between the two factors at that point. Points A, B, C, and D are such that the isoquants have the same slope for the two goods, i.e. they are all points of efficient production. Point E, in contrast, is inefficient in production because the MRTS is not the same for both goods. By shifting from point E to point A, production of good 2 is kept constant but production of good 1 is increased. The dashed line connecting O_1 and O_2 is the contract curve, i.e. the locus of points that are efficient in production. This contract curve can also be plotted directly in goods space as in Figure 8.4 In that figure, points B and D are both efficient in production, i.e. they both lie on the transformation curve in Figure 8.4 and on the contract curve in Figure 8.3.

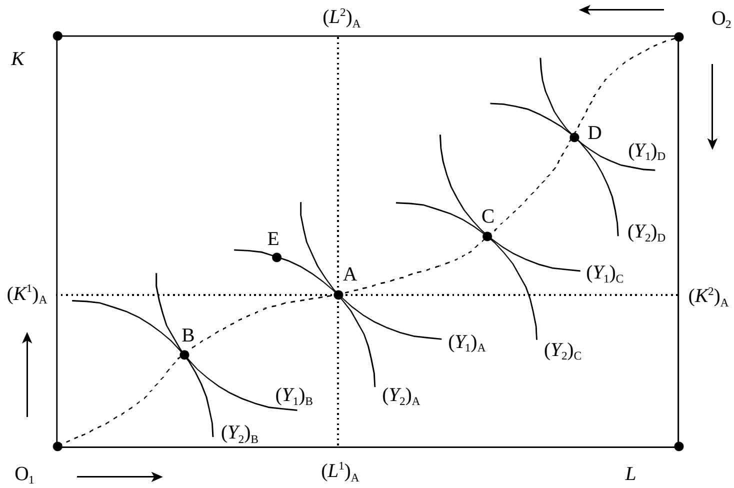

Figure 8.3: Efficient production

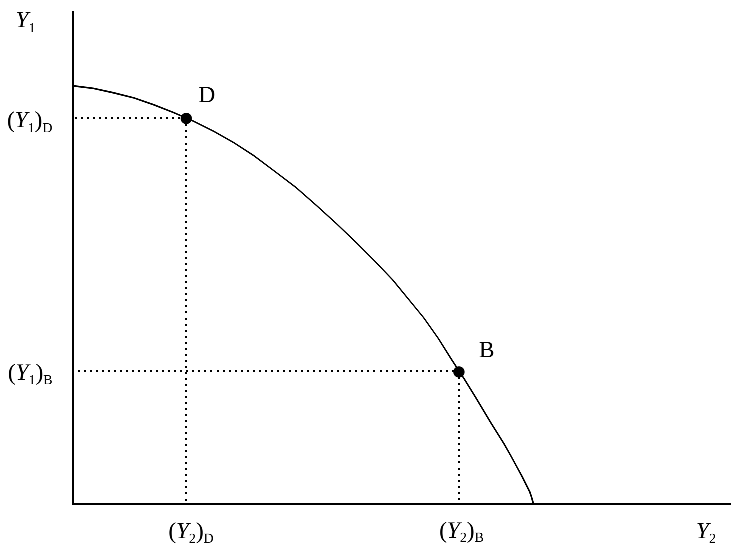

Figure 8.4: Transformation curve

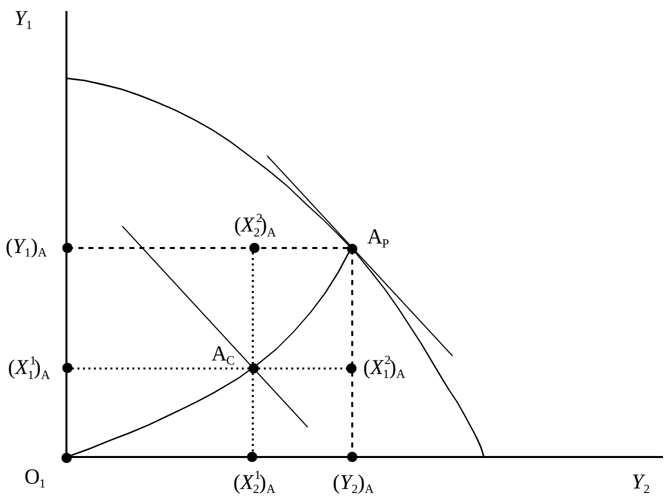

Figure 8.5: Efficient exchange and production

Third, we can now combine even further. By using (A5.16) and (A5.19) we find:

$$\left[\frac{\nu_{g_1}}{\nu_{g_2}} =\right] \quad \frac{\partial U^h / \partial X_{g_1}^h}{\partial U^h / \partial X_{g_2}^h} = \frac{\partial F^{g_2} / \partial Z_f^{g_2}}{\partial F^{g_1} / \partial Z_f^{g_1}}, \tag{A5.21}$$

i.e. the MRS between goods g_1 and g_2 (left-hand side of (A5.21)) must be equated to the *marginal rate of transformation* (MRT) between these two goods (right-hand side). Graphically, the marginal rate of transformation is represented by (minus) the slope of the transformation curve as drawn in Figure 8.4. Intuitively, it represents the marginal rate at which one good can be transformed into another good by re-allocating any factor.[4] The condition (A5.21) has been illustrated in Figure 8.5 for the case of two goods and two households. The efficient production point is on the transformation curve at point A_P and the efficient consumption point is on the contract curve O_1A_P, say at point A_C (the indifference curves tangent at that point are not shown to avoid cluttering the diagram). The lines through A_C and A_P are tangent (to each other and to the transformation curve). Total production is $(Y_1, Y_2)_A$, household 1 consumes $\left(X_1^1, X_2^1\right)_A$ and household 2 consumes $\left(X_1^2, X_2^2\right)_A$.

Fourth, from (A5.17) and (A5.20) we find:

$$\left[\frac{\xi_{f_1}}{\xi_{f_2}} =\right] \quad \frac{\partial U^h / \partial V_{f_1}^h}{\partial U^h / \partial V_{f_2}^h} = \frac{\partial F^g / \partial Z_{f_1}^g}{\partial F^g / \partial Z_{f_2}^g}, \tag{A5.22}$$

[4]Technically, we get from the first equality in (A5.19) that:

$$\frac{\partial F^{g_2} / \partial Z_f^{g_2}}{\partial F^{g_1} / \partial Z_f^{g_1}} = \frac{\mu_{g_1}}{\mu_{g_2}},$$

where the right-hand side is independent of which factor is shifted. Hence, the left-hand side is the MRT between goods g_1 and g_2.

i.e. for any good g, the common MRS between any two supplied factors f_1 and f_2 must be equated to the marginal rate of technical substitution (MRTS) between these two factors in production of that good.

Finally, from (A5.18) and (A5.14)-(A5.15) we find:

$$\left[\frac{\xi_f}{\nu_g} =\right] \quad \frac{\partial U^h / \partial V_f^h}{\partial U^h / \partial X_g^h} = -\frac{\partial F^g}{\partial Z_f^g}, \tag{A5.23}$$

i.e. the common MRS between any consumed good g and any supplied factor f is equal to the marginal product of that factor f in producing that good g.

Of course, the first-order necessary conditions stated above solve for just a single point on the UPF, namely that point conditional upon the utility levels held constant for households $h = 2, \ldots, H$. Hence, by changing one of these household's utility level we can re-derive the Pareto optimal allocation and obtain another point of the UPF (assuming feasibility etcetera). But there are infinitely many ways to do this so what we end up with is all the points on the UPF. The crucial thing to note is that the planning problem does not pin down a *single optimal point* on the UPF. By definition *all* allocations on the UPF are Pareto optimal. It is clear that Pareto optimality is *too weak* a criterion for public decision making. Using this principle one cannot even decide between points like D and A (or E and A) in Figure 8.1. Most people in the street would probably agree that giving some utility to both agents (point A) is to be preferred to giving all to either person 1 (point E) or person 2 (point D). But the Pareto principle does not give us that policy prescription.

8.2.2 Equity: SWF and the optimal distribution

By postulating a SWF, the social planning problem does provide conditions for both efficiency and optimal distributional choices. Assume that the SWF is written in general format as in (A5.1) above. The social planner now chooses X_g^h, V_f^h, Z_f^g, Y_g (for $f = 1, \ldots, F$, $g = 1, \ldots, G$, and $h = 1, \ldots, H$.) in order to maximize social welfare (A5.1) subject to the individual utility functions (A5.5), the production technology (A5.6), the goods market clearing conditions (A5.7), and the factor market clearing conditions (A5.8). The Lagrangian for this problem is:

$$\begin{aligned}
\mathcal{L} \equiv\ & \Psi\Big[U^1\left(X_1^1, \ldots, X_G^1, V_1^1, \ldots, V_F^1\right), \ldots, U^H\left(X_1^H, \ldots, X_G^H, V_1^H, \ldots, V_F^H\right)\Big] \\
& + \sum_{g=1}^{G} \mu_g \left[Y_g - F^g\left(Z_1^g, \ldots, Z_F^g\right)\right] \\
& + \sum_{g=1}^{G} \nu_g \left[\sum_{h=1}^{H} X_g^h - Y_g\right] + \sum_{f=1}^{F} \xi_f \left[\sum_{h=1}^{H} V_f^h - \sum_{g=1}^{G} Z_f^g\right],
\end{aligned} \tag{A5.24}$$

where the Lagrange multipliers are μ_g (for $g = 1, \ldots, G$), ν_g (for $g = 1, \ldots, G$), and ξ_f (for $f = 1, \ldots, F$), i.e. there are $2G + F$ Lagrange multipliers in all.

The first-order necessary conditions (assuming an interior solution) are the constraints and (a) for the good demands ($G \times H$ equations):

$$\frac{\partial \mathcal{L}}{\partial X_g^h} = \frac{\partial \Psi}{\partial U^h} \frac{\partial U^h}{\partial X_g^h} + \nu_g = 0, \tag{A5.25}$$

(b) for the factor supplies ($H \times F$ equations):

$$\frac{\partial \mathcal{L}}{\partial V_f^h} = \frac{\partial \Psi}{\partial U^h} \frac{\partial U^h}{\partial V_f^h} + \xi_f = 0, \tag{A5.26}$$

(c) for the output decisions (G equations):

$$\frac{\partial \mathcal{L}}{\partial Y_g} = \mu_g - \nu_g = 0, \tag{A5.27}$$

and (d) for the factor demands ($F \times G$ equations):

$$\frac{\partial \mathcal{L}}{\partial Z_f^g} = -\mu_g \frac{\partial F^g}{\partial Z_f^g} - \xi_f = 0. \tag{A5.28}$$

In principle we could redo the different pairings leading to (A5.16)-(A5.23). Fortunately, this is not necessary because these conditions are all still valid, *the social planner selects a point on the UPF*! This is what is meant by the notion of *first-best welfare analysis*. The policy maker has a sufficient number of suitable policy tools by which it can select any point along the UPF (Tresch, 2002, p. 67).

The only interpersonal equity conditions are obtained by using (A5.25) and (A5.26) and comparing different households, say $h_1$and h_2. From (A5.25) we find for any good g the following condition must hold:

$$[-\nu_g =] \quad \frac{\partial \Psi}{\partial U^{h_1}} \frac{\partial U^{h_1}}{\partial X_g^{h_1}} = \frac{\partial \Psi}{\partial U^{h_2}} \frac{\partial U^{h_2}}{\partial X_g^{h_2}}. \tag{A5.29}$$

According to (A5.29), interpersonal equity is achieved if all goods are distributed such that on the margin the increase in social welfare is the same no matter who consumes the last unit of the good. Similarly, by using (A5.26) we find that for any factor f the following condition must hold:

$$\left[-\xi_f =\right] \quad \frac{\partial \Psi}{\partial U^{h_1}} \frac{\partial U^{h_1}}{\partial V_f^{h_1}} = \frac{\partial \Psi}{\partial U^{h_2}} \frac{\partial U^{h_2}}{\partial V_f^{h_2}}. \tag{P10}$$

Hence, all factor supplies should be set such that on the margin the increase in social welfare is the same no matter who supplies the last unit of the factor.

In Figure 8.6 the determination of the first-best social optimum is illustrated for the two household case. Suppose that the minimal-government solution would be at point A, where the distribution of welfare is rather uneven. The associated level of social welfare is SW_A. Although point A is Pareto-

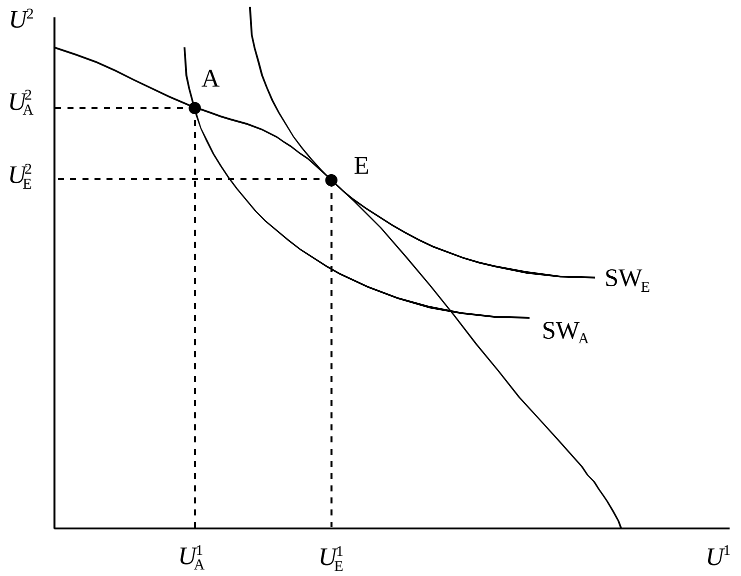

Figure 8.6: The first-best social optimum

optimal, the social planner can improve the social outcome by choosing a tangency between the UPF and a SWF. This tangency occurs at point E, where social welfare is SW_E (greater than SW_A). How is the redistribution engineered? The only way it can be done is if the social planner has access to *lump-sum redistribution instruments* (i.e. taxes/subsidies which do not themselves introduce inefficiencies into the economy and thus keep the economy on the UPF).

8.2.3 Basic theorems of welfare economics

The basic theorems of welfare economics provide important guidance concerning the link between the notion of Pareto efficiency and the outcome produced in the decentralized market economy. Atkinson & Stiglitz (1980, p. 343) state these theorems as follows:

First Theorem of Welfare Economics: If (i) households and firms are perfect competitors and thus take prices for all goods and factors as given, (ii) there is a full set of markets, and (iii) there is perfect information, then a competitive equilibrium (if it exists) is Pareto efficient.

Second Theorem of Welfare Economics: If (i) household indifference maps and firm production sets are convex, (ii) there is a full set of markets, (iii) there is perfect information, and (iv) lump-sum taxes/transfers can be carried out costlessly, then any Pareto-efficient allocation can be achieved as a competitive equilibrium with appropriate lump-sum transfers and taxes.

The importance of the availability of lump-sum taxes and transfers cannot be overstated. Indeed if such taxes and transfers are not available (or not sufficiently flexible) then one enters the much more complex (but also more realistic) realm of *second-best welfare economics*. In such an analysis there is generally a conflict between achieving efficiency (reaching the UPF) and equity (achieving a fair distribution

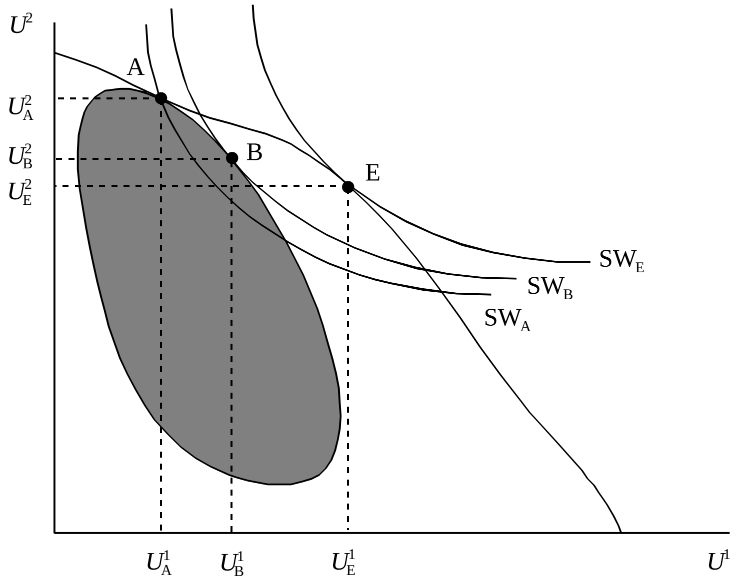

Figure 8.7: The second-best social optimum

of welfare). Figure 8.7, which is adapted from Tresch (2002, p. 73), shows in heuristic terms what is going on. Say the initial position is at point A. If the policy maker has the required lump-sum instruments he can ensure that the first-best equilibrium at E can be attained (first-best case). In contrast, if the policy maker does not have the (sufficiently flexible) lump-sum instruments, the effective set of allocations that can be reached is represented by the shaded area. Of course point A is in that area but point E is not. The *second-best social optimum* is then represented by point B, i.e. the point that can be reached from the initial situation A *and* has a higher social welfare level than at point A ($\text{SW}_B > \text{SW}_A$ but $\text{SW}_B < SW_E$). As we shall see time and again in the coming chapters, much of actual policy making is concerned with second-best situations.

We close this subsection by demonstrating the consequences of the second welfare theorem within the context of our simple general equilibrium model. The task at hand is to show how the competitive decentralized economy gives rise to exactly the same condition that determine Pareto efficiency. The decentralized economy has the following key features. First, household h has the utility function (A5.5) and faces the following budget restriction:

$$\sum_{g=1}^{G} P_g X_g^h + T^h = \sum_{f=1}^{F} W_f V_f^h, \tag{A5.30}$$

where P_g is the market price of good g, T^h is a household-specific lump-sum tax (or transfer), and W_f is the market price of factor f. Second, the representative firm producing good g is a price taker in its output market and all its input markets and has the following profit function:

$$\Pi_g \equiv P_g Y_g - \sum_{f=1}^{F} W_f Z_f^g, \tag{A5.31}$$

where Π_g is profit and where technology is given by (A5.6) above. Third, it is assumed that all markets clear, i.e. both (A5.7) and (A5.8) hold.

It can now be shown that the decentralized optimizing decisions by households and firms give rise to the first-order conditions (A5.16)-(A5.23) stated above. Household h chooses X_g^h and V_f^h in order to maximize utility (A5.5) subject to the budget constraint (A5.30), taking as given P_g, W_f, and T^h. The key first-order conditions are:

$$\frac{P_{g_1}}{P_{g_2}} = \frac{\partial U^h / \partial X_{g_1}^h}{\partial U^h / \partial X_{g_2}^h}, \tag{A5.32}$$

$$\frac{-W_{f_1}}{-W_{f_2}} = \frac{\partial U^h / \partial V_{f_1}^h}{\partial U^h / \partial V_{f_2}^h}, \tag{A5.33}$$

$$\frac{-W_f}{P_g} = \frac{\partial U^h / \partial V_f^h}{\partial U^h / \partial X_g^h}. \tag{A5.34}$$

Equations (A5.32)-(A5.34) are the market-based counterparts to (A5.16)-(A5.18) above, with P_g replacing ν_g and $-W_f$ replacing ξ_f.

The firm producing good g chooses Y_g and Z_f^g in order to maximize profit (A5.31) subject to the production function (A5.6) and taking as given P_g and W_f. The key first-order conditions for the firm producing good g are:

$$P_g \frac{\partial F^g}{\partial Z_f^g} = W_f. \tag{A5.35}$$

Using (A5.35) for a single factor and any two goods we obtain the counterpart to (A5.19):

$$\left[\frac{W_f}{W_f} = \right] 1 = \frac{P_{g_1} \partial F^{g_1} / \partial Z_f^{g_1}}{P_{g_2} \partial F^{g_2} / \partial Z_f^{g_2}}, \tag{A5.36}$$

where P_g again replaces ν_g. Similarly, by using (A5.35) for a single good and any two factors we obtain the counterpart to (A5.20):

$$\frac{-W_{f_1}}{-W_{f_2}} = \frac{\partial F^g / \partial Z_{f_1}^g}{\partial F^g / \partial Z_{f_2}^g}, \tag{A5.37}$$

where $-W_f$ replacing ξ_f. The counterparts to (A5.21)-(A5.23) are now already established. It follows that the competitive market solution satisfies exactly the efficiency conditions for a Pareto optimum. Note that the transfers, T^h, do not feature in any of the first-order conditions (A5.32)-(A5.37). Hence, the policy maker can use them for redistributive purposes to select a particular point on the UPF.

Key literature

- Atkinson & Stiglitz (1980, lecture 11) or Jha (1998, chs. 2-3) on theory.
- Tresch (2002) and Myles (1995, ch. 2) on theory.
- Bergson (1938), Little (1957), Samuelson (1947), and Stiglitz (1987) on the social welfare function.
- Leach (2004) on public goods, Barr (2004) on SWF.
- Second-best theory: Lipsey and Lancaster (1956), McManus (1959).

Chapter 9

The structure of indirect taxation

The purpose of this chapter is to discuss the following topics in the theory of optimal commodity taxation:

- How should different commodities be taxed? Uniformly or at different rates?
- If at different rates, which commodities should be taxed heavily and why? Which ones lightly?
- Can we derive clear and unambiguous prescriptions for policy reform?

9.1 Introduction

In this chapter we discuss the rather large literature on optimal commodity taxation. The chapter proceeds as follows. First, we build some intuition based on simple partial equilibrium reasoning. In a partial equilibrium setting there exists a simple and intuitive *inverse-elasticity result*, linking the optimal tax on a commodity to the demand elasticity of that commodity. Next we expand the model by taking into account the general equilibrium interactions between markets. This brings us in the realm of the *Ramsey optimal taxation* approach. The optimal tax rule is much less straightforward to interpret and we build intuition by looking at some special cases. Finally, we look at the issue of marginal tax reform and its effect on household welfare.

The maintained assumptions made throughout this chapter are as follows. First, we assume that the policy maker does not have access to lump-sum taxes/subsidies, i.e. the analysis here is an example of second-best welfare analysis (see Chapter 9). Second, throughout this chapter it is assumed that producer prices are fixed. Third, in most of this chapter we deal with the case of identical households. By focussing on this representative-agent case, we can ignore distributional issues ("equity") and focus on efficiency. Furthermore, we do not need to postulate an explicit social welfare function.

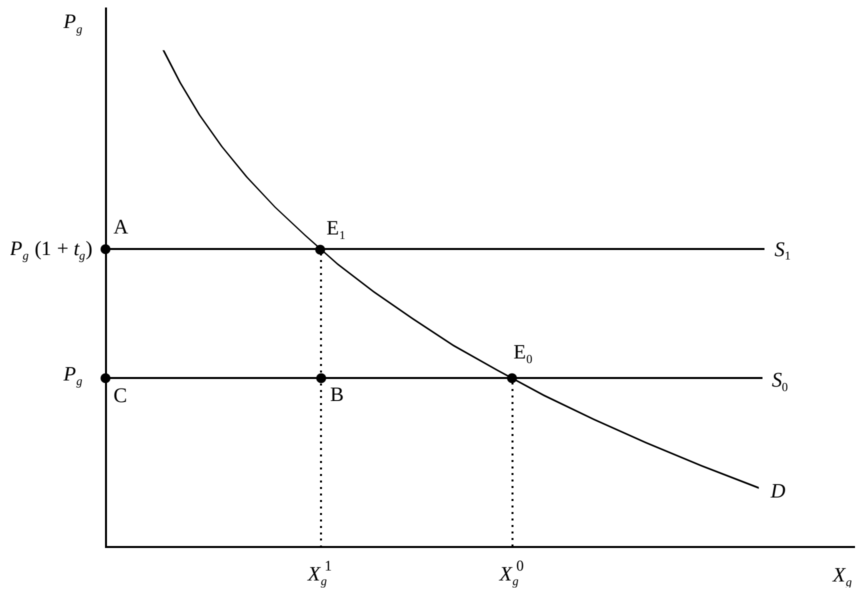

Figure 9.1: Excess burden of taxation

9.2 Indirect taxation in partial equilibrium

In order to build up some intuition behind the complex general equilibrium models to come, this section deals with optimal taxation in a partial equilibrium framework. We focus on a single commodity, assume away all cross-price effects (i.e. the demand for the good in question is independent of the prices of all other goods), ignore the income effect in demand, and assume that the producer price is fixed.

In Figure 9.1 the market for good g is illustrated. The quantity of good g is denoted by X_g and the (fixed) producer price of the good is P_g. The downward sloping demand curve, D, depends only on the price paid by the consumer. S_0 is the supply curve in the absence of the tax (horizontal), and the initial equilibrium is at point E_0. Next we consider the introduction of an *ad valorem* tax on good g, which we denote by t_g. This tax ensures that the price to consumers is $P_g\left(1+t_g\right)$ and the new (tax-inclusive) supply curve is S_1. The new equilibrium is at point E_1.

The welfare effects of the tax can be deduced as follows. On the one hand consumer surplus falls by the area AE_1E_0C, but on the other hand the government raises the tax revenue AE_1BC. It follows that the net effect of the tax, the so-called the *Excess Burden* (EB), is equal to the loss of consumer surplus minus the tax revenue raised.[1] The excess burden is equal to the area E_1E_0B in Figure 9.1. In formal terms we can measure the excess burden associated with the tax on good g as follows:

$$EB_g \equiv \int_{X_g^1}^{X_g^0} Q_g(X_g)dX_g - P_g\left[X_g^0 - X_g^1\right], \tag{A5.1}$$

where EB_g is the excess burden and $Q_g(X_g)$ is the *inverse* demand function, expressing the consumer price Q_g as a function of the quantity demanded. The first term on the right-hand side of (A5.1) repre-

[1]Supply is horizontal so there is no producer surplus.

sents the area under the demand curve between X_g^1 and X_g^0, whilst the second term on the right-hand side is the rectangular area $X_g^1 X_g^0 E_0 B$.

The effects of the tax on the excess burden can be determined by differentiating (A5.1):

$$\begin{aligned}\frac{\partial EB_g}{\partial t_g} &\equiv \frac{\partial}{\partial t_g}\left[\int_{X_g^1}^{X_g^0} Q_g(X_g)dX_g - P_g\left[X_g^0 - X_g^1\right]\right] \\ &= -Q_g(X_g^1)\frac{\partial X_g^1}{\partial t_g} + P_g\frac{\partial X_g^1}{\partial t_g} \\ &= \left[-P_g\left(1+t_g\right)+P_g\right]\frac{\partial X_g^1}{\partial t_g} \\ &= -P_g t_g \frac{\partial X_g^1}{\partial t_g}, \end{aligned} \tag{A5.2}$$

where we have used *Leibnitz' Rule*[2] for differentiating integrals in going from the first to the second line, and noted that $Q_g(X_g^1) = P_g\left(1+t_g\right)$ in going from the second to the third line. The expression in (A5.2) already gives us an important result, namely that the excess burden of a tax is zero for infinitesimal taxes (i.e. if the initial tax rate is zero). The distortion due to the introduction of a small tax is of second-order magnitude.

Armed with these insights, we can now present a first view of the policy maker's optimal tax problem. There are G taxable commodities in total and the policy maker wishes to raise an exogenously given tax revenue, R_0, in the least distorting fashion, i.e. such that the overall excess burden of the commodity tax system is minimized (recall that by assumption there are no lump-sum taxes available to the policy maker!). Formally, the revenue requirement constraint is given by:

$$R_0 = \sum_{g=1}^{G} t_g P_g X_g^1, \tag{A5.3}$$

where the left-hand side is the exogenous required revenue and the right-hand side is the revenue from commodity taxes. The objective function of the policy maker is the total excess burden:

$$EB \equiv \sum_{g=1}^{G} EB_g, \tag{A5.4}$$

where the expression for EB_g is given in (A5.1) above.

[2]Suppose that the function $f(x)$ is defined as follows:

$$f(x) \equiv \int_{u_1(x)}^{u_2(x)} g(t,x)dt, \qquad a \le x \le b.$$

Then, if (i) $g(t,x)$ and $\partial g/\partial x$ are continuous in both t and x (in some region including $u_1 \le t \le u_2$ and $a \le x \le b$) and (ii) $u_1(x)$ and $u_2(x)$ are continuous and have continuous derivatives (for $a \le x \le b$), then df/dx is given by:

$$\frac{df(x)}{dx} = \int_{u_1(x)}^{u_2(x)} \frac{\partial g(t,x)}{\partial x}dt + g(u_2,x)\frac{du_2}{dx} - g(u_1,x)\frac{du_1}{dx}.$$

Often u_1 and/or u_2 are constants so that one or both of the last two terms on the right-hand side of this expression vanish.

The policy maker chooses $t_1, \cdots, t_G$ such that EB is minimized subject to the revenue requirement constraint (A5.3) and taking into account that EB_g and X_g^1 depend on t_g (see (A5.2) above). The Lagrangian expression for this minimization problem is:

$$\mathcal{L} \equiv \sum_{g=1}^{G} EB_g + \lambda \left[\sum_{g=1}^{G} t_g P_g X_g^1 - R_0 \right],$$

where λ is the Lagrange multiplier associated with the revenue requirement constraint (A5.3). The first-order necessary conditions are the constraint (A5.3) and:

$$\frac{\partial \mathcal{L}}{\partial t_g} = \frac{\partial EB_g}{\partial t_g} + \lambda P_g \left[X_g^1 + t_g \frac{\partial X_g^1}{\partial t_g} \right] = 0, \tag{A5.5}$$

for all $g = 1, \cdots, G$. By using equation (A5.2) we can rewrite (A5.5) as follows:

$$\begin{aligned} -\frac{\partial EB_g}{\partial t_g} &= \lambda P_g \left[X_g^1 + t_g \frac{\partial X_g^1}{\partial t_g} \right] \quad \Leftrightarrow \\ P_g t_g \frac{\partial X_g^1}{\partial t_g} &= \lambda P_g \left[X_g^1 + t_g \frac{\partial X_g^1}{\partial t_g} \right] \quad \Leftrightarrow \\ t_g \frac{\partial X_g^1}{\partial t_g} &= \lambda X_g^1 \left[1 + \frac{t_g}{X_g^1} \frac{\partial X_g^1}{\partial t_g} \right] \quad \Leftrightarrow \\ -\frac{t_g}{X_g^1} \frac{\partial X_g^1}{\partial t_g} &= \theta, \end{aligned} \tag{A5.6}$$

where $\theta \equiv \lambda / (1 + \lambda) > 0$ is a constant (involving the optimized value of the Lagrange multiplier). Since θ is constant and the same for all goods g, the rewritten first-order condition (A5.6) thus calls for an equalization of the *tax elasticity* of demand for all goods.

We define the *price elasticity* of demand for good g in the usual way as:

$$\varepsilon_g^D \equiv -\frac{Q_g}{X_g^1} \frac{\partial X_g^1}{\partial Q_g}, \tag{A5.7}$$

where $Q_g \equiv P_g \left(1 + t_g\right)$ is the demand price. Using (A5.7), the elasticity expression (A5.6) can be rewritten in terms of the tax rate:

$$\begin{aligned} -\frac{P_g \left(1 + t_g\right)}{X_g^1} \frac{\partial X_g^1}{\partial P_g \left(1 + t_g\right)} \frac{t_g}{1 + t_g} = \theta \quad &\Leftrightarrow \\ \frac{t_g}{1 + t_g} = \frac{\theta}{\varepsilon_g^D}. & \end{aligned} \tag{A5.8}$$

Several things are worth noting about (A5.8). First, since θ is the same for all goods, uniform taxation is optimal if (and only if) all goods have the same elasticity (so that ε_g^D is the same for all g). In any other case tax rates should be different for different commodities. Second, the optimal tax on a good should be higher, the less elastic that good is, i.e.

$$\frac{\partial t_g}{\partial \varepsilon_g^D} = -\frac{\theta}{\left(\varepsilon_g^D - \theta\right)^2} < 0. \tag{A5.9}$$

By corollary, if there exists a commodity with a very low (or even zero) price elasticity of demand, then the optimal tax rule calls for the tax on that commodity to be very high. Indeed, in the limiting case, with a vertical demand curve for some good, all revenue should be raised by taxing only that commodity (and leaving all other commodities untaxed).[3] Effectively, the policy maker would have a lump-sum tax at his disposal in that case.

The inverse-elasticity result is intuitive and can be illustrated informally with the aid of Figure 9.2. Assume there are only two types of goods, namely price-elastic goods (with the demand curve D_E) and relatively price-inelastic goods (featuring a demand curve like D_I). For convenience the producer price of the two types of goods goods is the same (and equal to P in the figure). Under uniform taxation, the tax rate is t and the tax-inclusive supply curve is S_C. The excess burden is represented by the area ABE_0 for the inelastic good (EB_I) and CFE_0 for the elastic good (EB_E). Uniform taxation is suboptimal because the demand elasticities are unequal. Assume that the optimal taxes (as set according to (A5.8)) are denoted by, respectively, t_I and t_E. Excess burden is now equal to $A'B'E_0$ for the inelastic good and $C'F'E_0$ for the elastic good. The increase in excess burden EB_I (the area $A'ABB'$) is more than compensated by the decrease in EB_E (the area $CC'F_0'F$). (By definition total tax revenue is the same under the two cases.)

The advantage of the partial equilibrium approach is that the results are intuitive and relatively easy to visualize. It must be stressed, however, that the inverse-elasticity formula is based on a number of very special (and highly restrictive) assumptions. This prompts the question concerning how the results should be modified if we allow for income effects and non-zero cross-price effects in demand. This question was first studied by the Cambridge mathematician-economist, Frank Ramsey.[4]

9.3 Indirect taxation in general equilibrium

The general equilibrium approach to optimal taxation was pioneered by Frank Ramsey (1927) and further developed by Samuelson (1986), Corlett and Hague (1953-54), Boiteux (1971), and Diamond and Mirrlees (1971). The model is based on the following key assumptions. First, all consumers are identical and the argument proceeds on the basis of a representative agent. Second, the goods demand functions and labour supply are all derived from the (indirect) utility function so both income and cross-price

[3]With a vertical demand curve there is no excess burden associated with the tax, i.e. the loss in consumer surplus is exactly matched by the increase in tax revenue.

[4]In Chapter 8 we studied the Ramsey growth model. Frank Ramsey died after an operation in 1930, a month short of his 27th birthday. In his brief life, he nevertheless managed to write two classic papers in economic theory, namely Ramsey (1927, 1928).

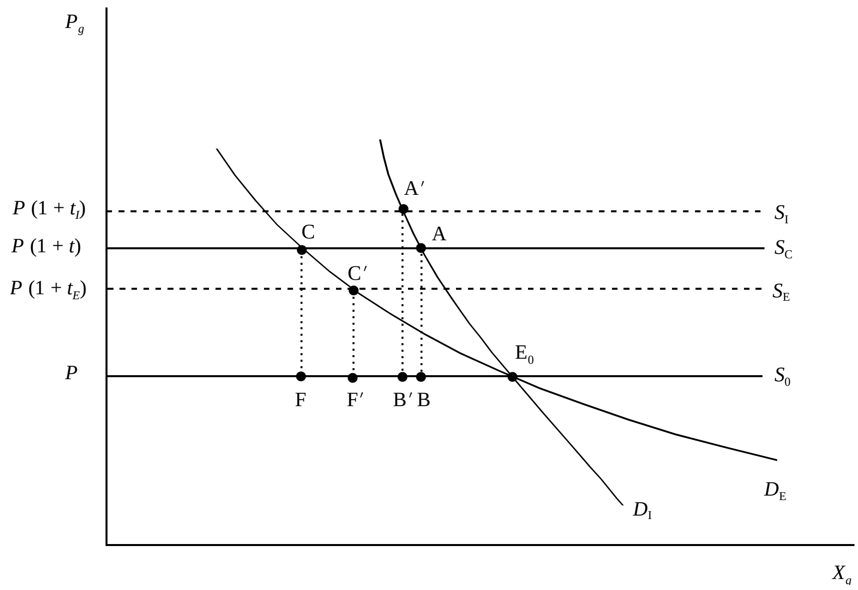

Figure 9.2: Optimal tax in partial equilibrium

effects are included. Third, labour is the only supplied factor of production and the representative household thus has no other source of income. Fourth, both the wage and all producer prices are fixed. Fifth, the objective of the policy maker is to maximize welfare of the representative household.

The representative household possesses the following (direct) utility function defined over leisure and goods:

$$U = U\left(1-L, X_1, X_2, \cdots, X_G\right), \tag{A5.10}$$

where X_g is the consumption of good g ($g = 1, \cdots, G$), L is labour supply (so $1-L$ is leisure, where the time endowment is unity), and $U(\cdot)$ has the usual properties, i.e. it features positive but diminishing marginal utility for leisure and all goods and is strictly quasi-concave in its arguments. Note that, by defining $X_0 \equiv -L$ we obtain $1-L = 1+X_0$ so that (A5.11) can be written in an alternative format as:

$$U = U\left(1+X_0, X_1, X_2, \ldots, X_G\right). \tag{A5.11}$$

(This format will yield some notational advantages later on.)

The household budget constraint is:

$$\sum_{g=1}^{G} Q_g X_g = WL, \tag{A5.12}$$

where $Q_g \equiv 1+t_g$ is the consumer price of good g (all producer prices are normalized to unity, i.e. $P_g = 1$ for $g = 1, \ldots, G$) and W is the wage rate. Note that there is no labour income tax.[5] By setting

[5] As Atkinson and Stiglitz (1980, p. 371-371) explain, this assumption is innocuous in the present setting. Here a labour income

$Q_0 = W$ and noting that $X_0 \equiv -L$ we can rewrite (BC) as follows:

$$\sum_{g=0}^{G} Q_g X_g = 0. \tag{A5.13}$$

Implicit in (A5.11) and (A5.13) is the notion that we treat $X_0 \equiv -L$ as just another good.

The indirect utility function associated with (A5.11) is defined in the usual way (see also Chapter 2):

$$V(Q_0, \ldots, Q_G) \equiv \max_{\{X_g\}} U(1 + X_0, X_1, \ldots, X_G) \text{ subject to } \sum_{g=0}^{G} Q_g X_g = 0. \tag{A5.14}$$

The Marshallian demands for good g and the Marshallian labour supply curve are obtained by using Roy's Identity (see the Intermezzo below):

$$X_g \equiv -\frac{\partial V / \partial Q_g}{\alpha}, \quad (\text{for } g = 1, \ldots, G), \tag{A5.15}$$

$$X_0 \equiv -L = -\frac{\partial V / \partial Q_0}{\alpha}, \tag{A5.16}$$

where α is the marginal utility of income.

We now have all the ingredients needed for our second view of the policy maker's optimal tax problem. The policy maker chooses tax rates t_g on the commodities (for $g = 1, \ldots, G$) such that utility of the representative household, $V(Q_0, \ldots, Q_G)$, is maximized subject to the revenue requirement restriction:

$$R_0 = \sum_{g=1}^{G} t_g X_g, \tag{A5.17}$$

where (A5.17) differs slightly from (A5.3) because producer prices are normalized to unity ($P_g = 1$ for all g). The Lagrangian expression associated with the maximization problem is:

$$\mathcal{L} \equiv V(Q_0, \ldots, Q_G) + \lambda \left[\sum_{g=1}^{G} t_g X_g - R_0 \right],$$

where λ is the Lagrange multiplier for the revenue requirement restriction (A5.17). The first-order necessary conditions are the constraint (A5.17) and:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial t_g} &= \frac{\partial V}{\partial Q_g} \frac{\partial Q_g}{\partial t_g} + \lambda \left[X_g + \sum_{j=1}^{G} t_j \frac{\partial X_j}{\partial Q_g} \frac{\partial Q_g}{\partial t_g} \right] \\ &= \frac{\partial V}{\partial Q_g} + \lambda \left[X_g + \sum_{j=1}^{G} t_j \frac{\partial X_j}{\partial Q_g} \right] = 0, \end{aligned} \tag{A5.18}$$

tax would be equivalent to a uniform tax on all goods. This result hinges on the absence of non-labour income and on the impossibility of taxing leisure.

for all $g = 1, \ldots, G$. By using equation (A5.15) (Roy's Identity) we can simplify (A5.18) even further:

$$\begin{aligned} \alpha X_g = \lambda \left[X_g + \sum_{j=1}^{G} t_j \frac{\partial X_j}{\partial Q_g} \right] \quad &\Leftrightarrow \\ \sum_{j=1}^{G} t_j \frac{\partial X_j}{\partial Q_g} = -\frac{\lambda - \alpha}{\lambda} X_g. & \end{aligned} \tag{A5.19}$$

Comparing (A5.19) with its partial equilibrium counterpart (A5.6), we find that $G-1$ additional cross-derivatives (of the type $\partial X_j / \partial Q_g$)) affect the optimal tax formula for good g in the general equilibrium model.

Although it is rather compact, equation (A5.19) is still a little difficult to interpret so we need to use some additional *useful results* to simplify it. The first of these results is the famous Slutsky decomposition. As we saw in Chapter 2, the Marshallian and Hicksian demands are related according the Slutsky equation:

$$\frac{\partial X_j}{\partial Q_g} = \left(\frac{\partial X_j}{\partial Q_g} \right)_{U_0} - X_g \frac{\partial X_j}{\partial M}, \tag{A5.20}$$

where $S_{jg} \equiv \left(\partial X_j / \partial Q_g \right)_{U_0}$ is the derivative of the Hicksian demand curve and $\partial X_j / \partial M$ is the income effect (evaluated at $M = 0$, since there is no lump-sum income in the model). The second useful result is the property of Slutsky symmetry:

$$S_{jg} \equiv \left(\frac{\partial X_j}{\partial Q_g} \right)_{U_0} = \left(\frac{\partial X_g}{\partial Q_j} \right)_{U_0} \equiv S_{gj}. \tag{A5.21}$$

According to (A5.21), for Hicksian demands the cross-derivative of good j with respect to the price of good g is equal to the cross-derivative of good g with respect to the price of good j. By using (A5.20) and (A5.21) in (A5.19) we find after some steps:

$$\begin{aligned} \sum_{j=1}^{G} t_j \left[\left(\frac{\partial X_j}{\partial Q_g} \right)_{U_0} - X_g \frac{\partial X_j}{\partial M} \right] = -\frac{\lambda - \alpha}{\lambda} X_g \quad &\Leftrightarrow \\ \sum_{j=1}^{G} t_j \left(\frac{\partial X_g}{\partial Q_j} \right)_{U_0} = -\frac{\lambda - \alpha}{\lambda} X_g + X_g \sum_{j=1}^{G} t_j \frac{\partial X_j}{\partial M} \quad &\Leftrightarrow \\ \sum_{j=1}^{G} \frac{t_j S_{gj}}{X_g} = -\theta \quad (\text{for } g = 1, \ldots, G), & \end{aligned} \tag{A5.22}$$

where $\theta > 0$ is defined as follows:[6]

$$\theta \equiv \frac{\lambda - \alpha}{\lambda} - \sum_{j=1}^{G} t_j \frac{\partial X_j}{\partial M}. \tag{A5.23}$$

Several things are worth noting about (A5.22) and (A5.23). First, the expression (A5.22) is the Ramsey optimal tax formula as it was first derived by Samuelson in 1951 in an unpublished memorandum for the U.S. Treasury (see Samuelson, 1986). He interprets the result as follows: "An optimal pattern of taxes is one which, if it were imposed but compensated for by giving the consumer enough lump-sum income to keep him on the same level of satisfaction, would then result in an equal percentage change in all goods and services." He hastens to add that as such the formula "consists substantially of 'empty boxes'" and thus provides little or no practical policy advice concerning taxes on actual goods (Samuelson, 1986, p. 139).

The second noteworthy feature of (A5.22) is that not the Marshallian but the *compensated* (Hicksian) derivatives feature in the optimal formula: *any* tax has income effects but the distorting effect of a tax originates from the *pure substitution effect.* Third, if we ignore all income effects ($\partial X_j / \partial M = 0$) and all cross-price effects ($S_{gj} = 0$ for $j \neq g$) then (A5.22) reduces to the partial equilibrium expression (A5.6).

As Samuelson warns us, the optimal tax formula (A5.22) has a rather "deceptive simplicity" about it (1986, p. 140). Indeed, the formula constitutes a system of G simultaneous equations in the tax rates $t_1, \ldots, t_G$ which must be inverted somehow to get expressions for these tax rates. In the presence of non-zero cross-price terms ($S_{gj} \neq 0$), this matrix inversion is far from straightforward. We therefore look at some special cases to build further intuition behind the optimal tax formula.

Atkinson and Stiglitz (1972, 1980) distinguish three special cases of the Ramsey optimal tax formula. In case 1, there are only two-goods ($G = 2$) and we reach the conclusion that the policy maker should tax more heavily the good that is complementary with leisure (conform the classic result by Corlett and Hague (1953-54)). Case 2 assumes that the utility function is *implicitly separable* between leisure and goods. In this case a uniform tax on all goods is optimal. Finally, in case 3 preferences are assumed to be *directly additive* and we obtain the result that necessities should be taxed more heavily than luxuries.

Intermezzo 9.1

Roy's Identity and labour supply. With variable labour supply, Roy's Identity can be proved as follows (see also Varian (1992, pp. 097-098)). The household maximizes (A5.10) subject to

[6] Atkinson and Stiglitz (1980, p. 373) prove that θ is positive, i.e. it has the same sign as required revenue R_0. The proof proceeds as follows. Taking X_g to the other side in (A5.22), multiplying by t_g, and summing over all g we get:

$$\sum_{g=1}^{G} \sum_{j=1}^{G} t_g S_{gj} t_j = -\theta \sum_{g=1}^{G} t_g X_g = -\theta R_0, \tag{A}$$

where we have used (A5.17) to get to the final equality. The left-hand side of (A) is negative because the Slutsky matrix is negative semi-definite) so that θR_0 must be positive.

(A5.12) and the Lagrangian is:

$$\mathcal{L} \equiv U(1-L, X_1, X_2, \ldots, X_G) + \alpha \left[WL - \sum_{g=1}^{G} Q_g X_g \right],$$

where α is the Lagrange multiplier (equal to the marginal utility of lump-sum income in the optimum). The first-order conditions are:

$$\frac{\partial U}{\partial X_g} = \alpha Q_g, \quad (\text{for } g = 1, \ldots, G), \tag{A}$$

$$\frac{\partial U}{\partial (1-L)} = \alpha W, \tag{B}$$

$$WL = \sum_{g=1}^{G} Q_g X_g. \tag{C}$$

These conditions, of course, define the Marshallian solutions for X_g and L which we write as:

$$X_g = X_g(Q_1, \ldots, Q_G, W), \quad (\text{for } g = 1, \ldots, G), \tag{D}$$

$$L = L(Q_1, \ldots, Q_G, W). \tag{E}$$

The indirect utility function can thus be written as follows:

$$V(Q_1, \ldots, Q_G, W) \equiv U\Big(X_1(Q_1, \ldots, Q_G, W), \ldots, X_G(Q_1, \ldots, Q_G, W), L(Q_1, \ldots, Q_G, W)\Big). \tag{F}$$

By differentiating (F) with respect to Q_j we obtain:

$$\begin{aligned} \frac{\partial V(\cdot)}{\partial Q_j} &= \frac{\partial U}{\partial X_1}\frac{\partial X_1}{\partial Q_j} + \ldots + \frac{\partial U}{\partial X_G}\frac{\partial X_G}{\partial Q_j} - \frac{\partial U}{\partial (1-L)}\frac{\partial L}{\partial Q_j} \\ &= \alpha Q_1 \frac{\partial X_1}{\partial Q_j} + \ldots + \alpha Q_G \frac{\partial X_G}{\partial Q_j} - \alpha W \frac{\partial L}{\partial Q_j} \\ &= \alpha \left[\sum_{g=1}^{G} Q_g \frac{\partial X_g}{\partial Q_j} - W \frac{\partial L}{\partial Q_j} \right], \end{aligned} \tag{G}$$

where we have used the first-order conditions (A)-(B) in going from the first to the second line. By differentiating the budget constraint (C) with respect to Q_j we obtain:

$$W \frac{\partial L}{\partial Q_j} = \sum_{g=1}^{G} Q_g \frac{\partial X_g}{\partial Q_j} + X_j. \tag{H}$$

By combining (G) and (H) we obtain Roy's Identity for good j:

$$\frac{\partial V(\cdot)}{\partial Q_j} = -\alpha X_j \quad \Leftrightarrow \quad X_j = -\frac{\partial V(\cdot)/\partial Q_j}{\alpha}. \tag{I}$$

Similarly, for labour supply we derive (by following analogous steps but differentiating with respect to the wage W):

$$\frac{\partial V(\cdot)}{\partial W} = \alpha \left[\sum_{g=1}^{G} Q_g \frac{\partial X_g}{\partial W} - W \frac{\partial L}{\partial W} \right], \tag{J}$$

$$L + W\frac{\partial L}{\partial W} = \sum_{g=1}^{G} Q_g \frac{\partial X_g}{\partial W}. \tag{K}$$

Finally, by combining (J) and (K) we find Roy's Identity for labour supply:

$$\frac{\partial V(\cdot)}{\partial W} = \alpha L \quad \Leftrightarrow \quad L = \frac{\partial V(\cdot)/\partial W}{\alpha}. \tag{L}$$

****

9.3.1 Special case 1: Two goods

In a classic paper, Corlett and Hague (1953-54) study the optimal taxation problem in a model with two goods. By setting $G = 2$ we derive from (A5.22) that the optimal taxes t_1 and t_2 are the solutions to the following matrix equation:

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = -\theta \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \tag{A5.24}$$

where $S_{ij} \equiv \left(\partial X_i / \partial Q_j\right)_{U_0}$ is the Hicksian cross-price elasticity and θ is defined as follows:

$$\theta \equiv \frac{\lambda - \alpha}{\lambda} - \left[t_1 \frac{\partial X_1}{\partial M} + t_2 \frac{\partial X_2}{\partial M} \right]. \tag{A5.25}$$

The matrix S on the left-hand side of (A5.24) is the so-called Slutsky matrix which is symmetric and negative semidefinite (Varian, 1992, pp. 123) and thus has a positive determinant, i.e. $|S| \equiv S_{11}S_{22} -$

$(S_{12})^2 > 0$.[7] By inverting (A5.24) we obtain:

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = -\frac{\theta}{|S|} \begin{bmatrix} S_{22} & -S_{12} \\ -S_{21} & S_{11} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \frac{\theta}{|S|} \begin{bmatrix} -S_{22}X_1 + S_{12}X_2 \\ S_{21}X_1 - S_{11}X_2 \end{bmatrix}. \tag{A5.26}$$

We define the elasticities of *compensated demand* as follows:

$$\varepsilon_{ij} \equiv \frac{Q_j}{X_i} \left(\frac{\partial X_i}{\partial Q_j} \right)_{U_0} = \frac{Q_j S_{ij}}{X_i}, \tag{A5.27}$$

and note that these elasticities satisfy:[8]

$$\varepsilon_{10} + \varepsilon_{11} + \varepsilon_{12} = 0, \tag{A5.28}$$

$$\varepsilon_{20} + \varepsilon_{21} + \varepsilon_{22} = 0. \tag{A5.29}$$

By using (A5.27), equation (A5.26) can be rewritten in terms of the compensated elasticities according to:

$$t_1 = \frac{\theta X_1 X_2}{Q_2 |S|} \left[\varepsilon_{12} - \varepsilon_{22} \right], \tag{A5.30}$$

$$t_2 = \frac{\theta X_1 X_2}{Q_1 |S|} \left[\varepsilon_{21} - \varepsilon_{11} \right], \tag{A5.31}$$

or (since $Q_g = 1 + t_g$):

$$\frac{t_1}{1+t_1} / \frac{t_2}{1+t_2} = \frac{\varepsilon_{12} - \varepsilon_{22}}{\varepsilon_{21} - \varepsilon_{11}}. \tag{A5.32}$$

---

[7] A square matrix S which has the property $x^T S x \leq 0$ for all x ($\neq 0$) is called negative semidefinite (and negative definite if the inequality is strict). A matrix is negative definite if and only if the determinants of the principal minors alternate in sign (starting negative). In the two-good case we $S_{11} < 0$, $S_{22} < 0$ and $|S| \equiv S_{11}S_{22} - (S_{12})^2 > 0$.

[8] The proof is as follows. The compensated demand for good g is defined as:

$$X_g^C \equiv \frac{\partial E}{\partial Q_g}$$

where X_g^C is the Hicksian demand and $E(\cdot)$ is the expenditure function. We know that X_g^C is homogeneous of degree zero in Q_0, Q_1 and Q_2 so by Euler's Theorem we have:

$$0 \times X_g^C = \frac{\partial X_g^C}{\partial Q_0} Q_0 + \frac{\partial X_g^C}{\partial Q_1} Q_1 + \frac{\partial X_g^C}{\partial Q_2} Q_2$$

Dividing both sides by X_g yields the required results.

Finally, by noting from (A5.28) that $\varepsilon_{12} - \varepsilon_{22} = -[\varepsilon_{10} + \varepsilon_{11} + \varepsilon_{22}]$ and from (A5.29) that $\varepsilon_{21} - \varepsilon_{11} = -[\varepsilon_{20} + \varepsilon_{11} + \varepsilon_{22}]$ we obtain the final expression for the (ratio of the) optimal tax rates:

$$\frac{t_1}{1+t_1} \Big/ \frac{t_2}{1+t_2} = \frac{-\varepsilon_{10} - (\varepsilon_{11} + \varepsilon_{22})}{-\varepsilon_{20} - (\varepsilon_{11} + \varepsilon_{22})}. \tag{A5.33}$$

Since the "own" compensated price elasticities are negative ($\varepsilon_{ii} < 0$) we conclude from (A5.33) that:

$$t_1 \underset{<}{\overset{>}{=}} t_2 \quad \Leftrightarrow \quad \varepsilon_{10} \underset{>}{\overset{<}{=}} \varepsilon_{20}. \tag{A5.34}$$

To interpret (A5.34) we recall from basic microeconomics that if $\varepsilon_{i0} > 0$ then good i is called a substitute for labour (i.e. a complement with leisure). Armed with this interpretation, it is clear that (A5.34) says that the good with the larger cross elasticity of compensated demand with the price of labour must in the optimum have the smaller tax rate. Or, as Corlett and Hague put it, the policy maker should tax more heavily that good which is more complementary with leisure (i.e. more of a substitute for labour). The intuition behind this results is as follows. If the policy maker could tax leisure, then he would have access to a lump-sum tax (first-best case). In the absence of leisure taxation (and with only two goods), it is second-best optimal to tax relatively more heavily that good which is complementary with leisure (Sandmo, 1976, p. 47).

9.3.2 Special case 2: Implicit separability between leisure and goods

In the second special case, we do not restrict the number of goods but instead we investigate under which conditions it is optimal (according to the tax formula (A5.22)) to tax all goods equally. This topic was studied by Angus Deaton (1979) who showed that uniform taxation is optimal when there is *implicit separability* between leisure and goods in the utility function.

For implicitly separable preferences the expenditure function can be written as follows:

$$E(Q_0, Q_1, \ldots, Q_G, U_0) = E(Q_0, e(Q_1, \ldots, Q_G, U_0), U_0), \tag{A5.35}$$

where $E(\cdot)$ is the overall expenditure function, $e(\cdot)$ is the (sub-)expenditure function relating to the goods ($g = 1, \ldots, G$), and U_0 is the level of utility which is held constant in the two expenditure functions. By making use of the derivative property (i.e. Shephard's Lemma) of the expenditure function we find:

$$X_g^C(Q_0, Q_1, \ldots, Q_G, U_0) \equiv \frac{\partial E}{\partial Q_g}, \tag{A5.36}$$

where X_g^C is the Hicksian (compensated) demand for good g.

Before we can investigate the consequences of implicit separability, we must redo the optimal tax problem with the expenditure function (rather than with the indirect utility function). The policy maker wishes to maximize utility, U_0, subject to the revenue requirement constraint:

$$R_0 = \sum_{g=1}^{G} t_g X_g^C \left(Q_0, Q_1, \ldots, Q_G, U_0\right), \tag{A5.37}$$

and the household's budget constraint:

$$E\left(Q_0, Q_1, \ldots, Q_G, U_0\right) = 0. \tag{A5.38}$$

The Lagrangian expression for this maximization problem is:

$$\mathcal{L} \equiv U_0 + \lambda E\left(Q_0, Q_1, \ldots, Q_G, U_0\right) + \mu \left[R_0 - \sum_{g=1}^{G} t_g X_g^C \left(Q_0, Q_1, \ldots, Q_G, U_0\right)\right],$$

where λ and μ are the Lagrange multipliers for, respectively, the household budget constraint (A5.38) and the revenue requirement constraint (A5.37). The first-order necessary conditions are the constraints and:

$$\frac{\partial \mathcal{L}}{\partial U_0} = 1 + \lambda \frac{\partial E}{\partial U_0} - \mu \sum_{g=1}^{G} t_g \frac{\partial X_g^C}{\partial U_0} = 0, \tag{A5.39}$$

$$\frac{\partial \mathcal{L}}{\partial t_g} = \lambda \frac{\partial E}{\partial Q_g} \frac{\partial Q_g}{\partial t_g} - \mu \left[X_g^C + \sum_{j=1}^{G} t_j \frac{\partial X_j^C}{\partial Q_g} \frac{\partial Q_g}{\partial t_g}\right] = 0, \tag{A5.40}$$

for $g = 1, \ldots, G$. By using (A5.36) and noting that $Q_g \equiv 1 + t_g$, the first-order conditions for the tax rates (A5.40) can be further simplified:

$$\lambda X_g^C = \mu \left[X_g^C + \sum_{j=1}^{G} t_j \frac{\partial X_j^C}{\partial Q_g}\right]$$

$$\frac{\lambda - \mu}{\mu} X_g = \sum_{j=1}^{G} t_j S_{gj}, \quad (\text{for } g = 1, \ldots, G), \tag{A5.41}$$

where we have used the fact that $X_g^C = X_g$ and $S_{jg} = S_{gj}$ in going from the first to the second line. Note that (A5.41) is formally identical to (A5.22).

For implicitly separable preferences the tax formula can be simplified even further. Indeed, by using (A5.35) we find that the compensated elasticities for good g with respect to the wage (Q_0) are the same for all goods:[9]

$$\varepsilon_{g0} \equiv \frac{Q_0 S_{g0}}{X_g} = \varepsilon_0. \tag{A5.42}$$

According to (A5.42) all goods are equally complementary with leisure. Using the reasoning explained above (for the case $G = 2$) we find that all goods should be taxed at the same rate, i.e. uniform commodity taxation is called for.

9.3.3 Special case 3: Directly-additive preferences

The third and final special case was studied by Atkinson and Stiglitz (1972, 1980) and is based on the assumption that preferences are *directly additive* (see Houthakker, 1960). This case is of interest because it shows the relationship between the optimal tax rates and the *income* elasticity of demand. With direct additivity, the direct utility function can be written as:

$$U = U^0(1-L) + U^1(X_1) + \ldots + U^G(X_G), \tag{A5.43}$$

where the sub-utility functions have the usual properties of positive and diminishing marginal utility, i.e. $U_g \equiv \partial U^g/\partial X_g > 0$, $U_{gg} \equiv \partial^2 U^g/\partial X_g^2 < 0$ (for $g = 1, \ldots, G$), $U_{1-L} \equiv \partial U^0/\partial(1-L) > 0$, and $U_{1-L,1-L} \equiv \partial^2 U^0/\partial(1-L)^2 < 0$.

Before we can investigate the consequences for optimal taxation of direct additivity of the utility function, we must redo the optimal tax problem using the *direct approach*,[10] i.e. by making use of the direct rather than the indirect utility function (see also Atkinson and Stiglitz (1980, pp. 376-379) for details). In the direct (or "primal") approach the policy maker uses quantities as the control variables. The *instruments* by which the policy maker affects quantities are the commodity tax rates, and the *constraints* of the optimization problem are the revenue requirement constraint and the first-order conditions of household optimization.

We must first derive a compact expression for the first-order conditions of household optimization.

[9]The expression in (A5.42) is derived as follows. From (A5.35) we derive:

$$X_g^C = \frac{\partial E}{\partial e}\frac{\partial e}{\partial Q_g}, \tag{a}$$

where by definition $\partial e/\partial Q_g$ is independent of Q_0. By differentiating X_g^C with respect to the wage rate, we find:

$$\frac{\partial X_g^C}{\partial Q_0} = \frac{\partial e}{\partial Q_g}\frac{\partial^2 E}{\partial e \partial Q_0} = \frac{X_g^C}{\partial E/\partial e}\frac{\partial^2 E}{\partial e \partial Q_0}, \tag{b}$$

where we have used (a) in the second step. It follows from (b) that:

$$\frac{Q_0}{X_g^C}\frac{\partial X_g^C}{\partial Q_0} = \frac{Q_0 \partial^2 E/\partial e \partial Q_0}{\partial E/\partial e} \equiv \varepsilon_0, \tag{c}$$

where the right-hand side does not depend on g.

[10]Note that Ramsey (1927) himself also used the direct approach.

The representative household maximizes direct utility (A5.43) subject to the budget constraint:

$$\sum_{g=1}^{G} Q_g X_g = WL. \tag{A5.44}$$

The first-order necessary conditions are:

$$U_g = \alpha Q_g, \quad (\text{for } g = 1, \ldots, G), \tag{A5.45}$$

$$U_{1-L} = \alpha W, \tag{A5.46}$$

where α is the Lagrange multiplier for the budget constraint (A5.44), equalling the marginal utility of lump-sum income in the optimum. By substituting these first-order conditions back into the budget constraint (A5.44) we obtain the following expression:

$$\sum_{g=1}^{G} Q_g X_g - WL = 0 \quad \Leftrightarrow \sum_{g=1}^{G} \frac{U_g}{\alpha} X_g - \frac{U_{1-L}}{\alpha} L = 0 \quad \Leftrightarrow$$

$$\sum_{g=1}^{G} U_g X_g - U_{1-L} L = 0, \tag{A5.47}$$

where we have used the fact that $\alpha > 0$ in the final step. Equation (A5.47) represents the household budget constraint expressed in quantities only (rather than in quantities and prices, as in (A5.44) above).

The optimal taxation problem now takes the following format. The policy maker chooses X_g (for $g = 1, \ldots, G$) and L in order to maximize (A5.43) subject to (A5.47) and the (rewritten version of the) revenue requirement constraint:

$$\begin{aligned} R_0 &= \sum_{g=1}^{G} t_g X_g = \sum_{g=1}^{G} \left(Q_g - 1\right) X_g \\ &= WL - \sum_{g=1}^{G} X_g, \end{aligned} \tag{A5.48}$$

where we have used (A5.44) to get to the second line. The Lagrangian expression for this maximization problem is:

$$\begin{aligned} \mathcal{L} \equiv\ & U^0 (1 - L) + U^1 (X_1) + \ldots + U^G (X_G) + \lambda \left[WL - \sum_{g=1}^{G} X_g - R_0 \right] \\ & + \mu \left[\sum_{g=1}^{G} U_g X_g - U_{1-L} L \right], \end{aligned}$$

where λ and μ are the Lagrange multipliers for, respectively, the revenue requirement constraint (A5.48) and the household budget constraint (A5.47). The first-order necessary conditions are the constraints and:

$$\frac{\partial \mathcal{L}}{\partial L} = -U_{1-L} + \lambda W - \mu \left[U_{1-L} - L U_{1-L,1-L} \right] = 0, \tag{A5.49}$$

$$\frac{\partial \mathcal{L}}{\partial X_g} = U_g - \lambda + \mu \left[U_g + X_g U_{gg} \right] = 0, \tag{A5.50}$$

where we have used the fact that $\partial U_g / \partial L = \partial U_{1-L} / \partial X_g = 0$ and $\partial U_g / \partial X_j = 0$ for $g \neq j$. We define the terms H^L and H^g (for $g = 1, \ldots, G$):

$$H^L \equiv -\frac{L U_{1-L,1-L}}{U_{1-L}} > 0, \qquad H^g \equiv -\frac{X_g U_{gg}}{U_g} > 0, \tag{A5.51}$$

and rewrite the first-order conditions (A5.49)-(A5.50) in a more compact format as:

$$\frac{\lambda - \alpha}{\alpha} = \mu \left(1 + H^L \right), \tag{A5.52}$$

$$\frac{\lambda}{\alpha} = \left(1 + t_g \right) \left[1 + \mu \left(1 - H^g \right) \right], \tag{A5.53}$$

where we have used (A5.45)-(A5.46) to obtain these expressions. In the final step we eliminate μ from these expressions and derive the optimal tax formula:[11]

$$\frac{t_g}{1 + t_g} = \frac{\lambda - \alpha}{\lambda} \frac{H^g + H^L}{1 + H^L}. \tag{A5.54}$$

Equation (A5.54) is yet another expression for the optimal commodity tax rates, this time expressed in terms of properties of the direct utility function. Atkinson and Stiglitz (1980, p. 378) discuss some interesting special cases. First, if labour supply is inelastic ($U_{1-L} = 0$ so that $H^L \rightarrow \infty$ by (A5.51)) then it follows from (A5.54) that there should be a uniform tax rate on all goods equal to $t_g = (\lambda - \alpha) / \alpha$. This tax is equivalent to a tax on labour income only. Intuitively, the tax is borne in that case by the factor which is inelastic in supply (labour). The uniform commodity tax is equivalent to a lump-sum tax. Second, for a perfectly elastic supply of labour (U_{1-L} is constant and $H^L = 0$), the optimal tax formula (A5.54) yields the partial equilibrium (inverse-elasticity) result :

$$\frac{t_g}{1 + t_g} = \frac{\lambda - \alpha}{\lambda} H^g = \frac{\lambda - \alpha}{\lambda} \frac{1}{e_g^D}, \tag{A5.55}$$

[11] By solving (A5.52) for μ and substituting the result into (A5.53) we find:

$$1 + t_g = \frac{\lambda \left(1 + H^L \right)}{\alpha \left(H^g + H^L \right) + \lambda \left(1 - H^g \right)},$$

and thus:

$$t_g = \frac{(\lambda - \alpha) \left(H^g + H^L \right)}{\alpha \left(H^g + H^L \right) + \lambda \left(1 - H^g \right)}.$$

By dividing these two expressions we obtain (A5.54).

where ε_g^D is the price elasticity of demand, and $\varepsilon_g^D = 1/H^g$ in this case.[12] Intuitively, the general equilibrium approach yields the same answers as the partial equilibrium approach because there are no income effects and the demand curves only depend on own prices.

In the case of directly additive preferences and elastic labour supply, Atkinson and Stiglitz (1980, p. 379) show that a third conclusion can be drawn, namely that the optimal tax rate depends inversely on the income elasticity of demand, i.e. necessities should be taxed more heavily than luxuries! Technically, this result follows from (A5.54) because H^g is inversely proportional with the income (semi-)elasticity of demand:

$$H^g \frac{1}{X_g}\frac{\partial X_g}{\partial M} = -\frac{1}{\alpha}\frac{\partial \alpha}{\partial M} > 0, \tag{A5.56}$$

where M is lump-sum (evaluated at $M = 0$).[13]

Intermezzo 9.2

Expenditure function and Slutsky terms. Here we present a simple proof of Shephard's Lemma with variable labour supply (see also Varian (1992, p. 74)). The expenditure function is defined as:

$$E(Q_1, Q_2, \ldots, Q_G, W, U_0) \equiv \min_{\{X_g, L\}} \sum_{g=1}^{G} Q_g X_g - WL$$
$$\text{s.t. } U_0 = U(X_1, X_2, \ldots, X_G, 1-L).$$

The Lagrangian associated with this problem is:

$$\mathcal{L} \equiv \sum_{g=1}^{G} Q_g X_g - WL + \lambda\Big[U_0 - U(X_1, X_2, \ldots, X_G, 1-L)\Big]$$

where λ is the Lagrange multiplier. The first-order conditions are:

$$Q_g = \lambda\frac{\partial U}{\partial X_g}, \quad (\text{for } g = 1, \ldots, G), \tag{A}$$

[12]In this case $U^0(1-L)$ is linear in leisure, say $U^0(1-L) = \gamma(1-L)$ so that (A5.46) implies $\gamma = \alpha W$. Since both γ and W are constant, so is the marginal utility of income, α. Differentiating (A5.45) with respect to Q_g we find $U_{gg}\partial X_g/\partial Q_g = \alpha$. Using this result in the second expression in (A5.51) we find the desired result:

$$H^g \equiv -\frac{X_g U_{gg}}{U_g} = -\frac{(\alpha/\partial X_g/\partial Q_g)\,X_g}{\alpha Q_g} = -\left[\frac{Q_g}{X_g}\frac{\partial X_g}{\partial Q_g}\right]^{-1} \equiv \frac{1}{\varepsilon_g^D}.$$

[13]Equation (A5.56) is derived as follows. By differentiating the household's first-order condition for good g (A5.45) with respect to income M we find:

$$U_{gg}\frac{\partial X_g}{\partial M} = Q_g\frac{\partial \alpha}{\partial M} = U_g\frac{1}{\alpha}\frac{\partial \alpha}{\partial M} \quad\Leftrightarrow\quad H^g\frac{1}{X_g}\frac{\partial X_g}{\partial M} = -\frac{1}{\alpha}\frac{\partial \alpha}{\partial M}.$$

The marginal utility of income falls with income, i.e. $\partial\alpha/\partial M < 0$.

$$W = \lambda \frac{\partial U}{\partial (1 - L)}, \tag{B}$$

$$U_0 = U(X_1, X_2, \ldots, X_G, 1 - L). \tag{C}$$

These conditions define the Hicksian (or compensated) solutions for X_g and L which we write as follows:

$$X_g^C = X_g^C(Q_1, \ldots, Q_G, W, U_0), \quad \text{(for } g = 1, \ldots, G\text{)}, \tag{D}$$

$$L^C = L^C(Q_1, \ldots, Q_G, W, U_0), \tag{E}$$

where the superscript "C" stands for compensated.

Shephard's Lemma (or the derivative property) now takes the following form:

$$X_g^C = \frac{\partial E}{\partial Q_g}, \tag{F}$$

$$L^C = -\frac{\partial E}{\partial W}. \tag{G}$$

The proof proceeds as follows. Let Q_g^* and W^* be the actual prices and the wage rate, respectively, and let X_g^C and L^C be the optimal choices associated with these prices and that wage rate. Next, define the following function:

$$\phi(Q_1, \ldots, Q_G, W) \equiv E(Q_1, Q_2, \ldots, Q_G, W, U_0) - \left[\sum_{g=1}^{G} Q_g X_g^C - WL^C\right], \tag{H}$$

where X_g^C and L^C are the expenditure-minimizing choices. Obviously, since $E(Q_1, Q_2, \ldots, Q_G, W, U_0)$ is the optimal expenditure-minimizing solution, it follows that $\phi(\cdot) \leq 0$ and that $\phi(Q_1^*, \ldots, Q_G^*, W^*) = 0$. It follows that $\phi(\cdot)$ attains a maximum at the point $(Q_1^*, \ldots, Q_G^*, W^*)$, i.e. its derivatives are zero there:

$$\frac{\partial \phi(\cdot)}{\partial Q_g} = \frac{\partial E}{\partial Q_g} - X_g^C = 0, \quad \text{(for } g = 1, \ldots, G\text{)}, \tag{I}$$

$$\frac{\partial \phi(\cdot)}{\partial W} = \frac{\partial E}{\partial W} + L^C = 0. \tag{J}$$

Equations (I)-(J) are the same as (F)-(G). Q.E.D.

****

9.4 Many-person Ramsey rule

Up to this point we have restricted attention to the case of identical households. The objective of this section is to generalize the Ramsey taxation formula to a setting with many heterogeneous households who differ in ability, tastes, and endowments. The key references in this literature are Diamond (1975), Mirrlees (1975), and Atkinson and Stiglitz (1976).. Here we follow the exposition of Atkinson and Stiglitz (1976, pp. 59-63).

There are H households indexed by h. Each household has an indirect utility function of the following form:

$$V^h = V^h(Q_0, \ldots, Q_G), \tag{A5.57}$$

where $Q_0 = W$ and $Q_g \equiv 1 + t_g$ (for $g = 1, \ldots, G$). The policy maker has an individualistic social welfare function (see Chapter 9) featuring the (indirect) utility levels of the households:

$$SW \equiv \Psi\left(V^1, V^2, \ldots, V^H\right), \tag{A5.58}$$

where $\Psi_h \equiv \partial\Psi/\partial V^h > 0$ for all $h = 1, 2, \ldots, H$. The social planner chooses the commodity taxes, $t_1, t_2, \ldots, t_G$ in order to maximize social welfare subject to the following revenue requirement constraint:

$$R_0 = \sum_{g=1}^{G} t_g \sum_{h=1}^{H} X_g^h, \tag{A5.59}$$

where X_g^h is the demand for good g by household h. The Lagrangian expression associated with the social optimization program is:

$$\mathcal{L} \equiv \Psi\left(V^1, V^2, \ldots, V^H\right) + \lambda\left[\sum_{g=1}^{G} t_g \sum_{h=1}^{H} X_g^h - R_0\right],$$

where λ is the Lagrange multiplier for the revenue requirement constraint (A5.59). The first-order necessary conditions are the constraint and:

$$\frac{\partial\mathcal{L}}{\partial t_g} = \sum_{h=1}^{H} \frac{\partial\Psi}{\partial V^h}\frac{\partial V^h}{\partial Q_g} + \lambda\left[\sum_{h=1}^{H} X_g^h + \sum_{j=1}^{G} t_j \sum_{h=1}^{H} \frac{\partial X_j^h}{\partial Q_g}\frac{\partial Q_g}{\partial t_g}\right] = 0. \tag{A5.60}$$

To simplify these expressions we note Roy's identity in the multi-person setting:

$$\frac{\partial V^h}{\partial Q_g} = -\alpha^h X_g^h, \tag{A5.61}$$

where α^h is the marginal utility of income for household h. Furthermore, the Slutsky equation is given by:

$$\frac{\partial X_j^h}{\partial Q_g} = S_{jg}^h - X_g \frac{\partial X_j^h}{\partial M^h}, \tag{A5.62}$$

where $S_{jg}^h \equiv \left(\partial X_j^h / \partial Q_g\right)_{U_0^h}$ is the derivative of the Hicksian demand curve, $\partial X_j^h / \partial M^h$ is the income effect (evaluated at $M^h = 0$, since there is no lump-sum income in the model), and $S_{jg}^h = S_{gj}^h$ (Slutsky symmetry). By using (A5.61) and (A5.62) in (A5.60) we find after some steps:

$$\begin{aligned}
&\sum_{h=1}^{H} \Psi_h \alpha^h X_g^h = \lambda \left[H\bar{X}_g + \sum_{j=1}^{G} t_j \sum_{h=1}^{H} \frac{\partial X_j^h}{\partial t_g} \right] \quad \Leftrightarrow \\
&\sum_{j=1}^{G} t_j \sum_{h=1}^{H} S_{gj}^h = \sum_{h=1}^{H} \frac{\Psi_h \alpha^h}{\lambda} X_g^h + \sum_{j=1}^{G} t_j \sum_{h=1}^{H} X_g^h \frac{\partial X_j^h}{\partial M^h} - H\bar{X}_g,
\end{aligned} \tag{A5.63}$$

where $\bar{X}_g \equiv \sum_{h=1}^{H} X_g^h / H$ is the average demand for good g and $\Psi_h \alpha^h$ is the *gross* social marginal utility of income (or consumption) by household h. Next we define the *net* social marginal utility of income for household h:

$$\beta^h \equiv \Psi_h \alpha^h + \lambda \sum_{j=1}^{G} t_j \frac{\partial X_j^h}{\partial M^h}. \tag{A5.64}$$

As Diamond (1975, p. 338) points out, β^h represents the gain to society that results if household h receives additional income. It contains two terms. The first term ($\Psi_h \alpha^h$) measures by how much social welfare increases if household h attains a higher utility level due to the additional income. The second part ($\lambda \sum_{j=1}^{G} t_j \partial X_j^h / \partial M^h$) is the social evaluation of the additional tax revenue resulting from the fact that household h's income has gone up. By using (A5.64) it is possible to rewrite (A5.63) in a more compact format:

$$\begin{aligned}
\sum_{j=1}^{G} t_j \sum_{h=1}^{H} S_{gj}^h &= \sum_{h=1}^{H} \frac{\beta^h}{\lambda} X_g^h - H\bar{X}_g \quad \Leftrightarrow \\
\frac{1}{H\bar{X}_g} \sum_{h=1}^{H} \sum_{j=1}^{G} t_j S_{gj}^h &= -\left[1 - \sum_{h=1}^{H} \frac{\beta^h}{\lambda} \frac{X_g^h}{H\bar{X}_g} \right], \quad (\text{for } g = 1, 2, \ldots, G),
\end{aligned} \tag{A5.65}$$

where we have changed the order of summation in going from the first to the second line. The left-hand side of (A5.65) is the proportional reduction in the consumption of good g along the compensated demand curves. The right-hand side is not necessarily the same for all commodities.[14] Indeed, it is only the same for all commodities if (i) β^h is the same for all households (so that the right-hand side simplifies to $-(\lambda - \beta)/\lambda$), or (ii) if $X_g^h / \left(H\bar{X}_g\right)$ is the same for all commodities (there are no goods that are consumed disproportionately by the rich or the poor). In any other case, the right-hand side of

[14] Recall that for the representative-agent case, the right-hand side is the same for all commodities–see equation (A5.22) above.

(A5.65) will be different for each commodity. The term in square brackets is smaller, the higher is β^h/λ and/or $X_g^h/\left(H\bar{X}_g\right)$.

The optimal tax formula (A5.65) is often written in the following form:

$$\frac{1}{H\bar{X}_g}\sum_{h=1}^{H}\sum_{j=1}^{G}t_j S_{gj}^h = -\left[1-\frac{\bar{\beta}}{\lambda}-\frac{\bar{\beta}}{\lambda}\text{cov}\left(\beta^h/\lambda, X_g^h\right)\right], \tag{A5.66}$$

where $\bar{\beta} \equiv \sum_{h=1}^{H}\beta^h/H$ and $\text{cov}\left(\beta^h/\lambda, X_g^h\right)$ is the normalized covariance between β^h/λ and X_g^h:[15]

$$\text{cov}\left(\beta^h/\lambda, X_g^h\right) \equiv \frac{1}{H}\sum_{h=1}^{H}\frac{\left(\beta^h/\lambda\right)X_g^h}{\left(\bar{\beta}/\lambda\right)\bar{X}_g} - 1. \tag{A5.67}$$

The first term on the right-hand side of (A5.67) is called the *distributional characteristic* of good g.[16]

9.5 Marginal tax reform

- optimal taxation literature deals with *tax design*: what would the system look like if we could design it from the bottom up
- perhaps more relevant issue is that of *tax reform*: can we identify Pareto-improving changes in the tax system?
- this is the inherently difficult subject of second-best economics
- pessimistic reaction:
  - starting from a sub-optimal situation you cannot formulate simple/general rules about tax reform ["you cannot say anything"]
  - for example, a switch from distortionary to lump-sum taxation is not guaranteed to raise welfare
- constructive/realistic reaction:
  - starting from a sub-optimal situation there are many directions in which tax changes are welfare improving

[15] In general the normalized covariance between x_h and y_h is defined as follows:

$$\text{cov}\left(x_h, y_h\right) \equiv \frac{1}{H}\sum_{h=1}^{H}\frac{x_h-\bar{x}}{\bar{x}}\frac{y_h-\bar{y}}{\bar{y}} = \frac{1}{H}\sum_{h=1}^{H}\frac{x_h y_h}{\bar{x}\bar{y}} - 1,$$

where $\bar{x} \equiv \sum_{h=1}^{H}x_h/H$ and $\bar{y} \equiv \sum_{h=1}^{H}y_h/H$ are the respective means.

[16] This concept was proposed by Feldstein (1972a, 1972b) in the context of optimal public sector pricing with heterogeneous households.

- the correct direction of tax reform depends on the structure of preferences and the pre-existing situation
- welfare improving moves can be identified theoretically [see Dixit (1975) and Dixit and Munk (1977) for some examples]...
- ....and empirically, e.g. by using calibrated computable general equilibrium (CGE) models.

Key literature

- Atkinson & Stiglitz (1980, lecture 12), Jha (1998, chs. 13), or Myles (1995, ch. 4) on theory.
- Auerbach (1985) and Stern (1987) on theory.
- Auerbach and Hines (2002), on recent empirics and theory.
- Classics: Corlett and Hague (1953-54), Diamond (1975), Mirrlees (1976), Sandmo (1975, 1976), Dixit (1975), Dixit and Munk (1977), Atkinson and Stern (1974), and Atkinson and Stiglitz (1972, 1976).
- Mirrlees (1972), Samuelson (1986), Diamond and McFadden (1974), Dixit (1970), Deaton (1981), Sandmo (1974), Martina (2000).
- Further entries from reading list Poterba (see also Chapter 2): Hausman (1981a), Akerlof (1978), Diamond (1998), MaCurdy (1992), Eissa (1995), Gruber and Saez (2002), Slemrod (1998), Eissa and Liebman (1996).
- Empirical examples for many-person Ramsey taxation: Deaton (1977) and Harris and MacKinnon (1979).

Chapter 10

The structure of income taxation

The purpose of this chapter is to discuss the following topics:

- The classical economists on income taxation.
- Optimal linear income taxation.
- Optimal nonlinear income taxation.
- Income taxation and/or commodity taxation?

10.1 Introduction

In this chapter we present a brief overview of the main theories of income taxation. The overview is far from complete, and the interested reader is referred to Atkinson and Stiglitz (1980, lectures 13-14), Stiglitz (1987), and Tuomola (1990) for further details. There are three key questions that are dealt with in the literature on income taxation, namely (a) should there be an income tax?; (b) if so, should the income tax be graduated with income?; and (c) if so, should it be progressive, regressive, or something else?

The answers to question (a) have been mixed. Most modern authors answer it in the affirmative on the grounds that in the absence of income taxation, some people may not pay any tax at all (which seems unjust). On the other hand, the notion of income taxation was by no means uncontroversial throughout history. In the United Kingdom it was felt that the income tax is "hostile to every sense of freedom, revolting to the feelings of Englishmen" (this led to the 1816 abolishment of the tax in that country). Furthermore, as late as 1894 the United States Supreme Court ruled the income tax to be "unconstitutional."

In a similar fashion, questions (b) and (c) have also received mixed answers. On the one hand, a widely accepted mid-20th century sentiment was that the tax should be progressive because it serves to redistribute income (the *equity* aspect). On the other hand, some classical economists felt that "graduation is not an evil to be paltered with. Adopt it and you will effectively paralyse industry..." (J.R. McCulloch). Also, a widely accepted late-20th century sentiment is that high marginal tax rates are bad for "economic incentives" (the *efficiency* aspect).

What is clear from these views is that there is a potential conflict between equity and efficiency. On the one hand the policy maker may want to have a graduated tax system in order to redistribute incomes. On the other hand he may want to minimize the distorting aspects of income taxation. The objective of this chapter is to study the appropriate system of redistributive income taxation and to show how is it affected by (i) differences in distributive objectives, (ii) endogeneity of labour supply, and (iii) inequality of the pre-tax income distribution.

10.2 The sacrifice theory of income taxation

The Classical theory of income taxation is based on the *equal sacrifice approach*. Adam Smith for example argued that "subjects should contribute in proportion to their respective abilities" whilst in John Stuart Mill's view "whatever sacrifices the government requires...should be made to bear as nearly as possible with the same pressure upon all." A more recent version of the equal sacrifice theory says that the income tax system should be designed such that it maximizes a Benthamite social welfare function (consisting of the sum of individual utilities; see Chapter 9).

Consider the following simple model. Individuals differ in their earning ability, n, and before-tax earning of an individual of type n is denoted by $Z(n)$. The tax paid by this individual is denoted by $T(n)$ and the (indirect) utility of the n-type individual depends on after-tax earnings:

$$U^n = U^n(Y(n)), \tag{A5.1}$$

where $U^n(\cdot)$ has the usual properties ($\partial U^n/\partial Y(n) > 0$ and $\partial^2 U^n/\partial Y(n)^2 < 0$) and $Y(n)$ is after-tax income:

$$Y(n) \equiv Z(n) - T(n). \tag{A5.2}$$

Implicit in this formulation is the notion that the household consumes his entire after-tax income, i.e. the model is static and there is no saving.

The *cumulative distribution* of people of type n is denoted by $F(n)$ and the *density function* is $f(n) \equiv F'(n)$. The policy maker's objective function is the "sum" of individual utilities which (in this continuous-type model) amounts to:

$$SW \equiv \int_0^\infty U^n(Z(n) - T(n)) f(n)\, dn. \tag{A5.3}$$

The instrument of the policy maker is the tax schedule, $T(n)$ (which may be positive or negative), and the revenue requirement constraint is:

$$\int_0^\infty T(n)\, f(n)\, dn = R_0, \tag{A5.4}$$

where R_0 is exogenous. The policy maker chooses a tax schedule $T(n)$ such that (A5.3) is maximized subject to (A5.4). The Lagrangian for this maximization problem is:

$$\mathcal{L} \equiv \int_0^\infty U^n\left(Z(n) - T(n)\right) f(n)\, dn + \lambda\left[\int_0^\infty T(n)\, f(n)\, dn - R_0\right],$$

where λ is the Lagrange multiplier for the constraint (A5.4). The first-order necessary conditions are the constraint and:

$$\frac{\partial \mathcal{L}}{\partial T(n)} = -\frac{\partial U^n}{\partial Y(n)} f(n) + \lambda f(n) = 0, \tag{A5.5}$$

or:

$$\lambda = \frac{\partial U^n}{\partial Y(n)}. \tag{A5.6}$$

Since the same λ applies to all individuals, the optimal income tax system calls for an equalization of the *marginal utility* of after-tax earnings for all individuals. In the special case of identical utility functions ($U^n(\cdot) = U(\cdot)$) we obtain the completely egalitarian solution, i.e. after-tax incomes should be the same for all individuals!

As an example, consider the case in which the utility function is iso-elastic:

$$U^n = \begin{cases} \frac{Y(n)^{1-1/\sigma}-1}{1-1/\sigma} & \text{if } \sigma \geq 0, \quad \sigma \neq 1 \\ \\ \ln Y(n) & \text{if } \sigma = 1 \end{cases}, \tag{A5.7}$$

where σ is the substitution elasticity ($\sigma > 0$). The first-order condition (A5.7) calls for:

$$Y(n)^{-1/\sigma} = \lambda, \tag{A5.8}$$

so that it follows from (A5.2) that:

$$Z(n) - T(n) = \lambda^{-\sigma}. \tag{A5.9}$$

By substituting (A5.9) into the revenue requirement constraint (A5.4) we find the equilibrium value for

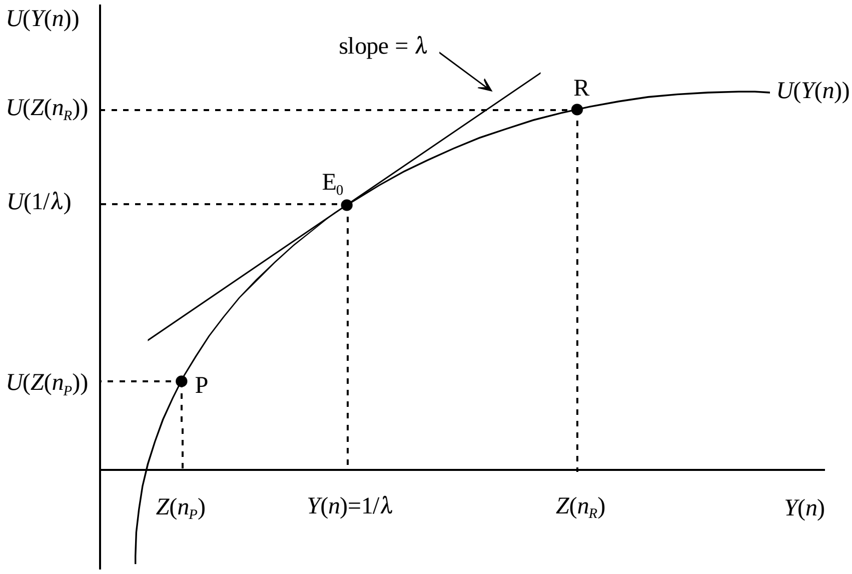

Figure 10.1: Utilitarian optimal income tax ($\sigma = 1$ case)

λ:

$$
\begin{aligned}
R_0 &= \int_0^\infty T(n) f(n)\, dn \\
&= \int_0^\infty \left[Z(n) - \lambda^{-\sigma}\right] f(n)\, dn \\
&= \bar{Z} - \lambda^{-\sigma},
\end{aligned}
\tag{A5.10}
$$

where $\bar{Z} \equiv \int_0^\infty Z(n) f(n)\, dn$ is average pre-tax earnings in the economy. Furthermore, by substituting (A5.10) in (A5.9) we find the tax schedule:[1]

$$
\begin{aligned}
T(n) &= Z(n) - \lambda^{-\sigma} \\
&= [Z(n) - \bar{Z}] + R_0.
\end{aligned}
\tag{A5.11}
$$

According to (A5.11), individuals with higher than average ability ($Z(n) > \bar{Z}$) pay more taxes and the tax for everybody is higher the larger is the required revenue, R_0. In Figure 10.1 we illustrate the pre-tax and after-tax situation for a poor and a rich individual (P and R) for the unit-elastic case ($\sigma = 1$). Under the optimal tax system, the after-tax point is at E_0, where $Y(n) = 1/\lambda$ for both types of agents. Figure 10.2 illustrates what happens to the optimal point if either required revenue increases (R_0 up) or average pre-tax earnings fall ($\bar{Z}$ down). In both cases the egalitarian after-tax income level falls.

The sacrifice theory of income taxation has been criticized for the following reasons. First, it ignores the disincentive effects of taxation, i.e. that $Z(n)$ itself may be affected by $T(n)$, for example by

[1] This tax function is quite simple in form because household utility functions are assumed to be identical. If households differ, say in their parameter σ, then we obtain a term like $\int_0^\infty \lambda^{-\sigma_n} f(n)\, dn$ in (A5.10) and we are forced to make an assumption about the shape of the distribution of inherent abilities.

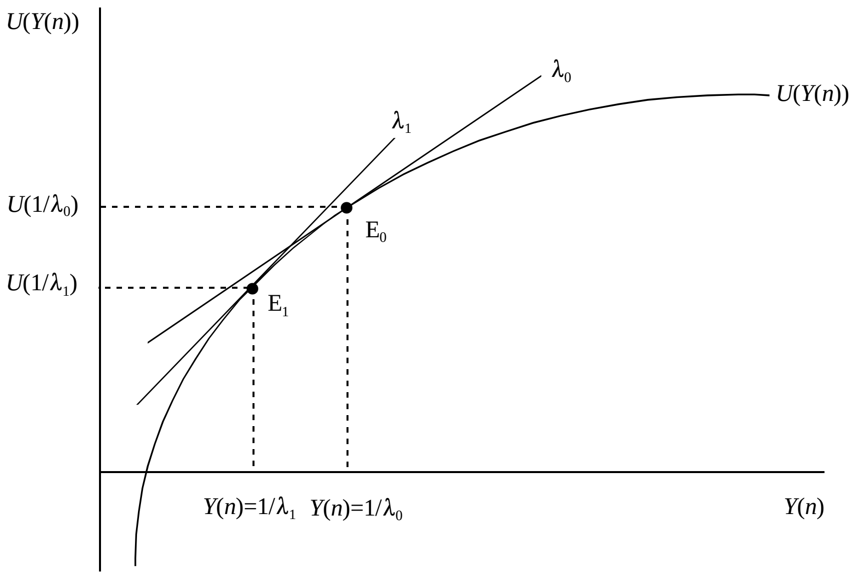

Figure 10.2: An increase in R_0 or decrease in $\bar{Z}$

labour supply reactions (see below). Second, it uses a rather restrictive utilitarian framework, i.e. the social welfare function may not be Benthamite (or may even be non-individualistic). Third, it ignores restrictions on the types of taxes that may be levied. The shape of tax system may have very important consequences if labour supply effects are allowed for.

10.3 The optimal linear income tax

The modern theory of income taxation was pioneered by James Mirrlees (1971). The key element of the modern approach is the central role it reserves for the (adverse) incentive effects of taxation under information asymmetry. An attractive aspect of the modern approach is that the conflict between equity and efficiency can be studied jointly in a single model. Because the modern theory of optimal income taxation is quite complex, we adopt a gradual method of attack (going from the easy to the more difficult material). The remainder of this section presents the (relatively straightforward) optimal linear income tax which was first studied by Sheshinski (1972). Attention is focused on the special case of a Benthamite social welfare function, and some intuitive remarks are made for the more general case.

In the next section we shift gear and study the optimal non-linear income tax of Mirrlees (1971). Again we approach the material in a number of steps. First, we study the special case with a Rawlsian social welfare function. Next, we deal with a general social welfare function but quasi-linear preferences. Finally, we deal with the most general case in an intuitive fashion and by means of some simulation results.

10.3.1 A simple model

The optimal linear income tax was first studied by Sheshinski (1972). In this subsection we use the formulation of Dixit and Sandmo (1977) which has the attractive feature that it stresses the similarity between the income taxation literature and the commodity taxation literature of the previous chapter. The model features the following basic assumptions. First, individuals differ according to innate ability but have identical utility functions. Second, the wage rate is exogenous and producer prices are fixed. Third, the social welfare function is Benthamite (utilitarian).

There are H households indexed by h and the direct utility function of household h is:

$$U^h \equiv U\left(C^h, 1 - L^h\right), \tag{A5.12}$$

where U^h, C^h, and $1 - L^h$ are, respectively, utility, consumption, and leisure of household h (L^h is labour supply). The utility function, $U(\cdot)$, takes the same form for all households. It has the usual properties of positive but diminishing marginal utility and indifference curves bulge towards the origin [REFER TO CHAPTER 2] ($U_C > 0$, $U_{1-L} > 0$, $U_{CC} < 0$, $U_{1-L,1-L} < 0$, and $U_{CC}U_{1-L,1-L} - U_{C,1-L} > 0$). The consumption good is the numeraire commodity and we set its price equal to unity ($P = 1$). Just as in the previous section, households are assumed to differ in their labour productivity (innate/exogenous skill differences). The effective labour supply in *efficiency units* is denoted by $n^h L^h$, and the representative firm maximizes profit,

$$\Pi \equiv \sum_{h=1}^{H} n^h L^h - \sum_{h=1}^{H} W^h L^h, \tag{A5.13}$$

by choice of L^h. It follows that the wage rate of household h is equal to that household's marginal product:

$$W^h = n^h. \tag{A5.14}$$

There is no non-labour income so the household budget constraint is:

$$C^h = W^h L^h - T^h, \tag{A5.15}$$

where T^h is the tax paid by household h.

The policy maker acts on the basis of an individualistic Benthamite (utilitarian) social welfare function:

$$SW \equiv \sum_{h=1}^{H} U^h, \tag{A5.16}$$

where SW is an index of social welfare. The policy maker's instrument consists of a *tax schedule* which relates the tax to the agent's economic performance, i.e. to the household's *labour income* ($W^h L^h$). In this section we focus on the simple case, where the tax schedule is linear in the tax base:

$$T^h = -S + t_L W^h L^h, \tag{A5.17}$$

where S is the lump-sum subsidy (or tax, if $S < 0$) and t_L is the (constant) marginal tax rate (obviously $0 < t_L < 1$). Note that since $t_L > 0$, the tax schedule is progressive (regressive) if $S > 0$ ($S < 0$). The informational asymmetry mentioned above originates from the implicit assumption that the policy maker cannot directly observe (or infer) the household's innate ability n^h (e.g. by IQ tests, higher degrees, diploma's etcetera) so that a tax on innate ability is not feasible. As a result, the policy maker faces a *second-best* social optimization problem! He can only observe $W^h L^h$ but not L^h and W^h separately.[2] (Of course, if an ability tax would be feasible, then the policy maker would face a first-best social optimization problem. See the Intermezzo below.)

Household h chooses C^h and L^h in order to maximize utility (A5.12) subject to the budget constraint (A5.15) and the linear tax schedule (A5.17). The Lagrangian for this maximization problem is:

$$\mathcal{L}^h \equiv U\left(C^h, 1 - L^h\right) + \alpha^h \left[S + (1 - t_L) W^h L^h - C^h\right],$$

where α^h is the Lagrange multiplier for the budget constraint (equalling the marginal utility of lump-sum income to household h in the optimum). The first-order conditions are:

$$\frac{\partial \mathcal{L}^h}{\partial C^h} = \frac{\partial U}{\partial C^h} - \alpha^h = 0, \tag{A5.18}$$

$$\frac{\partial \mathcal{L}^h}{\partial L^h} = -\frac{\partial U}{\partial \left(1 - L^h\right)} + \alpha^h (1 - t_L) W^h = 0. \tag{A5.19}$$

An implication of (A5.18)-(A5.19) is that the household equates the marginal rate of substitution between consumption and leisure to the after-tax wage rate ($U_{1-L}/U_C = W^h (1 - t_L)$), i.e. the household's labour supply decision is distorted if the marginal tax rate is non-zero.

Equations (A5.18)-(A5.19) and the budget constraint implicitly define Marshallian consumption demand and labour supply which we write as follows:

$$C^h = C\left(S, (1 - t_L) W^h\right), \tag{A5.20}$$

$$L^h = L\left(S, (1 - t_L) W^h\right). \tag{A5.21}$$

By substituting these expressions into the direct utility function (A5.12) we obtain the indirect utility

[2] In contrast, the firm is able to observe innate ability of its workers–see equation (A5.14) above.

function:

$$V^h = V\left(S, (1-t_L)\,W^h\right). \tag{A5.22}$$

Recall that the indirect utility function has the following properties:[3]

$$\frac{\partial V^h}{\partial S} = \frac{\partial V(\cdot)}{\partial S} = \alpha^h, \tag{A5.23}$$

$$\frac{\partial V^h}{\partial (1-t_L)} = \frac{\partial V(\cdot)}{\partial (1-t_L)} = \alpha^h W^h L^h. \tag{A5.24}$$

10.3.2 The second-best social optimum

In the second-best social optimum, the policy maker chooses S and t_L in order to maximize social welfare,

$$SW \equiv \sum_{h=1}^{H} V\left(S, (1-t_L)\,W^h\right), \tag{A5.25}$$

subject to the revenue requirement constraint:

$$\sum_{h=1}^{H}\left[t_L W^h L^h(\cdot) - S\right] = R_0, \tag{A5.26}$$

where R_0 is the exogenous (net) revenue requirement and $L^h(\cdot)$ is the Marshallian labour supply curve (A5.21). The Lagrangian expression for this maximization problem is:

$$\mathcal{H} \equiv \sum_{h=1}^{H} V\left(S, (1-t_L)\,W^h\right) + \lambda\left[\sum_{h=1}^{H}\left[t_L W^h L^h\left(S, (1-t_L)\,W^h\right) - S\right] - R_0\right],$$

where λ is the Lagrange multiplier for the revenue requirement constraint (A5.26). The two interesting first-order necessary conditions are:

$$\frac{\partial \mathcal{H}}{\partial S} = \sum_{h=1}^{H}\frac{\partial V}{\partial S} + \lambda t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial S} - \lambda H = 0, \tag{A5.27}$$

$$\frac{\partial \mathcal{H}}{\partial t_L} = -\sum_{h=1}^{H}\frac{\partial V}{\partial (1-t_L)} + \lambda\left[\sum_{h=1}^{H} W^h L^h - t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial (1-t_L)}\right] = 0. \tag{A5.28}$$

[3] Roy's Lemma for labour supply says:

$$L^h = \frac{\partial V/\partial\left[(1-t_L)\,W^h\right]}{\partial V/\partial S} = \frac{1}{W^h}\frac{\partial V/\partial(1-t_L)}{\alpha^h},$$

from which the result in (A5.24) follows.

By using (A5.23)-(A5.24) these expressions can be further simplified to:

$$0 = \sum_{h=1}^{H} \alpha^h + \lambda t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial S} - \lambda H, \tag{A5.29}$$

$$0 = \sum_{h=1}^{H} \alpha^h W^h L^h - \lambda \sum_{h=1}^{H} W^h L^h + \lambda t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial (1-t_L)}. \tag{A5.30}$$

Equation (A5.29) is the condition determining the optimal lump-sum subsidy (S, the revenue raising instrument) whilst (A5.30) is the condition determining the optimal marginal tax rate (t_L, the redistribution instrument). Together with (A5.26) these conditions determine (S, t_L, λ).

Before developing the various expressions for the optimal tax formulae suggested by Dixit and Sandmo (1977) we test our intuition by asking what would be the optimum choice of (S, t_L) if all households were identical. Intuitively one would expect that, since there is no need for redistribution, there should be no distorting labour income tax and all revenue should raised by means of the lump-sum subsidy. Does the theory also give that answer? Technically, with identical households we have $W^h = W$, $L^h = L$, and $\alpha^h = \alpha$ (for all $h = 1, \cdots, H$) so that (A5.29), (A5.30), and (A5.26) simplify to:

$$0 = H\left[(\alpha - \lambda) + \lambda t_L W \frac{\partial L}{\partial S}\right], \tag{A5.31}$$

$$0 = HW\left[(\alpha - \lambda) L + \lambda t_L \frac{\partial L}{\partial (1-t_L)}\right], \tag{A5.32}$$

$$R_0 = H\left[t_L W L - S\right]. \tag{A5.33}$$

The Slutsky equation for labour supply is:

$$\frac{\partial L}{\partial (1-t_L)} = S_{LL} + WL\frac{\partial L}{\partial S}, \tag{A5.34}$$

where $S_{LL} \equiv (\partial L / \partial (1-t_L))_{U_0} > 0$ is the pure substitution effect. By using this expression in (A5.32) we obtain:

$$\left[(\alpha - \lambda) + \lambda t_L W \frac{\partial L}{\partial S}\right] L + \lambda t_L S_{LL} = 0, \tag{A5.35}$$

and it follows from (A5.31) and (A5.35) that the unique solution is $\alpha = \lambda$ and $t_L = 0$. To summarize, the socially optimal solution is:

$$\alpha = \lambda, \quad t_L = 0, \quad S = -\frac{R_0}{H}. \tag{A5.36}$$

These solutions accord with intuition. The policy maker raises revenue in a non-distorting fashion by ensuring that the marginal dollar taken from the household costs privately as much as it yields socially ($\alpha = \lambda$). The policy maker does not distort the labour supply decision because there is no need for redistribution ($t_L = 0$), and all households are identical so they all pay the same lump-sum tax. The intuition is this borne out by the technical expressions. Note that in this case we get exactly the same solution as in the sacrifice theory discussed above.

Intermezzo 10.1

The first-best optimal linear income tax. Assume for the sake of argument that the policy maker can observe innate ability, n^h. What kind of a tax would the policy maker choose? Clearly, in such a setting the instruments can be individualized, i.e. the policy maker's instruments are S^h and t_L^h (for $h = 1, \cdots, H$). In the first-best social optimum, the policy maker's objective function is:

$$SW \equiv \sum_{h=1}^{H} V\left(S^h, \left(1 - t_L^h\right) W^h\right), \tag{A}$$

and the revenue requirement is:

$$\sum_{h=1}^{H} \left[t_L^h W^h L^h - S^h\right] = R_0, \tag{B}$$

where R_0 is the exogenous (net) revenue requirement. The policy maker chooses S^h and t_L^h in order to maximize (A) subject to (B). The Lagrangian expression for the maximization problem is:

$$\begin{aligned}\mathcal{H} \equiv & \sum_{h=1}^{H} V\left(S^h, \left(1 - t_L^h\right) W^h\right) \\ & + \lambda\left[\sum_{h=1}^{H}\left[t_L^h W^h L^h\left(S^h, \left(1 - t_L^h\right) W^h\right) - S^h\right] - R_0\right],\end{aligned}$$

where λ is the Lagrange multiplier for the revenue requirement constraint (B). The interesting first-order necessary conditions are

$$\frac{\partial \mathcal{H}}{\partial S^h} = \frac{\partial V}{\partial S^h} + \lambda t_L^h W^h \frac{\partial L^h}{\partial S^h} - \lambda = 0, \tag{C}$$

$$\frac{\partial \mathcal{H}}{\partial t_L^h} = -\frac{\partial V}{\partial\left(1 - t_L^h\right)} + \lambda\left[W^h L^h - t_L^h W^h \frac{\partial L^h}{\partial\left(1 - t_L^h\right)}\right] = 0, \tag{D}$$

for $h = 1, \ldots, H$.

Equations (A5.23) and (A5.24) are modified to:

$$\frac{\partial V^h}{\partial S^h} = \frac{\partial V(\cdot)}{\partial S^h} = \alpha^h, \tag{E}$$

$$\frac{\partial V^h}{\partial\left(1-t_L^h\right)} = \frac{\partial V(\cdot)}{\partial\left(1-t_L^h\right)} = \alpha^h W^h L^h, \tag{F}$$

where α^h is the marginal utility of income to household h. Using (E)-(F), the first-order conditions (C)-(D) can be simplified to:

$$0 = \alpha^h + \lambda t_L^h W^h \frac{\partial L^h}{\partial S} - \lambda, \tag{G}$$

$$0 = \alpha^h W^h L^h - \lambda W^h L^h + \lambda t_L^h W^h \frac{\partial L^h}{\partial\left(1-t_L^h\right)}. \tag{H}$$

Equation (G) is the condition determining the optimal lump-sum subsidies (S^h), i.e. the revenue raising instruments. Equation (H) is the condition determining the optimal marginal tax rates (t_L^h), i.e. the redistribution instruments. Together with (B) these conditions jointly determine $\left(S^h, t_L^h, \lambda\right)$ (note that there are $2H+1$ independent equations). The solution is:

$$\alpha^h = \lambda, \quad (\text{for } h = 1, \ldots, H), \tag{I}$$

$$t_L^h = 0, \quad (\text{for } h = 1, \ldots, H), \tag{J}$$

$$\frac{\partial V\left(S^h, n^h\right)}{\partial S^h} = \alpha^h. \tag{K}$$

Intuitively, the policy maker raises revenue in a non-distorting fashion ($t_L^h = 0$) by choosing S^h appropriately such that the marginal utility of income is equated for all agents, taking into account that $W^h = n^h$. It follows that the first-best optimal income tax is a lump-sum ability tax.

****

10.3.3 Formulae for the second-best optimal tax

Dixit and Sandmo (1977, p. 419) show that (for the general case with heterogeneous households) the two first-order conditions (A5.29) and (A5.30) can be rewritten in various alternative formats which stress the similarity with the optimal commodity tax formulae studied in the previous chapter. We start with (A5.30), the condition determining the optimal marginal income tax. It can be rewritten as follows:

$$t_L \sum_{h=1}^{H} \frac{\partial W^h L^h}{\partial\left(1-t_L\right)} = -\frac{1}{\lambda} \sum_{h=1}^{H} \alpha^h W^h L^h + \sum_{h=1}^{H} W^h L^h \quad \Leftrightarrow$$

$$t_L \frac{\partial \sum_{h=1}^{H} W^h L^h}{\partial\left(1-t_L\right)} = -\frac{1}{\lambda} \sum_{h=1}^{H} \alpha^h W^h L^h + \sum_{h=1}^{H} W^h L^h \quad \Leftrightarrow$$

$$\frac{t_L}{1-t_L}\frac{\partial\sum_{h=1}^{H}W^hL^h}{\partial\left(1-t_L\right)}\frac{1-t_L}{\sum_{h=1}^{H}W^hL^h}=\frac{1}{\lambda}\left[\lambda-\frac{\sum_{h=1}^{H}\alpha^hW^hL^h}{\sum_{h=1}^{H}W^hL^h}\right]\quad\Leftrightarrow$$

$$\frac{t_L}{1-t_L}=\frac{\lambda-\bar{\alpha}}{\lambda}\frac{1}{\sigma_L},\tag{A5.37}$$

where σ_L and $\bar{\alpha}$ are defined as:

$$\sigma_L\equiv\frac{\partial\sum_{h=1}^{H}n^hL^h}{\partial\left(1-t_L\right)}\frac{1-t_L}{\sum_{h=1}^{H}n^hL^h},\tag{A5.38}$$

$$\bar{\alpha}\equiv\frac{\sum_{h=1}^{H}\alpha^hn^hL^h}{\sum_{h=1}^{H}n^hL^h},\tag{A5.39}$$

and where we have used the fact that $W^h=n^h$ in the final expressions. The interpretation of (A5.37) is as follows. The parameter σ_L is the elasticity of *aggregate* labour supply in efficiency units with respect to the after-tax wage. Under the assumption that individual labour supply is increasing in the wage ($\partial L^h/\partial\left(1-t_L\right)>0$), this elasticity is positive ($\sigma_L>0$). Assuming that t_L is non-negative, it follows from (A5.37) that λ is greater than $\bar{\alpha}$.[4] In this scenario, since $\sigma_L>0$ and $\lambda>\bar{\alpha}$, the optimal marginal tax is higher, the lower is σ_L. Equation (A5.37) is thus like an inverse-elasticity formula one often finds for optimal commodity taxation (see Chapter 10 for examples). Redistributional concerns enter via the term, $\bar{\alpha}$, involving the weighted average of marginal utilities of income, with tax bases acting as weights. Ceteris paribus, the optimal tax is higher, the lower is $\bar{\alpha}$, i.e. the stronger the tendency for high-income individuals is to have a low marginal utility of income (α^h).[5]

Dixit and Sandmo (1977, p. 420) also develop an alternative formula which is expressed in terms of the tax elasticity of consumption demand. The aggregate production constraint is:

$$C\equiv\sum_{h=1}^{H}C^h=\sum_{h=1}^{H}n^hL^h,\tag{A5.40}$$

where C is aggregate consumption. By using (A5.40) and (A5.38) in (A5.37) we find the alternative formula for the optimal labour income tax:

$$\frac{t_L}{1-t_L}=\frac{\lambda-\bar{\alpha}}{\lambda}\frac{1}{\frac{\partial C}{\partial(1-t_L)}\frac{1-t_L}{C}}.\tag{A5.41}$$

A third formula developed by Dixit and Sandmo (1977, pp. 421-422) makes use of a covariance term, not unlike the one we studied in the many-person Ramsey rule in Chapter 10. We return to (A5.29)-

[4]Note that it follows from (A5.29) that λ is certainly positive if t_L is non-negative and leisure is a normal good ($\partial L^h/\partial S<0$).

[5]In the special case with identical households ($\alpha^h=\alpha$ for all h), equation (A5.37) simplifies to:

$$\frac{t_L}{1-t_L}=\frac{\lambda-\alpha}{\lambda}\frac{1}{\sigma_L}.$$

This expression holds regardless of (A5.31) (which determines S). If S is set optimally, then $\lambda=\alpha$ and it is optimal to set $t_L=0$: the distorting tax is not used for revenue raising. If, in contrast, S is not set optimally, then t_L is also used to raise revenue.

(A5.30) and define the following auxiliary variables:

$$\beta^h \equiv \alpha^h + \lambda t_L W^h \frac{\partial L^h}{\partial S}, \tag{A5.42}$$

$$\bar{\beta} \equiv \sum_{h=1}^{H} \frac{\beta^h}{H}, \tag{A5.43}$$

where β^h is called the *social marginal utility of income* to household h. The intuition behind this term is as follows. A marginal increase in the lump-sum subsidy increases private utility by α^h. It also reduces labour supply (provided leisure is a normal good), erodes the tax base, and reduces tax revenue by $t_L W^h \frac{\partial L^h}{\partial S}$. This is valued at the marginal utility of income to the policy maker (i.e. λ). Note that in (A5.43), $\bar{\beta}$ is the average value of β^h over the population.

Using these definitions, the first-order conditions (A5.29)-(A5.30) can be rewritten as follows:

$$\bar{\beta} = \lambda, \tag{A5.44}$$

$$t_L = -\frac{1}{\lambda} \frac{\text{cov}\left(\beta^h, W^h L^h\right)}{\frac{1}{H}\sum_{h=1}^{H} W^h S_{LL}^h}, \tag{A5.45}$$

where $\text{cov}\left(\beta^h, W^h L^h\right)$ is the covariance between β^h and $W^h L^h$, $S_{LL}^h \equiv \left(\partial L^h / \partial\left(1 - t_L\right)\right)_{U_0} > 0$ is the pure substitution effect in the labour supply of household h. Note that equation (A5.44) follows directly from (A5.42)-(A5.43) and (A5.29). The derivation of (A5.45), however, is non-trivial–see the Intermezzo. The interpretation of (A5.44) and (A5.45) is as follows. According to (A5.44), the lump-sum element of the tax (i.e. the subsidy S) should be adjusted in such a way that the *average* social marginal utility of the subsidy to the households ($\bar{\beta}$) equals the cost to the policy maker of making that transfer (λ). In equation (A5.45), the marginal tax rate contains an *equity element* (the numerator) and an *efficiency element* (the denominator). To interpret this expression, note that since $S_{LL}^h > 0$ we know for sure that the denominator is positive. The sign of the numerator is somewhat more problematic to establish although the presumption is that it is negative. There are two effects in operation. First, it is reasonable to assume that $\text{cov}\left(\alpha^h, W^h L^h\right) < 0$, i.e. that private marginal utility of income falls as income rises. This effect tends to make the covariance between β^h and $W^h L^h$ negative. Second, if leisure is a normal good, $\partial L^h / \partial S < 0$ so we expect $\text{cov}\left(W^h \partial L^h / \partial S, W^h L^h\right) < 0$ also. Hence, as Atkinson and Stiglitz (1980, p. 408) point out, only if leisure becomes inferior at high wage rates is it is possible for this term to cause ambiguity for the sign of $\text{cov}\left(\beta^h, W^h L^h\right)$.

10.3.4 Closing remarks

We conclude this discussion on the optimal linear income tax with a number of closing remarks. First, referring to the optimal tax formula (A5.45) it is easy to establish the link between the first-best and second best policies. In the first-best policy, the policy maker sets $t_L = 0$ and $\alpha^h = \lambda$ for all h (by using a lump-sum ability tax) so that it follows from (A5.42) that $\beta^h = \lambda$, i.e. cov $\left(\beta^h, W^h L^h\right) = 0$ in that case. Equation (A5.45) confirms that there is no need for distorting taxation in that case.

The second remark deals with the implications of assuming a more general social welfare function than the Benthamite formulation given in (A5.16) above. Suppose that the social planner has a general social welfare function of the form:

$$SW \equiv \Psi\left(U^1, U^2, \ldots, U^H\right), \tag{A5.46}$$

where $\Psi_h \equiv \partial\Psi/\partial U^h$ is the marginal weight of household h's utility in social welfare. Using the same steps as before we obtain the following first-order conditions:

$$0 = \sum_{h=1}^{H} \alpha^h \Psi_h + \lambda t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial S} - \lambda H, \tag{A5.47}$$

$$0 = \sum_{h=1}^{H} \alpha^h \Psi_h W^h L^h - \lambda \sum_{h=1}^{H} W^h L^h + \lambda t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial (1-t_L)}. \tag{A5.48}$$

Hence, compared to (A5.29)-(A5.30), the only change that occurs is that private marginal utility of income (α^h) is now weighted by the policy maker's marginal weight for household h (Ψ_h). By suitably modifying the definition of β^h in (A5.42) we still obtain the formulae (A5.44)-(A5.45).

Intermezzo 10.2

Derivation of the covariance formula (A5.45). Derivations for the covariance formula are also given in Stiglitz (1987) for the continuous case and in Dixit and Sandmo (1977) for the discrete case. *Step 1*: We recall the Slutsky equation for labour supply:

$$\frac{\partial L^h}{\partial (1-t_L)} = W^h L^h \frac{\partial L^h}{\partial S} + S_{LL}^h, \tag{A}$$

where $S_{LL}^h \equiv \left(\partial L^h/\partial (1-t_L)\right)_{U_0} > 0$ is the pure substitution effect. Using (A) we can write (A5.30) in terms of β^h:

$$\begin{aligned}
0 &= \sum_{h=1}^{H} \left[\alpha^h W^h L^h - \lambda W^h L^h + \lambda t_L W^h \left(W^h L^h \frac{\partial L^h}{\partial S} + S_{LL}^h\right)\right] \\
0 &= \sum_{h=1}^{H} W^h L^h \left[\alpha^h - \lambda + \lambda t_L W^h \frac{\partial L^h}{\partial S} + \lambda t_L \frac{S_{LL}^h}{L^h}\right] \\
0 &= \sum_{h=1}^{H} W^h L^h \left[\beta^h - \bar{\beta} + \lambda t_L \frac{S_{LL}^h}{L^h}\right]
\end{aligned} \tag{B}$$

where we have used the fact that $\lambda = \bar{\beta}$ (see (A5.44)) in the final step.

Step 2: We solve (B) for t_L and obtain:

$$-\lambda t_L \sum_{h=1}^{H} W^h S_{LL}^h = \sum_{h=1}^{H} W^h L^h \left(\beta^h - \bar{\beta}\right) \quad \Leftrightarrow$$

$$t_L = -\frac{1}{\lambda}\frac{\frac{1}{H}\sum_{h=1}^{H} W^h L^h \left(\beta^h - \bar{\beta}\right)}{\frac{1}{H}\sum_{h=1}^{H} W^h S_{LL}^h}. \tag{C}$$

The covariance between variables x_h and y_h is defined as follows:

$$\text{cov}\left(x_h, y_h\right) \equiv \frac{1}{H}\sum_{h=1}^{H}\left(x_h - \bar{x}\right)\left(y_h - \bar{y}\right) = \frac{1}{H}\sum_{h=1}^{H} x_h y_h - \bar{x}\bar{y}, \tag{D}$$

where $\bar{x} \equiv \sum_{h=1}^{H} x_h / H$ and $\bar{y} \equiv \sum_{h=1}^{H} y_h / H$ are the respective means. By using this definition in (C) we obtain:

$$\begin{aligned} t_L &= -\frac{1}{\lambda}\frac{\frac{1}{H}\sum_{h=1}^{H}\left(W^h L^h - \overline{WL} + \overline{WL}\right)\left(\beta^h - \bar{\beta}\right)}{\frac{1}{H}\sum_{h=1}^{H} W^h S_{LL}^h} \\ &= -\frac{1}{\lambda}\frac{\text{cov}\left(W^h L^h, \beta^h\right)}{\frac{1}{H}\sum_{h=1}^{H} W^h S_{LL}^h}, \end{aligned} \tag{E}$$

where we have used the fact that:

$$\frac{1}{H}\sum_{h=1}^{H}\overline{WL}\left(\beta^h - \bar{\beta}\right) = \frac{\overline{WL}}{H}\sum_{h=1}^{H}\left(\beta^h - \bar{\beta}\right) = 0. \tag{F}$$

****

10.4 The optimal non-linear income tax

The theory of optimal non-linear income taxation was pioneered by Mirrlees (1971, 1976, 1986). Further noteworthy contributions were made by Sadka (1976), Seade (1977, 1982), Ebert (1992), and Diamond (1998). The key element in this approach is the link between the income tax system and the labour supply behaviour of households with different abilities (and wage incomes). The material is technically rather advanced, and in this section we therefore first develop the general model and then study some special cases to build intuition. The key question we wish to answer is: should the *marginal* tax rate rise or fall with income?

The model features the following main assumptions. Just as in the classical sacrifice theory discussed above, household skills are distributed according to some continuous distribution. Individual earning ability is n and the before-tax earning of a household of type n is defined as:

$$
\begin{aligned}
Z(n) &\equiv W(n)\,L(n) \\
&= nL(n),
\end{aligned}
\tag{A5.49}
$$

where $L(n)$ is labour supply (in hours) and $W(n)$ is the real wage rate (which equals n, i.e. firms can observe the individual's productivity). The tax schedule is rather general in form, but is conditioned on what the policy maker can actually observe, i.e. income. The budget constraint for a household of type n is thus:

$$
C(n) = Z(n) - T(Z(n)),
\tag{A5.50}
$$

where $C(n)$ is consumption and $T(Z(n))$ is the income tax. In addition the household has a time endowment of unity and (obviously) cannot supply a negative amount of labour. Hence, the time constraint is:

$$
0 \le L(n) \le 1.
\tag{A5.51}
$$

The *cumulative distribution* of people of type n is denoted by $F(n)$ and the *density function* is $f(n) \equiv F'(n)$.

The utility function is assumed to be the same for all households:

$$
U^n \equiv U(C(n), L(n)),
\tag{A5.52}
$$

where $U(\cdot)$ has the usual properties: $U_C > 0$, $U_{CC} < 0$, $U_L < 0$, $U_{LL} < 0$, and $U_{CC}U_{LL} - U_{CL}^2 > 0$. We need some more assumptions in order to keep the analysis tractable. First, in order to rule out zero-leisure or zero-consumption corner solutions we assume:

$$
\lim_{L(n)\to 1} U_L(\cdot) \to -\infty,
\tag{A5.53}
$$

$$
\lim_{C(n)\to 0} U_C(\cdot) \to +\infty.
\tag{A5.54}
$$

According to (A5.53), if the household is nearly fully employed ($L(n) \approx 1$) then the marginal disutility of labour supply becomes infinite. Similarly, (A5.54) says that at low consumption levels the marginal utility of consumption is very high. In Figure 10.3 we show indifference curves in (C, L)-space.[6] Note

[6] The slope of the indifference curve is $dC/dL = -U_L/U_C$. We write the implicit function $C = f(L)$ with $f'(L) = -U_L/U_C$. By differentiating the slope of the indifference curve we get:

$$
\frac{d^2C}{dL^2} = \frac{U_C\left[U_{LC}f' + U_{LL}\right] - U_L\left[U_{CC}f' + U_{CL}\right]}{-U_C^2} = \frac{-U_C^2U_{LL} + 2U_LU_CU_{CL} - U_L^2U_{CC}}{U_C^3} > 0,
$$

where the sign follows from the fact that the numerator is a positive definite quadratic form in U_C and U_L.

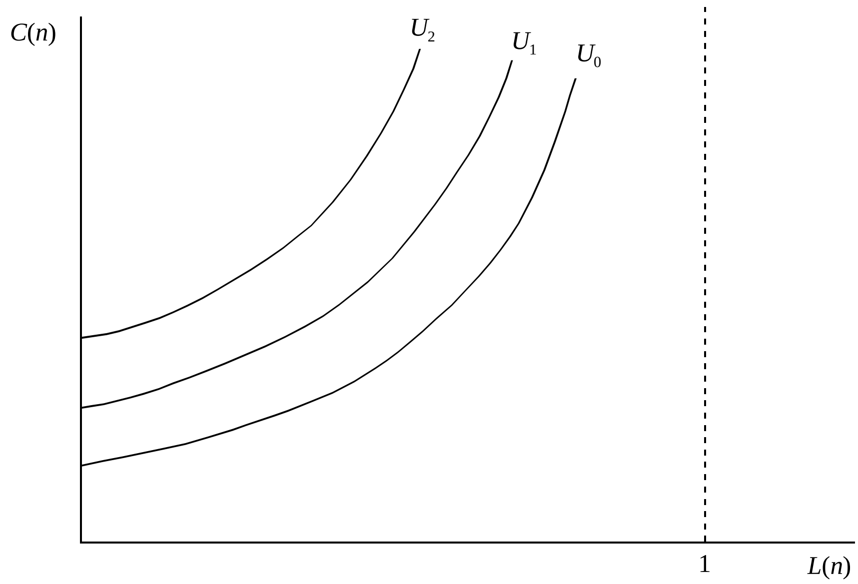

Figure 10.3: Indifference curves in $C(n), L(n)$ space

that utility increases in north-westerly direction, i.e. $U_2 > U_1 > U_0$.

It turns out to be useful to write utility in terms of consumption and income $Z(n)$ (rather than labour supply) by noting (A5.49):

$$U^n = U\left(C(n), \frac{Z(n)}{n}\right) \equiv u\left(C(n), Z(n), n\right). \qquad \text{(A5.55)}$$

In Figure 10.4 we illustrate household indifference curves in (C, Z)-space for three different skill levels. Intuitively, holding constant Z, $U_0[n > n_0]$ exceeds $U_0[n = n_0]$ because the more skillful person does not have to work so hard to get that particular income level (and thus get a higher utility level).

Armed with the utility specification (A5.55) we can formulate an important additional assumption, called the *Agent Monotonicity Assumption* (or *Spence-Mirrlees Single Crossing Property*). Technically we have agent monotonicity if the marginal rate of substitution between consumption and pre-tax income is a decreasing function of n:

$$\frac{\partial}{\partial n}\left[-\frac{u_Z}{u_C}\right] < 0, \qquad \text{(A5.56)}$$

where $u_Z \equiv \partial u/\partial Z$ and $u_C \equiv \partial u/\partial C$. A number of useful results can now be established (see Myles (1995, pp. 136-138) or the Intermezzo for proofs):

(R1) The condition (A5.56) is equivalent to the following condition expressed in terms of the original

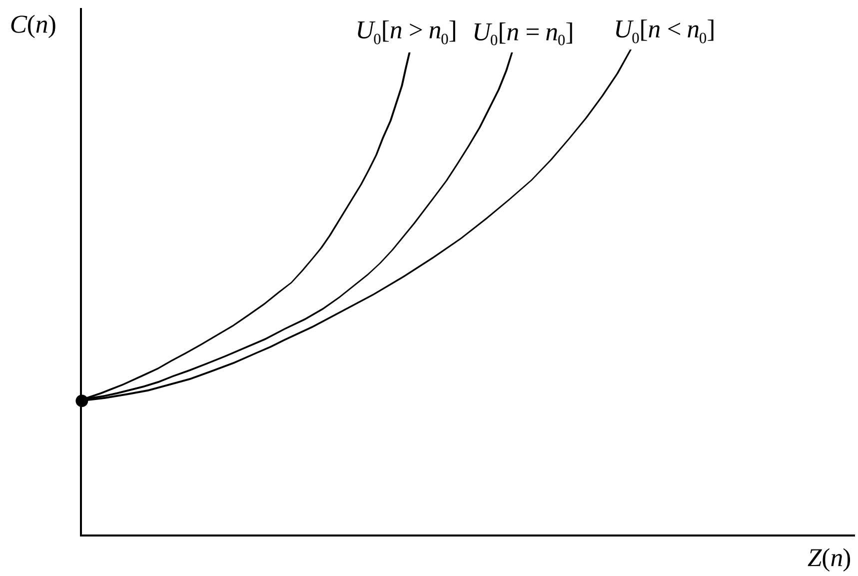

Figure 10.4: Indifference curves in $C(n), Z(n)$ space

utility function:

$$\frac{\partial}{\partial L}\left[-\frac{LU_L}{U_C}\right] > 0. \tag{A5.57}$$

(R2) The agent monotonicity assumption implies that in the absence of taxation, consumption increases as the wage increases (a sufficient condition is that consumption is not inferior).

(R3) When $L > 0$ and the sufficient conditions for utility maximization are satisfied then gross income, $Z(n)$, is increasing in ability, n, i.e. $Z'(n) > 0$.

(R4) The agent monotonicity assumption implies that any two indifference curves of households of different abilities only cross once. See Figure 10.5 for this so-called "single crossing" result.

Intermezzo 10.3

Derivation of results (R1)-(R4). The proof of result (R1) proceeds as follows. We define:

$$\Phi(n) \equiv -\frac{u_Z(C,Z,n)}{u_C(C,Z,n)}, \tag{A}$$

and differentiate with respect to n:

$$\frac{\partial \Phi(n)}{\partial n} = -\frac{u_C u_{Zn} - u_Z u_{Cn}}{u_C^2} < 0. \tag{B}$$

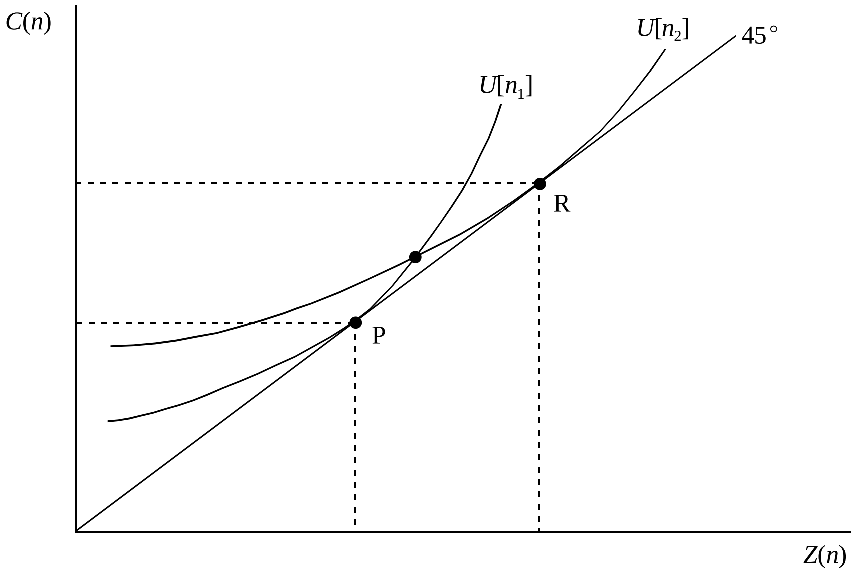

Figure 10.5: Spence-Mirrlees single-crossing assumption

From the definition of $u\,(C, Z, n)$ in (A5.55) we know that:

$$\begin{aligned} u_C &\equiv \frac{\partial U}{\partial C} = U_C, \quad u_Z \equiv \frac{\partial U}{\partial Z} = \frac{U_L}{n}, \\ u_n &\equiv \frac{\partial U}{\partial n} = -\frac{U_L}{n^2}, \quad u_{Cn} \equiv \frac{\partial^2 U}{\partial C \partial n} = -U_{CL}\frac{Z}{n^2}, \\ u_{Zn} &\equiv \frac{\partial^2 U}{\partial Z \partial n} = \frac{n\frac{\partial U_L}{\partial n} - U_L}{n^2} = -\frac{LU_{LL} + U_L}{n^2}. \end{aligned} \tag{C}$$

We find by differentiation:

$$\frac{\partial}{\partial L}\left(\frac{LU_L}{U_C}\right) = \frac{U_C\,[LU_{LL} + U_L] - LU_L U_{CL}}{U_C^2}. \tag{D}$$

By using the relevant results from (C) in (D) we obtain:

$$\frac{\partial}{\partial L}\left(\frac{LU_L}{U_C}\right) = -n^2\left[\frac{u_C u_{Zn} - u_Z u_{Cn}}{u_C^2}\right] = n^2\frac{\partial \Phi\,(n)}{\partial n} < 0. \tag{E}$$

Hence, (A5.56) and (A5.57) are indeed equivalent.

The proofs of results (R2) and (R3) proceeds as follows. We postulate the existence of a "consumption function" which relates $C\,(n)$ to income $Z\,(n)$, i.e. $C \equiv \Gamma(Z)$. We assume also that $\Gamma\,(Z)$ is twice differentiable. In the absence of taxation, the household budget constraint is $C = Z$ and the household chooses Z to maximize:

$$U\,(Z) \equiv U\left(\Gamma\,(Z)\,, \frac{Z}{n}\right). \tag{F}$$

The first-order condition for this problem is:

$$\begin{aligned}\frac{\partial U(Z)}{\partial Z} &= U_C \Gamma'(Z) + \frac{U_L}{n} = 0 \quad \Leftrightarrow \\ &= \frac{U_C}{Z}\left[Z\Gamma'(Z) + \frac{LU_L}{U_C}\right] = 0.\end{aligned} \tag{G}$$

We define the following function:

$$\Psi(Z,n) \equiv Z\Gamma'(Z) + \frac{LU_L}{U_C}, \tag{H}$$

so that in the optimum, Z is chosen such that $\Psi(Z^*,n) = 0$. The second-order condition for a maximum is:

$$\left[\frac{\partial^2 U(Z)}{\partial Z^2}\right]_{Z=Z^*} < 0,$$

or, equivalently:

$$\left[\frac{\partial \Psi(Z,n)}{\partial Z}\right]_{Z=Z^*} < 0. \tag{I}$$

By totally differentiating the first-order condition $\Psi(Z^*,n) = 0$ we obtain:

$$\begin{aligned} 0 &= d\Psi(Z^*,n) \quad \Leftrightarrow \\ 0 &= \left[\frac{\partial \Psi(Z,n)}{\partial Z}\right]_{Z=Z^*} dZ^* + \frac{\partial \Psi(Z^*,n)}{\partial n} dn \quad \Leftrightarrow \\ \frac{dZ^*}{dn} &= \frac{\partial \Psi(Z^*,n)/\partial n}{-\left[\partial \Psi(Z,n)/\partial Z\right]_{Z=Z^*}},\end{aligned} \tag{J}$$

where the denominator is positive by (I). We know that the numerator is also positive:

$$\frac{\partial \Psi(Z^*,n)}{\partial n} = \frac{\partial}{\partial L}\left[\frac{LU_L}{U_C}\right]\frac{\partial L}{\partial n} = n^2 \frac{\partial \Phi(n)}{\partial n}\frac{-Z}{n^2} > 0, \tag{K}$$

where we have used $L = Z/n$ and (A5.57). It thus follows from (J) and (K) that:

$$\frac{dZ^*}{dn} > 0, \quad \frac{dC^*}{dn} > 0, \tag{L}$$

where we have used the fact that $C^* = Z^*$ to get the second expression.

Finally, the proof of result (R4) is as follows. The first-order condition for utility maximization can be written as:

$$-\frac{u_Z}{u_C} = 1. \tag{M}$$

In Figure 10.5 the optimum is at the point where the indifference curve is tangent to the 45-degree line. In that figure we show the optimum for two skill levels, where $n_2 > n_1$. We have already shown that $dZ^*/dn > 0$ and $dC^*/dn > 0$ so it must be the case that the two optima are at points P and R, respectively. It follows that the indifference curves cross only once.

****

10.4.1 The self-selection constraint

The policy maker's choice problem can be viewed in two equivalent ways. The problem can be cast as a choice of a particular income tax function, $T(Z)$. Equivalently, one can cast the optimal tax problem as a choice of pre-tax income-consumption pairs $(C(n), Z(n))$ for different households. In computing the optimal tax function the so-called *self-selection constraints* are of vital importance. These constraints say that households must find it in their own self-interest to actually choose the $(C(n), Z(n))$ pair that the policy maker intends for *them* rather than a pair meant for another type of household. If the self-selection constraints are satisfied then the tax policy is called *implementable*.

Let $C(n)$ and $Z(n)$ be the consumption-income combination that the policy maker intends for n-type households. Then the self-selection constraints are satisfied if and only if:

$$u(C(n), Z(n), n) \geq u\left(C\left(n'\right), Z\left(n'\right), n\right), \quad \text{for all } n, n'. \tag{A5.58}$$

The technical problem with constraints like (A5.58) is that they involve a double-infinity of constraints, i.e. there is a continuum of types and each type must satisfy (A5.58). This makes it difficult to incorporate (A5.58) as a constraint on the optimization problem.

The solution to this problem was proposed by Mirrlees (1976). It consists of looking for an alternative but equivalent set of constraints which ensures that the self-selection constraints are satisfied. It turns out that the following two conditions ensure that (A5.58) holds:

$$u_C(C(n), Z(n), n) C'(n) + u_Z(C(n), Z(n), n) Z'(n) = 0, \tag{A5.59}$$

$$Z'(n) \geq 0. \tag{A5.60}$$

These conditions are derived in an Intermezzo below. Note that (A5.59)-(A5.60) must hold for all n. Equation (A5.59) is the "first-order" necessary condition whilst (A5.60) is the "second-order" sufficient condition. Intuitively, (A5.59)-(A5.60) represent the optimality conditions of the maximization problem: "it is utility maximizing to pose as oneself rather than to pose as someone else."

Intermezzo 10.4

Derivation of (A5.59)-(A5.60). The **derivation of (A5.59)-(A5.60)** proceeds as follows. The self-selection constraints (A5.58) are reproduced here:

$$u\left(C\left(n\right), Z\left(n\right), n\right) \geq u\left(C\left(n'\right), Z\left(n'\right), n\right), \quad \text{for all } n, n'. \tag{A}$$

We define the following function:

$$V\left(n', n\right) \equiv u\left(C\left(n'\right), Z\left(n'\right), n\right), \tag{B}$$

where $V\left(n', n\right)$ represents the utility level of an n-type household who claims to be of type n' (who poses as somebody else). The self-selection constraints imply that $V\left(n, n'\right)$ must attain a maximum for $n' = n$. The first-order necessary condition for this to be the case is:

$$\frac{\partial V\left(n, n'\right)}{\partial n'} = 0, \quad \text{for } n' = n, \tag{C}$$

whilst the second-order sufficient condition is:

$$\frac{\partial^2 V\left(n, n'\right)}{\partial\left(n'\right)^2} \leq 0, \quad \text{for } n' = n. \tag{D}$$

We first work on the first-order condition (C). From (B) we derive:

$$\frac{\partial V\left(n, n'\right)}{\partial n'} = u_C C'\left(n'\right) + u_Z Z'\left(n'\right). \tag{E}$$

By using (E) and (C) we find (A5.59):

$$u_C\left(C\left(n\right), Z\left(n\right), n\right) C'\left(n\right) + u_Z\left(C\left(n\right), Z\left(n\right), n\right) Z'\left(n\right) = 0. \tag{F}$$

Next we work on the second-order condition (D). By differentiating (C) with respect to $n' = n$ we find:

$$\frac{d}{dn}\left[\frac{\partial V\left(n, n'\right)}{\partial n'}\right] = \frac{\partial^2 V\left(n, n'\right)}{\partial n \partial n'} + \frac{\partial^2 V\left(n, n'\right)}{\partial\left(n'\right)^2} = 0. \tag{G}$$

It follows that (D) can be written as:

$$\frac{\partial^2 V\left(n, n'\right)}{\partial n \partial n'} \geq 0, \quad \text{for } n' = n. \tag{H}$$

From (E) we obtain by differentiation:

$$\frac{\partial V^2(n,n')}{\partial n \partial n'} = u_{Cn} C'(n') + u_{Zn} Z'(n'), \tag{I}$$

(since $C'(n')$ and $Z'(n')$ do not depend on n). Finally, by using (E), (H) and (I) we obtain:

$$\begin{aligned} \frac{\partial V^2(n,n')}{\partial n \partial n'} &= Z'(n') \left[u_{Cn} \frac{C'(n')}{Z'(n')} + u_{Zn} \right] \\ &= Z'(n') \left[-u_{Cn} \frac{u_Z}{u_C} + u_{Zn} \right] \geq 0, \quad \text{for } n' = n. \end{aligned} \tag{J}$$

The single-crossing condition implies that the term in square brackets on the right-hand side is positive, so that (J) reduces to:

$$Z'(n) \geq 0. \tag{K}$$

Hence, (A5.58) is equivalent to (A5.59)-(A5.60) holding for every n.

****

10.4.2 Characterizing the optimal tax function

We are now in the position to characterize the optimal non-linear tax problem. With a continuum of household types, the social welfare function takes the following form:

$$SW \equiv \int_0^\infty \Psi\big(u(C(n), Z(n), n)\big) f(n)\, dn. \tag{A5.61}$$

The revenue requirement constraint is:

$$\int_0^\infty [nL(n) - C(n)]\, f(n)\, dn = R_0, \tag{A5.62}$$

where R_0 is exogenous. On the left-hand side of (A5.62), $nL(n)$ is wage income and $C(n)$ is consumption, and their difference represents the tax paid by an n-type household. By integrating over all household types the aggregate government budget constraint is obtained. The choice problem of the policy maker is to choose *functions* $C(n)$ and $Y(n)$ such that social welfare (A5.61) is maximized subject to the revenue requirement constraint (A5.62) and the incentive constraints (A5.59)-(A5.60). Unfortunately, the general problem is very difficult and few robust analytical results can be obtained (see, for example, Ebert (1992)). For that reason we first focus on two special cases that were also studied by Salanié (2003, ch. 4). In both cases household utility is assumed to be quasi-linear so that there is no income effect in labour supply. In the first special case, the social welfare function is Rawlsian (maximin) whereas the second special cases uses a general functional form for the social welfare function. We close this section by means of some numerical simulation results for the general case.

10.4.2.1 Special case I: Quasi-linear-Rawlsian SWF

This subsection makes use of two *very special* assumptions regarding private and public preferences. First, the private utility function is assumed to be quasi-linear:

$$U^n = U\left(C\left(n\right), L\left(n\right)\right) \equiv C\left(n\right) - \frac{L\left(n\right)^{1+1/\sigma}}{1+1/\sigma}, \quad \sigma > 0, \tag{A5.63}$$

where the marginal utility of consumption is constant (at unity) and thus independent of labour supply, and where the marginal disutility of labour supply is $L\left(n\right)^{1/\sigma}$ and thus independent of consumption. The parameter σ represents the labour supply elasticity. The second special assumption is that the social welfare function takes a Rawlsian form:

$$SW \equiv \min_{n} U^n. \tag{A5.64}$$

Equation (A5.64) implies that the policy maker wishes to maximize the welfare of the least well-off individual. This is, of course, the individual with the lowest productivity level (see below).

The indirect utility function for the household of type n is defined as follows:

$$V^n\left(n\right) \equiv \max_{\{L(n),C(n)\}} U^n \text{ subject to: } C\left(n\right) = nL\left(n\right) - T\left(nL\left(n\right)\right). \tag{A5.65}$$

By substituting the household budget constraint it can be rewritten in terms of $L\left(n\right)$ as:

$$V^n\left(n\right) \equiv \max_{\{L(n)\}} U\left(nL\left(n\right) - T\left(nL\left(n\right)\right), L\left(n\right)\right). \tag{A5.66}$$

The first-order necessary condition associated with (A5.66) is thus:

$$n\left[1 - T'\left(\cdot\right)\right] U_C = -U_L. \tag{A5.67}$$

By applying the envelope theorem we obtain from (A5.66) and (A5.67):

$$\begin{aligned} \frac{dV^n\left(n\right)}{dn} &= \left[n\left[1 - T'\left(\cdot\right)\right] U_C + U_L\right] \frac{dL\left(n\right)}{dn} + U_C L\left[1 - T\left(\cdot\right)\right] \\ &= L\left(n\right) U_C\left[1 - T\left(\cdot\right)\right] \geq 0, \end{aligned} \tag{A5.68}$$

where the sign follows from the fact that $L \geq 0$ and the assumption that $T'\left(0\right) \leq 1$. It follows from (A5.68) that (indirect) utility is increasing in the productivity level, i.e. the individual with $n = 0$ is the least well-off individual in the economy. This person supplies no labour (see below) and lives off transfers received from the government, $-T(0)$.

In view of (A5.63) and (A5.67) the labour supply function is given by:

$$L(n) = \left(n\left[1 - T'(\cdot)\right]\right)^{\sigma}. \tag{A5.69}$$

There is no income effect in labour supply and σ represents the labour supply elasticity with respect to the after-tax wage.

With the Rawlsian social welfare function (A5.64), the social planner's objective function coincides with the utility function of the least-productive individual (who does not work and lives off transfers):

$$SW \equiv V^{n}(0) \equiv U(-T(0), 0). \tag{A5.70}$$

In order to maximize social welfare, the policy maker must make transfers ($-T(0)$) as large as possible by taxing agents who actually work as much as possible, subject to the incentive constraint (that it does not pay such agents to pose as somebody else). In an Intermezzo we show that the optimal marginal income tax on those who actually work (i.e. all individuals with $n > 0$; see (A5.69)) equals:

$$\frac{T'(nL(n))}{1 - T'(nL(n))} = \left[1 + \frac{1}{\sigma}\right]\Phi(n), \tag{A5.71}$$

$$\Phi(n) \equiv \frac{1 - F(n)}{nf(n)} \geq 0, \tag{A5.72}$$

where $T'(nL(n))$ is the marginal tax rate of a household with income level $nL(n)$, $1 - F(n)$ is the proportion of the population with productivity level higher than n, and σ is the labour supply elasticity ($\sigma > 0$). It follows from (A5.71) that the marginal income tax rate is positive but less than unity. It is lower, the more elastic is labour supply (i.e. the larger is σ).

The optimal marginal income tax also depends critically on the distribution of skills in society as summarized by the $\Phi(n)$ function. To build some intuition behind the $\Phi(n)$ term we consider two examples. *Example 1*: if n is distributed according to the *exponential distribution*, then it follows that the optimal marginal income tax declines with n, i.e. the rich face a lower marginal tax rate than the poor do. Indeed, for the exponential distribution we have $f(n) \equiv \zeta e^{-\zeta n}$ and $F(n) = 1 - e^{-\zeta n}$ ($\zeta > 0$) so that *hazard rate* is constant:

$$\frac{f(n)}{1 - F(n)} = \zeta, \tag{A5.73}$$

and $\Phi(n) = 1/(\zeta n)$. It follows from (A5.71) that the optimal marginal tax rate is then declining in n:

$$\frac{T'(nL(n))}{1 - T'(nL(n))} = \frac{1}{\zeta n}\left[1 + \frac{1}{\sigma}\right]. \tag{A5.74}$$

Example 2: if n follows the *Weibull distribution*,[7] then the optimal marginal tax also declines with n. For the Weibull distribution we have:

$$f(n) \equiv \zeta\theta e^{-(\zeta n)\theta} (\zeta n)^{\theta-1}, \qquad \zeta > 0, \quad \theta > 0, \tag{A5.75}$$

$$F(n) = 1 - e^{-(\zeta n)^{\theta}}, \tag{A5.76}$$

so that the *hazard rate* depends on n ($f(n) / (1 - F(n)) = \zeta\theta(\zeta n)^{\theta-1}$) and the $\Phi(n)$ term is:

$$\Phi(n) \equiv \frac{1}{\theta(\zeta n)^{\theta}}. \tag{A5.77}$$

The optimal marginal tax rate is then declining in n also.

10.4.2.2 Special case II: Quasi-linear-general SWF

In this subsection we continue to assume that utility is quasi-linear utility (as in (A5.63) but we replace (A5.64) by a general social welfare function of the following form:

$$SW \equiv \int_0^{\infty} \Psi(U^n) f(n)\, dn. \tag{A5.78}$$

In the Intermezzo below, it is shown that the optimal marginal income tax on working households equals:

$$\frac{T'(nL(n))}{1 - T'(nL(n))} = \left[1 + \frac{1}{\sigma}\right] \Delta(n)\Phi(n), \tag{A5.79}$$

$$\Delta(n) \equiv \frac{\int_n^{\infty}\left[\mu - \Psi'(U(n))\right] f(n)\, dn}{\mu\left[1 - F(n)\right]}, \tag{A5.80}$$

where $\Phi(n)$ is defined in (A5.72) above. By writing it in the form (A5.79), Diamond (1998, p. 87) is able to provide the economic intuition behind the three constituent components. We consider a particular skill level, n_0. The first term appearing on the right-hand side of (A5.79) is $1 + 1/\sigma$.[8] Increasing $T'(n_0L(n_0))$ increases the deadweight burden of individuals at this skill level n_0. In the absence of income effects in labour supply, the (uncompensated) labour supply elasticity σ is important in this effect.

The intuition behind the $\Delta(n)$ term is as follows. Increasing $T'(n_0L(n_0))$ also transfers income from all individuals with higher skills to the government without changing their labour supply distortions. It can be shown that μ (the Lagrange multiplier of the government budget constraint) is the average of

[7] Note that the Weibull distribution is the same as the exponential distribution for $\theta = 1$.

[8] Diamond also considers the case in which (A5.63) depends on leisure:

$$U^n \equiv C(n) + \frac{(1 - L(n))^{1+1/\eta}}{1 + 1/\eta}, \quad \eta > 0.$$

For this case, the first term on the right-hand side of (A5.79) depends on n also.

the marginal social utilities:

$$\mu = \int_0^\infty \Psi' (U(n)) f(n) \, dn, \tag{A5.81}$$

so that it follows from (A5.80) that $\Delta(0) = 0$. Since $\Psi(\cdot)$ is concave, it follows that $\Delta'(n) > 0$ (see Lemma B in Diamond (1998, p. 87)). Taken in isolation, the second term thus argues in favour of a marginal tax rate that increases with n.

Finally, the intuition behind the $\Phi(n)$ term is as follows. The weight applied to the first two terms on the right-hand side of (A5.79) ($(1 + 1/\sigma)\Delta(n_0)$) is the *ratio* of individuals with skills above n_0 ($1 - F(n_0)$) to individuals with this skill level ($f(n_0)$). The $1/n$ term features in $\Phi(n)$ because taxes are levied on income rather than on hours (recall that n_0 is also the gross wage rate).

10.4.2.3 Closing remarks on the quasi-linear model:

It is clear from our discussion of the quasi-linear model that the shape of the optimal marginal income tax depends on three major factors, namely the labour supply elasticity, the shape of the skills distribution, and the policy maker's taste for redistribution. Diamond (1998) presents some further theoretical results regarding the $\Delta(n)$ and $\Phi(n)$ terms. Saez (2001) presents some interesting simulations for the US economy. He typically finds a U-shaped pattern for optimal marginal tax rates.

10.4.3 General case: simulation results

In this subsection we present some illustrative numerical simulation results taken from Mirrlees (1971, pp. 193-207).[9] Mirrlees (in his Case I) makes use of the following specification. Household utility is a Cobb-Douglas function in consumption and leisure:

$$U^n = \ln C(n) + \ln(1 - L(n)), \tag{A5.82}$$

where $C(n)$ is consumption and $L(n)$ is labour supply. The social welfare function is exponential (or linear, if $\beta = 0$):

$$\Psi(U^n) = \begin{cases} -\frac{1}{\beta} e^{-\beta U^n} & (\text{for } \beta > 0) \\ \\ U^n & (\text{for } \beta = 0) \end{cases}. \tag{A5.83}$$

Finally, the skill distribution is assumed to be lognormal, with density function:

$$f(n) = \frac{1}{n} \exp\left[-\frac{[\ln(n+1)]^2}{2} \right]. \tag{A5.84}$$

[9] Other simulations are found in Stern (1976), Tuomala (1990, pp. 93-99), and Kanbur and Tuomala (1994).

$Z(n)$	$C(n)$	$\frac{T(Z(n))}{Z(n)}$ in %	$T'(Z(n))$ in %
0	0.03	n.a.	23
0.05	0.07	−34	26
0.10	0.10	−5	24
0.20	0.18	9	21
0.30	0.26	13	19
0.40	0.34	14	18
0.50	0.43	15	16

Table 10.1: Mirrlees simulations for the optimal income tax schedule

	Benthamite SWF		*Rawlsian SWF*	
$1-F(n)$ in %	$\frac{T(Z(n))}{Z(n)}$ in %	$T'(Z(n))$ in %	$\frac{T(Z(n))}{Z(n)}$ in %	$T'(Z(n))$ in %
50	6	21	10	52
10	14	20	28	34
1	16	17	28	26

Table 10.2: Bentham versus Rawls and the optimal income tax schedule

Mirrlees presents six different cases but we focus on his first case (see Tables I-II in Mirrlees (1971, p. 202)) for which the social welfare function is Benthamite (or "utilitarian", i.e. $\beta = 0$), and the revenue requirement absorbs 0.013 of total labour supply in efficiency units, i.e. $C = Z - 0.013$, where $C \equiv \int C(n) f(n)\, dn$ is aggregate consumption and $Z \equiv \int Z(n) f(n)\, dn$ is aggregate income. The results obtained by Mirrlees are shown in Table 10.1. Based on these results, Mirrlees draws a number of conclusions. First, he argues that the marginal tax rates are rather low: "I must confess that I had expected the rigorous analysis of income-taxation in the utilitarian framework to provide an argument for high tax rates. It has not done so." (1971, p. 207). Second, he finds that the tax schedule is very close to linear. Finally, he argues that "the income tax is not as effective a weapon for redistributing income, under the assumptions we have made, as one might have expected" (1971, p. 206).

Atkinson and Stiglitz (1980, p. 421) show that the first two conclusions are critically affected by the shape of the social welfare function (embodying the government's taste for redistribution). Their results are summarized in Table 10.2 (which is adapted from Atkinson and Stiglitz (1980, p. 412, Table 13-3)). Two extreme cases are considered, namely a Benthamite social welfare function (columns 2 and 3) and a Rawlsian social welfare function (columns 4 and 5). It is clear from the simulations that the marginal tax rates are much higher for the Rawlsian case. In addition the tax schedule is not close to linear for the Rawlsian case.

Intermezzo 10.5

Derivation of optimal non-linear income tax. We follow the heuristic derivation by Diamond (1998, pp. 93-94). The social welfare function is:

$$SW \equiv \int_0^\infty \Psi\big(U\left(C\left(n\right), L\left(n\right)\right)\big) f\left(n\right) dn, \tag{A}$$

the resource constraint is:

$$\int_0^\infty \left[nL\left(n\right) - C\left(n\right)\right] f\left(n\right) dn = R_0, \tag{B}$$

and household utility is quasi-linear:

$$U\left(C\left(n\right), L\left(n\right)\right) \equiv C\left(n\right) - \frac{L\left(n\right)^{1+1/\sigma}}{1+1/\sigma}, \quad \sigma > 0. \tag{C}$$

The incentive compatibility constraint (first-order condition of household optimum) is:

$$n\big[1 - T'\left(nL\left(n\right)\right)\big] = -U_L\left(L\left(n\right)\right). \tag{D}$$

The optimization problem. The planner maximizes (A) subject to (B) and (D). The problem is first rewritten in a more convenient format. Recall that consumption is $C\left(n\right) \equiv nL\left(n\right) - T\left(nL\left(n\right)\right)$ so that:

$$\begin{aligned} C'\left(n\right) &\equiv L\left(n\right) + nL'\left(n\right) - T'\left(nL\left(n\right)\right)\left[L\left(n\right) + nL'\left(n\right)\right] \\ &= \left[1 - T'\left(nL\left(n\right)\right)\right]\left[L\left(n\right) + nL'\left(n\right)\right] \\ &= -\frac{U_L\left(L\left(n\right)\right)}{n}\left[L\left(n\right) + nL'\left(n\right)\right], \end{aligned} \tag{E}$$

where we have used (D) in the final step. By differentiating (C) with respect to n we find:

$$\begin{aligned} U'\left(n\right) &= C'\left(n\right) + U_L\left(L\left(n\right)\right) L'\left(n\right) \\ &= -\frac{U_L\left(L\left(n\right)\right)}{n}\left[L\left(n\right) + nL'\left(n\right)\right] + \frac{U_L\left(L\left(n\right)\right)}{n} nL'\left(n\right) \\ &= -\frac{L\left(n\right) U_L\left(L\left(n\right)\right)}{n}. \end{aligned} \tag{F}$$

The rewritten optimum control problem now features the following objective function:

$$SW \equiv \int_0^\infty \Psi\left(U\left(n\right)\right) f\left(n\right) dn, \tag{G}$$

and the constraints are:

$$R_0 = \int_0^\infty \left[nL\left(n\right) - \left(U\left(n\right) + \frac{L\left(n\right)^{1+1/\sigma}}{1+1/\sigma}\right)\right] f\left(n\right) dn, \tag{H}$$

$$U'(n) = -\frac{L(n)\,U_L(L(n))}{n}. \tag{I}$$

The control variable is $L(n)$, the state variable is $U(n)$, and the co-state variable is $\lambda(n)$. (Note that (H) is obtained by combining (B) and (C) whilst (I) is just (F).). The Hamiltonian expression for the optimization problem is:

$$\begin{aligned}\mathcal{H} \equiv & \left[\Psi(U(n)) + \mu\left[nL(n) - U(n) - \frac{L(n)^{1+1/\sigma}}{1+1/\sigma}\right]\right] f(n) \\ & - \lambda(n)\frac{L(n)\,U_L(L(n))}{n},\end{aligned} \tag{J}$$

where μ is the Lagrange multiplier for the government budget constraint (H). Note that μ is constant, a result which has been already been incorporated in (J) (see my own notes on Diamond (1998)).

The first-order necessary conditions are the following. For the control variable we have $\partial\mathcal{H}/\partial L(n) = 0$ or:

$$\mu\left[n + U_L(L(n))\right] f(n) = \lambda(n)\frac{U_L(L(n)) + L(n)\,U_{LL}(L(n))}{n}. \tag{K}$$

For the state variable we have $-\partial\mathcal{H}/\partial U(n) = \lambda'(n)$ or:

$$\lambda'(n) = -\left[\Psi'(U(n)) - \mu\right] f(n). \tag{L}$$

The transversality condition is:

$$\lambda(0) = \lim_{n\to\infty}\lambda(n) = 0. \tag{M}$$

Equations (K)-(M) can be used to derive the solution used in the main text.

First we integrate (L) in the interval $[n,\infty)$:

$$\begin{aligned}\frac{d\lambda(s)}{ds} &= -\left[\Psi'(U(s)) - \mu\right] f(s) \\ d\lambda(s) &= -\left[\Psi'(U(s)) - \mu\right] f(s)\,ds \\ \int_n^\infty d\lambda(s) &= -\int_n^\infty \left[\Psi'(U(s)) - \mu\right] f(s)\,ds \\ \lim_{s\to\infty}\lambda(s) - \lambda(n) &= -\int_n^\infty \left[\Psi'(U(s)) - \mu\right] f(s)\,ds \\ \lambda(n) &= \int_n^\infty \left[\Psi'(U(s)) - \mu\right] f(s)\,ds,\end{aligned} \tag{N}$$

where we have used (M) in the final step. From (K) we derive:

$$\mu U_L(\cdot)\left[\frac{n}{U_L(\cdot)}+1\right]f(n) = \frac{\lambda(n)}{n}U_L(\cdot)\left[1+\frac{L(\cdot)U_{LL}(\cdot)}{U_L(\cdot)}\right]$$
$$\left[\frac{n}{U_L(\cdot)}+1\right] = \frac{\lambda(n)}{n\mu f(n)}\left[1+\frac{L(\cdot)U_{LL}(\cdot)}{U_L(\cdot)}\right]. \tag{O}$$

But from (C) and (D) we obtain:

$$\frac{L(\cdot)U_{LL}(\cdot)}{U_L(\cdot)} = \frac{1}{\sigma}, \tag{P}$$
$$\frac{n}{U_L(\cdot)} = -\frac{1}{1-T'(\cdot)}. \tag{Q}$$

By using these results in (O) we obtain the required optimal marginal tax formula:

$$\frac{T'(\cdot)}{1-T'(\cdot)} = -\frac{\lambda(n)}{n\mu f(n)}\left[1+\frac{1}{\sigma}\right]. \tag{R}$$

Note that (N) and (R) are the expressions found in the text.

For *special case I* we have:

$$\Psi'(U(0)) = 1$$
$$\Psi'(U(n)) = 0, \qquad \text{(for } n > 0\text{)}.$$

Using these properties in (N) we find:

$$\begin{aligned}\lambda(n) &= -\int_n^\infty \mu f(s)\,ds \\ &= -\mu\int_n^\infty dF(s) \\ &= -\mu\left[1-F(n)\right],\end{aligned} \tag{S}$$

and by using this result in (R) we find:

$$\frac{T'(\cdot)}{1-T'(\cdot)} = \frac{1-F(n)}{nf(n)}\left[1+\frac{1}{\sigma}\right]. \tag{T}$$

Equation (T) coincides with (A5.71) in the text. Note also that (M) and (N) together imply that:

$$0 = \lambda(0) = \int_0^\infty \left[\Psi'(U(n))-\mu\right]f(n)\,dn$$
$$\mu = \int_0^\infty \Psi'(U(n))\,f(n)\,dn,$$

i.e. the Lagrange multiplier on the government budget constraint (μ) is equal to the population average of $\Psi'(U(n))$ in this quasi-linear case (see also Diamond (1998, p. 87)). Intuitively, a uniform transfer to all workers leaves labour supply unchanged (no income effect) and just boosts consumption and thus utility. Hence, μ is the average of the marginal social utilities.

The sign of $\Delta'(n)$ is obtained as follows. By differentiating (A5.80) with respect to n we obtain:

$$\begin{aligned}\Delta'(n) &\equiv \frac{-[1-F(n)]\,[\mu-\Psi'(U(n))]\,f(n)+\mu\,[1-F(n)]\,\Delta(n)\,f(n)}{\mu\,[1-F(n)]^2} \\ &= \frac{f(n)}{\mu\,[1-F(n)]}\left[\mu\Delta(n)-\left[\mu-\Psi'(U(n))\right]\right]>0,\end{aligned}$$

where the sign follows from the fact that $\Psi'(U(n))$ is decreasing in n. The proof for this final claim is as follows. If $\Psi'(U(n))$ were constant, then $\mu\Delta(n)$ would equal:

$$\mu\Delta(n)=\frac{[\mu-\Psi'(U(n))]\int_n^\infty f(s)\,ds}{1-F(n)}=\mu-\Psi'(U(n)).$$

But since $\Psi'(U(n))$ is decreasing in n it follows that $\mu\Delta(n)>\mu-\Psi'(U(n))$.

****

10.5 Key literature

- Atkinson & Stiglitz (1980, lectures 13-14), Jha (1998, chs. 14), or Myles (1995, ch. 5) on theory.
- Stiglitz (1987) and Stern (1987a) on theory.
- Auerbach (1985) on theory.
- Auerbach and Hines (2002) on recent theory and empirics.
- Haveman (1994) for critical remarks.
- Classics: Mirrlees (1971, 1976), Atkinson (1973), Seade (1977), Sadka (1976), Diamond (1998), Sheshinski (1972), Phelps (1973), Fair (1971), Stiglitz (1982),
- Linear income tax: Stokey (1980), Slemrod et al. (1994), Mirrlees (1990), Sandmo (1983), Hellwig (1986).
- Recent papers: Konrad (2001), Saez (2002a,2002b), Jones et al. (1993), Corsetti and Roubini (1996), Plug et al. (1999), Piketty (1997), Brito et al. (1991), Kanbur and Tuomola (1994), Saez (2001), Cremer et al. (2001).
- Technical: Lollivier and Rochet (1983), Guesnerie and Laffont (1984).
- Survey: Stern (1987c), Atkinson (1990), Ebert (1992), Tuomola (1990), and Fagan (1938)(old literature).
- Varian (1980). Optimal taxation: Luck and income differences. Also: optimal taxation under uncertainty: Diamond et al. (1980), Eaton and Rosen (1980a,b).

Chapter 11

Public goods and externalities

The purpose of this chapter is to discuss the following topics:

- What do we mean by (pure) public goods?
- What should the supply of public goods be in a first-best world? The Samuelson Rule.
- Second-best supply of public goods: the Modified Samuelson Rule.
- Publicly provided private goods.
- Private provision of public goods: subscription goods.
- Production and consumption externalities and corrective taxation.

11.1 Public goods

In this section we study some basic issues surrounding the public provision of public goods. There are two important aspects relating to a good or a service, namely the notions of excludability and rivalry. With *excludability* the key question is whether it is possible to charge for the use of a particular good or service, i.e. whether or not agents can be excluded from consuming the good or service (Atkinson and Stiglitz, 1980, p. 483). Excludability is essentially a technical property. With public television broadcasts it is technically virtually impossible to exclude reception or measure actual viewing (as signal scramblers can easily be circumvented). In the case of public freeways it is technically quite possible (though politically often difficult to get through parliament) to install toll booths and electronic measuring devices to measure actual usage (and charge on that basis). With national defense it is impossible to measure individual benefits so exclusion is impossible. If a good or service is non-excludable then consumption cannot be controlled efficiently by a price system.

The second aspect of a good or service has to do its *rivalry* in consumption. Does one agent's consumption reduce the amount available to the other agents? For ordinary goods we have that if one agent consumes them, the other agents can no longer do so (e.g. bread, beer, peanuts, etcetera). In contrast, for public television broadcast the fact that one agent watches does not reduce in any way the possibility for all other agents to watch the same broadcast.

Based on the two concepts, *pure public goods* are those goods are services which are non-excludable and non-rival in consumption. Pure private goods are those goods and services which are excludable and rival in consumption. In more formal terms, a pure private good satisfies:

$$\sum_{h=1}^{H} X_g^h = Y_g, \tag{A5.1}$$

where Y_g is the supply of good g (where $g = 1, \ldots, G$), and X_g^h is the consumption demand by household h of good g. According to (A5.1), one household's consumption of the good reduces the amount available to other households by exactly that amount. For a pure public good we have:

$$X_g^h = Y_g, \quad \text{for all } h = 1, \ldots, H. \tag{A5.2}$$

In the formulation (A5.2), all households consume same the quantity Y_g. Implicitly it thus assumes the absence of *free disposal*, i.e. households cannot avoid to consume the public goods as in the case of national defense. If free disposal is possible (e.g. public TV broadcasts) we have the alternative formulation:

$$X_g^h \leq Y_g, \quad \text{for all } h = 1, \ldots, H. \tag{A5.3}$$

Pure public and pure private goods can be illustrated with the aid of Figure 11.1 for the case with two households, h_1 and h_2. If X_g is a pure public good, then the consumption possibility frontier is given by ABC, and if X_g is a pure private good, the frontier is given by AC. If X_g is somewhere in between a pure public good and a pure private good, the consumption possibility frontier is the concave dashed line AEC. This case describes an impure public good, for example highways subject to congestion.

In general terms the consumption possibility frontier can be defined formally as follows:

$$\Gamma\left(X_g^1, X_g^2, \ldots, X_g^H, Y_g\right) = 0. \tag{A5.4}$$

With this formulation, pure public goods are characterized as follows:

$$\frac{\partial \Gamma}{\partial X_g^h} = 0, \quad \text{for all } h = 1, \ldots, H, \tag{A5.5}$$

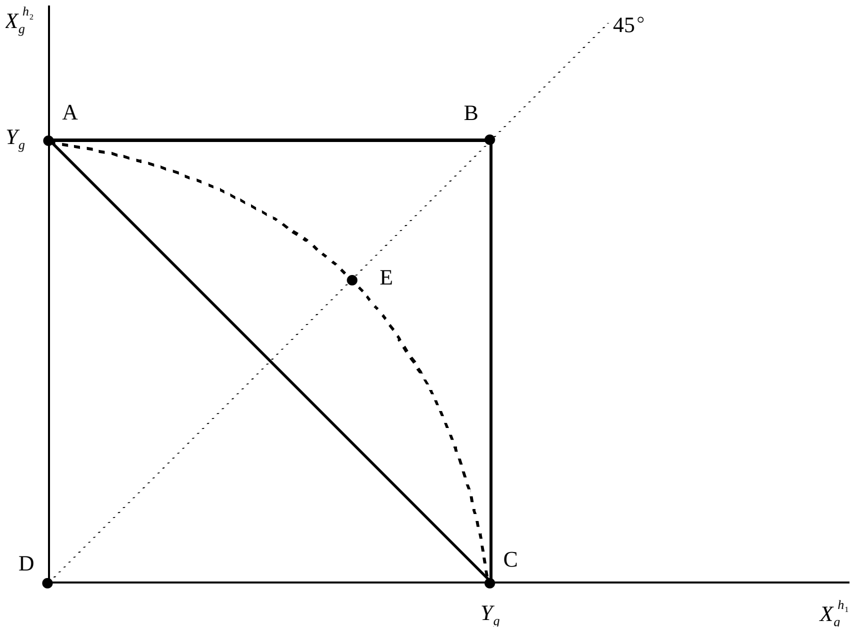

Figure 11.1: Private and public goods

whereas for pure private goods we have:

$$\frac{\partial\Gamma/\partial X_g^{h_1}}{\partial\Gamma/\partial X_g^{h_2}} = 1, \quad \text{for all } h_1, h_2 = 1, \ldots, H. \tag{A5.6}$$

For impure public goods the condition is:

$$0 < \frac{\partial\Gamma/\partial X_g^{h_1}}{\partial\Gamma/\partial X_g^{h_2}} < 1, \quad \text{for all } h_1, h_2 = 1, \ldots, H. \tag{A5.7}$$

An example of a convenient functional form is the Constant Elasticity of Transformation (CET) function:

$$\begin{aligned} Y_g &= M\left(X_g^1, X_g^2, \ldots, X_g^H\right) \\ &\equiv \left[\varepsilon_1^{-1/\sigma}\left(X_g^1\right)^{(1+\sigma)/\sigma} + \varepsilon_2^{-1/\sigma}\left(X_g^2\right)^{(1+\sigma)/\sigma} + \ldots + \varepsilon_H^{-1/\sigma}\left(X_g^H\right)^{(1+\sigma)/\sigma}\right]^{\sigma/(1+\sigma)} \\ &= \left[\sum_{h=1}^{H} \varepsilon_h^{-1/\sigma}\left(X_g^h\right)^{(1+\sigma)/\sigma}\right]^{\sigma/(1+\sigma)}, \qquad \sigma \geq 0 \end{aligned} \tag{A5.8}$$

Depending on the magnitude of σ, equation (A5.8) nests the cases of a pure private good ($\sigma \to \infty$), a pure public good ($\sigma \to 0$ and $\varepsilon_h = 1$ for all h), and an impure pure public good ($0 < \sigma \ll \infty$).

11.1.1 Samuelson rule

In this subsection we study the optimal provision of a pure public consumption good in the context of the general equilibrium welfare theoretic model introduced in Chapter 9. The utility function of household h is:

$$U^h = U^h \left(X_1^h, \ldots, X_G^h, X_{G+1}, V_1^h, \ldots, V_F^h \right), \tag{A5.9}$$

where U^h is utility of household h (where $h = 1, 2, \ldots, H$), X_g^h is consumption of the private good g by household h (where $g = 1, 2, \ldots, G$), and V_f^h is the supply of production factor f by household h (where $f = 1, 2, \ldots, F$). Note that X_{G+1} is the (single) public consumption good supplied by the policy maker. It features in every household's utility function and is thus a pure public good.

The production technology for good g (for $g = 1, \ldots, G+1$) is written as:

$$Y_g = F^g \left(Z_1^g, \ldots, Z_F^g \right), \tag{A5.10}$$

where Y_g is aggregate production, $F^g(\cdot)$ is the production function, and Z_f^g is factor f used in the production of good g. The market clearing conditions for the private goods markets ($g = 1, 2, \ldots, G$) are thus:

$$Y_g = \sum_{h=1}^{H} X_g^h. \tag{A5.11}$$

For the public good (and in the absence of free disposal) we have:

$$Y_{G+1} = X_{G+1}, \tag{A5.12}$$

and for the factor markets ($f = 1, 2, \ldots, F$) the clearing conditions are:

$$\sum_{h=1}^{H} V_f^h = \sum_{g=1}^{G+1} Z_f^g. \tag{A5.13}$$

The general equilibrium model is fully described by equations (A5.9)-(A5.13).

In order to derive the necessary conditions for Pareto efficiency we once again focus on an arbitrary household (say household $h = 1$), hold every other household's utility constant (i.e. $U^h = U_0^h$ for $h = 2, \ldots, H$), and maximize household 1's utility subject to the restrictions, i.e. the social planner chooses X_g^h, V_f^h, Z_f^g, Y_g in order to maximize:

$$U^1 = U^1 \left(X_1^1, \ldots, X_G^1, X_{G+1}, V_1^1, \ldots, V_F^1 \right), \tag{A5.14}$$

subject to:

$$U_0^h = U^h\left(X_1^h, \ldots, X_G^h, X_{G+1}, V_1^h, \ldots, V_F^h\right), \qquad \text{(for } h = 2, \ldots, H\text{)}, \tag{A5.15}$$

and (A5.10)-(A5.13). The Lagrangian expression for this social optimization problem is:

$$\begin{aligned}
\mathcal{L} \equiv{} & U^1\left(X_1^1, \ldots, X_G^1, X_{G+1}, V_1^1, \ldots, V_F^1\right) \\
& + \sum_{h=2}^{H} \lambda_h \left[U^h\left(X_1^h, \ldots, X_G^h, X_{G+1}, V_1^h, \ldots, V_F^h\right) - U_0^h\right] \\
& + \sum_{g=1}^{G+1} \mu_g \left[Y_g - F^g\left(Z_1^g, \ldots, Z_F^g\right)\right] + \sum_{f=1}^{F} \xi_f \left[\sum_{h=1}^{H} V_f^h - \sum_{g=1}^{G+1} Z_f^g\right] \\
& + \sum_{g=1}^{G} \nu_g \left[\sum_{h=1}^{H} X_g^h - Y_g\right] + \nu_{G+1}\left[X_{G+1} - Y_{G+1}\right],
\end{aligned}$$

where the Lagrange multipliers are λ_h (for $h = 2, \ldots, H$), μ_g and ν_g (for $g = 1, \ldots, G+1$), and ξ_f (for $f = 1, \ldots, F$), i.e. there are $(H-1) + 2(G+1) + F$ Lagrange multipliers in all. To simplify the notation, we set $\lambda_1 = 1$. The first-order necessary conditions (assuming an interior solution) are the constraints and (i) for the private good demands ($G \times H$ equations):

$$\frac{\partial \mathcal{L}}{\partial X_g^h} = \lambda_h \frac{\partial U^h}{\partial X_g^h} + \nu_g = 0, \tag{A5.16}$$

(ii) for the public good (1 equation):

$$\frac{\partial \mathcal{L}}{\partial X_{G+1}} = \sum_{h=1}^{H} \lambda_h \frac{\partial U^h}{\partial X_{G+1}} + \nu_{G+1} = 0, \tag{A5.17}$$

(iii) for the factor supplies ($H \times F$ equations):

$$\frac{\partial \mathcal{L}}{\partial V_f^h} = \lambda_h \frac{\partial U^h}{\partial V_f^h} + \xi_f = 0, \tag{A5.18}$$

(iv) for the output decisions ($G+1$ equations):

$$\frac{\partial \mathcal{L}}{\partial Y_g} = \mu_g - \nu_g = 0, \tag{A5.19}$$

and (v) for the factor demands ($F \times (G+1)$ equations):

$$\frac{\partial \mathcal{L}}{\partial Z_f^g} = -\mu_g \frac{\partial F^g}{\partial Z_f^g} - \xi_f = 0. \tag{A5.20}$$

Just as in Chapter 9, we can eliminate the various Lagrange multipliers and derive the condensed statement of the necessary conditions for the Pareto optimum. We look at the following pairings. For two private goods, g_1 and g_2 we find from (A5.16):

$$\frac{\nu_{g_1}}{\nu_{g_2}} = \frac{\partial U^h / \partial X^h_{g_1}}{\partial U^h / \partial X^h_{g_2}}, \tag{A5.21}$$

i.e. the marginal rate of substitution (MRS) between any two consumption goods g_1 and g_2 is the same for all households h. From (A5.19)-(A5.20) we find for those goods from the production side:

$$\frac{\nu_{g_1}}{\nu_{g_2}} = \frac{\partial F^{g_2} / \partial Z^{g_2}_f}{\partial F^{g_1} / \partial Z^{g_1}_f}, \tag{A5.22}$$

where the right-hand side is the marginal rate of transformation (MRT) between goods g_1 and g_2. Combining (A5.21) and (A5.22) we obtain the efficient provision condition for *private* goods:

$$\frac{\partial U^h / \partial X^h_{g_1}}{\partial U^h / \partial X^h_{g_2}} = \frac{\partial F^{g_2} / \partial Z^{g_2}_f}{\partial F^{g_1} / \partial Z^{g_1}_f} \tag{A5.23}$$

Just as in Chapter 9, the MRS between goods g_1 and g_2 (left-hand side) must be equated to the MRT (right-hand side).

For the pure public good $g = G + 1$ and any private good g_2 we find from (A5.16)-(A5.17):[1]

$$\frac{\nu_{G+1}}{\nu_{g_2}} = \sum_{h=1}^{H} \frac{\partial U^h / \partial X^h_{G+1}}{\partial U^h / \partial X^h_{g_2}}. \tag{A5.24}$$

From the production side we find from (A5.19)-(A5.20):

$$\frac{\nu_{G+1}}{\nu_{g_2}} = \frac{\partial F^{g_2} / \partial Z^{g_2}_f}{\partial F^{G+1} / \partial Z^{G+1}_f}, \tag{A5.25}$$

where the right-hand side is the MRT between the public good $g = G + 1$ and the private good g_2. Combining (A5.24) and (A5.25) we find the famous *Samuelson Rule* for the efficient provision of pure public goods:

$$\sum_{h=1}^{H} \frac{\partial U^h / \partial X^h_{G+1}}{\partial U^h / \partial X^h_{g_2}} = \frac{\partial F^{g_2} / \partial Z^{g_2}_f}{\partial F^{G+1} / \partial Z^{G+1}_f}. \tag{A5.26}$$

This rule is perhaps best understood by comparing it with the efficient provision condition for private goods (A5.23). For private goods the MRT must be equated to the MRS. An extra unit of the private good only affects the welfare of a single recipient. In the optimum, the marginal benefits should be the same for all households and equal to marginal cost of production–see equation (A5.23). In contrast, for

[1] This expression is obtained by solving (A5.16) for $\lambda_h = -\nu_{g_2} / \left(\partial U^h / \partial X^h_{g_2} \right)$ and substituting the result into (A5.17).

the public good the MRT must be equated to the *summation* of the MRS's over all households. In the optimum, the sum of the marginal benefits over all households must be equated to marginal cost of production–see equation (A5.26).

Figure 11.2 presents a two-household-two-good ($F = G = 2$) illustration of the Samuelson Rule.[2] In that figure, all private goods are aggregated into one composite good G and the public good is good $G+1$. In panel (a), $T(Y_G, Y_{G+1})$ is the production possibility frontier and U_0^2 is household 2's fixed indifference curve. In panel (b), the concave curve BC^1C is the vertical difference between $T(Y_G, Y_{G+1})$ and U_0^2, and constitutes the consumption possibility locus for household 1. In the bottom panel, household 1 chooses the optimum at point C^1 where there is a tangency between its consumption possibility locus and an indifference curve. In the top panel, household 2's consumption bundle is located at point C^2, and the optimal production point is at P. The vertical summation of the slopes at points C^1 and C^2 equals the slope at point P, i.e.

$$MRS_{G,G+1}^1 + MRS_{G,G+1}^2 = MRT_{G,G+1}, \tag{A5.27}$$

where $MRS_{G,G+1}^h$ and $MRT_{G,G+1}$ are, respectively, the MRS between goods G and $G+1$ for household $h\ (=1,2)$ and the MRT between these two goods.

We close this subsection on the Samuelson rule with a number of pertinent remarks. First, the property of non-excludability has not been used in deriving the Samuelson Rule. Excludability is relevant for designing feasible (private or public) provision mechanisms (Oakland, 1987, p. 491). Second, the Samuelson rule is a first-best rule and is thus not easily implemented in a decentralized setting (it requires lump-sum taxes for redistribution and financing purposes). Third, it is relatively straightforward to model congestion (e.g. excessive road use) or public inputs (e.g. infrastructure) and to re-derive the appropriate Samuelson rule. These extensions are left as an exercise to the reader.

11.1.2 Modified Samuelson rule

What is the optimal public good provision rule if distorting taxes have to be used? This is a question which clearly falls within the realm of second-best welfare economics. In this subsection we discuss the seminal paper by Atkinson and Stern (1974) which provides a thorough analysis of the issues involved (see also Atkinson and Stiglitz (1980, pp. 490-492)). The main assumptions made in this subsection are the following. First, we assume identical households and thus abstract from equity considerations. The social welfare function is the utility of the representative household. Second, there is a single mobile production factor (labour), a single *private* consumption good (C), and one pure *public* consumption good (G). The labour supply decision by households (L) is endogenous. Since the problem is completely symmetric, there is no need to distinguish individual households and we can simplify notation by rec-

[2]This figure was suggested by Samuelson (1955) and can also be found in Atkinson and Stiglitz (1980, p. 489) and in Oakland (1987, p. 489).

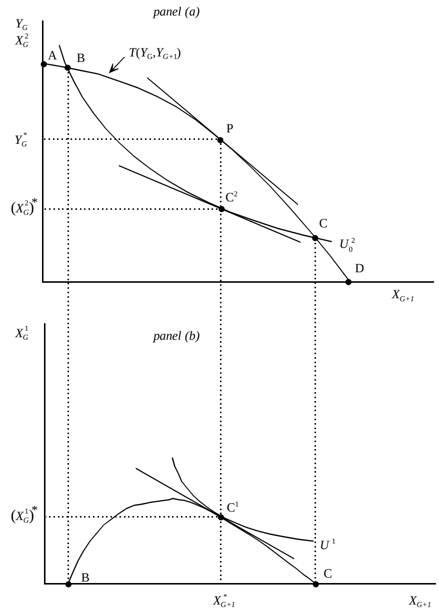

Figure 11.2: The Samuelson rule for public good provision

ognizing that $C^h = C$ and $L^h = L$ (for all $h = 1, \ldots, H$). Third, the production technologies for the two goods are both linear. Fourth, the policy maker possesses two fiscal instruments, namely the (distorting) labour income tax (t_L) and (the non-distorting) lump-sum tax (T, which is used for interpretation purposes only).

The representative household h ($h = 1, \ldots, H$) has the following (direct) utility function:

$$U = U\left(1 - L, C, G\right), \tag{A5.28}$$

where C is consumption of the private good, L is labour supply (so that $1 - L$ is leisure as the time endowment is unity), and G is the pure public consumption good. The utility function features the same functional form for all households and has the usual properties (as far as the private arguments are concerned), i.e. $U_C > 0$, $U_{1-L} > 0$, $U_{CC} < 0$, $U_{1-L,1-L} < 0$ and $U_{CC}U_{1-L,1-L} - U_{C,1-L}^2 > 0$. In addition, it is assumed that $U_G > 0$ and $U_{GG} < 0$. The budget constraint is:

$$PC = (1 - t_L)\, WL - T, \tag{A5.29}$$

where P is the consumer price of the private good (which is untaxed), W is the wage rate, t_L is the labour income tax, and T is the lump-sum tax (the same for all households).

The household chooses L and C in order to maximize (A5.28) subject to (A5.29), taking as given the level of public good provision (G), the prices (W and P), and the taxes (t_L and T). The indirect utility function is defined in the usual way as (see also Chapter 2):

$$V \equiv V\left(\left(1-t_L\right)W, P, T, G\right)$$
$$\equiv \max_{\{L,C\}} U\left(1-L, C, G\right) \text{ subject to } \left(1-t_L\right)WL - T - PC = 0, \tag{A5.30}$$

and the Marshallian consumption demand and labour supply are obtained by using Roy's Identity:

$$C \equiv -\frac{\partial V/\partial P}{\alpha}, \tag{A5.31}$$
$$L = \frac{\partial V/\partial\left[\left(1-t_L\right)W\right]}{\alpha}, \tag{A5.32}$$

where α is the marginal utility of lump-sum income (i.e. $-\alpha$ is the marginal disutility associated with the lump-sum tax).

The production function for the private good features constant returns to scale:

$$Y = \frac{L_Y}{k_Y}, \tag{A5.33}$$

where Y is aggregate output, L_Y is labour used, and k_Y is an exogenous productivity index (marginal and average production cost is thus equal to Wk_Y). Under perfect competition the price is equal to $P = Wk_Y$ and there are zero pure profits. The production function for the public good is also linear in labour:

$$G = \frac{L_G}{k_G}, \tag{A5.34}$$

where G and L_G are, respectively, output and labour used in the public sector, and where k_G is an exogenous productivity index. Since labour is mobile across the two production sectors, private and public employers face the same wage rate so that marginal production cost in the public sector is Wk_G. Under efficient production the cost of one unit of G is Wk_G.

The market clearing conditions are as follows. For the consumption goods market aggregate consumption demand equals total supply:

$$HC = Y, \tag{A5.35}$$

whereas for the labour market aggregate labour supply equals the total demand for labour from the private and public sectors:

$$HL = L_Y + L_G. \tag{A5.36}$$

Using (A5.33)-(A5.36), the aggregate production possibility constraint can be derived:

$$HL = k_Y HC + k_G G. \tag{A5.37}$$

Holding constant aggregate labour supply (left-hand side), equation (A5.36) defines a linear relationship between aggregate consumption demand (HC) and the level of public good provision (G). An increase in aggregate labour supply shifts this relationship in a parallel fashion to the right.

The government budget constraint is:

$$t_L WHL + HT = Wk_G G, \tag{A5.38}$$

where the left-hand side is total tax revenue and the right-hand side is the total (wage) cost of the public good provision. The objective function of the policy maker takes the utilitarian (Benthamite) form:

$$SW \equiv HV, \tag{A5.39}$$

where V is the indirect utility function of the representative household (see (A5.30) above). The policy maker chooses G, T (in the first-best) or t_L (in the second-best), in order to maximize social welfare (A5.39) subject to the government budget constraint (A5.38). Labour is used as the numeraire, so W is taken as given.

The Lagrangian expression for the social optimization problem is:

$$\mathcal{L} \equiv HV\left((1-t_L)\, W, P, T, G\right) + \lambda\left[t_L WHL + HT - Wk_G G\right].$$

The first-order necessary condition for public good provision is:

$$\frac{\partial \mathcal{L}}{\partial G} = H\frac{\partial V}{\partial G} + \lambda W\left[t_L H\frac{\partial L}{\partial G} - k_G\right] = 0. \tag{A5.40}$$

If the lump-sum tax could be varied freely (*first-best case*), then the associated first-order condition would be:

$$\frac{\partial \mathcal{L}}{\partial T} = H\left[\frac{\partial V}{\partial T} + \lambda\left[t_L W\frac{\partial L}{\partial T} + 1\right]\right] = 0. \tag{A5.41}$$

In contrast, if the lump-sum tax is fixed and financing is by means of t_L (*second-best case*) we obtain instead:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial t_L} &= H\frac{\partial V}{\partial\left[(1-t_L)\, W\right]}\frac{\partial\left[(1-t_L)\, W\right]}{\partial t_L} + \lambda WH\left[t_L\frac{\partial L}{\partial t_L} + L\right] \\ &= HW\left[-\frac{\partial V}{\partial\left[(1-t_L)\, W\right]} + \lambda\left[t_L\frac{\partial L}{\partial t_L} + L\right]\right] = 0. \end{aligned} \tag{A5.42}$$

In order to understand the implications of these first-order conditions we first look at the (relatively straightforward) first-best case. Once that case is fully understood, we proceed to study the complications popping up in the second-best case.

11.1.2.1 First-best case

In the first-best social optimum, the policy maker can use the lump-sum tax and thus employ the optimal tax structure. By using equation (A5.32), the first-order condition for t_L (given in equation (A5.42) above) can be simplified to:

$$\frac{\alpha-\lambda}{\lambda}=\frac{t_L}{L}\frac{\partial L}{\partial t_L}. \tag{A5.43}$$

Recall (from Chapter 11) that the Slutsky equation for labour supply can be written as:

$$\frac{1}{L}\frac{\partial L}{\partial(1-t_L)}=\frac{S_{LL}}{L}-W\frac{\partial L}{\partial T}, \tag{A5.44}$$

where $S_{LL}\equiv(\partial L/\partial(1-t_L))_{U_0}>0$ is the pure substitution effect. By using (A5.44) and noting that $\partial V/\partial T=-\alpha$, the first-order condition for T (given in (A5.41) above) can be simplified to:

$$\begin{aligned}\frac{\alpha-\lambda}{\lambda}&=t_L W\frac{\partial L}{\partial T}\\&=\frac{t_L}{L}\frac{\partial L}{\partial t_L}+\frac{t_L S_{LL}}{L},\end{aligned} \tag{A5.45}$$

where we have also used the fact that $\partial L/\partial t_L=-\partial L/\partial(1-t_L)$ to facilitate the comparison between (A5.43) and (A5.45). Since S_{LL} is strictly positive, it follows from (A5.43) and (A5.45) that under the optimal (first-best) tax system the policy maker sets $t_L=0$ so that $\lambda=\alpha$, i.e. it costs one euro to raise one euro of public funds (there is no excess burden associated with the lump-sum tax). By using these results in (A5.40) and noting that $\partial V/\partial G=\partial U/\partial G$ we find:

$$H\frac{\partial U}{\partial G}=\lambda W k_G, \tag{A5.46}$$

The first-order condition for the household optimum is $\partial U/\partial C=\alpha P$ so that (A5.46) can be rewritten in the familiar format of the Samuelson rule:

$$\frac{H\,\partial U/\partial G}{\partial U/\partial C}=\frac{Wk_G}{P}=\frac{k_G}{k_Y}. \tag{A5.47}$$

The sum of the marginal rates of substitution between G and C (left-hand side) is equated to the marginal rate of transformation (k_G/k_Y on the right-hand side). Equation (A5.47) is thus the exact counterpart of (A5.26) in the context of the context of the two-good model.

11.1.2.2 Second-best case

In the second-best social optimum the policy maker does not have access to the lump-sum tax and only (A5.40) and (A5.42) are relevant. After some manipulations these expressions can be rewritten as:

$$\frac{H\,\partial U/\partial G}{\partial U/\partial C} = \frac{\lambda}{\alpha}\frac{1}{k_Y}\left[k_G - t_L H\frac{\partial L}{\partial G}\right], \tag{A5.48}$$

$$\frac{\alpha}{\lambda} = 1 + \frac{t_L}{L}\frac{\partial L}{\partial t_L}. \tag{A5.49}$$

Several things are worth noting about these expressions. First, it follows from (A5.48) that, if G is complementary with leisure (so that $\partial(1-L)/\partial G > 0$ and $\partial L/\partial G < 0$) then (for $t_L > 0$) the right-hand side contains an additional revenue term $-t_L H\partial L/\partial G > 0$. Second, it follows from (A5.49) that for an upward sloping (uncompensated) labour supply curve $((t_L/L)(\partial L/\partial t_L) < 0)$ the social cost of raising one dollar of public funds (λ) exceeds the private marginal utility of income (α), i.e. $\lambda > \alpha$ and there is an excess burden associated with the labour income tax.

By substituting (A5.49) into (A5.48) we obtain the formula for the modified Samuelson rule:

$$\frac{H\,\partial U/\partial G}{\partial U/\partial C} = \frac{\frac{k_G}{k_Y} - t_L H\frac{1}{k_Y}\frac{\partial L}{\partial G}}{1 + \frac{t_L}{L}\frac{\partial L}{\partial t_L}}. \tag{A5.50}$$

This modified Samuelson rule can be interpreted as follows. The left-hand side represents the sum of the marginal rates of substitution between the public good and the private consumption good. The right-hand side is the social cost of the public good in terms of the commodity. This social cost differs from marginal rate of transformation ($\equiv k_G/k_Y$ here) for two reasons: (i) if $\partial L/\partial G < 0$ then an increase in G increases the excess burden because of an erosion of the labour income tax base, and (ii) if $(t_L/L)(\partial L/\partial t_L) < 0$ then $\lambda/\alpha > 1$ which also increases the social cost of the public good.

Intermezzo 11.1

Alternative expression for (A5.49). One often finds slightly different (though equivalent) expressions for (A5.49) in the literature which make use of the concept of social marginal utility of income (β):

$$\frac{\beta-\lambda}{\lambda} = -t_L\frac{S_{LL}}{L} \tag{A}$$

$$= -\sigma_{LL}\frac{t_L}{1-t_L}, \tag{B}$$

where $S_{LL} \equiv (\partial L/\partial(1-t_L))_{U_0} > 0$ is (proportional to) the pure substitution effect in labour supply and β and σ_{LL} are defined as follows:

$$\beta \equiv \alpha - \lambda t_L W \frac{\partial L}{\partial T}, \tag{C}$$

$$\sigma_{LL} \equiv \left(\frac{\partial L}{\partial \left[(1 - t_L) W \right]} \right)_{U_0} \frac{(1 - t_L) W}{L} = \frac{(1 - t_L) S_{LL}}{L} > 0. \tag{D}$$

In equation (C), β represents the social marginal utility of income (see also Chapter 10) and $-\partial L / \partial T$ is the income effect in labour supply (which is negative if leisure is normal in consumption). In equation (D), σ_{LL} is the slope of the Hicksian labour supply curve (this slope is of course positive).

The expressions (A) and (B) are obtained from (A5.49) by using the Slutsky equation (A5.44):

$$\begin{aligned} \frac{\alpha - \lambda}{\lambda} &= -\frac{t_L}{L} \frac{\partial L}{\partial (1 - t_L)} \\ &= -t_L \left[\frac{S_{LL}}{L} - W \frac{\partial L}{\partial T} \right] \qquad \Leftrightarrow \\ \frac{\beta - \lambda}{\lambda} &= -\frac{t_L S_{LL}}{L} = -\sigma_{LL} \frac{t_L}{1 - t_L}. \end{aligned}$$

It is easy to demonstrate that for positive required government revenue, λ exceeds β. We multiply equation (A) by $t_L WH$ to obtain:

$$\begin{aligned} \frac{\lambda - \beta}{\lambda} t_L WHL &= S_{LL} WH t_L^2 \\ \frac{\lambda - \beta}{\lambda} \left[W k_G G - HT \right] &= WH S_{LL} t_L^2 > 0, \end{aligned} \tag{E}$$

where we have used the government budget constraint (A5.38) to get to the second line. The sign on the right-hand side of (E) follows from that fact that $S_{LL} > 0$ and $t_L^2 > 0$. So the left-hand side of (E) confirms that for positive required revenue ($W k_G G - HT > 0$) it must be the case that $\lambda > \beta$.

****

We close this subsection on the modified Samuelson rule with a cautionary remark due to Atkinson and Stern (1974). Although it is rather tempting to do so, one cannot infer anything about the *levels* of public goods provision under the first-best and second-best cases by just comparing first order conditions (A5.47) and (A5.50). Even if the right-hand side of (A5.50) exceeds the right-hand side of (A5.47) it is not automatically the case that G is lower if t_L is used for financing than if T is used. Despite the fact that the left-hand sides of (A5.47) and (A5.50) look the same, they are nevertheless different because they depend on the triple (t_L, T, G). In the first-best case, this triple is set according to $\left(0, T^F, G^F\right)$ where T^F and G^F are the optimal choices for T and G. In the second-best case, the triple is set according to $\left(t_L^S, 0, G^S\right)$, where T^S and G^S are the second-best optimal choices for t_L and G. Since there are three things that differ between the first-best and second best cases, no unambiguous statement concerning the level of G under the two scenarios is possible. We return to this issue in an exercise to this chapter.

11.1.3 Redistribution

Up to this point we have abstracted from distributional issues by restricting attention to efficiency consideration in subsection 11.1.1 and by adopting a symmetric utilitarian approach in subsection 11.1.2. In this subsection the horizon is broadened by including redistributional considerations into the analysis.[3] Once the assumption of identical households is abandoned, one could pose the question whether the poor value public goods more highly than the rich do. And if so, how does this phenomenon affect the (modified) Samuelson rule? We study these questions in the context of a model in which households differ in ability and the policy maker must determine the optimal provision of a pure public consumption good. In concrete terms, we use the linear income tax model (studied in Chapter 11 above) and extend it by incorporating a pure public consumption good.

The (direct) utility function of household h ($h = 1, \ldots, H$) is:

$$U^h = U\left(1 - L^h, C^h, G\right), \tag{A5.51}$$

where the functional form $U(\cdot)$ is the same for all households. Households differ in their labour productivity (due to innate/exogenous skill differences) and the effective labour supply in *efficiency units* is denoted by $n^h L^h$, where L^h is labour supply in hours of work and n^h is an indicator of efficiency. There is no non-labour income so the household budget constraint is given by:

$$C^h = W^h L^h - T^h, \tag{A5.52}$$

where W^h is the wage of household h and T^h is the tax paid by that household. We focus on the simple case, where the tax schedule is linear:

$$T^h = -S^h + t_L W^h L^h, \tag{A5.53}$$

where S^h is the lump-sum subsidy (or tax, if $S^h < 0$) and t_L is the (constant) marginal tax rate (satisfying $0 < t_L < 1$). Household h chooses C^h and L^h in order to maximize (A5.51) subject to (A5.52) and (A5.53),

[3] This subsection draws on Atkinson and Stiglitz (1980, pp. 494-497).

taking as given W^h, S^h, t_L, and G. The Lagrangian for this optimization problem is:

$$\mathcal{L}^h \equiv U\left(C^h, 1-L^h, G\right) + \alpha^h \left[S^h + (1-t_L) W^h L^h - C^h\right],$$

where α^h is the Lagrange multiplier (equal to the marginal utility of lump-sum income to household h in the optimum). The first-order necessary conditions are the constraint and:

$$\frac{\partial U}{\partial C^h} = \alpha^h, \tag{A5.54}$$

$$\frac{\partial U}{\partial (1-L^h)} = \alpha^h (1-t_L) W^h. \tag{A5.55}$$

It follows from (A5.54)-(A5.55) that the household's labour supply decision is distorted if the marginal tax rate is non-zero. The first-order condition (A5.54)-(A5.55) and the constraint implicitly define the Marshallian consumption demand and labour supply which we write in general terms as:

$$C^h = C\left(S^h, (1-t_L) W^h, G\right), \tag{A5.56}$$

$$L^h = L\left(S^h, (1-t_L) W^h, G\right). \tag{A5.57}$$

Finally, by substituting (A5.56)-(A5.57) into the direct utility function (A5.51) we find the indirect utility function:

$$V^h = V\left(S^h, (1-t_L) W^h, G\right). \tag{A5.58}$$

From duality theory we recall the following useful properties of the indirect utility function:[4]

$$\frac{\partial V^h}{\partial S^h} = \frac{\partial V(\cdot)}{\partial S^h} = \alpha^h, \tag{A5.59}$$

$$\frac{\partial V^h}{\partial (1-t_L)} = \frac{\partial V(\cdot)}{\partial (1-t_L)} = \alpha^h W^h L^h, \tag{A5.60}$$

$$\frac{\partial V}{\partial G} = \frac{\partial U}{\partial G}. \tag{A5.61}$$

The objective function of the policy maker consists of an individualistic social welfare function:

$$SW \equiv \Psi\left(U^1, U^2, \ldots, U^H\right), \tag{A5.62}$$

where $\Psi_h \equiv \partial \Psi / \partial U^h > 0$ for all h. The policy maker cannot directly observe (or indirectly infer) the

[4]Roy's Lemma for labour supply says:

$$L^h = \frac{\partial V / \partial \left[(1-t_L) W^h\right]}{\partial V / \partial S^h} = \frac{1}{W^h} \frac{\partial V / \partial (1-t_L)}{\alpha^h},$$

from which the result in (A5.60) follows.

household's innate ability n^h so a tax on innate ability is infeasible. The policy maker thus faces a *second-best* social optimization problem! He can observe $W^h L^h$ but not W^h and L^h separately. For this reason the tax function (A5.53) is conditioned on $W^h L^h$ (rather than on n^h).

The production function for the private consumption good is linear:

$$Y = \frac{1}{k_Y} \sum_{h=1}^{H} n^h L_Y^h, \tag{A5.63}$$

where Y is aggregate output, k_Y is an exogenous productivity index, L_Y^h is labour of type h used in the Y-sector, n^h is productivity of type-h labour, and $n^h L_Y^h$ is thus the labour input measured in efficiency units. The profit function of the representative firm is defined as follows:

$$\begin{aligned} \Pi &\equiv PY - \sum_{h=1}^{H} W^h L_Y^h \\ &= \frac{P}{k_Y} \sum_{h=1}^{H} n^h L_Y^h - \sum_{h=1}^{H} W^h L_Y^h, \end{aligned} \tag{A5.64}$$

where P is the output price and n^h is assumed to be observable to the firm.[5] Under perfect competition there are zero pure profits and the wage paid to type-h workers is:

$$W^h = \frac{n^h}{k_Y}, \tag{A5.65}$$

where we have incorporated the assumption that the consumption good is used as the numeraire commodity, i.e. we have set $P = 1$. Not surprisingly, the real wage of type-h labour is proportional to innate ability n^h.

The production function for the public consumption good is:

$$G = \frac{1}{k_G} \sum_{h=1}^{H} n^h L_G^h, \tag{A5.66}$$

where G is output of the public good, L_G^h is type-h labour used in the public sector, and k_G is an exogenous productivity index. (In its role of employer of civil servants, the government is thus assumed to be able to observe the innate ability of its employees!) The government is assumed to produce the public good in an efficient, cost-minimizing manner, i.e. it minimizes labour cost $\sum_{h=1}^{H} W^h L_G^h$ subject to the technology (A5.66). The first-order condition is:

$$\frac{W^h}{P_G} = \frac{n^h}{k_G}, \tag{A5.67}$$

where P_G is marginal (and average) production cost for G. Note that from (A5.65) and (A5.67) we obtain

[5] The firm chooses the different types of labour in a profit maximizing manner. We obtain from (A5.64) that $\partial \Pi / \partial L_Y^h = P n^h / k_Y - W^h = 0$ or $W^h / P = n^h / k_Y$.

the expression for the *relative* price of the public good $P_G = k_G/k_Y$ (relative, that is, to the consumption good).

The market clearing conditions are:

$$Y = \sum_{h=1}^{H} C^h, \tag{A5.68}$$

$$L^h = L_Y^h + L_G^h, \qquad (\text{for } h = 1, \ldots, H). \tag{A5.69}$$

According to (A5.68), aggregate supply of the private good (left-hand side) equals aggregate demand for the consumption good by all households (right-hand side). Similarly, (A5.69) equates the supply of type-h labour (left-hand side) to the demand for that labour type in the private and public sectors.

Finally, the aggregate production constraint implied by (A5.63), (A5.66), and (A5.68) is:

$$\sum_{h=1}^{H} n^h L^h = k_Y \sum_{h=1}^{H} C^h + k_G G, \tag{A5.70}$$

whilst the government budget constraint is:

$$t_L \sum_{h=1}^{H} W^h L^h = \sum_{h=1}^{H} S^h + P_G G. \tag{A5.71}$$

The left-hand side is the total revenue raised by means of the labour income tax. The right-hand side is total outlays of the government, consisting of lump-sum transfers and outlays on the public good.

11.1.3.1 First-best social optimum

In order to build up intuition, we first study the first-best social optimization problem. Based on the previous analysis in subsection 11.1.2.1 one would expect the policy maker to set $t_L = 0$ and to raise all required revenue in a non-distorting fashion. Does the analysis give us this conclusion? By using (A5.58) in (A5.62), the policy maker's objective function can be written as follows:

$$\begin{aligned} SW \equiv \Psi\Big(& V\left(S^1, (1-t_L)\, W^1, G\right), V\left(S^2, (1-t_L)\, W^2, G\right), \ldots, \\ & V\left(S^H, (1-t_L)\, W^H, G\right)\Big). \end{aligned} \tag{A5.72}$$

The revenue requirement constraint is given in (A5.71) and the policy instruments are G, t_L, and S^h (for $h = 1, \ldots, H$). The policy maker sets these instruments in such a way as to maximize (A5.72) subject to (A5.71). The associated Lagrangian is:

$$\mathcal{L} \equiv SW + \lambda \left[\sum_{h=1}^{H} \left[t_L W^h L^h \left(S^h, (1-t_L)\, W^h, G\right) - S^h \right] - P_G G \right],$$

where λ is the Lagrange multiplier for the government budget constraint. The first-order necessary conditions consist of the constraint and:

$$\frac{\partial \mathcal{L}}{\partial G} = \sum_{h=1}^{H} \Psi_h \frac{\partial V}{\partial G} + \lambda \left[t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial G} - P_G \right] = 0, \tag{A5.73}$$

$$\frac{\partial \mathcal{L}}{\partial S^h} = \Psi_h \frac{\partial V}{\partial S^h} + \lambda \left[t_L W^h \frac{\partial L^h}{\partial S^h} - 1 \right] = 0, \quad (\text{for } h = 1, \ldots, H), \tag{A5.74}$$

$$\frac{\partial \mathcal{L}}{\partial t_L} = - \sum_{h=1}^{H} \Psi_h \frac{\partial V}{\partial (1 - t_L)} + \lambda \left[\sum_{h=1}^{H} W^h L^h - t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial (1 - t_L)} \right] = 0. \tag{A5.75}$$

By using (A5.59)-(A5.61) we can simplify these expressions to:

$$\sum_{h=1}^{H} \Psi_h \frac{\partial U}{\partial G} = \lambda \left[P_G - t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial G} \right], \tag{A5.76}$$

$$0 = \Psi_h \alpha^h + \lambda \left[t_L W^h \frac{\partial L^h}{\partial S^h} - 1 \right], \quad (\text{for } h = 1, \ldots, H), \tag{A5.77}$$

$$0 = \sum_{h=1}^{H} \Psi_h \alpha^h W^h L^h - \lambda \sum_{h=1}^{H} W^h L^h + \lambda t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial (1 - t_L)}. \tag{A5.78}$$

Equation (A5.76) is the condition determining the optimal level of public goods provision, (A5.77) is the condition determining the optimal lump-sum subsidy/tax *for each household* (S^h), and (A5.78) is the condition determining the optimal marginal tax rate. Jointly with (A5.71), equations (A5.76)-(A5.78) determine $\left(G, S^1, \ldots, S^H, t_L, \lambda\right)$.

Let us test our intuition (mentioned above) by rewriting the formulae dealing with the tax system (i.e. equations (A5.77) and (A5.78)). The Slutsky equation for household h is:

$$\frac{\partial L^h}{\partial (1 - t_L)} = S_{LL}^h + W^h L^h \frac{\partial L^h}{\partial S^h}, \tag{A5.79}$$

where $S_{LL}^h \equiv \left(\partial L^h / \partial (1 - t_L)\right)_{U_0} > 0$ is the pure substitution effect for household h.[6] We define the social marginal utility of income as:

$$\beta^h \equiv \Psi_h \alpha^h + \lambda t_L W^h \frac{\partial L^h}{\partial S^h}. \tag{A5.80}$$

By using (A5.79)-(A5.80), equations (A5.77) and (A5.78) can now be written in a rather compact format as:

$$\beta^h = \lambda, \tag{A5.81}$$

[6]Note that (A5.79) differs from (A5.44) because (i) the income and pure substitution effects are household specific and (ii) because S^h is a lump-sum transfer (rather than a tax, as in (A5.44)).

$$0 = \sum_{h=1}^{H} \left[\beta^h - \lambda + \frac{\lambda t_L S_{LL}^h}{L^h} \right] W^h L^h. \tag{A5.82}$$

If the policy maker has access to h-specific lump-sum transfers, he will set the marginal social utility of income equal to λ for all households (see (A5.81)). It follows from (A5.82) that the labour income tax will be set equal to zero (as the pure substitution effect is strictly positive). hence, in the first-best optimum, we have:

$$t_L = 0, \quad \beta^h \equiv \Psi_h \alpha^h = \lambda, \tag{A5.83}$$

i.e. redistribution and public good financing is achieved by means of the non-distorting tax. No distorting labour income tax is needed.

In view of these results, the first-order condition for public good provision (A5.76) reduces to the standard Samuelson rule:

$$\sum_{h=1}^{H} \frac{\Psi_h \frac{\partial U}{\partial G}}{\lambda} = P_G \quad \Leftrightarrow \quad \sum_{h=1}^{H} \frac{\Psi_h \frac{\partial U}{\partial G}}{\Psi_h \alpha^h} = P_G \quad \Leftrightarrow$$
$$\sum_{h=1}^{H} \frac{\partial U / \partial G}{\partial U / \partial C^h} = \frac{k_G}{k_Y}, \tag{A5.84}$$

where we have used the fact that $P_G = k_G / k_Y$ in the final step. Equation (A5.84) generalizes (A5.47) to the case of heterogeneous households.

11.1.3.2 Second-best social optimum, I

Depending on the kind of additional restriction faced by the policy maker, the social optimization problem becomes a second-best problem. Here we study a *restricted* scenario in which it is assumed that the only tax instrument available is a *uniform* lump-sum tax, i.e. $S^h = S$ for all h and there is no labour income tax ($t_L = 0$). What does the second-best optimal provision of public goods rule look like under this scenario?

The policy maker's objective function is now:

$$SW \equiv \Psi\left(V\left(S, W^1, G\right), V\left(S, W^2, G\right), \ldots, V\left(S, W^H, G\right) \right), \tag{A5.85}$$

the revenue requirement constraint is:

$$-HS = P_G G, \tag{A5.86}$$

and the instruments are G and S. Writing the Lagrangian as $\mathcal{L} \equiv SW - \lambda \left[HS + P_G G \right]$, we obtain the following first-order conditions:

$$\frac{\partial \mathcal{L}}{\partial G} = 0 \quad \Leftrightarrow \quad \lambda P_G = \sum_{h=1}^{H} \Psi_h \frac{\partial U}{\partial G}, \tag{A5.87}$$

$$\frac{\partial \mathcal{L}}{\partial S} = 0 \quad \Leftrightarrow \quad \frac{1}{H} \sum_{h=1}^{H} \Psi_h \alpha^h = \lambda, \tag{A5.88}$$

where we have used (A5.59) and (A5.61).

Defining the social marginal utility of income as $\beta^h \equiv \Psi_h \alpha^h$ and its mean value as $\bar{\beta} \equiv \sum_{h=1}^{H} \beta^h / H$, we can rewrite (A5.88) in compact format as:

$$\bar{\beta} = \lambda. \tag{A5.89}$$

Whereas in the first-best optimum, the policy maker can manipulate S^h in such a way as to equate β^h to λ for all households (see equation (A5.83) above), in this version of the second-best optimum, the policy maker can only manipulate S and can thus only make the *average* of all β^h's equal to λ. By using the definitions of β^h and $\bar{\beta}$ and noting (A5.89), equation (A5.87) can be rewritten as:

$$\sum_{h=1}^{H} \frac{\Psi_h \frac{\partial U}{\partial G}}{\bar{\beta}} = P_G \quad \Leftrightarrow \quad \sum_{h=1}^{H} \frac{\alpha^h \Psi_h \frac{\partial U}{\partial G}}{\bar{\beta} \alpha^h} = P_G \quad \Leftrightarrow$$
$$\sum_{h=1}^{H} \frac{\beta^h}{\bar{\beta}} \frac{\partial U / \partial G}{\partial U / \partial C^h} = \frac{k_G}{k_Y}. \tag{A5.90}$$

Equation (A5.90) is the counterpart to (A5.84) in a second-best setting. In contrast to the (first-best) Samuelson Rule, here it is the *weighted sum* of marginal rates of substitution that features on the left-hand side, with the weight applied to household h equalling $\beta^h / \bar{\beta}$.

It is straightforward (though a little tedious) to show that (A5.90) can also be written in terms of a covariance term (see the Intermezzo below):

$$\frac{k_G}{k_Y} = \sum_{h=1}^{H} MRS^h \left[1 + \text{cov}\left(\frac{\beta^h}{\bar{\beta}} /, \frac{MRS^h}{\overline{MRS}} \right) \right], \tag{A5.91}$$

where $MRS^h \equiv (\partial U / \partial G) \, / \left(\partial U / \partial C^h \right)$ is the marginal rate of substitution between G and C^h for household h, cov $\left(\beta^h / \bar{\beta}, MRS^h / \overline{MRS} \right)$ is the covariance between $\beta^h / \bar{\beta}$ and $MRS^h / \overline{MRS}$ representing the *distributional characteristic* of the public good. Compared to the first-best rule (A5.84), the second-best rule takes into account the covariance between the social marginal utility of income and the marginal rate of substitution. Consider the case where β^h declines with income of household h. Then for a public good that is valued more highly by the poor than by the rich (cov $\left(\beta^h / \bar{\beta}, MRS^h / \overline{MRS} \right) > 0$), the covariance term is positive and the term in round brackets on the left-hand side of (A5.91) is greater than unity, i.e. $\sum_{h=1}^{H} MRS^h < MRT$ (see Atkinson and Stiglitz, 1980, p. 496).

11.1.3.3 Second-best social optimum, II

We now expand the set of tax instruments available to the policy maker by assuming that both the uniform lump-sum tax ($S^h = S$ for all h) and the labour income tax (t_L) can be used for financing and redistribution purposes. In this version of the second-best problem the policy maker's objective function is:

$$SW \equiv \Psi\left(V\left(S, (1-t_L)W^1, G\right), \ldots, V\left(S, (1-t_L)W^H, G\right)\right), \tag{A5.92}$$

the revenue requirement is:

$$t_L \sum_{h=1}^{H} W^h L^h - HS = P_G G, \tag{A5.93}$$

and the instruments are G, S and t_L. The Lagrangian expression is:

$$\mathcal{L} \equiv SW + \lambda \left[t_L \sum_{h=1}^{H} W^h L^h \left(S, (1-t_L)W^h, G\right) - HS - P_G G \right],$$

and the (most interesting) first-order conditions are:

$$\frac{\partial \mathcal{H}}{\partial G} = 0 \quad \Leftrightarrow \quad \sum_{h=1}^{L} \Psi_h \frac{\partial V}{\partial G} = \lambda \left[P_G - t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial G} \right], \tag{A5.94}$$

$$\frac{\partial \mathcal{H}}{\partial S} = 0 \quad \Leftrightarrow \quad \sum_{h=1}^{L} \Psi_h \frac{\partial V}{\partial S} = \lambda \left[H - t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial S} \right], \tag{A5.95}$$

$$\frac{\partial \mathcal{H}}{\partial t_L} = 0 \quad \Leftrightarrow \quad \sum_{h=1}^{L} \Psi_h \frac{\partial V}{\partial (1-t_L)} = \lambda \left[\sum_{h=1}^{H} W^h L^h - t_L \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial (1-t_L)} \right]. \tag{A5.96}$$

By using (A5.59)-(A5.61) these expressions can be further simplified:

$$\lambda P_G = \sum_{h=1}^{H} \beta^h MRS^h + \lambda t_L H \frac{\partial \overline{WL}}{\partial G}, \tag{A5.97}$$

$$\bar{\beta} = \lambda \left[1 - t_L \frac{\partial \overline{WL}}{\partial S} \right], \tag{A5.98}$$

$$0 = \sum_{h=1}^{H} \beta^h W^h L^h - \lambda H \overline{WL} \left[1 - \frac{t_L}{\overline{WL}} \frac{\partial \overline{WL}}{\partial t_L} \right], \tag{A5.99}$$

where we have defined $\beta^h \equiv \Psi_h \alpha^h$, $\bar{\beta} \equiv \sum_{h=1}^{H} \beta^h / H$ and $\overline{WL} \equiv \sum_{h=1}^{H} W^h L^h / H$.[7] Equation (A5.97) is the condition determining the optimal level of public goods provision, (A5.97) determines the optimal

[7]Note that, since W^h is constant, we have:

$$H \frac{\partial \overline{WL}}{\partial G} \equiv \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial G}, \quad H \frac{\partial \overline{WL}}{\partial S} \equiv \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial S}, \quad H \frac{\partial \overline{WL}}{\partial (1-t_L)} \equiv \sum_{h=1}^{H} W^h \frac{\partial L^h}{\partial (1-t_L)}.$$

lump-sum subsidy/tax, and (A5.99) determines the optimal labour income tax rate. Jointly with the revenue requirement constraint (A5.93), these conditions determine (G, S, t_L, λ). The modified Samuelson rule is obtained by combining (A5.97)-(A5.98) and rewriting slightly:

$$\frac{\lambda}{\bar{\beta}}\left[P_G - t_L H \frac{\partial \overline{WL}}{\partial G}\right] = \sum_{h=1}^{H} \frac{\beta^h}{\bar{\beta}} MRS^h \quad \Leftrightarrow$$

$$\frac{\frac{k_G}{k_Y} - t_L H \frac{\partial \overline{WL}}{\partial G}}{1 - t_L \frac{\partial \overline{WL}}{\partial S}} = \sum_{h=1}^{H} MRS^h \left[1 + \text{cov}\left(\frac{\beta^h}{\bar{\beta}}, \frac{MRS^h}{\overline{MRS}}\right)\right], \tag{A5.100}$$

where we have used the fact that $P_G = k_G/k_Y$. Compared to (A5.91), this expression only differs on the left-hand side. An "inverse-elasticity type" expression for the optimal labour income tax formula can be derived from (A5.99):

$$\frac{t_L}{1 - t_L} = \frac{1}{\sigma_L} \frac{\lambda - \hat{\beta}}{\lambda}, \tag{A5.101}$$

where $\hat{\beta}$ and σ_L are defined as follows:

$$\hat{\beta} \equiv \frac{\sum_{h=1}^{H} \beta^h W^h L^h}{H\overline{WL}}, \tag{A5.102}$$

$$\sigma_L \equiv \frac{\partial \sum_{h=1}^{H} W^h L^h}{\partial (1 - t_L)} \frac{1 - t_L}{\sum_{h=1}^{H} W^h L^h} = \frac{\partial \overline{WL}}{\partial (1 - t_L)} \frac{1 - t_L}{\overline{WL}}. \tag{A5.103}$$

The derivation of (A5.101) is left as exercise to the reader (Hint: see the derivation of equation (11.37) in Chapter 11).

Intermezzo 11.2

Derivation of the covariance formula (A5.91). We write the expression in (A5.90) as follows:

$$\sum_{h=1}^{H} \frac{\beta^h}{\bar{\beta}} MSR^h = \frac{k_G}{k_Y}, \tag{A}$$

where $MSR^h \equiv (\partial U/\partial G) / \left(\partial U/\partial C^h\right)$. It follows that:

$$\begin{aligned} \frac{k_G}{k_Y} &= \sum_{h=1}^{H} \frac{\beta^h + \bar{\beta} - \bar{\beta}}{\bar{\beta}} MSR^h \\ &= \sum_{h=1}^{H} MSR^h + \sum_{h=1}^{H} \frac{\beta^h - \bar{\beta}}{\bar{\beta}} MSR^h \\ &= \sum_{h=1}^{H} MSR^h + \sum_{h=1}^{H} \frac{\beta^h - \bar{\beta}}{\bar{\beta}} \left(MSR^h - \overline{MRS}\right) \end{aligned} \tag{B}$$

where we have used the fact that:

$$\sum_{h=1}^{H} \frac{\beta^h - \bar{\beta}}{\bar{\beta}} \overline{MRS} = \frac{\overline{MRS}}{\bar{\beta}} \left[\sum_{h=1}^{H} \left(\beta^h - \bar{\beta} \right) \right] = 0. \tag{C}$$

Recall the definition of the covariance between variables x_h and y_h:

$$\text{cov}\,(x_h, y_h) \equiv \frac{1}{H} \sum_{h=1}^{H} (x_h - \bar{x})\,(y_h - \bar{y}) \tag{D}$$

$$= \frac{1}{H} \sum_{h=1}^{H} x_h y_h - \bar{x}\bar{y}, \tag{E}$$

where $\bar{x} \equiv \sum_{h=1}^{H} x_h / H$ and $\bar{y} \equiv \sum_{h=1}^{H} y_h / H$ are the respective means. By using (D) in (B) we find:

$$\frac{k_G}{k_Y} = \sum_{h=1}^{H} MSR^h + H\text{cov}\left(\frac{\beta^h}{\bar{\beta}}, MRS^h \right). \tag{F}$$

An alternative way to write (F) is:

$$\frac{k_G}{k_Y} = \sum_{h=1}^{H} MSR^h \left[1 + \text{cov}\left(\frac{\beta^h}{\bar{\beta}}, \frac{MRS^h}{\overline{MRS}} \right) \right], \tag{G}$$

where we have used the fact that:

$$H\text{cov}\left(\frac{\beta^h}{\bar{\beta}}, MRS^h \right) = H\overline{MRS}\text{cov}\left(\frac{\beta^h}{\bar{\beta}}, \frac{MRS^h}{\overline{MRS}} \right)$$

$$= \sum_{h=1}^{H} MSR^h \text{cov}\left(\frac{\beta^h}{\bar{\beta}}, \frac{MRS^h}{\overline{MRS}} \right). \tag{H}$$

****

11.2 Private provision of public goods

Up to this point attention has been restricted to publicly provided public goods. The *private* provision of public goods is however quite significant in most countries. One can think, for example, of contributions to national and international charities, donations to political parties / pressure groups, public goods (or public "bads") as by-product of private goods, and foundations for arts, sciences, health care, e.g. Ford Foundation, Carnegie-Mellon, Bill & Melinda Gates Foundation, Getty Museum. In this section we study two models dealing with the private provision of public goods, namely the *private subscription* model and the *externality-based* model.

11.2.1 Private subscriptions

In the subscription model of public goods provision, individual households value the public good and make individual contributions in order to increase its production. This approach was studied *inter alia* by Malinvaud (1972, pp. 211-218) and Atkinson and Stiglitz (1980, pp. 505-507). In this subsection we present a simple subscription model and study its key properties. The key assumptions of the model are the following. There are H identical households who each possess selfish motives for contributing to the cost of public goods production. Households take the contributions by other households as given (Cournot-Nash assumption). The technologies for the private and public goods are both linear, and labour is the only production factor. Labour is used as the numeraire ($W = 1$). The key question to be answered is whether the subscription equilibrium gives rise to excessive or deficient provision of public goods (from a social point of view).

The representative household h ($h = 1, \ldots, H$) has a (direct) utility function defined over the private good and the *total* supply of the pure public good:

$$U^h = U\left(C^h, G^h + G^{others}\right), \tag{A5.104}$$

where C^h is consumption of the private good, G^h is the amount of the public good provided by household h, and G^{others} is the total supply of the public good by all other households:

$$G^{others} \equiv \sum_{\substack{i=1 \\ i \neq h}}^{H} G^i \tag{A5.105}$$

The utility function $U[\cdot]$ has the usual properties: $U_C > 0$, $U_G > 0$, $U_{CC} < 0$, $U_{GG} < 0$ and $U_{CC}U_{GG} - U_{CG}^2 > 0$ (i.e., the indifference curves are downward sloping in $\left(C^h, G^h\right)$ space and bulge towards the origin). [REFER TO CHAPTER 2] The budget constraint of household h is:

$$PC^h = W - P_G G^h, \tag{A5.106}$$

where P is the consumer price of the private good, W is wage income ($L^h = 1$ is the exogenous labour supply), and P_G is the price of the public good (see below). The household chooses C^h and G^h in order to maximize (A5.104) subject to (A5.106), taking as given the wage rate and both prices (W, P, P_G) and the donations by other households (G^{others}; the Cournot-Nash assumption). Clearly, the household cannot give a negative contribution, i.e. an additional constraint is:

$$G^h \geq 0 \tag{A5.107}$$

Production conditions are exactly the same as in the model of subsection 11.1.2, i.e. the technologies are $Y = L_Y/k_Y$ (for the private good) and $G = L_G/k_G$ (for the public good), where L_Y and L_G are the respective labour inputs. The prices are $P = Wk_Y$ (for the private good) and $P_G = Wk_G$ (for the public good). The private goods market clearing condition is:

$$\sum_{h=1}^{H} C^h = Y, \tag{A5.108}$$

and the labour market clearing condition is:

$$\left(\sum_{h=1}^{H} L^h = \right) \quad H = L_Y + L_G. \tag{A5.109}$$

The decision problem for household h is best solved by transforming it slightly. The utility function (A5.104) is rewritten as:

$$U^h = U\left(C^h, G\right), \tag{A5.110}$$

where G is the total public good provision:

$$G \equiv G^h + G^{others}. \tag{A5.111}$$

Using (A5.111) the budget equation (A5.106) can be rewritten as:

$$k_Y C^h + k_G G = 1 + k_G G^{others}, \tag{A5.112}$$

where we have used $P = Wk_Y$, $P_G = Wk_G$, and we have set $W = 1$. Finally, the non-negativity constraint (A5.107) is rewritten as:

$$G \geq G^{others}. \tag{A5.113}$$

In the transformed model, the household chooses C^h and G in order to maximize (A5.110) subject to (A5.112) and (A5.113). The Lagrangian for this problem is:

$$\mathcal{L} \equiv U\left(C^h, G\right) + \alpha\left[1 + k_G G^{others} - k_Y C^h - k_G G\right] + \eta\left[G^{others} - G\right],$$

where α and η are the Lagrange multipliers for, respectively, the budget constraint and the non-negativity constraint. The first-order necessary conditions for C^h and G are:

$$\frac{\partial \mathcal{L}}{\partial C^h} = \frac{\partial U}{\partial C^h} - \alpha k_Y = 0, \tag{A5.114}$$

$$\frac{\partial \mathcal{L}}{\partial G} = \frac{\partial U}{\partial G} - \alpha k_G - \eta = 0, \tag{A5.115}$$

whilst the first-order conditions for α and η are

$$\frac{\partial \mathcal{L}}{\partial \alpha} = 0, \tag{A5.116}$$

$$\frac{\partial \mathcal{L}}{\partial \eta} = G^{others} - G \geq 0, \quad \eta \geq 0, \quad \eta \frac{\partial \mathcal{L}}{\partial \eta} = 0. \tag{A5.117}$$

In (A5.117) the so-called *Kuhn-Tucker conditions* are used because it may very well be the case that household h chooses not to contribute to the public good provision. The interpretation of the first-order conditions is facilitated by looking at two cases. In case 1, assume that $\eta > 0$. Then it follows from the third expression in (A5.117) (the complementary slackness condition) that $\partial \mathcal{L}/\partial \eta = 0$, i.e. $G = G^{others}$ and household h does not contribute to the public goods supply. In case 2, we assume that the household contributes ($G > G^{others}$). Then it follows from (A5.117) that $\partial \mathcal{L}/\partial \eta > 0$ and $\eta = 0$. In case 2 we have an interior solution with:

$$\frac{\frac{\partial U(C^h,G)}{\partial G}}{\frac{\partial U(C^h,G)}{\partial C^h}} = \frac{k_G}{k_Y}, \qquad (\text{ for all } h = 1, \ldots, H). \tag{A5.118}$$

According to (A5.118), the *own* marginal rate of substitution between the public good and the private good of household h is equated to the marginal rate of transformation between these two goods. (Note that (A5.118) holds for all households because the model is completely symmetric, i.e. all households face the same conditions.)

In Figure 11.3 the model is illustrated graphically. In that figure, HBC_0 is the household budget line for the case that the other households do not contribute, i.e. $G^{others} = 0$. The household chooses point E_0 and contributes G_0. Similarly, HBC_1 is the budget constraint for $G^{others} = G_1^{others} > 0$. The slope of the budget line is unchanged (and equal to k_G/k_Y) and the household now chooses point E_1 and contributes G_1. At point C we have $C^h = 1/k_Y$ so that (A5.112) implies that $G = G_1^{others}$ there. Since G_1 is the optimum choice, the line segment CD measures the household's own contribution. Note that for simplicity we assume *homothetic preferences* so that the tangencies between HBC_0 and HBC_1 and the indifference curves lie along a straight line through the origin. We interpret this Engel curve as a *Nash reaction curve*. It plots the optimal choices of household h as a function of the collective decisions by all other households.

Since all households are (by assumption) identical, they all set $\left(C^h, G\right)$ according to (A5.118). The aggregate economy-wide resource constraint can then be written as:

$$H = L_Y + L_G = k_Y H C^h + k_G G \quad \Leftrightarrow$$

$$1 = k_Y C^h + k_G \frac{G}{H}. \tag{A5.119}$$

In the Nash Equilibrium, the economy is at the intersection of the Nash reaction curve and the resource

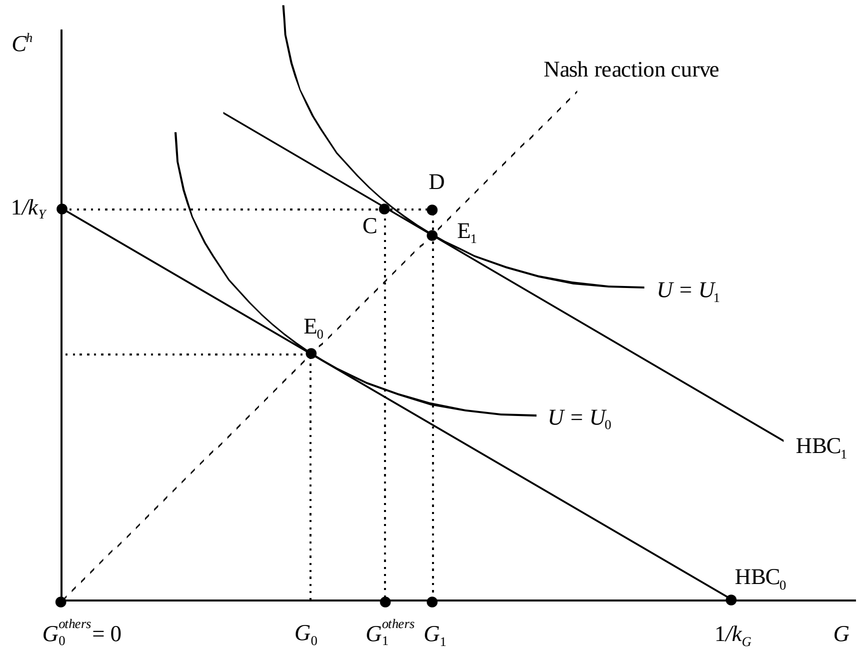

Figure 11.3: The Nash reaction curve

constraint. In Figure 11.4, the private subscription Nash equilibrium is at point E_M and the equilibrium public good supply is G_M. As we recall from subsection 11.1.2.1, the first-best social optimum calls for:

$$H\frac{\frac{\partial U(C^h,G)}{\partial G}}{\frac{\partial U(C^h,G)}{\partial C^h}} = \frac{k_G}{k_Y}, \tag{A5.120}$$

which is a *tangency* between the resource constraint (A5.119) and an indifference curve: see point E_S in Figure 11.4. It follows that there is an under-supply of public goods in the private subscription equilibrium. Intuitively, individual households fail to take into account other agents' beneficial utility effects when choosing their own contribution level. (The reader is asked to verify that the first-best equilibrium can be decentralized as a private subscription equilibrium by means of a Pigouvian subsidy on public goods plus a lump-sum tax.)

11.2.2 Externalities and public goods

For pure private goods one agent's consumption of it does not directly affect any other agents' enjoyment. For pure public goods, in contrast, if one agent provides some of it, then all other agents are also able to enjoy it. Some private goods can be seen as impure public goods, in that there may be *consumption externalities* or *production externalities*. (Note we are talking about *direct* externalities, not pecuniary externalities.)

Examples of external effects are easy to come by. For vaccinations, it is typically the case that if one agent is vaccinated against smallpox then all other agents are (a little) safer as a result. Another

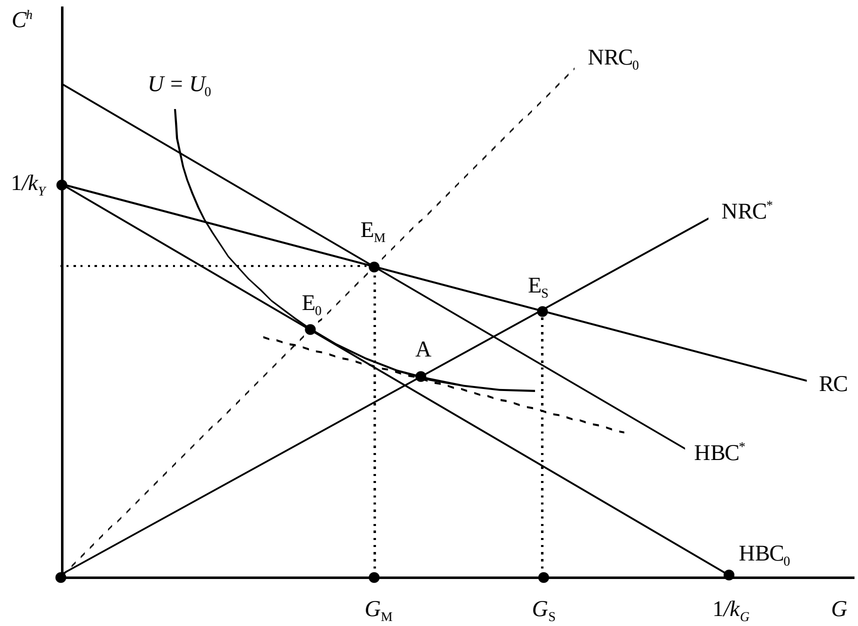

Figure 11.4: The private subscription equilibrium for public good provision

example of an external effect deals with the interaction between beekeepers and orchard farmers: the bees fertilize the flowers of the trees and thus increase productivity of the orchard. Education may also cause significant external effects: a society with an educated population functions more smoothly and there may be external effects in R&D activities (see Chapter 8). All of the above are examples of positive consumption / production externalities. Examples of negative externalities are the over-exploitation of a common property resource (the "tragedy of the commons") and air- and water pollution as a by-product of production activities. All these examples can be interpreted as private provision of impure public goods (or public "bads").

The objective of this subsection is to demonstrate that the market equilibrium is generally not efficient in the presence of external effects. In addition we show how the policy maker can use corrective taxation to restore the economy to the Pareto frontier. We use a simple two-good-two-factor general equilibrium model which is loosely based on the seminal paper by Meade (1952). The key assumptions of the model are as follows. There are two produced goods (Y_1 and Y_2) and the output in sector 1 (Y_1) negatively affects household utility, say because its production results in pollution as a by-product. We use the notation of the general equilibrium model introduced in Chapter 9: there are H households, $F = 2$ production factors, and $G = 2$ goods.

The utility function of household h takes the following form:

$$U^h = U^h\left(X_1^h, X_2^h, V_1^h, V_2^h, Y_1\right), \tag{A5.121}$$

where U^h is utility of household h $(1, 2, \ldots, H)$, X_g^h is consumption of good g by household h $(g = 1, 2)$, V_f^h is the supply of production factor f by household h $(f = 1, 2)$, and $\partial U^h / \partial Y_1 < 0$, i.e. the *aggregate* output of sector 1 negatively affects the household's utility (say due to pollution). The technology for good g is summarized by the production functions

$$Y_g = F^g\left(Z_1^g, Z_2^g\right), \qquad \text{(for } g = 1,2\text{)}, \tag{A5.122}$$

where Y_g is aggregate production of good g, $F^g(\cdot)$ is the production function for good g, and Z_f^g is factor f used in the production of good g. The market clearing conditions in the goods and factor markets are:

$$Y_g = \sum_{h=1}^{H} X_g^h, \qquad \text{(for } g = 1,2\text{)}, \tag{A5.123}$$

$$\sum_{h=1}^{H} V_f^h = \sum_{g=1}^{2} Z_f^g, \qquad \text{(for } f = 1,2\text{)}. \tag{A5.124}$$

The general equilibrium model is fully characterized by equations (A5.121)-(A5.124).

11.2.2.1 Pareto optimum

In order to find the necessary conditions for Pareto efficiency, we follow the usual procedure by focusing on an arbitrary household, say household $h = 1$, holding every other household's utility constant (i.e. $U^h = U_0^h$ for $h = 2, \ldots, H$), and maximizing household 1's utility subject to the restrictions. The social planner thus chooses X_g^h, V_f^h, Z_f^g, Y_g in order to maximize:

$$U^1 = U^1\left(X_1^1, X_2^1, V_1^1, V_2^1, Y_1\right), \tag{A5.125}$$

subject to (A5.122)-(A5.124) and:

$$U_0^h = U^h\left(X_1^h, X_2^h, V_1^h, V_2^h, Y_1\right), \qquad \text{(for } h = 2, \ldots, H\text{)}. \tag{A5.126}$$

The Lagrangian for this problem is:

$$\begin{aligned}
\mathcal{L} \equiv\ & U^1\left(X_1^1, X_2^1, V_1^1, V_2^1, Y_1\right) + \sum_{h=2}^{H} \lambda_h \left[U^h\left(X_1^h, X_2^h, V_1^h, V_2^h, Y_1\right) - U_0^h\right] \\
& + \sum_{g=1}^{2} \mu_g \left[Y_g - F^g\left(Z_1^g, \ldots, Z_F^g\right)\right] + \sum_{g=1}^{2} \nu_g \left[\sum_{h=1}^{H} X_g^h - Y_g\right] \\
& + \sum_{f=1}^{2} \xi_f \left[\sum_{h=1}^{H} V_f^h - \sum_{g=1}^{2} Z_f^g\right],
\end{aligned}$$

where the Lagrange multipliers are λ_h (for $h = 2, \ldots, H$), μ_g (for $g = 1,2$), ν_g (for $g = 1,2$), and ξ_f (for $f = 1,2$), i.e. there are $(H-1)+4+2$ Lagrange multipliers in all. To simplify notation we set $\lambda_1 = 1$. The first-order necessary conditions (assuming interior solutions) are the constraints and: (i) for the good demands (2H equations):

$$\frac{\partial \mathcal{L}}{\partial X_1^h} = \lambda_h \frac{\partial U^h}{\partial X_1^h} + \nu_1 = 0, \tag{A5.127}$$

$$\frac{\partial \mathcal{L}}{\partial X_2^h} = \lambda_h \frac{\partial U^h}{\partial X_2^h} + \nu_2 = 0, \tag{A5.128}$$

(ii) for the factor supplies (2H equations):

$$\frac{\partial \mathcal{L}}{\partial V_1^h} = \lambda_h \frac{\partial U^h}{\partial V_1^h} + \xi_1 = 0, \tag{A5.129}$$

$$\frac{\partial \mathcal{L}}{\partial V_2^h} = \lambda_h \frac{\partial U^h}{\partial V_2^h} + \xi_2 = 0, \tag{A5.130}$$

(iii) for the output decisions (2 equations):

$$\frac{\partial \mathcal{L}}{\partial Y_1} = \sum_{h=1}^{H} \lambda_h \frac{\partial U^h}{\partial Y_1} + \mu_1 - \nu_1 = 0, \tag{A5.131}$$

$$\frac{\partial \mathcal{L}}{\partial Y_2} = \mu_2 - \nu_2 = 0, \tag{A5.132}$$

and (iv) for the factor demands (4 equations):

$$\frac{\partial \mathcal{L}}{\partial Z_1^1} = -\mu_1 \frac{\partial F^1}{\partial Z_1^1} - \xi_1 = 0, \tag{A5.133}$$

$$\frac{\partial \mathcal{L}}{\partial Z_2^1} = -\mu_1 \frac{\partial F^1}{\partial Z_2^1} - \xi_2 = 0, \tag{A5.134}$$

$$\frac{\partial \mathcal{L}}{\partial Z_1^2} = -\mu_2 \frac{\partial F^2}{\partial Z_1^2} - \xi_1 = 0, \tag{A5.135}$$

$$\frac{\partial \mathcal{L}}{\partial Z_2^2} = -\mu_2 \frac{\partial F^2}{\partial Z_2^2} - \xi_2 = 0. \tag{A5.136}$$

Compared to the case without external effects, the new terms appear in equation (A5.131). In the pollution causing industry it is no longer the case that $\mu_1 = \nu_1$. Indeed, since $\lambda_h > 0$ and $\partial U^h / \partial Y_1 < 0$ it will be the case that $\mu_1 > \nu_1$.

Next we perform the usual trick by eliminating the various Lagrange multipliers. For an arbitrary household h we derive the usual expressions:

$$\frac{\nu_1}{\nu_2} = \frac{\partial U^h / \partial X_1^h}{\partial U^h / \partial X_2^h}, \tag{A5.137}$$

$$\frac{\xi_1}{\xi_2} = \frac{\partial U^h / \partial V_1^h}{\partial U^h / \partial V_2^h}, \tag{A5.138}$$

$$\frac{\xi_f}{\nu_g} = \frac{\partial U^h / \partial V_f^h}{\partial U^h / \partial X_g^h}. \tag{A5.139}$$

On the production side we obtain from (A5.133)-(A5.136) that:

$$\frac{\xi_1}{\xi_2} = \frac{\partial F^1/\partial Z_1^1}{\partial F^1/\partial Z_2^1} = \frac{\partial F^2/\partial Z_1^2}{\partial F^2/\partial Z_2^2}, \tag{A5.140}$$

and from (A5.131)-(A5.136) that:

$$1 = \frac{\nu_1 + \Phi}{\nu_2} \frac{\partial F^1/\partial Z_f^1}{\partial F^2/\partial Z_f^2}, \tag{A5.141}$$

where Φ is defined as:

$$\Phi \equiv -\sum_{h=1}^{H} \lambda_h \frac{\partial U^h}{\partial Y_1} = \nu_1 \sum_{h=1}^{H} \frac{\partial U^h/\partial Y_1}{\partial U^h/\partial X_1^h} > 0, \tag{A5.142}$$

and we have also used (A5.127) to get to the second expression on the right-hand side of (A5.142).

We can now combine even further. From (A5.137) and (A5.141) we find:

$$\left[\frac{\nu_1}{\nu_2} =\right] \quad \frac{\partial U^h/\partial X_1^h}{\partial U^h/\partial X_2^h} = \frac{\nu_1}{\nu_1 + \Phi} \frac{\partial F^2/\partial Z_f^2}{\partial F^1/\partial Z_f^1}, \tag{A5.143}$$

from (A5.138) and (A5.140) we obtain:

$$\left[\frac{\xi_1}{\xi_2} =\right] \quad \frac{\partial U^h/\partial V_1^h}{\partial U^h/\partial V_2^h} = \frac{\partial F^g/\partial Z_1^g}{\partial F^g/\partial Z_2^g}, \tag{A5.144}$$

and from (A5.139) and (A5.131)-(A5.136) we get:

$$\left[-\frac{\xi_f}{\nu_1} =\right] \quad -\frac{\partial U^h/\partial V_f^h}{\partial U^h/\partial X_1^h} = \frac{\nu_1 + \Phi}{\nu_1} \frac{\partial F^1}{\partial Z_f^1}, \tag{A5.145}$$

$$\left[-\frac{\xi_f}{\nu_2} =\right] \quad -\frac{\partial U^h/\partial V_f^h}{\partial U^h/\partial X_2^h} = \frac{\partial F^2}{\partial Z_f^2}. \tag{A5.146}$$

11.2.2.2 Market equilibrium

Does the decentralized economy satisfy Pareto efficiency? The *intuitive* answer is "no, probably not, unless the government intervenes" because competitive firms do not take their negative pollution effect on households into account when making their production decisions. The *technical* answer is obtained by spelling out the institutional setting in the decentralized economy and comparing the conditions characterizing the market equilibrium with the conditions for Pareto optimality. The utility function of household h is given in (A5.121) and the budget constraint is:

$$\sum_{g=1}^{2} P_g X_g^h + T^h = \sum_{f=1}^{2} W_f V_f^h, \tag{A5.147}$$

where P_g is the market (consumer) price of good g, W_f is the market price of factor f, and T^h is the lump-sum tax levied on household h. The representative competitive firm g faces technology (A5.122) and maximizes profit Π_g (taking P_g and W_f as given):

$$\Pi_g \equiv P_g \left(1 - t_g\right) Y_g - \sum_{f=1}^{2} W_f Z_f^g, \tag{A5.148}$$

where t_g is an ad valorem output tax on good g. Prices and wages are flexible and all markets are assumed to clear.

The (decentralized) optimizing decisions by households and firms give rise to the following first-order conditions:

$$\frac{P_1}{P_2} = \frac{\partial U^h / \partial X_1^h}{\partial U^h / \partial X_2^h}, \tag{A5.149}$$

$$\frac{W_1}{W_2} = \frac{\partial U^h / \partial V_1^h}{\partial U^h / \partial V_2^h}, \tag{A5.150}$$

$$\frac{W_f}{P_g} = -\frac{\partial U^h / \partial V_f^h}{\partial U^h / \partial X_g^h}, \tag{A5.151}$$

$$\frac{W_f}{W_f} = \frac{P_1 (1 - t_1)}{P_2 (1 - t_2)} \frac{\partial F^1 / \partial Z_f^1}{\partial F^2 / \partial Z_f^2}, \tag{A5.152}$$

$$\frac{W_1}{W_2} = \frac{\partial F^1 / \partial Z_1^1}{\partial F^1 / \partial Z_2^1} = \frac{\partial F^2 / \partial Z_1^2}{\partial F^2 / \partial Z_2^2}, \tag{A5.153}$$

$$\left[\frac{P_1}{P_2} =\right] \quad \frac{\partial U^h / \partial X_1^h}{\partial U^h / \partial X_2^h} = \frac{1 - t_2}{1 - t_1} \frac{\partial F^2 / \partial Z_f^2}{\partial F^1 / \partial Z_f^1}, \tag{A5.154}$$

$$\left[\frac{W_1}{W_2} =\right] \quad \frac{\partial U^h / \partial V_1^h}{\partial U^h / \partial V_2^h} = \frac{\partial F^g / \partial Z_1^g}{\partial F^g / \partial Z_2^g}, \tag{A5.155}$$

$$\left[\frac{W_f}{P_1} =\right] \quad -\frac{\partial U^h / \partial V_f^h}{\partial U^h / \partial X_1^h} = (1 - t_1) \frac{\partial F^1}{\partial Z_f^1}, \tag{A5.156}$$

$$\left[\frac{W_f}{P_2} =\right] \quad -\frac{\partial U^h / \partial V_f^h}{\partial U^h / \partial X_2^h} = (1 - t_2) \frac{\partial F^2}{\partial Z_f^2}. \tag{A5.157}$$

We observe that (A5.149)-(A5.157) exactly match the Pareto optimality conditions if and only if the output taxes are set according to:

$$t_1 = -\sum_{h=1}^{H} \frac{\partial U^h / \partial Y_1}{\partial U^h / \partial X_1^h} > 0, \tag{A5.158}$$

$$t_2 = 0. \tag{A5.159}$$

(The reader is invited to verify that $\nu_g = P_g$, $-\xi_f = W_f$, in combination with (A5.158)-(A5.159) indeed gives an exact matching of the two sets of first-order conditions.) We close this subsection with a number of remarks. First, the expression for t_1 in (A5.158) is called the *Pigouvian Tax* (in honour of the great Cambridge economist Arthur C. Pigou who lived from 1877-1959). This tax restores Pareto efficiency by internalizing the damaging effects of the pollution. The Pigouvian tax is equal to the *sum* of the marginal rates of substitution between the polluting good Y_1 in its role of public "bad" and that same good in its role of private good (X_1^h). Second, since good 2 does not cause an external effect there is no reason to tax it–see equation (A5.159). Third, it cannot be overstressed that the decentralization only works if there are lump-sum taxes at the disposal of the policy maker. If this is not the case, then there is an interaction between distorting taxation for revenue raising (and possibly redistribution) purposes and the correction of externalities. This issue is studied in detail by Sandmo (1975).

Key literature

- Atkinson & Stiglitz (1980, lecture 16), Jha (1998, ch. 5), or Myles (1995, ch. 9) on theory.
- Laffont (1987) and Oakland (1987) on theory.
- Classics: Meade (1952), Samuelson (1954, 1955, 1958, 1969), Tiebout (1956).
- In folder of literature to be mentioned: Atkinson and Stern (1974), Sandmo (1975), Cornes and Sandler (2000), Cremer et al. (1998), Wilson (1991a, 1991b), Munro (1991), Milleron (1972), Brito et al. (1991), Diamond (1973), Varian (1994a, 1994b), Arrow (1971a), Berstrom et al. (1986), Sandmo (1972, 1998).
- Club goods: Buchanan (1965).
- Public inputs: Tawada (1980), Manning et al. (1980), Feehan (1998).
- Externalities: Chipman (1970), Buchanan and Stubblebine (1962).

Part III

Selected topics

Chapter 12

Intergenerational economics

The purpose of this chapter is to discuss the following topics:

- Definition of pension systems: PAYG, fully funded.
- Reform of the pension system.
- Ageing.
- Labour supply effects.
- Empirics.

12.1 The Diamond-Samuelson model

In this chapter we discuss a number of issues in intergenerational economics. Central in the discussion is a simple model that was formulated by Diamond (1965) using the earlier insights of Samuelson (1958).[1] This model captures both the finite-horizon and life-cycle aspects of household behaviour. The Diamond-Samuelson model is formulated in discrete time and has been *the* workhorse model in various fields of economics for almost four decades. In the remainder of this section we describe (a simplified version of) the Diamond (1965) model in detail.

12.1.1 Households

Individual agents live for two periods. During the first period (their "youth") they work and in their second period (their "old age") they are retired from the labour force. Since they want to consume in both periods, agents save during youth and dissave during old age. We abstract from bequests and assume that the population grows at a constant rate n.

[1] An even earlier overlapping-generations model was developed by Allais (1947). Unfortunately, due to the non-trivial language barrier, it was not assimilated into the Anglo-Saxon literature.

A representative young agent at time t has the following CES lifetime utility function:[2]

$$\Lambda_t^Y = U\left(C_t^Y, C_{t+1}^O\right) \equiv \left[\varepsilon\left(C_t^Y\right)^{1-1/\sigma} + (1-\varepsilon)\left(C_{t+1}^O\right)^{1-1/\sigma}\right]^{1/(1-1/\sigma)}, \tag{A5.1}$$

where σ is the substitution elasticity between current and future consumption ($\sigma > 0$), and $0.5 < \varepsilon < 1$ captures the notion of pure time preference (see below).[3] In (A5.1) the subscript identifies the time period and the superscript the period of life the agent is in, with "Y" and "O" standing for, respectively, youth and old age. Hence, C_t^Y and C_{t+1}^O denote consumption by an agent born in period t during youth and old age, respectively, and Λ_t^Y is lifetime utility of a young agent from the perspective of his birth.

During the first period the agent inelastically supplies one unit of labour and receives a wage W_t which is spent on consumption, C_t^Y, and savings, S_t. In the second period, the agent does not work but receives interest income on his savings, $r_{t+1}S_t$. Principal plus interest are spent on consumption during old age, C_{t+1}^O. The household thus faces the following budget identities:

$$C_t^Y + S_t = W_t, \tag{A5.2}$$

$$C_{t+1}^O = (1 + r_{t+1})S_t. \tag{A5.3}$$

By substituting (A5.3) into (A5.2) we obtain the consolidated (or lifetime) budget constraint:

$$W_t = C_t^Y + \frac{C_{t+1}^O}{1 + r_{t+1}}. \tag{A5.4}$$

The young agent chooses C_t^Y and C_{t+1}^O to maximize (A5.1) subject to (A5.4). The first-order conditions for consumption in the two periods can be combined after which we obtain the familiar consumption Euler equation:

$$\frac{\partial U(C^Y, C^O)/\partial C^Y}{\partial U(C^Y, C^O)/\partial C^O} = 1 + r_{t+1} \quad \Leftrightarrow \quad \frac{C_{t+1}^O}{C_t^Y} = \left(\frac{1 + r_{t+1}}{1 + \rho}\right)^{\sigma}, \tag{A5.5}$$

where $\rho \equiv \varepsilon/(1-\varepsilon) - 1 > 0$ is the pure rate of time preference. By combining (A5.4) and the second expression in (A5.5), the following solutions for C_t^Y and C_{t+1}^O (and thus S_t) are obtained:

$$C_t^Y = [1 - s(r_{t+1})] W_t, \tag{A5.6}$$

$$\frac{C_{t+1}^O}{1 + r_{t+1}} = S_t = s(r_{t+1}) W_t, \tag{A5.7}$$

[2] The two-period consumption-saving model was studied in a partial equilibrium setting in Chapter 3 above. In this chapter, general equilibrium repercussions are taken into account.

[3] Recall that, for the special case with $\sigma = 1$, equation (A5.1) reduces to the Cobb-Douglas form: $\Lambda_t^Y \equiv (C_t^Y)^{\varepsilon} (C_{t+1}^O)^{1-\varepsilon}$.

where the savings propensity is defined as:

$$s\left(r_{t+1}\right) \equiv \frac{\left(1+r_{t+1}\right)^{\sigma-1}}{(1+\rho)^{\sigma}+\left(1+r_{t+1}\right)^{\sigma-1}}. \tag{A5.8}$$

With homothetic preferences, present and future consumption are both normal goods, i.e. $\partial C_t^Y/\partial W_t = 1 - s\left(r_{t+1}\right) > 0$ and $\partial C_{t+1}^O/\partial W_t = (1+r_{t+1})s\left(r_{t+1}\right) > 0$. The response of savings with respect to the interest rate is ambiguous as the income and substitution effects work in opposite directions (see Chapter 3 for a detailed discussion). On the one hand an increase in r_{t+1} reduces the relative price of future goods which prompts the agent to substitute future for present consumption and to increase savings. On the other hand, the increase in r_{t+1} expands the budget available for present and future consumption which prompts the agent to increase both present and future consumption and to decrease savings.[4] Equation (A5.8) shows that the substitution (income) effect dominates and the savings rate depends positively (negatively) on the interest rate if the substitution elasticity exceeds (falls short of) unity:

$$s_r\left(k_{t+1}\right) \equiv \frac{\partial s\left(r_{t+1}\right)}{\partial r_{t+1}} = \frac{(\sigma-1)\left(1-s\left(r_{t+1}\right)\right)\left(1+r_{t+1}\right)^{\sigma-2}}{(1+\rho)^{\sigma}+\left(1+r_{t+1}\right)^{\sigma-1}} \gtreqless 0 \quad \Leftrightarrow \quad \sigma \gtreqless 1. \tag{A5.9}$$

12.1.2 Firms

The perfectly competitive firm sector produces output, Y_t, by hiring capital, K_t, from the currently old agents, and labour, L_t, from the currently young agents. The production function is linearly homogeneous:

$$Y_t = F(K_t, L_t), \tag{A5.10}$$

and profit maximization ensures that the production factors receive their respective marginal physical products (and that pure profits are zero):

$$W_t = F_L(K_t, L_t), \tag{A5.11}$$

$$r_t + \delta = F_K(K_t, L_t), \tag{A5.12}$$

where $0 < \delta < 1$ is the depreciation rate of capital. The crucial thing to note about (A5.12) concerns the timing. Capital that was accumulated by the currently old, K_t, commands the rental rate $r_t + \delta$. It follows that the rate of interest upon which the currently young agents base their savings decisions (i.e.

[4] See Figure 3.1 in Chapter 3 for an illustration of the income and substitution effects associated with an interest rate change. Note that there is no human wealth effect because the agent does not work during old age.

r_{t+1}) depends on the *future* aggregate capital stock and labour force:

$$r_{t+1} + \delta = F_K(K_{t+1}, L_{t+1}). \tag{A5.13}$$

Since the labour force grows at a constant rate and we ultimately wish to study an economy which possesses a well-defined steady-state equilibrium, it is useful to rewrite (A5.10)-(A5.11) and (A5.13) in per capita form (see Chapter 8 for details):

$$y_t = f(k_t), \tag{A5.14}$$

$$W_t = f(k_t) - k_t f'(k_t), \tag{A5.15}$$

$$r_{t+1} + \delta = f'(k_{t+1}), \tag{A5.16}$$

where $y_t \equiv Y_t/L_t$, $k_t \equiv K_t/L_t$, and $f(k_t) \equiv F(k_t, 1)$.

12.1.3 Market equilibrium

The resource constraint for the economy as a whole can be written as follows:

$$Y_t + (1-\delta)K_t = K_{t+1} + C_t, \tag{A5.17}$$

where C_t represents aggregate consumption in period t. Equation (A5.17) says that output plus the undepreciated part of the capital stock (left-hand side) can be either consumed or carried over to the next period in the form of capital (right-hand side). Alternatively, (A5.17) can be written as $Y_t = C_t + I_t$ with $I_t \equiv \Delta K_{t+1} + \delta K_t$ representing gross investment.

Aggregate consumption is the sum of consumption by the young and the old agents in period t:

$$C_t \equiv L_{t-1}C_t^O + L_t C_t^Y. \tag{A5.18}$$

Since the old, as a group, own the capital stock, their total consumption in period t is the sum of the undepreciated part of the capital stock plus the rental payments received from the firms, i.e. $L_{t-1}C_t^O = (r_t + \delta)K_t + (1-\delta)K_t$. For each young agent consumption satisfies (A5.2) so that total consumption by the young amounts to: $L_t C_t^Y = W_t L_t - S_t L_t$. By substituting these two results into (A5.18), we obtain:

$$\begin{aligned} C_t &= (r_t + \delta)K_t + (1-\delta)K_t + W_t L_t - S_t L_t \\ &= Y_t + (1-\delta)K_t - S_t L_t, \end{aligned} \tag{A5.19}$$

where we have used the fact that $Y_t = (r_t + \delta)K_t + W_t L_t$ in going from the first to the second line. Output is fully exhausted by factor payments and pure profits are zero.

Finally, by combining (A5.17) and (A5.19) we obtain the expression linking this period's savings decisions by the young to next period's capital stock:

$$S_t L_t = K_{t+1}. \tag{A5.20}$$

The population is assumed to grow at a constant rate,

$$L_t = L_0(1+n)^t, \quad n > -1, \tag{A5.21}$$

so that (A5.20), in combination with (A5.7), can be rewritten in per capita form as:

$$s(r_{t+1})W_t = (1+n)k_{t+1}. \tag{A5.22}$$

The capital market is represented by the demand for capital by entrepreneurs (equation (A5.16)) and the supply of capital by households (equation (A5.22)).

12.1.4 Dynamics and stability

The dynamical behaviour of the economy can be studied by substituting the expressions for W_t and r_{t+1} (given in, respectively, (A5.15) and (A5.16)) into the capital supply equation (A5.22):

$$(1+n)k_{t+1} = s\left(f'(k_{t+1}) - \delta\right)\left[f(k_t) - k_t f'(k_t)\right]. \tag{A5.23}$$

This expression represents an implicit relationship between the present and future capital stocks per worker. It is thus suitable to study the stability of the model. By totally differentiating (A5.23) we obtain:

$$\frac{dk_{t+1}}{dk_t} = \frac{-s\left(r_{t+1}\right)k_t f''(k_t)}{1+n-s_r\left(k_{t+1}\right)W_t f''(k_{t+1})}, \tag{A5.24}$$

where $s\left(r_{t+1}\right)$, $s_r\left(k_{t+1}\right)$, and W_t are defined in, respectively, (A5.8), (A5.9), and (A5.15). A steady state (if it exists) is *locally stable* if and only if $|dk_{t+1}/dk_t| < 1$. It is clear from (A5.24) that clear-cut results are hard to come by in the most general version of our model. Although we know that the numerator of (A5.24) is positive (because $s\left(r_{t+1}\right) > 0$ and $f''\left(k_t\right) < 0$), the sign of the denominator is indeterminate (because s_r is ambiguous).

Referring the interested reader to Galor and Ryder (1989) for a rigorous analysis of the most general case, we take the practical way out by illustrating the existence and stability issues with the *unit-elastic* model. Specifically, we assume that technology is Cobb-Douglas, so that $y_t = k_t^{1-\epsilon_L}$, and that the utility function is logarithmic ($\sigma = 1$), so that $s\left(r_{t+1}\right) = 1/(2+\rho)$, the wage rate is $W_t = \epsilon_L k_t^{1-\epsilon_L}$, and (A5.23)

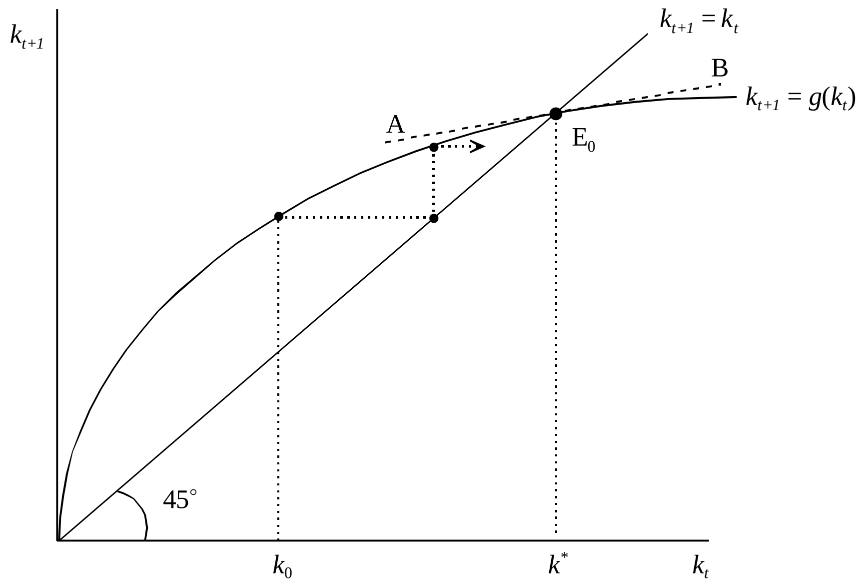

Figure 12.1: The unit-elastic Diamond-Samuelson model

becomes:

$$k_{t+1} = g(k_t) \equiv \frac{\epsilon_L}{(1+n)(2+\rho)} k_t^{1-\epsilon_L}. \tag{A5.25}$$

Equation (A5.25) has been drawn in Figure 12.1. Since $\lim_{k \to 0} g'(k) = \infty$ and $\lim_{k \to \infty} g'(k) = 0$, the steady state, satisfying $k^* = g(k^*)$, is unique and stable. The diagram illustrates one stable trajectory from k_0. The tangent of $g(\cdot)$ passing through the steady-state equilibrium point E_0 is the dashed line AB. It follows from the diagram (and indeed from (A5.25)) that the unit-elastic Diamond-Samuelson model satisfies the stability condition with a positive slope for $g(\cdot)$, i.e. $0 < g'(k^*) < 1$.

12.1.5 Efficiency

It is clear from the discussion surrounding Figure 12.1 that there is a perfectly reasonable setting in which the Diamond-Samuelson model possesses a stable and unique steady-state equilibrium. We now assume for convenience that our most general model also has this property and proceed to study its welfare properties. To keep things simple, and to prepare for the discussion of taxations and social security issues below, we restrict attention to a steady-state analysis. Indeed, following Diamond (1965) we compare the market solution to the *optimal steady state*.

In the steady state, the capital-labour ratio is constant over time, i.e. $k_{t+1} = k_t = k$ and K_t, L_t, and Y_t all grow at the rate of population growth (n). Such a balanced growth path is called *optimal* if (i) each individual agent has the highest possible utility, and (ii) all agents have the same utility level (Diamond, 1965, p. 1128). Formally, the optimal balanced growth path maximizes the lifetime utility of a "representative" individual,

$$\Lambda^Y \equiv U(C^Y, C^O), \tag{A5.26}$$

subject to the economy-wide steady-state resource constraint:

$$f(k) - (n+\delta)k = C^Y + \frac{C^O}{1+n}. \tag{A5.27}$$

Note that we have dropped the time subscripts in (A5.26)-(A5.27) in order to stress the fact that we are looking at a steady-state situation only.[5] An important thing to note about this formulation is the following. In (A5.26) C^Y and C^O refer, respectively, to consumption during youth and retirement *of a particular individual*. In contrast, in (A5.27) C^Y and C^O refer to consumption levels of young and old agents, respectively, at a *particular moment in time*. This does, of course, not mean that we are comparing apples and oranges–for the purposes of selecting an optimal balanced growth path we can ignore these differences because all individuals are treated symmetrically.

The first-order conditions for the optimal golden-age path consist of the steady-state resource constraint (A5.27) and:

$$\frac{\partial U(C^Y, C^O)/\partial C^Y}{\partial U(C^Y, C^O)/\partial C^O} = 1+n, \tag{A5.28}$$

$$f'(k) = n+\delta. \tag{A5.29}$$

Samuelson (1968a) calls these conditions, respectively, the biological-interest-rate consumption golden rule and the production golden rule. Comparing (A5.28)-(A5.29) with their respective market counterparts (A5.5) and (A5.16) reveals that they coincide if the market rate of interest equals the rate of population growth:

$$r = f'(k) - \delta = n \qquad \text{(golden rule)}$$

As is stressed by Samuelson (1968a, p. 87) the two conditions (A5.28)-(A5.29) are analytically independent: even if k is held constant at some suboptimal level, so that production is inefficient as $f'(k) \neq n+\delta$, the optimum consumption pattern must still satisfy (A5.28). Similarly, if the division of output among generations is suboptimal (e.g. due to a badly designed pension system), condition (A5.28) no longer holds but the optimal k still follows from the production golden rule (A5.29).

If the steady-state interest rate is less than the rate of population growth ($r < n$) then there is overaccumulation of capital, k is too high, and the economy is *dynamically inefficient*. A quick inspection of the unit-elastic model reveals that such a situation is theoretically quite possible for reasonable parameter

[5]The steady-state resource constraint (A5.27) is obtained as follows. First, (A5.18) is substituted in (A5.17) and the resulting expression is divided by L_t. Then (A5.14) is inserted, the steady state is imposed ($k_{t+1} = k_t = k$), and all time indexes are dropped.

values. Indeed, by computing the steady-state capital-labour ratio from (A5.25) and using the result in (A5.16) we find that the steady-state interest rate for the unit-elastic model is:

$$r = \frac{(1-\epsilon_L)(2+\rho)(1+n)}{\epsilon_L} - \delta. \tag{A5.30}$$

Blanchard and Fischer (1989, p. 147) suggest the following numbers. Each period of life is 30 years and the labour share is $\epsilon_L = 3/4$. Population grows at 1% per annum so $n = 1.01^{30} - 1 = 0.348$. Capital depreciates at 5% per annum so $\delta = 1 - (0.95)^{30} = 0.785$. With relatively impatient agents, the pure discount rate is 3% percent per annum, so $\rho = (1.03)^{30} - 1 = 1.427$ and (A5.30) shows that $r = 0.754$ which exceeds n by quite a margin. With more patient agents, whose pure discount rate is 1% percent per annum, $\rho = (1.01)^{30} - 1 = 0.348$ and $r = 0.269$ which is less than n.

Although dynamic inefficiency cannot be ruled out *a priori*, empirical studies of the issue typically find that actual economies are not likely to suffer from this oversaving phenomenon. See, for example, the study by Abel et al. (1989) for the United States. Unless stated otherwise we will therefore focus throughout this chapter on the case in which the net interest rate $(r - n)$ is positive.

12.2 Applications of the basic model

In this section we show how the standard Diamond-Samuelson model can be used to study the macroeconomic and welfare effects of old-age pensions. A system of social security was introduced in Germany during the 1880s by Otto von Bismarck, purportedly to stop the increasingly radical working class from overthrowing his conservative regime. It did not help poor Otto–he was forced to resign from office in 1890–but the system he helped create stayed. Especially following the Second World War, most developed countries have similarly adopted social security systems. Typically such a system provides benefit payments to the elderly which continue until the recipient dies.

In the first subsection we show how the method of financing old-age pensions critically determines the effects of such pensions on resource allocation and welfare. In the second subsection we study the effects of a demographic shock, such as an ageing population, on the macroeconomy.

12.2.1 Pensions

In order to study the effects of public pensions we must introduce the government into the Diamond-Samuelson model. Assume that, at time t, the government provides lump-sum transfers, Z_t, to old agents and levies lump-sum taxes, T_t, on the young. It follows that the budget identities of a young household at time t are changed from (A5.2)-(A5.3) to:

$$C_t^Y + S_t = W_t - T_t, \tag{A5.31}$$

$$C_{t+1}^O = (1 + r_{t+1})S_t + Z_{t+1}, \tag{A5.32}$$

so that the consolidated lifetime budget constraint of such a household is now:

$$W_t - T_t + \frac{Z_{t+1}}{1+r_{t+1}} = C_t^Y + \frac{C_{t+1}^O}{1+r_{t+1}}. \tag{A5.33}$$

The left-hand side of (A5.33) shows that lifetime wealth consists of after-tax wages during youth plus the present value of pension receipts during old age.

Depending on the way in which the government finances its transfer scheme, we can distinguish two prototypical social security schemes. In a *fully funded* system the government invests the contributions of the young and returns them with interest in the next period in the form of transfers to the then old agents. In such a system we have:

$$Z_{t+1} = (1+r_{t+1})T_t. \tag{A5.34}$$

In contrast, in an unfunded or *pay-as-you-go* (PAYG) system, the transfers to the old are covered by the taxes of the young *in the same period*. Since, at time t, there are L_{t-1} old agents (each receiving Z_t in transfers) and L_t young agents (each paying T_t in taxes) a PAYG system satisfies $L_{t-1}Z_t = L_t T_t$ which can be rewritten by noting (A5.21) as:

$$Z_t = (1+n)T_t. \tag{A5.35}$$

12.2.1.1 Fully funded pensions

A striking property of a fully funded social security system is its neutrality. With this we mean that an economy with a fully funded system is identical in all relevant aspects to an economy without such a system. This important neutrality result can be demonstrated as follows.

First, we note that, by substituting (A5.34) into (A5.33), the fiscal variables, T_t and Z_{t+1}, disappear from the lifetime budget constraint of the household. Consequently, these variables also do not affect the household's optimal life-cycle consumption plan, i.e. C_t^Y and C_{t+1}^O are exactly as in the pension-less economy described in section 12.1.1 above. It follows, by a comparison of (A5.2) and (A5.31), that with a fully funded pension system saving plus tax payments are set according to:

$$S_t + T_t = s(r_{t+1})W_t, \tag{A5.36}$$

where $s(r_{t+1})$ is the same function as the one appearing in (A5.8).

As a second preliminary step we must derive an expression linking savings of the young to next period's stock of productive capital. The key aspect of a fully funded system is that the government puts the tax receipts from the young to productive use by renting them out in the form of capital goods to firms. Hence, the economy-wide capital stock, K_t, is:

$$K_t = K_t^H + K_t^G, \tag{A5.37}$$

where K_t^H and $K_t^G \equiv L_{t-1}T_{t-1}$ denote capital owned by households and the government, respectively. The economy-wide resource constraint is still as given in (A5.17) but the expression for total consumption is changed from (A5.19) to:[6]

$$C_t = Y_t + (1-\delta)K_t - L_t(S_t + T_t). \tag{A5.38}$$

Finally, by using (A5.17), (A5.36), and (A5.38) we find that the capital market equilibrium condition is identical to (A5.22). Since the factor prices, (A5.15)-(A5.16), are also unaffected by the existence of the social security system, economies with and without such a system are essentially the same. Intuitively, with a fully funded system the household knows that its contributions, T_t, attract the same rate of return as its own private savings, S_t. As a result, the household only worries about its total saving, $S_t + T_t$, and does not care that some of this saving is actually carried out on its behalf by the government.[7]

12.2.1.2 Pay-as-you-go pensions

Under a PAYG system there is a transfer from young to old in each period according to (A5.35). Assuming that the contribution rate per person is held constant over time (so that $T_{t+1} = T_t = T$), (A5.35) implies that $Z_{t+1} = (1+n)T$ so that consolidation of (A5.31)-(A5.32) yields the following lifetime budget constraint of a young household:

$$\hat{W}_t \equiv W_t - \frac{r_{t+1} - n}{1 + r_{t+1}} T = C_t^Y + \frac{C_{t+1}^O}{1 + r_{t+1}}. \tag{A5.39}$$

This expression is useful because it shows that, ceteris paribus the factor prices, the existence of a PAYG system contracts (expands) the consumption possibility frontier for young agents if the interest rate exceeds (falls short of) the growth rate of the population. Put differently, if $r_{t+1} > n$ ($r_{t+1} < n$) the contribution rate is seen as a lump-sum tax (subsidy) by the young household.

[6]Equation (A5.38) is derived as follows. Consumption by the old agents is $L_{t-1}C_t^O = (r_t + \delta)K_t^H + (1-\delta)K_t^H + L_{t-1}Z_t$. For young agents we have $L_t C_t^Y = L_t [W_t - S_t - T_t]$ so that aggregate consumption is:

$$\begin{aligned} C_t &= (r_t + \delta)K_t^H + (1-\delta)K_t^H + L_{t-1}Z_t + L_t [W_t - S_t - T_t] \\ &= Y_t + (1-\delta)K_t^H - (r_t + \delta)K_t^G + L_{t-1}Z_t - L_t(S_t + T_t) \\ &= Y_t + (1-\delta)K_t - L_t(S_t + T_t) + \left[L_{t-1}Z_t - (1+r_t)K_t^G\right]. \end{aligned}$$

This final expression collapses to (A5.38) because the term in square brackets on the right-hand side vanishes:

$$L_{t-1}Z_t - (1+r_t)K_t^G = L_{t-1}[Z_t - (1+r_t)T_{t-1}] = 0.$$

[7]An important proviso for the neutrality result to hold is that the social security system should not be too severe, i.e. it should not force the household to save more than it would in the absence of social security. In terms of the model we must have that $T_t < (1+n)k_{t+1}$ (see Blanchard and Fischer, 1989, p. 111).

The household maximizes lifetime utility (A5.1) subject to its lifetime budget constraint (A5.39). Current and future consumption are set according to:

$$C_t^Y = [1 - s(r_{t+1})]\,\hat{W}_t, \tag{A5.40}$$

$$\frac{C_{t+1}^O}{1+r_{t+1}} = s(r_{t+1})\,\hat{W}_t, \tag{A5.41}$$

where $s(r_{t+1})$ is defined in (A5.8) above. By using (A5.32), (A5.39), and (A5.41) we find that the savings function can be written as follows:

$$\begin{aligned} S_t &= s(r_{t+1})\,\hat{W}_t - \frac{1+n}{1+r_{t+1}}T \\ &= s(r_{t+1})\,W_t - \frac{s(r_{t+1})(1+r_{t+1}) + [1-s(r_{t+1})](1+n)}{1+r_{t+1}}T. \end{aligned} \tag{A5.42}$$

Since T, $1+r_{t+1}$, and $1+n$ are all positive and the savings rate satisfies $0 < s(r_{t+1}) < 1$, it follows that the term in square brackets on the right-hand side of (A5.42) is positive. To keep matters as simple as possible we now restrict attention to the simple unit-elastic model for which utility is logarithmic (and technology is Cobb-Douglas). In that case, the savings rate is constant ($s(r_{t+1}) = 1/(2+\rho)$) and the savings function simplifies to:

$$S_t = \frac{W_t}{2+\rho} - \frac{1+r_{t+1}+(1+\rho)(1+n)}{1+r_{t+1}}\frac{T}{2+\rho} \equiv S(W_t, r_{t+1}, T). \tag{A5.43}$$

It is easy to verify that the partial derivatives of the savings function satisfy $0 < S_W < 1$, $S_r > 0$, $-1 < S_T < 0$ (if $r_{t+1} > n$), and $S_T < -1$ (if $r_{t+1} < n$).

Since the PAYG pension is a pure transfer from co-existing young to old generations it does not itself lead to the formation of capital in the economy. Since only private saving augments the capital stock, equation (A5.20) is still relevant.[8] By combining (A5.20) with (A5.43) we obtain the expression linking the future capital stock to current saving plans:

$$S(W_t, r_{t+1}, T) = (1+n)k_{t+1}. \tag{A5.44}$$

With Cobb-Douglas technology ($y_t \equiv k_t^{1-\epsilon_L}$) equations (A5.15) and (A5.16) reduce to, respectively, $W_t \equiv W(k_t) = \epsilon_L k_t^{1-\epsilon_L}$ and $r_{t+1} \equiv r(k_{t+1}) = (1-\epsilon_L)k_{t+1}^{-\epsilon_L} - \delta$. By using these expressions in (A5.44) we obtain the fundamental difference equation (in implicit form) characterizing the economy under a

[8]Consumption by the old agents is $L_{t-1}C_t^O = (r_t+\delta)K_t + (1-\delta)K_t + L_{t-1}Z_t$. For young agents we have $L_tC_t^Y = L_t[W_t - S_t - T_t]$ so that aggregate consumption is:

$$\begin{aligned} C_t &= (r_t+\delta)K_t + (1-\delta)K_t + L_{t-1}Z_t + L_t[W_t - S_t - T_t] \\ &= Y_t + (1-\delta)K_t + [L_{t-1}Z_t - L_tT_t] - L_tS_t. \end{aligned}$$

This final expression collapses to (A5.19) because the term in square brackets on the right-hand side vanishes under the PAYG scheme. Combining (A5.17) and (A5.19) yields (A5.20).

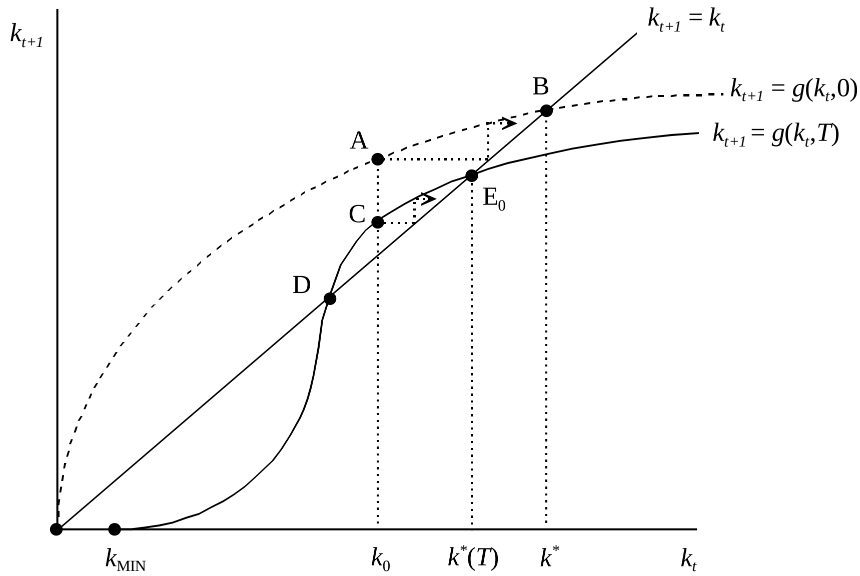

Figure 12.2: PAYG pensions in the unit-elastic model

PAYG system, $k_{t+1} = g(k_t, T)$. The partial derivatives of this function are:

$$g_k \equiv \frac{\partial g}{\partial k_t} = \frac{S_W W'(k_t)}{1 + n - S_r r'(k_{t+1})} > 0, \tag{A5.45}$$

$$g_T \equiv \frac{\partial g}{\partial T} = \frac{S_T}{1 + n - S_r r'(k_{t+1})} < 0, \tag{A5.46}$$

where S_W and S_r are obtained from (A5.43). We illustrate the fundamental difference equation in Figure 12.2.[9]

In Figure 12.2, the dashed line, labelled "$k_{t+1} = g(k_t, 0)$" characterizes the standard unit-elastic Diamond-Samuelson model without social security, i.e. it reproduces Figure 12.1 and point B is the steady state to which the economy converges in the absence of social security. Suppose now that the PAYG system is introduced at time $t = 0$ when the economy has an initial (non-steady-state) capital-labour ratio of k_0. Members of the old generation at time $t = 0$ cannot believe their luck. They have not contributed anything to the PAYG system but nevertheless receive a pension of $Z = (1 + n)T$ (see equation (A5.35)). Since the old do not save this *windfall gain* is spent entirely on additional consumption.

[9]The fundamental difference equation can be written as:

$$(1+n)k_{t+1} = \frac{W(k_t) - T}{2+\rho} - \frac{1+\rho}{2+\rho}\frac{(1+n)T}{1+r(k_{t+1})}.$$

The second term on the right-hand side vanishes as $k_{t+1} \to 0$ (since $r(k_{t+1}) \to +\infty$ in that case). Hence, $W(k_{\text{MIN}}) = T$. For $k_t < k_{\text{MIN}}$ the wage rate is too low ($W(k_t) < T$) and the PAYG scheme is not feasible. By differentiating the fundamental difference equation we obtain:

$$\frac{dk_{t+1}}{dk_t} = \frac{W'(k_t)}{(1+n)\left[2+\rho+(1+\rho)T\psi(k_{t+1})\right]} \geq 0, \qquad \psi(k_{t+1}) \equiv \frac{-r'(k_{t+1})}{\left[1+r(k_{t+1})\right]^2}.$$

It is straightforward to show that $\psi(k_{t+1}) \to +\infty$ for $k_{t+1} \to 0$, $\psi(k_{t+1}) \to 0$ for $k_{t+1} \to \infty$, $W'(k_t) \to 0$ for $k_t \to \infty$, and $W'(k_{\text{MIN}}) > 0$. It follows that $g(k_t, T)$ is horizontal in $k_t = k_{\text{MIN}}$, is upward sloping for larger values of k_t, and becomes horizontal as k_t gets very large. Provided T is not too large there exist two intersections with the $k_{t+1} = k_t$ line.

Consumption by each old household at time $t = 0$ is now:

$$C_0^O = (1+n)\left[(1+r(k_0))k_0 + T\right], \tag{A5.47}$$

and, since k_0 is predetermined, so is the interest rate and $dC_0^O/dT = 1+n$.

In contrast, members of the young generation at time $t = 0$ are affected by the introduction of the PAYG system in a number of different ways. On the one hand, they must pay T in the current period in exchange for which they receive a pension $(1+n)T$ in the next period. Since the wage rate at time $t = 0$, $W(k_0)$, is predetermined, the net effect of these two transactions is to change the value of lifetime resources ($\hat{W}_0$ defined in (A5.39)) according to:

$$\frac{\partial \hat{W}_0}{\partial T} = -\frac{r(k_1) - n}{1 + r(k_1)} \lesseqgtr 0, \tag{A5.48}$$

where the sign is ambiguous because $r(k_1)$ may exceed or fall short of the population growth rate n. (Of course, in the dynamically efficient case, we have that $\partial \hat{W}_0/\partial T > 0$.) Note that equation (A5.48) only shows a partial effect because the interest rate depends on the capital stock in the next period (k_1), which is itself determined by the savings behaviour of the young in period $t = 0$. It follows from (A5.44) and (A5.46), however, that the total effect of the introduction of the PAYG system is to reduce saving by the young and thus to reduce next period's capital stock, i.e. $dk_1/dT = g_T < 0$. This adverse effect on the capital stock is represented in Figure 12.2 by the vertical difference between points A and C.

As a result of the policy shock, the economy now follows the convergent path from C to the ultimate steady state E_0. It follows from Figure 12.2 that k_t is less than it would have been without the PAYG pension, both during transition and in the new steady state (i.e. the path from C to E_0 lies below the path from A to B). Hence, since $W'(x) > 0$ and $r'(x) < 0$, the steady-state wage is lower and the interest rate is higher than it would have been. The long-run effect on the capital-labour ratio is obtained by using (A5.44) and imposing the steady state ($k_{t+1} = k_t$):

$$\frac{dk}{dT} = \frac{g_T}{1 - g_k} < 0, \tag{A5.49}$$

where $0 < g_k < 1$ follows from the stability condition.

The upshot of the discussion so far is that, unlike a fully funded pension system, a PAYG system is not neutral but leads to crowding out of capital, a lower wage rate, and a higher interest rate in the long run. Is that good or bad for households? To answer that question we now study the welfare effect on a *steady-state generation* of a change in the contribution rate, T. As in our discussion of dynamic efficiency above we thus continue to ignore transitional dynamics for the time being by only looking at the steady state.

To conduct the welfare analysis we need to utilize two helpful tools, i.e. the *indirect utility function* and the *factor price frontier*. The indirect utility function is defined in formal terms by:

$$\bar{\Lambda}^Y = V(W, r, T) \equiv \max_{\{C^Y, C^O\}} U(C^Y, C^O) \text{ subject to } \hat{W} = C^Y + \frac{C^O}{1+r}, \tag{A5.50}$$

where $U(C^Y, C^O)$ is the direct utility function (i.e. equation (A5.1)). The lack of subscripts indicates steady-state values and $\hat{W}$ represents lifetime household resources under the PAYG system:

$$\hat{W} = W - \frac{r-n}{1+r}T. \tag{A5.51}$$

For example, for the Cobb-Douglas utility function (employed regularly in this chapter) the indirect utility function takes the following form:

$$\bar{\Lambda}^Y = \omega_0 (1+r)^{1/(2+\rho)} \hat{W}, \tag{A5.52}$$

where $\omega_0 \equiv \varepsilon^{\varepsilon} (1-\varepsilon)^{1-\varepsilon}$ is a positive constant (and we have used the fact that $1 - \varepsilon = 1/(2+\rho)$).

The indirect utility function (A5.50) has a number of properties which will prove to be very useful below:[10]

$$\frac{\partial \bar{\Lambda}^Y}{\partial W} = \frac{\partial \Lambda^Y}{\partial C^Y} > 0, \tag{A5.53}$$

$$\frac{\partial \bar{\Lambda}^Y}{\partial r} = \frac{S}{1+r} \frac{\partial \Lambda^Y}{\partial C^Y} > 0, \tag{A5.54}$$

$$\frac{\partial \bar{\Lambda}^Y}{\partial T} = -\frac{r-n}{1+r} \frac{\partial \Lambda^Y}{\partial C^Y} \gtreqless 0. \tag{A5.55}$$

According to (A5.53)-(A5.54), steady-state welfare depends positively on both the wage rate and the interest rate. Since we saw above that the wage falls ($dW/dT = W'(k)dk/dT < 0$) but the interest rate rises ($dr/dT = r'(k)dk/dT > 0$) in the long run, the effects of factor prices on welfare work in opposite directions even in the absence of a PAYG system (if $T = 0$).

But both W and r depend on the capital-labour ratio (as in the standard neoclassical model) and are

[10] These properties are derived as follows. We start with the identity $V(W, r, T) \equiv U\left(C^Y(W, r, T), C^O(W, r, T)\right)$, where $C^i(W, r, T)$ are the optimal consumption levels during the two periods of life. By using this identity, partially differentiating (A5.1), and using (A5.5) we obtain:

$$\frac{\partial V}{\partial W} = \frac{\partial U}{\partial C^Y} \left[\frac{\partial C^Y}{\partial W} + \frac{1}{1+r} \frac{\partial C^O}{\partial W} \right].$$

It follows from the constraint in (A5.50) that the term in square brackets is equal to unity. Using the same steps we obtain for $\partial V / \partial r$:

$$\frac{\partial V}{\partial r} = \frac{\partial U}{\partial C^Y} \left[\frac{\partial C^Y}{\partial r} + \frac{1}{1+r} \frac{\partial C^O}{\partial r} \right] = \frac{\partial U}{\partial C^Y} \frac{C^O - (1+n)T}{(1+r)^2}.$$

Using $C^O - (1+n)T = (1+r)S$ we obtain (A5.54). Finally, we obtain for $\partial V / \partial T$:

$$\frac{\partial V}{\partial T} = \frac{\partial U}{\partial C^Y} \left[\frac{\partial C^Y}{\partial T} + \frac{1}{1+r} \frac{\partial C^O}{\partial T} \right] = -\frac{r-n}{1+r} \frac{\partial U}{\partial C^Y},$$

where the final result follows from the constraint in (A5.50).

thus not independent of each other. By exploiting this dependency we obtain the factor price frontier, $W_t = \phi(r_t)$, which has a very useful property:

$$W_t = \phi(r_t), \qquad \frac{dW_t}{dr_t} \equiv \phi'(r_t) = -k_t. \tag{A5.56}$$

The slope of the factor price frontier is obtained as follows. In general, by differentiating (A5.15) and (A5.16) (for r_t) we get $dr_t = f''(k_t)dk_t$ and $dW_t = -k_t f''(k_t)dk_t$ so that $dW_t/dr_t = -k_t$. From this it follows that $d^2W_t/dr_t^2 = -dk_t/dr_t = -1/f''(k_t)$.[11]

We now have all the necessary ingredients to perform our welfare analysis. By differentiating the indirect utility function with respect to T we obtain in a few steps:

$$\begin{aligned}\frac{d\bar{\Lambda}^Y}{dT} &= \frac{\partial\bar{\Lambda}^Y}{\partial W}\frac{dW}{dT} + \frac{\partial\bar{\Lambda}^Y}{\partial r}\frac{dr}{dT} + \frac{\partial\bar{\Lambda}^Y}{\partial T} \\ &= \frac{\partial\Lambda^Y}{\partial C^Y}\left[\frac{dW}{dT} + \frac{S}{1+r}\frac{dr}{dT} - \frac{r-n}{1+r}\right] \\ &= -\frac{r-n}{1+r}\frac{\partial\Lambda^Y}{\partial C^Y}\left[1 + k\frac{dr}{dT}\right] \propto \text{sgn}(n-r),\end{aligned} \tag{A5.57}$$

where we have used (A5.51) and (A5.53)-(A5.55) in going from the first to the second line and (A5.56) as well as $S = (1+n)k$ in going from the second to the third line. The term in square brackets on the right-hand side of (A5.57) shows the two channels by which the PAYG pension affects welfare. The first term is the partial equilibrium effect of T on lifetime resources and the second term captures the general equilibrium effects that operate via factor prices.

The expression in (A5.57) is important because it illustrates in a transparent fashion the intimate link that exists between, on the one hand, the steady-state welfare effect of a PAYG pension and, on the other hand, the dynamic (in)efficiency of the initial steady-state equilibrium. If the economy happens to be in the golden-rule equilibrium (so that $r = n$) then it follows from (A5.57) that a *marginal* change in the PAYG contribution rate has no effect on steady-state welfare (i.e. $d\bar{\Lambda}^Y/dT = 0$ in that case). Since the yield on private saving and the PAYG pension are the same in that case, a small change in T does not produce a first-order welfare effect on steady-state generations despite the fact that it causes crowding out of capital (see (A5.49)) and thus an increase in the interest rate (since $r'(k) < 0$).

Matters are different if the economy is initially not in the golden-rule equilibrium (so that $r \neq n$) because the capital crowding out does produce a first-order welfare effect in that case. For example, if the economy is initially dynamically inefficient ($r < n$), then an increase in the PAYG contribution rate actually raises steady-state welfare! The intuition behind this result, which was first demonstrated in the pensions context and with a partial equilibrium model by Aaron (1966), is as follows. In a dynamically

[11] The factor price frontier for the Cobb-Douglas technology is given by:

$$W = \epsilon_L \left(\frac{1-\epsilon_L}{r+\delta}\right)^{(1-\epsilon_L)/\epsilon_L},$$

where the reader should verify the property stated in (A5.56).

inefficient economy there is oversaving by the young generations as a result of which the market rate of interest is low. By raising T the young partially substitute private saving for saving via the PAYG pension. The latter has a higher yield than the former because the biological interest rate, n, exceeds the market interest rate, r. The reduction in the capital stock lowers the wage but this adverse effect on welfare is offset by the increase in the interest rate in a dynamically inefficient economy. To put it bluntly, capital crowding out is good in such an economy.

12.2.1.3 Equivalence PAYG and deficit financing government debt

As was shown by Auerbach and Kotlikoff, a PAYG social security scheme can also be reinterpreted as a particular kind of government debt policy (1987, pp. 149-150). In order to demonstrate this equivalency result, we now introduce government debt into the model. This model extension also allows us to further clarify the link between the pension insights of Aaron (1966) and the macroeconomic effects of debt as set out by Diamond (1965).

Assume that the government taxes the young generations, provides transfers to the old generations, and issues one-period (indexed) debt which yields the same rate of interest as capital. Ignoring government consumption, the government budget identity is now:

$$B_{t+1} - B_t = r_t B_t + L_{t-1} Z_t - L_t T_t, \tag{A5.58}$$

where B_t is the stock of public debt at the beginning of period t. Interest payments on existing debt ($r_t B_t$) plus transfers to the old are covered by the revenues from the tax on the young and/or additional debt issues ($B_{t+1} - B_t$).

Because government debt and private capital attract the same rate of return, the household is indifferent about the *composition* of its savings over these two assets. Consequently, the young choose consumption in the two periods and *total* saving in order to maximize lifetime utility (A5.1) subject to the budget identities (A5.31) and (A5.32). The savings function that results takes the following form:

$$S_t = S(\hat{W}_t, r_{t+1}), \tag{A5.59}$$

where $\hat{W}_t$ is given by the left-hand side of (A5.33) which is reproduced here for convenience:

$$\hat{W}_t = W_t - T_t + \frac{Z_{t+1}}{1 + r_{t+1}}. \tag{A5.60}$$

It remains to derive the expression linking private savings plans and aggregate capital formation. There are L_t young agents who each save S_t so that aggregate saving is $S_t L_t$. Saving can be in the form of private capital or public debt. Hence the capital market equilibrium condition is now:[12]

$$L_t S_t = B_{t+1} + K_{t+1}. \tag{A5.61}$$

We are now in the position to present an important equivalence result which was proved *inter alia* by Wallace (1981), Sargent (1987), and Calvo and Obstfeld (1988). Buiter and Kletzer state the equivalence result as follows: "...any equilibrium with government debt and deficits can be replicated by an economy in which the government budget is balanced period-by-period (and the stock of debt is zero) by appropriate age-specific lump-sum taxes and transfers" (1992, pp. 27-28). A corollary of the result is that if the policy maker has access to unrestricted age-specific taxes and transfers then public debt is redundant in the sense that it does not permit additional equilibria to be supported (1992, p. 28).

The model developed in this subsection is fully characterized (for $t \geq 0$) by the following equations:

$$C_t^O = (1 + r(k_t))(1+n)(k_t + b_t) + Z_t, \tag{A5.62}$$

$$\frac{C_{t+1}^O}{C_t^Y} = \left(\frac{1 + r\left(k_{t+1}\right)}{1+\rho} \right)^{\sigma}, \tag{A5.63}$$

$$W(k_t) - T_t - C_t^Y = (1+n)\left[k_{t+1} + b_{t+1}\right], \tag{A5.64}$$

$$(1+n) b_{t+1} = (1 + r(k_t)) b_t + \frac{Z_t}{1+n} - T_t, \tag{A5.65}$$

where $b_t \equiv B_t / L_t$ is per capita government debt and where k_0 and b_0 are both given. Equation (A5.62) is consumption of an old household, (A5.63) is the consumption Euler equation for a young household (see also (A5.5)), (A5.64) is (A5.31) combined with (A5.61), and (A5.65) is the government budget identity (A5.58) expressed in per capita form. Finally, we have substituted the rental expressions $W_t = W(k_t)$ and $r_t = r(k_t)$ in the various equations (see equations (A5.15) and (A5.16) above).

The first thing we note is that the fiscal variables only show up in two places in the dynamical system. In (A5.62) there is a resource transfer from the government to each old household (Γ_t^{GO}) consisting of debt service and transfers:

$$\Gamma_t^{GO} \equiv (1 + r(k_t))(1+n) b_t + Z_t. \qquad \text{(government to old)}$$

Similarly, in (A5.64) there is a resource transfer from each young household to the government (Γ_t^{YG}) in

[12] Consumption by old agents is $L_{t-1} C_t^O = (r_t + \delta) K_t + (1-\delta) K_t + (1 + r_t) B_t + L_{t-1} Z_t$. For young agents we have $L_t C_t^Y = L_t \left[W_t - T_t - S_t\right]$ so that aggregate consumption is:

$$\begin{aligned} C_t &= (r_t + \delta) K_t + (1-\delta) K_t + (1 + r_t) B_t + L_{t-1} Z_t + L_t \left[W_t - T_t - S_t\right] \\ &= Y_t + (1-\delta) K_t + \left[(1 + r_t) B_t + L_{t-1} Z_t - L_t T_t\right] - L_t S_t \\ &= Y_t + (1-\delta) K_t + B_{t+1} - L_t S_t. \end{aligned}$$

By combining the final expression with the resource constraint (A5.17) we obtain (A5.61).

the form of purchases of government debt plus taxes:

$$\Gamma_t^{YG} \equiv (1+n)b_{t+1} + T_t. \qquad \text{(young to government)}$$

Since there are L_{t-1} old and L_t young households, the net resource transfer to the government is $L_t\Gamma_t^{YG} - L_{t-1}\Gamma_t^{GO} = 0$, where the equality follows from the government budget constraint (A5.65). Hence, in the absence of government consumption, what the government takes from the young it must give to the old. Put differently, once you know Γ_t^{YG} you also know $\Gamma_t^{GO} \equiv (1+n)\Gamma_t^{YG}$ (and vice versa) and the individual components appearing in the government budget identity (such as b_{t+1}, b_t, Z_t, and T_t) are irrelevant for the determination of the paths of consumption and the capital stock (Buiter and Kletzer, 1992, p. 17).

The equivalence result is demonstrated by considering two paths of the economy which, though associated with different paths for bonds, taxes, and transfers, nevertheless give rise to the same paths for the real variables, namely the capital stock and consumption by the young and the old. For the reference path, the sequence $\{\hat{b}_t, \hat{Z}_t, \hat{T}_t\}_{t=0}^{\infty}$ gives rise to a sequence for the real variables denoted by $\{\hat{C}_t^Y, \hat{C}_t^O, \hat{k}_t\}_{t=0}^{\infty}$ given k_0 and b_0. We can then show that for any other debt sequence $\{\check{b}_t\}_{t=1}^{\infty}$ we can always find sequences for taxes and transfers $\{\check{Z}_t, \check{T}_t\}_{t=0}^{\infty}$ such that the resulting sequences for the real variables are the same as in the reference path, i.e. $\{\hat{C}_t^Y\}_{t=0}^{\infty} = \{\check{C}_t^Y\}_{t=0}^{\infty}$, $\{\hat{C}_t^O\}_{t=0}^{\infty} = \{\check{C}_t^O\}_{t=0}^{\infty}$, and $\{\hat{k}_t\}_{t=0}^{\infty} = \{\check{k}_t\}_{t=0}^{\infty}$.

The key ingredient of the proof is to construct the alternative path such that the resource transfers from the young to the government (Γ_t^{YG}) and from the government to the old (Γ_t^{GO}) are the same for the two paths. These requirements give rise to the following expressions:

$$\hat{Z}_t - \check{Z}_t = (1+n)\left[(1+r(\check{k}_t))\check{b}_t - (1+r(\hat{k}_t))\hat{b}_t\right], \tag{A5.66}$$

$$\check{b}_{t+1} - \hat{b}_{t+1} = \frac{1}{1+n}\left[\hat{T}_t - \check{T}_t\right]. \tag{A5.67}$$

By using (A5.66) in (A5.62) and (A5.67) in (A5.64) we find that these equations solve for the same real variables. As a result, the Euler equation (A5.63) is the same for both paths. Obviously the government budget identity still holds. Finally, if the reference path satisfies the government solvency condition then so will the alternative path.

As a special case of the equivalence result we can take as the reference path the PAYG system (studied above), which has $\hat{b}_t = 0$, $\hat{T}_t = T$, and $\hat{Z}_t = (1+n)T$ for all t. One (of many) alternative paths is the deficit path in which there are only taxes on the young generations, i.e. $\check{Z}_t = 0$, $\check{b}_t = (1+n)T/(1+r_t)$, and $\check{T}_t = T - (1+n)\check{b}_{t+1}$ for all t.

12.2.1.4 From PAYG to a funded system

In the previous subsection we have established the equivalence between traditional deficit financing and a PAYG social security system. As a by-product of the analysis there we showed how public debt affects the equilibrium path of the economy. In this section we continue our analysis of the welfare effects of a PAYG system, first without and then with bond policy.

Up to this point we have only unearthed the welfare effect of a PAYG system on steady-state generations (see (A5.57)) and we have ignored the initial conditions facing the economy, i.e. we have not yet taken into account the costs associated with the transition from the initial growth path to the golden-rule path. As both Diamond (1965, pp. 1128-1129) and Samuelson (1975b, p. 543) stress, ignoring transitional welfare effects is not a very good idea.

As we argued above, the introduction of a PAYG system (or the expansion of a pre-existing one) affects different generations differently. The welfare of old generations at the time of the shock unambiguously rises because of the windfall gain the shock confers on them. From the perspective of their last period of life, they gain utility to the tune of $U'(C_1^O)dC_0^O/dT = U'(C_1^O) > 0$ (see (A5.47)). The welfare effect on generations born in the new steady state is ambiguous as it depends on whether or not the economy is dynamically efficient (see (A5.57)). In a dynamically inefficient economy, $r < n$, all generations, including those born in the new steady state, gain from the pension shock. Intuitively, the PAYG system acts like a "chain letter" system which ensures that each new generation passes resources to the generation immediately preceding it. In such a situation a PAYG system which moves the economy in the direction of the golden-rule growth path is surely "desirable" for society as a whole.

As was mentioned above, however, actual economies are not likely to be dynamically inefficient and "free lunches" are not for the taking. If the economy is dynamically efficient, so that $r > n$, then it follows from, respectively, (A5.47) and (A5.57) that whilst an increase in T still makes the old initial generation better off, it leaves steady-state generations worse off than they would have been in the absence of the shock. Since some generations gain and other lose out, it is no longer obvious whether a pension-induced move in the direction of the golden-rule growth path is "socially desirable" at all.

There are two ways in which the concept of social desirability, which we have deliberately kept vague up to now, can be made operational. The first approach, which was pioneered by Bergson (1938) and Samuelson (1947), makes use of a so-called *social welfare function* (see also Chapter 9 above). In this approach, a functional form is typically postulated which relates an indicator for social welfare (SW) to the welfare levels experienced by the different generations. Using our notation, an example of a social welfare function would be:

$$SW_t = \Psi(\Lambda_{t-1}^Y, \Lambda_t^Y, ..., \Lambda_\infty^Y), \tag{A5.68}$$

where $\Psi_s \equiv \partial\Psi/\partial\Lambda_s^Y > 0$ for $s = t-1, t, \cdots, \infty$. The currently alive generations and all future generations feature in the social welfare function with a positive weight. Once a particular form for the social welfare function is adopted, the social desirability of different policies can be ranked. If policy A is such that it yields a higher indicator of social welfare than policy B, then it follows that policy A is *socially preferred* to policy B (i.e. $SW_t^A > SW_t^B$). Note that, depending on the form of the social welfare function $\Psi(\cdot)$, it may very well be the case that some generations are worse off under policy A than under policy B despite the fact that A is socially preferred to B. What the social welfare function does is establish marginal rates of substitution between lifetime utility levels of different generations (i.e. $(\partial w/\partial \Lambda_{t-1}^Y)/(\partial w/\partial \Lambda_t^Y)$, etc.).[13]

The second approach to putting into operation the concept of social desirability makes use of the concept of *Pareto-efficiency*. Recall that an allocation of resources in the economy is called Pareto-optimal (or Pareto-efficient) if there is no other feasible allocation which (i) makes no individual in the economy worse off and (ii) makes at least one individual strictly better off than he/she was. Similarly, a policy is called *Pareto-improving* vis-à-vis the initial situation if it improves welfare for at least one agent and leaves all other agents equally well off as in the status quo.

Recently, a number of authors have applied the Pareto-criterion to the question of pension reform. Specifically, Breyer (1989) and Verbon (1989) ask themselves the question whether it is possible to abolish a pre-existing PAYG system (in favour of a fully funded system) in a Pareto-improving fashion in a dynamically efficient economy. This is a relevant question because in such an economy, steady-state generations gain if the PAYG system is abolished or reduced (since $r > n$ it follows from equation (A5.57) that $d\Lambda^Y/dT < 0$ in that case) but the old generation at the time of the policy shock loses out (see (A5.47)). This generation paid into the PAYG system when it was young in the expectation that it would receive back $1+n$ times its contribution during old age. Taken in isolation, the policy shock is clearly not Pareto-improving.

Of course bond policy constitutes a mechanism by which the welfare gains and losses of the different generations can be redistributed. This is the case because it breaks the link between the contributions of the young (L_tT_t) and the pension receipts by the old in the same period ($L_{t-1}Z_t$)–compare (A5.35) and (A5.58). The key issue is thus whether it is possible to find a bond path such that the reduction in the PAYG contribution is Pareto-improving. As it turns out, no such path can be found. It is thus not possible to compensate the old generation at the time of the shock without making at least one future generation worse off (Breyer, 1989, p. 655).

12.2.2 PAYG pensions and endogenous retirement

In a very influential article, Feldstein (1974) argued that a PAYG system not only affects a household's savings decisions (as is the case in the model studied up to this point) but also its decision to retire from the labour force. We now augment the model in order to demonstrate the implications for allocation and welfare of endogenous retirement. Following the literature, we capture the notion of retirement by

[13] An application of the social welfare function approach is given in the next subsection.

assuming that labour supply during the first period of life is endogenous. To keep the model as simple as possible, we continue to assume that households do not work at all during the second period of life. To bring the model closer to reality, we assume furthermore that the contribution to the PAYG system is levied in the form of a proportional tax on labour income and that the pension is *intragenerationally fair,* i.e. an agent who works a lot during youth gets a higher pension during old age than an agent who has been lazy during youth. Within the augmented model it is possible that the PAYG system distorts the labour supply decisions by households.

12.2.2.1 Households

The lifetime utility function of a (representative) young agent who is born at time t is given in general form by:

$$\Lambda_t^Y \equiv U(C_t^Y, 1 - N_t, C_{t+1}^O,), \tag{A5.69}$$

where C_t^Y, C_{t+1}^O, and N_t are, respectively, consumption during youth, consumption during old age, and labour supply ($1 - N_t$ is leisure). The utility function $U(\cdot)$ features positive first-order derivatives and is strictly quasi-concave (see Silberberg and Suen, 2001, p. 260). The agent faces the following budget identities:

$$C_t^Y + S_t = W_t N_t - T_t, \tag{A5.70}$$

$$C_{t+1}^O = (1 + r_{t+1}) S_t + Z_{t+1}, \tag{A5.71}$$

where T_t and Z_{t+1} are defined as follows:

$$T_t = t_L W_t N_t, \tag{A5.72}$$

$$Z_{t+1} = t_L W_{t+1} L_{t+1} N_{t+1} \frac{N_t}{NL_t}, \tag{A5.73}$$

where $0 < t_L < 1$. According to (A5.72), the individual agent's contribution to the PAYG system is equal to a proportion of his labour income, where the proportional tax, t_L, is assumed to be the same for all individuals and constant over time. Equation (A5.73) shows that the pension is intragenerationally fair (as in Breyer and Straub, 1993, p. 81). The term in round brackets on the right-hand side of (A5.73) is the total tax revenue that is available for pension payments in the *next* period. Each old agent get a share of this revenue which depends on his *relative* labour supply effort, i.e. NL_t represents aggregate labour supply in period t. We assume that the young household takes all variables appearing in (A5.73) as given, *except* his own labour supply N_t. Put differently, the household realizes the link between working hard during youth and receiving a high pension payment during old age.[14]

[14] In the symmetric equilibrium, all households supply the same amount of labour and $NL_t = N_t L_t$ so that each household's share of the pension revenue is equal to $1/L_t$. Working directly with this expression would obscure the hypothesized link between

The household is fully aware of the features of the pension system (as formalized in (A5.72)-(A5.73)) so that the consolidated lifetime budget constraint is given by:

$$(1 - t_{Et})\, W_t N_t = C_t^Y + \frac{C_{t+1}^O}{1 + r_{t+1}}, \tag{A5.74}$$

where t_{Et} is defined as follows:

$$t_{Et} \equiv t_L \left[1 - \frac{W_{t+1}}{W_t\,(1 + r_{t+1})} \frac{L_{t+1} N_{t+1}}{NL_t} \right]. \tag{A5.75}$$

The key thing to note about (A5.74)-(A5.75) is that in the current setting the household's pension depends not only on future wages but also on the aggregate supply of labour *by future young agents*. To solve its optimization problem, the household must thus form expectations regarding these variables and, as usual, by suppressing the expectations operator we have implicitly assumed that the agent is blessed with perfect foresight.

In the interior optimum, the first-order conditions for consumption during the two periods and labour supply are:

$$\frac{\partial U}{\partial C_{t+1}^O} = \frac{1}{1 + r_{t+1}} \frac{\partial U}{\partial C_t^Y}, \tag{A5.76}$$

$$\left[-\frac{\partial U}{\partial N_t} = \right] \frac{\partial U}{\partial (1 - N_t)} = (1 - t_{Et}) W_t \frac{\partial U}{\partial C_t^Y}. \tag{A5.77}$$

Equation (A5.76) is the familiar consumption Euler equation in general functional form. The optimal labour supply decision is characterized by (A5.77) and (A5.75). Equation (A5.77) is the usual condition calling for an equalization of the after-tax wage rate and the marginal rate of substitution between leisure and consumption during youth. Equation (A5.75) shows to what extent the PAYG system has the potential to distort the labour supply decision. It is not the *statutory* tax rate, t_L, which determines whether or not the labour supply decision is distorted but rather the (potentially time-varying) *effective* tax rate, t_{Et}. By paying the PAYG premium during youth one obtains the right to a pension. Ceteris paribus labour supply, the effective tax rate may actually be negative, i.e. it may in fact be an employment subsidy (Breyer and Straub, 1993, p. 82).

Since all agents of a particular generation are identical in all aspects we can now simplify the model. In the symmetric equilibrium we have $NL_t = N_t L_t$ and with a constant growth rate of the population it follows that $L_{t+1} = (1 + n) L_t$. Hence, equation (A5.75) simplifies to:

$$t_{Et} \equiv t_L \left[1 - \frac{W_{t+1}}{W_t} \frac{N_{t+1}}{N_t} \frac{1 + n}{1 + r_{t+1}} \right]. \tag{A5.78}$$

Holding constant labour supply we find that the pension system acts like an employment subsidy (and

pension payments and receipts.

$t_{Et} < 0$) if the so-called *Aaron condition* holds, i.e. if the growth of the population and wages exceeds the rate of interest (Aaron, 1966).

In the symmetric equilibrium, equations (A5.74) and (A5.76)-(A5.78) define the optimal values of C_t^Y, C_{t+1}^O, and N_t as a function of the variables that are exogenous to the representative agent (namely, W_t, r_{t+1}, and t_{Et}). We write these solutions as $C_t^Y = C^Y(W_t^N, r_{t+1})$, $C_{t+1}^O = C^O(W_t^N, r_{t+1})$, and $N_t = N(W_t^N, r_{t+1})$, where $W_t^N \equiv W_t(1 - t_{Et})$. The (partial-equilibrium) effect of a change in the statutory tax rate, t_L, on the household's labour supply decision can thus be written in elasticity format as:

$$\frac{t_L}{N_t}\frac{\partial N_t}{\partial t_L} = -\epsilon_{W^N}^N \frac{t_{Et}}{1 - t_{Et}}, \qquad \epsilon_{W^N}^N \equiv \frac{W^N}{N}\frac{\partial N}{\partial W^N}, \tag{A5.79}$$

where $\epsilon_{W^N}^N$ is the uncompensated elasticity of labour supply. It follows from (A5.79) that the effect of the contribution rate on labour supply is ambiguous for two reasons. First, it depends on whether the Aaron-condition is satisfied (so that $t_{Et} < 0$) or violated (so that $t_{Et} > 0$). Second, it also depends on the sign of $\epsilon_{W^N}^N$. We recall that $\epsilon_{W^N}^N > 0$ (< 0) if the substitution effect in labour supply dominates (is dominated by) the income effect. If labour supply is upward sloping *and* the Aaron condition is satisfied then, for given factor prices, an increase in the statutory tax rate decreases labour supply.

12.2.2.2 The macroeconomy

We must now complete the description of the model and derive the fundamental difference equation for the economic system. We follow the approach of Ihori (1996, pp. 36-37). With endogenous labour supply, the number of agents (L_t) no longer coincides with the amount of labour used in production (L_tN_t). By redefining the capital-labour ratio as $k_t \equiv K_t/(L_tN_t)$, however, the expressions for the wage and the interest rate are still as in (A5.15)-(A5.16) and the factor price frontier is still as given in (A5.56). Current savings leads to the formation of capital in the next period, i.e. $L_tS_t = K_{t+1}$. In terms of the redefined capital-labour ratio we get:

$$S_t = (1+n)N_{t+1}k_{t+1}. \tag{A5.80}$$

To characterize this fundamental difference equation we note that the labour supply and savings equations can be written in general functional form as:

$$N_t = N(W_t(1 - t_{Et}), r_{t+1}), \tag{A5.81}$$

$$S(\cdot) \equiv \frac{C^O(W_t(1 - t_{Et}), r_{t+1}) - (1+n)t_L W_{t+1} N_{t+1}}{1 + r_{t+1}}. \tag{A5.82}$$

By using these expressions in (A5.80) we obtain the following expression:

$$S\left(W_t(1 - t_{Et}), r_{t+1}, t_L W_{t+1} N_{t+1}\right) = (1+n)N(W_{t+1}(1 - t_{Et+1}), r_{t+2})k_{t+1}. \tag{A5.83}$$

Clearly, since $W_t = W(k_t)$ and $r_t = r(k_t)$, this expression contains the terms k_t, k_{t+1}, and k_{t+2} so one is tempted to conclude that it is a second-order difference equation in the capital stock. As Breyer and Straub (1993, p. 82) point out, however, the presence of future pensions introduces an infinite regress into the model, i.e. since t_{Et} depends on N_{t+1} (see (A5.78)), it follows that t_{Et+1} depends on N_{t+2} which itself depends on k_{t+2}, k_{t+3}, *and* t_{Et+2}. As a result, (A5.83) depends on the entire sequence of present and future capital stocks, $\{k_{t+\tau}\}_{\tau=0}^{\infty}$ so that, even though we assume perfect foresight, the model has a continuum of equilibria.[15] Since we assume that the population growth rate is constant, however, we can skip over the indeterminacy issue by first studying the steady state.

12.2.2.3 The steady state

We study two pertinent aspects of the steady state. First, we show how the endogeneity of labour supply affects the welfare effect of the PAYG pension. Second, we show that in the unit-elastic model the pension crowds out capital in the long run. In the steady state, we have that $W_{t+1} = W_t = W$, $N_{t+1} = N_t = N$, and $r_{t+1} = r$ so that $t_{Et} = (r-n)\, t_L / \, (1+r)$. As before, the long-run welfare analysis makes use of the indirect utility function which is defined as follows:

$$\bar{\Lambda}^Y(W, r, t_L) \equiv \max_{\{C^Y, C^O, N\}} U(C^Y, C^O, 1-N)$$

$$\text{subject to:} \quad WN\left[1 - t_L \frac{r-n}{1+r}\right] = C^Y + \frac{C^O}{1+r}. \tag{A5.84}$$

Retracing our earlier derivation we can derive the following properties of the indirect utility function:

$$\frac{\partial \bar{\Lambda}^Y}{\partial W} = N \frac{\partial \Lambda^Y}{\partial C^Y}\left[1 - t_L \frac{r-n}{1+r}\right], \tag{A5.85}$$

$$\frac{\partial \bar{\Lambda}^Y}{\partial r} = \frac{S}{1+r}\frac{\partial \Lambda^Y}{\partial C^Y}, \tag{A5.86}$$

$$\frac{\partial \bar{\Lambda}^Y}{\partial t_L} = -WN \frac{r-n}{1+r}\frac{\partial \Lambda^Y}{\partial C^Y}. \tag{A5.87}$$

The effect of a marginal change in the statutory tax rate on steady-state welfare is now easily computed:

$$\begin{aligned}
\frac{d\Lambda^Y}{dt_L} &= \frac{\partial \bar{\Lambda}^Y}{\partial W}\frac{dW}{dt_L} + \frac{\partial \bar{\Lambda}^Y}{\partial r}\frac{dr}{dt_L} + \frac{\partial \bar{\Lambda}^Y}{\partial t_L} \\
&= \frac{\partial \Lambda^Y}{\partial C^Y}\left[N\left[1 - t_L \frac{r-n}{1+r}\right]\frac{dW}{dt_L} + \frac{S}{1+r}\frac{dr}{dt_L} - \frac{r-n}{1+r}WN\right] \\
&= -N\frac{r-n}{1+r}\frac{\partial \Lambda^Y}{\partial C^Y}\left[W + (1-t_L)k\frac{dr}{dt_L}\right],
\end{aligned} \tag{A5.88}$$

[15] Indeterminacy and multiple equilibria are quite common phenomena in overlapping-generations models of the Diamond-Samuelson type. Azariadis (1993) gives a general discussion and Reichlin (1986) deals specifically with the case of endogenous labour supply.

where we have used (A5.85)-(A5.87) in going from the first to the second line and (A5.56) and (A5.80) in going from the second to the third line. There are two noteworthy conclusions that can be drawn on the basis of (A5.88). First, if the economy is initially in the golden-rule equilibrium ($r = n$), then a marginal change in t_L does not produce a first-order welfare effect on steady-state generations. Intuitively, the labour supply decision is not distorted because the effective tax on labour is zero in that case ($t_E = t_L(r-n)/(1+r) = 0$). Second, if the economy is not in the golden-rule equilibrium ($r \neq n$), then the sign of the welfare effect is determined by the sign of the term in square brackets on the right-hand side of (A5.88). Just as for the case with lump-sum contributions (see (A5.57)), the PAYG pension affects welfare through lifetime resources (first term in brackets) and via the capital-labour ratio (second term). It turns out, however, that with endogenous labour supply the sign of dr/dt_L (and thus the sign of $d\bar{\Lambda}^Y/dt_L$) is ambiguous (Ihori, 1996, p. 237).

Next we return to the analysis of the model outside the steady state. Matters are simplified quite a lot if Cobb-Douglas preferences are assumed, i.e. if (A5.69) is specialized to:

$$\Lambda_t^Y \equiv \ln C_t^Y + \lambda_C \ln(1-N_t) + \frac{1}{1+\rho} \ln C_{t+1}^O, \tag{A5.89}$$

where ρ is the rate of time preference and λ_C (≥ 0) regulates the strength of the labour supply effect. The following solutions for the decision variables are then obtained by maximizing (A5.89) subject to (A5.74):

$$C_t^Y = \frac{1+\rho}{2+\rho+\lambda_C(1+\rho)} W_t^N, \tag{A5.90}$$

$$C_{t+1}^O = \frac{1+r_{t+1}}{2+\rho+\lambda_C(1+\rho)} W_t^N, \tag{A5.91}$$

$$N_t = \frac{2+\rho}{2+\rho+\lambda_C(1+\rho)}, \tag{A5.92}$$

where $W_t^N \equiv W_t(1-t_{Et})$ is the effective after-tax wage. In the unit-elastic model, consumption during youth and old age are both normal goods and labour supply is constant because income and substitution effects cancel out. Since the current workers know that future workers will also supply a fixed amount of labour ($N_{t+1} = N_t = N$), the expression for the after-tax wage simplifies to:

$$W_t^N \equiv W_t(1-t_{Et}) \equiv W_t \left[1 - t_L \left[1 - \frac{W_{t+1}}{W_t} \frac{1+n}{1+r_{t+1}}\right]\right]. \tag{A5.93}$$

Note furthermore that in (A5.90) the presence of pension payments during old age ensures that consumption during youth depends negatively on the interest rate–via the effective tax rate–despite the fact that logarithmic preferences are used. According to (A5.91) old-age consumption depends positively on the interest rate and negatively (positively) on the tax rate if the Aaron condition is violated (holds) $t_{Et} > 0$ ($t_{Et} < 0$). Finally, in (A5.92) the standard model is recovered by setting $\lambda_C = 0$, in which case labour supply is exogenous and equal to unity ($N_t = 1$).

We can now determine the extent to which capital is crowded out by the PAYG system. By using (A5.80), (A5.82), and (A5.91)-(A5.93), the fundamental difference equation for the model can be written as follows:

$$(1+n)k_{t+1} = \frac{W_t(1-t_L)}{2+\rho} - \frac{1+\rho}{2+\rho}\frac{t_L\,(1+n)\,W_{t+1}}{1+r_{t+1}}. \tag{A5.94}$$

Since $W_t = W(k_t)$ and $r_t = r(k_t)$, equation (A5.94) constitutes a first-order difference equation in the capital-labour ratio. Hence, in the unit-elastic model the indeterminacy of the transition path (that was mentioned above) disappears because the uncompensated labour supply elasticity is zero.

The stability condition and the long-run effect of the PAYG system on the capital-labour ratio are derived in the usual manner by finding the partial derivatives of the implicit function, $k_{t+1} = g(k_t, t_L)$, around the steady state. After some manipulation we obtain:

$$g_k \equiv \frac{\partial k_{t+1}}{\partial k_t} = \frac{(1-t_L)W'}{(1+n)(2+\rho)\left[1+t_L\frac{1+\rho}{2+\rho}\frac{(1+r)W'-Wr'}{(1+r)^2}\right]} > 0, \tag{A5.95}$$

$$g_t \equiv \frac{\partial k_{t+1}}{\partial t_L} = -\frac{W\left[1+r+(1+\rho)(1+n)\right]}{(1+r)(1+n)(2+\rho)\left[1+t_L\frac{1+\rho}{2+\rho}\frac{(1+r)W'-Wr'}{(1+r)^2}\right]} < 0. \tag{A5.96}$$

Since g_k is positive (as $W' > 0 > r'$), stability requires it to be less than unity ($0 < g_k < 1$). As a result, the long-run effect on the capital-labour ratio is unambiguously negative in the unit-elastic model:

$$\frac{dk}{dt_L} = \frac{g_t}{1-g_k} < 0. \tag{A5.97}$$

12.2.2.4 Welfare effects

We are now in a position to compare and contrast the key results of this subsection to those that hold when labour supply is exogenous and the pension contribution is levied in a lump-sum fashion (see subsection 12.2.1.2). At first view, the assumption of a distorting pension contribution does not seem to change the principal conclusions very much–at least in the unit-elastic model. In both cases, the PAYG contribution leads to long-run crowding out of the capital-labour ratio (compare (A5.49) and (A5.97)) and a reduction (increase) in steady-state welfare for a dynamically efficient (inefficient) economy (compare (A5.57) and (A5.88)). Intuitively, this similarity is only moderately surprising in view of the fact that in the unit-elastic model (optimally chosen) labour supply is constant (see (A5.92)).

There is a very important difference between the two cases, however, because the pension contribution, t_L, causes a distortion of the labour supply decision of households which is absent if the contribution is levied in a lump-sum fashion. The resulting loss to the economy of using a distorting rather than a non-distorting tax is often referred to as the *deadweight loss* (or burden) of the distorting tax (Diamond and McFadden, 1974, p. 5). Following Diamond and McFadden we define the deadweight loss (DWL)

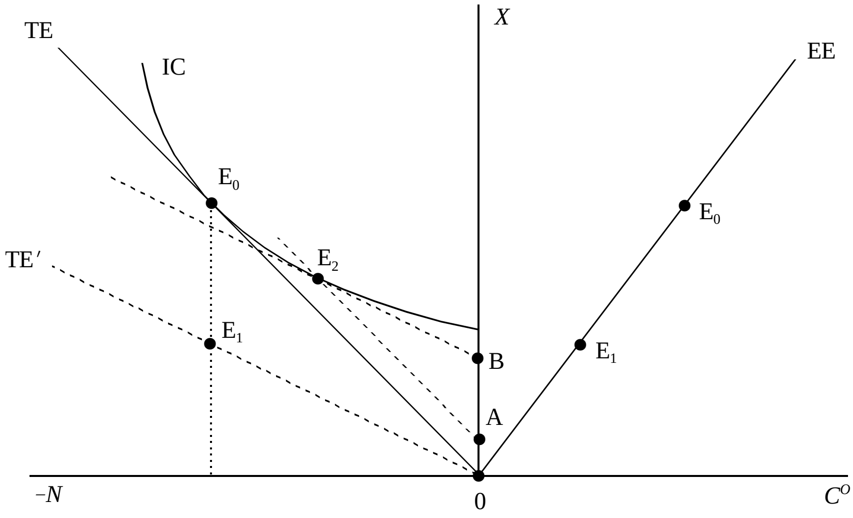

Figure 12.3: Deadweight loss of taxation

associated with t_L as the difference between, on the one hand, the income one must give a young household to restore it to its pre-tax indifference curve and, on the other hand, the tax revenue collected from it (1974, p. 5).

In Figure 12.3 we illustrate the DWL of the pension contribution for a steady-state generation in the unit-elastic model. We hold factor prices (W and r) constant and assume that the economy is dynamically efficient ($r > n$). We follow the approach of Belan and Pestieau (1999) by solving the model in two stages. In the first stage we define *lifetime income* as:

$$X \equiv WN\left[1 - t_L\frac{r-n}{1+r}\right] \equiv WN(1-t_E), \tag{A5.98}$$

and let the household choose current and future consumption in order to maximize:

$$\ln C^Y + \frac{1}{1+\rho}\ln C^O, \tag{A5.99}$$

subject to the constraint $C^Y + C^O/(1+r) = X$. This yields the following expressions:

$$C^Y = \frac{1+\rho}{2+\rho}X, \qquad C^O = \frac{1+r}{2+\rho}X. \tag{A5.100}$$

In the right-hand panel of Figure 12.3 the line EE relates old-age consumption to lifetime income. In that panel the value of consumption during youth can be deduced from the fact that it is proportional to lifetime income.

By substituting the expressions (A5.100) into, respectively, the utility function (A5.89) and the budget constraint (given in (A5.84)) we obtain:

$$\Lambda^Y \equiv \frac{2+\rho}{1+\rho}\ln X + \lambda_C \ln(1-N_t) + \ln\left(\frac{1+\rho}{2+\rho}\left(\frac{1+r}{2+\rho}\right)^{1/(1+\rho)}\right), \tag{A5.101}$$

$$X = WN(1-t_E). \tag{A5.102}$$

In the second stage, the household chooses its labour supply and lifetime income in order to maximize (A5.101) subject to (A5.102). The solution to this second-stage problem is, of course, that N takes the value indicated in (A5.92) and X follows from the constraint. The second-stage optimization problem is shown in the left-hand panel of Figure 12.3. In that panel, TE represents the budget line (A5.102) in the absence of taxation ($t_E = 0$). It is downward sloping because we measure *minus* N on the horizontal axis. The indifference curve which is tangent to the pre-tax budget line is given by IC and the initial equilibrium is at E_0. In the right-hand panel E_0 on the EE line gives the corresponding optimal value for old-age consumption.

Now consider what happens if a positive effective tax is levied ($t_L^E > 0$). Nothing happens in the right-hand panel but in the left-hand panel the budget line rotates in a counter-clockwise fashion. The new budget line is given by the dashed line TE′ from the origin. We know that in the unit-elastic model income and substitution effects in labour supply cancel out so that labour supply does not change (see (A5.92)). Hence, the new equilibrium is at E_1 in the two panels. By shifting the new budget line in a parallel fashion and finding a tangency along the pre-tax indifference curve we find that the pure substitution effect of the tax change is given by the shift from E_0 to E_2 (the income effect is thus the shift from E_2 to E_1). Hence, the vertical distance OB represents the income one would have to give the household to restore it to its pre-tax indifference curve. We call this hypothetical transfer Z_0. What is the tax revenue which is collected from the agent? To answer that question we draw a line, that is parallel to the pre-tax budget line TE, through the compensated point E_2. This line has an intercept with the vertical axis at point A. We now have two expressions for lines that both pass through the compensated point E_2, namely $X + W(1-t_L^E)(-N) = Z_0$ and $X + W(-N) = Z_0 - T$, where T is the vertical distance AB in Figure 12.3. By deducting the two lines we find that $T = t_E WN$ so that AB represents the tax revenue collected from the agent. Since the required transfer is 0B the DWL of the tax is given by the distance 0A.

12.2.2.5 Reform

As a number of authors have recently pointed out, the distorting nature of the pension system has important implications for the possibility of designing Pareto-improving reform (see e.g. Homburg, 1990, Breyer and Straub, 1993, and the references to more recent literature in Belan and Pestieau, 1999). Recall from the discussion at the end of section 12.2.1.4 that a Pareto-improving transition from PAYG to a fully funded system is not possible in the standard model because the resources cannot be found to compensate the old generations at the time of the reform without making some future generation worse off. Matters are different if the PAYG system represents a distorting system. In that case, as Breyer and Straub (1993) point out, provided lump-sum (non-distorting) contributions can be used during the transition phase, a gradual move from a PAYG to a fully funded system can be achieved in a Pareto-improving manner. Intuitively, by moving from a distortionary to a non-distortionary scheme, additional resources are freed up which can be used to compensate the various generations (Belan and Pestieau, 1999).[16]

12.2.3 The macroeconomic effects of ageing

Up to this point we have assumed that the rate of population growth is constant and equal to n (see equation (A5.21) above). This simplifying assumption of course means that the age composition of the population is constant also. A useful measure to characterize the economic impact of demography is the so-called (old-age) *dependency ratio*, which is defined as the number of retired people divided by the working-age population. In our highly stylized two-period overlapping-generations model the number of old and young people at time t are, respectively, L_{t-1} and $L_t = (1+n)L_{t-1}$ so that the dependency ratio is $1/(1+n)$.

Of course, as all members of the baby-boom generation will surely know, the assumption of a constant population composition, though convenient, is not a particularly realistic one. Table 12.1, which is taken from Weil (1997, p. 970), shows that significant demographic changes have taken place between 1950 and 1990 and are expected to take place between 1990 and 2025.

The figures in Table 12.1 graphically illustrate that throughout the world, and particularly in the group of OECD countries and in the US, the proportion of young people (0-20 years of age) is on the decline whilst the fraction of old people (65 and over) steadily increases. Both of these phenomena are tell-tale signs of an ageing population.

In this subsection we show how the macroeconomic effects of demographic composition changes can be analysed with the aid of a simple overlapping-generations model. We only stress some of the key results, especially those relating to the interaction between demography and the public pension system. The interested reader is referred to Weil (1997) for an excellent survey of the literature on the economics of ageing.

In the absence of immigration from abroad, population ageing can result from two distinct sources, namely a decrease in *fertility* and a decrease in *mortality*. In the two-period overlapping-generations model used so far the length of life is exogenously fixed but we can nevertheless capture the notion of ageing by reducing the rate of population growth, n. In order to study the effects on allocation and

[16]The distortive nature of the PAYG scheme does not have to result from endogenous labour supply. Demmel and Keuschnigg (2000), for example, assume that union wage-setting causes unemployment which is exacerbated by the pension contribution. Efficiency gains then materialize because pension reform reduces unemployment. In a similar vein, Belan et al. (1998) use a Romer-style (1986, 1989) endogenous growth model and show that reform may be Pareto-improving because it helps to internalize a positive externality in production. See also Corneo and Marquardt (2000).

Table 12.1: Age composition of the population

	1950	1990	2025
World			
0-19	44.1	41.7	32.8
20-65	50.8	52.1	57.5
65+	5.1	6.2	9.7
OECD			
0-19	35.0	27.2	24.8
20-64	56.7	59.9	56.6
65+	8.3	12.8	18.6
US			
0-19	33.9	28.9	26.8
20-65	57.9	58.9	56.0
65+	8.1	12.2	17.2

welfare of such a demographic shock we first reformulate the model of subsection 12.2.1.2 in terms of a variable growth rate of the population, n_t. Hence, instead of (A5.21) we use:

$$L_t = (1 + n_t)L_{t-1}, \quad n_t > -1. \tag{A5.103}$$

Assuming a constant contribution rate per person ($T_t = T$), the pension at time t equals $Z_t = (1 + n_t)T$. Redoing the derivations presented in subsection 12.2.1.2 yields the following fundamental difference equation of the model:

$$S(W_t, r_{t+1}, n_{t+1}, T) = (1 + n_{t+1})k_{t+1}, \tag{A5.104}$$

where the savings function is the same as in (A5.43) but with n_{t+1} replacing n. Ceteris paribus, saving by the young depends negatively on the (expected) rate of population growth, n_{t+1}, because the pension they receive when old depends on it (as $Z_{t+1} = (1 + n_{t+1})T$). An anticipated reduction in fertility reduces the expected pension and lifetime income, and causes the agent to cut back on both present and future consumption and to increase saving. Hence, $S_n \equiv \partial S/\partial n_{t+1} < 0$. The right-hand side of (A5.104) shows that a decrease in the population growth rate makes it possible to support a higher capital-labour ratio for a given amount of per capita saving.

Following the solution method discussed in subsection 12.2.1.2, we can derive that (A5.104) defines an implicit function, $k_{t+1} = g(k_t, n_{t+1}, T)$, with partial derivatives $0 < g_k < 1$ (see equation (A5.45)) and $g_n < 0$:

$$g_n \equiv \frac{\partial g}{\partial n_{t+1}} = \frac{S_n - k_{t+1}}{1 + n_{t+1} - S_r r'(k_{t+1})} < 0. \tag{A5.105}$$

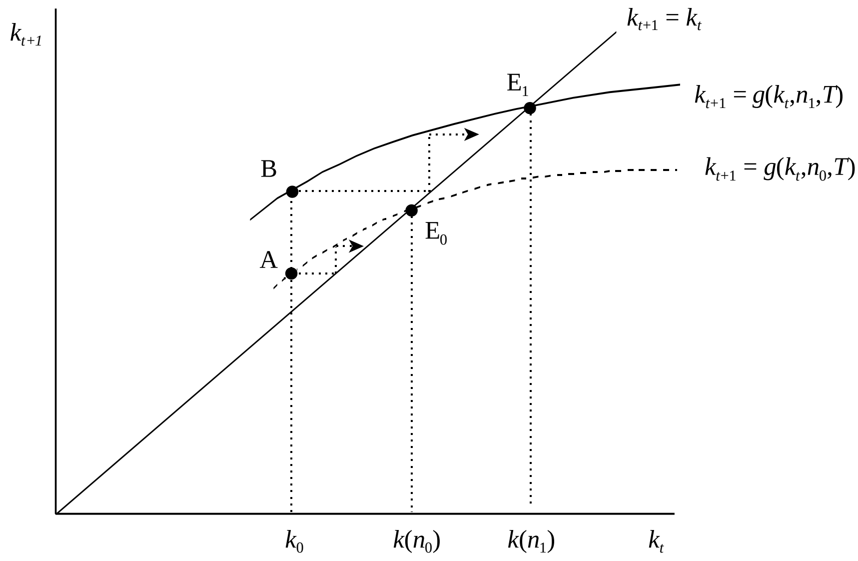

Figure 12.4: The effects of ageing

It follows that a permanent reduction in the population growth rate, say from n_0 to n_1, gives rise to an increase in the long-run capital stock, i.e. $dk/dn = g_n/(1 - g_k) < 0$. The transition path of the economy to the steady state is illustrated in Figure 12.4. In that figure, the dashed line labelled "$k_{t+1} = g(k_t, n_0)$" reproduces the initial transition path with social security in Figure 12.2 (note that only the stable segment has been drawn). The reduction in fertility boosts saving at impact so that, if the economy starts out with a capital stock k_0, the new transition path is the dotted line from B to the new equilibrium at E_1. During transition the wage rate gradually rises and the interest rate falls. The intuition behind the long-run increase in the capital-labour ratio is straightforward. As a result of the demographic shock there are fewer young households, who own no assets, and more old households, who own a lot of assets which they need to provide income for their retirement years (Auerbach and Kotlikoff, 1987, p. 163).

The effect of a permanent reduction in fertility on steady-state welfare can be computed by differentiating the indirect utility function (A5.50) with respect to n, using (A5.53)-(A5.54) and (A5.56), and noting that $\partial\bar{\Lambda}^Y/\partial n = T\partial\bar{\Lambda}^Y/\partial C^Y/(1+r)$:

$$\begin{aligned}\frac{d\bar{\Lambda}^Y}{dn} &= \frac{\partial\bar{\Lambda}^Y}{\partial W}\frac{dW}{dn} + \frac{\partial\bar{\Lambda}^Y}{\partial r}\frac{dr}{dn} + \frac{\partial\bar{\Lambda}^Y}{\partial n} \\ &= \frac{\partial\Lambda^Y}{\partial C^Y}\left[\frac{dW}{dn} + \frac{S}{1+r}\frac{dr}{dn} + \frac{T}{1+r}\right] \\ &= \frac{\partial\Lambda^Y}{\partial C^Y}\left[-k\frac{r-n}{1+r}\frac{dr}{dn} + \frac{T}{1+r}\right] \lesseqgtr 0. \end{aligned} \quad \text{(A5.106)}$$

In a dynamically efficient economy (for which $r > n$ holds) there are two effects which operate in opposite directions. The first term in square brackets on the right-hand side of (A5.106) represents the effect of fertility on the long-run interest rate. Since $dr/dn = r'dk/dn > 0$, a fall in fertility raises long-run welfare on that account. The second term in square brackets on the right-hand side of (A5.106) is the PAYG-yield effect. If fertility falls so does the rate of return on the PAYG contribution. Since the yield effect works in the opposite direction to the interest rate effect, the overall effect of a fertility change is ambiguous. If the PAYG contribution is very small ($T \approx 0$) and the economy is not close to the golden-rule point ($r \gg n$), then a drop in fertility raises long-run welfare.

Although our results are based on a highly stylized (and perhaps oversimplified) model, they nevertheless seem to bear some relationship to reality. Indeed, Auerbach and Kotlikoff (1987, ch. 11) simulate a highly detailed computable general equilibrium model for the US economy and find qualitatively very similar results: wages rise, the interest rate falls, and long-run welfare increases strongly (see their Table 11.3). In their model, households live for 75 years, labour supply is endogenous, productivity is age-dependent, households' retirement behaviour is endogenous, taxes are distorting, and demography is extremely detailed.

12.3 Extensions

12.3.1 Human capital accumulation

12.3.1.1 Human capital and growth

Following the early contributions by Arrow (1962) and Uzawa (1965), a number of authors have drawn attention to the importance of human capital accumulation for the theory of economic growth. The key papers that prompted the renewed interest in human capital in the 1980s are Romer (1986) and Lucas (1988). In this subsection we show how the Diamond-Samuelson overlapping-generations model can be extended by including the purposeful accumulation of human capital by households. We show how this overlapping-generations version of the celebrated Lucas (1988) model can give rise to endogenous growth in the economy (see also Chapter 8 above).

As in the standard model, we continue to assume that households live for two periods, but we deviate from the standard model by assuming that the household works full-time during the second period of life and divides its time between working and training during youth. Following Lucas (1988) human capital is equated to the worker's level of skill at producing goods. We denote the human capital of worker i at time t by H_t^i and assume that producers can observe each worker's skill level and will thus pay a skill-dependent wage (just as in the continuous-time model discussed in Chapter 8 above).

The lifetime utility function of a young agent who is born at time t is given in general terms by:

$$\Lambda_t^{Y,i} \equiv \Lambda^Y(C_t^{Y,i}, C_{t+1}^{O,i}). \tag{A5.107}$$

This expression incorporates the notion that the household does not value leisure and attaches no utility value to training *per se*. The household is thus only interested in improving its skills because it will improve its income later on in life. The budget identities facing the agent are:

$$C_t^{Y,i} + S_t^i = W_t H_t^i N_t^i, \tag{A5.108}$$

$$C_{t+1}^{O,i} = (1 + r_{t+1}) S_t^i + W_{t+1} H_{t+1}^i, \tag{A5.109}$$

where W_t denotes the going wage rate for an efficiency unit of labour at time t, and N_t^i is the amount of time spent working (rather than training) during youth. Since the agent has one unit of time available in each period we have by assumption that $N_{t+1}^i = 1$ (there is no third period of life so there is no point in training during the second period). The amount of training during youth is denoted by E_t^i and equals:

$$E_t^i = 1 - N_t^i \geq 0. \tag{A5.110}$$

To complete the description of the young household's decision problem we must specify how training augments the agent's skills. As a first example of a *training technology* we consider the following specification:

$$H_{t+1}^i = G(E_t^i) H_t^i, \tag{A5.111}$$

where $G' > 0 \geq G''$ and $G(0) = 1$. This specification captures the notion that there are positive but non-increasing returns to training in the production of human capital and that zero training means that the agent keeps his initial skill level.

The household chooses $C_t^{Y,i}$, $C_{t+1}^{O,i}$, S_t^i, N_t^i, and E_t^i in order to maximize lifetime utility $\Lambda_t^{Y,i}$ (given in (A5.107)) subject to the constraints (A5.108)-(A5.110), and given the training technology (A5.111), the expected path of wages W_t, and its own initial skill level H_t^i. The optimization problem can be solved in two steps. In the first step the household chooses its training level, E_t^i, in order to maximize its lifetime income, I_t^i, i.e. the present value of wage income:

$$I_t^i(E_t^i) \equiv H_t^i \left[W_t (1 - E_t^i) + \frac{W_{t+1} G(E_t^i)}{1 + r_{t+1}} \right]. \tag{A5.112}$$

The first-order condition for this optimal human capital investment problem, taking explicit account of the inequality constraint (A5.110), is:

$$\frac{dI_t^i}{dE_t^i} = H_t^i \left[-W_t + \frac{W_{t+1} G'(E_t^i)}{1 + r_{t+1}} \right] \leq 0, \quad E_t^i \geq 0, \quad E_t^i \frac{dI_t^i}{dE_t^i} = 0. \tag{A5.113}$$

This expression shows that it may very well be in the best interest of the agent not to pursue any training at all during youth. Indeed, this no-training solution will hold if the first inequality in (A5.113) is strict. Since there are non-increasing returns to training (so that $G'(0) \geq G'(E_t^i)$ for $E_t^i \geq 0$) we derive the following implication from (A5.113):

$$G'(0) < \frac{W_t(1+r_{t+1})}{W_{t+1}} \quad \Rightarrow \quad E_t^i = 0. \tag{A5.114}$$

If the training technology is not very productive ($G'(0)$ low) then the corner solution will be selected.

An internal solution with a strictly positive level of training is such that $dI_t^i/dE_t^i = 0$. After some rewriting we obtain the investment equation in arbitrage format:

$$E_t^i > 0 \quad \Rightarrow \quad 1 + r_{t+1} = \frac{W_{t+1}}{W_t} G'(E_t^i). \tag{A5.115}$$

This expression shows that in the interior optimum the agent accumulates physical and human capital such that their respective yields are equalized. By investing in physical capital during youth the agent receives a yield of $1 + r_{t+1}$ during old age (left-hand side of (A5.115)). By working a little less and training a little more during youth, the agent upgrades his human capital and gains $W_{t+1}G'(E_t^i)$ during old age. Expressed in terms of the initial investment (foregone wages in the first period) we get the yield on human capital (right-hand side of (A5.115)).

In the second step of the optimization problem the household chooses consumption for the two periods and its level of savings in order to maximize lifetime utility (A5.107) subject to its lifetime budget constraint:

$$C_t^{Y,i} + \frac{C_{t+1}^{O,i}}{1 + r_{t+1}} = I_t^i, \tag{A5.116}$$

where I_t^i is now maximized lifetime income. The savings function which results from this stage of the optimization problem can be written in general form as:

$$S_t^i = S(r_{t+1}, (1 - E_t^i)W_t H_t^i, W_{t+1}H_{t+1}^i). \tag{A5.117}$$

In order to complete the description of the decision problem of household i we must specify its initial level of human capital at birth, i.e. H_t^i in the training technology (A5.111). Following Azariadis and Drazen (1990, p. 510) we assume that each household born in period t "inherits" (is born with) the average stock of currently available knowledge at that time, i.e. $H_t^i = H_t$ on the right-hand side of (A5.111). With this final assumption it follows that all individuals in the model face the same interest rate and learning technology so that they will choose the same consumption, saving, and investment plans. We can thus drop the individual index i from here on and study the symmetric equilibrium.

We assume that there is no population growth and normalize the size of the young and old populations to unity ($L_{t-1} = L_t = 1$). Total labour supply in efficiency units is defined as the sum of efficiency units supplied by the young and the old, i.e. $N_t = (1 - E_t)H_t + H_t$. For convenience we summarize the key expressions of the (simplified) Azariadis-Drazen model below.

$$N_{t+1}k_{t+1} = S(r_{t+1}, (1-E_t)W_tH_t, W_{t+1}H_{t+1}) \tag{A5.118}$$

$$r_{t+1} + \delta = f'(k_{t+1}) \tag{A5.119}$$

$$W_t = f(k_t) - k_t f'(k_t) \tag{A5.120}$$

$$N_t = (2 - E_t)H_t \tag{A5.121}$$

$$1 + r_{t+1} = \frac{W_{t+1}}{W_t} G'(E_t) \tag{A5.122}$$

$$H_{t+1} = G(E_t)H_t, \tag{A5.123}$$

Equation (A5.118) relates saving by the representative young household to next period's stock of physical capital. Note that the capital-labour ratio is defined in terms of efficiency units of labour, i.e. $k_t \equiv K_t/N_t$. With this definition, the expressions for the wage rate and the interest rate are, respectively (A5.119) and (A5.120). Equation (A5.121) is labour supply in efficiency units, (A5.122) is the investment equation for human capital (assuming an internal solution), and (A5.123) is the accumulation for aggregate human capital in the symmetric equilibrium.

It is not difficult to show that the model allows for endogenous growth in the steady state. In the steady-state growth path the capital-labour ratio, the wage rate, the interest rate, and the proportion of time spent training during youth, are all constant over time (i.e. $k_t = k$, $W_t = W$, $r_t = r$, and $E_t = E$). The remaining variables grow at a common growth rate $\gamma \equiv G(E) - 1$. Referring the reader for a general proof to Azariadis (1993, p. 231), we demonstrate the existence of a unique steady-state growth path for the unit-elastic model for which technology is Cobb-Douglas ($y_t = k_t^{1-\epsilon_L}$) and the utility function (A5.107) is loglinear ($\Lambda_t^Y = \ln C_t^Y + (1/(1+\rho)) \ln C_{t+1}^O$). For the unit-elastic case the savings function can be written as:

$$S_t = \left[\frac{1}{2+\rho}(1-E_t)W_t - \frac{1+\rho}{2+\rho} \frac{W_{t+1}G(E_t)}{1+r_{t+1}} \right] H_t. \tag{A5.124}$$

By using (A5.124), (A5.121), and (A5.123) in (A5.118) and imposing the steady state we get an implicit relationship between E and k for which savings equals investment:

$$(2+\rho)\frac{k}{W(k)} = \frac{1}{2-E}\left[\frac{1-E}{G(E)} - \frac{1+\rho}{1+r(k)}\right]. \tag{A5.125}$$

Similarly, by using (A5.119) and (A5.121) in the steady-state we get a second expression, again relating E and k, for which the rates of return on human and physical capital are equalized:

$$[1 + r(k) =] \quad G'(E) = f'(k) + 1 - \delta. \tag{A5.126}$$

The joint determination of E and k in the steady-state growth path is illustrated in the upper panel of Fig-

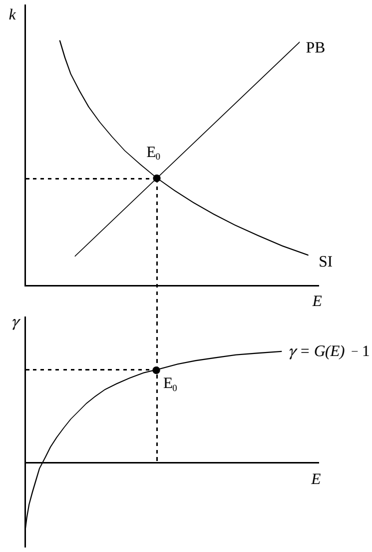

Figure 12.5: Endogenous growth and human capital formation

ure 12.5. The portfolio-balance (PB) line is upward sloping because both the production technology and the training technology exhibit diminishing returns ($f''(k) < 0$ and $G''(E) < 0$). The savings-investment (SI) line is downward sloping with Cobb-Douglas technology. The right-hand side of (A5.125) is downward sloping in both k and E. With Cobb-Douglas technology we have that $k/W(k) = (1/\epsilon_L)k^{\epsilon_L}$ which ensures that the left-hand side of (A5.125) is increasing in k. Together these result imply that SI slopes down. In the upper panel the steady state is at E_0. In the bottom panel we relate the equilibrium growth rate to the level of training.

The *engine of growth* in the Azariadis-Drazen model is clearly the training technology (A5.123) which ensures that a given steady-state *level* of training allows for a steady-state *rate of growth* in the stock of human capital. Knowledge and technical skills are *disembodied*, i.e. they do not die with the individual agents but rather they are passed on in an automatic fashion to the newborns. The newborns can then add to the stock of knowledge by engaging in training. It should be clear that endogenous growth would disappear from the model if skills were *embodied* in the agents themselves. In that case young agents would have to start all over again and "re-invent the wheel" the moment they are born.

12.3.1.2 Human capital and education

Whilst it is undoubtedly true that informal social interactions can give rise to the transmission of knowledge and skills (as in the Azariadis-Drazen (1990) model) most developed countries have had formal educational systems for a number of centuries. A striking aspect of these systems is that they are compulsory, i.e. children up to a certain age are forced by law to undergo a certain period of basic training. This prompts the question why the adoption of compulsory education has been so widespread, even in countries which otherwise strongly value their citizens' right to choose.

Eckstein and Zilcha (1994) have recently provided an ingenious answer to this question which stresses the role of parents in the transmission of human capital to their offspring. They use an extended version of the Azariadis-Drazen model and show that compulsory education may well be welfare-enhancing to the children if the parents do not value the education of their offspring to a sufficient extent. The key insight of Eckstein and Zilcha (1994) is thus that there may exist a significant *intra-family* external effect which causes parents to underinvest in their children's human capital. Note that such an effect is not present in the Azariadis-Drazen model because in that model the agent *himself* bears the cost of training during youth and reaps the benefits during old age.

We now develop a simplified version of the Eckstein-Zilcha model to demonstrate their important underinvestment result. We assume that all agents are identical. The representative parent consumes goods during youth and old age (C_t^Y and C_{t+1}^O, respectively), enjoys leisure during youth (Z_t), is retired during old age, and has $1+n$ children during the first period of life. Fertility is exogenous so that the number of children is exogenously given ($n \geq 0$). The lifetime utility function of the young agent at time t is given in general form as:

$$\Lambda_t^Y \equiv \Lambda^Y(C_t^Y, C_{t+1}^O, Z_t, O_{t+1}), \tag{A5.127}$$

where $O_{t+1} \equiv (1+n)H_{t+1}$ represents the total human capital of the agent's offspring. Since the agent has $1+n$ kids, each child gets H_{t+1} in human capital (knowledge) from its parent. There is no formal schooling system so the parent cannot purchase education services for its offspring in the market. Instead, the parent must spend (part of its) leisure time during youth to educate its children and the training function is given by:

$$H_{t+1} = G(E_t)H_t^{\beta}, \tag{A5.128}$$

where E_t is the educational effort per child, $G(\cdot)$ is the training curve (satisfying $0 < G(0) \leq 1$, $G(1) > 1$, $G' > 0 \geq G''$) and $0 < \beta \leq 1$. Equation (A5.128) is similar in format to (A5.123) but its interpretation is different. In (A5.123) H_{t+1} and E_t are chosen by and affect the same agent. In contrast, in (A5.128) the parent chooses H_{t+1} and E_t and the consequences of this choice are felt by both the parent and his/her offspring.

The agent has two units of time available during youth, one of which is supplied inelastically to the labour market (Eckstein and Zilcha, 1994, p. 343), and the other of which is spent on leisure and educational activities:

$$Z_t + (1+n)E_t = 1. \tag{A5.129}$$

The household's consolidated budget constraint is of a standard form:

$$C_t^Y + \frac{C_{t+1}^O}{1+r_{t+1}} = W_t H_t, \tag{A5.130}$$

where the left-hand side represents the present value of consumption and the right-hand side is labour income. Competitive firms hire capital, K_t, and efficiency units of labour, $N_t \equiv L_t H_t$, from the households, and the aggregate production function is $Y_t = F(K_t, N_t)$. The wage and interest rate then satisfy, respectively, $W_t = F_N(K_t, N_t)$ and $r_t + \delta = F_K(K_t, N_t)$.

The representative parent chooses C_t^Y, C_{t+1}^O, Z_t, E_t, and H_{t+1} in order to maximize lifetime utility (A5.127) subject to the training technology (A5.128), the time constraint (A5.129), and the consolidated budget constraint (A5.130). By substituting the constraints into the objective function and optimizing with respect to the remaining choice variables (C_t^Y, C_{t+1}^O, and E_t) we obtain the following first-order conditions:

$$\frac{\partial \Lambda^Y / \partial C_t^Y}{\partial \Lambda^Y / \partial C_{t+1}^O} = 1 + r_{t+1} \tag{A5.131}$$

$$\frac{\partial \Lambda^Y}{\partial O_t} G'(E_t) H_t^{\beta} - \frac{\partial \Lambda^Y}{\partial Z_t} < 0 \quad \Rightarrow \quad E_t = 0 \tag{A5.132}$$

$$\frac{\partial \Lambda^Y}{\partial O_t} G'(E_t) H_t^{\beta} - \frac{\partial \Lambda^Y}{\partial Z_t} = 0 \quad \Leftarrow \quad E_t > 0 \tag{A5.133}$$

Equation (A5.131) is the standard consumption Euler equation, which we encountered time and again, and (A5.132)-(A5.133) characterizes the optimal educational activities of the parent. The left-hand side appearing in (A5.132)-(A5.133) represents the net marginal benefit of child education. If the (marginal) costs outweigh the benefits this term is negative and the parent chooses not to engage in educational activities at all (see (A5.132)). Conversely, a strictly positive (interior) choice of E_t implies that the net marginal benefit of child education is zero. In the remainder we assume that conditions are such that $E_t > 0$ is chosen by the representative parent.

A notable feature of the parent's optimal child education rule (A5.133) is that it only contains the costs and benefits as they accrue to the parent. But if a child receives a higher level of human capital from its parents, then it will have a higher labour income and will thus be richer and enjoy a higher level of welfare. By assumption, however, the parent only cares about the level of education it passes on to its children and therefore disregards any welfare effects that operate directly on its offspring. This is the first hint of the under-investment problem. Loosely put, by disregarding some of the positive welfare effects its own educational activities have on its children, the parent does not provide "enough" education.

As was explained above, in our discussion regarding pension reform, there are several ways in which we can tackle the efficiency issue of under-investment in a more formal manner. One way would be to look for Pareto-improving policy interventions. For example, in the present context one could investigate whether a system of financial transfers to parents could be devised which (a) would induce parents to raise their child-educational activities and (b) would make no present or future generation worse off and at least one strictly better off. If such a transfer system can be found we can conclude that the status quo is inefficient and there is underinvestment.

An alternative approach, one which we pursue here, makes use of a social welfare function. Following Eckstein and Zilcha (1994, pp. 344-345) we postulate a specific form for the social welfare function (A5.68) which is linear in the lifetime utilities of present and future agents:

$$SW_0 \equiv \sum_{t=0}^{\infty} \lambda_t \Lambda_t^Y = \sum_{t=0}^{\infty} \lambda_t \Lambda^Y(C_t^Y, C_{t+1}^O, Z_t, O_{t+1}), \tag{A5.134}$$

where SW_0 is social welfare in the planning period ($t = 0$) and $\{\lambda_t\}_{t=0}^{\infty}$ is a positive monotonically decreasing sequence of weights attached to the different generations which satisfies $\sum_{t=0}^{\infty} \lambda_t < \infty$. In the social optimum, the social planner chooses sequences for consumption ($\{C_t^Y\}_{t=0}^{\infty}$ and $\{C_{t+1}^O\}_{t=0}^{\infty}$), the stocks of physical and human capital ($\{K_{t+1}\}_{t=0}^{\infty}$ and $\{H_{t+1}\}_{t=0}^{\infty}$), and the educational effort ($\{E_t\}_{t=0}^{\infty}$) in order to maximize (A5.134) subject to the training technology (A5.128), the time constraint (A5.129), and the following resource constraint:

$$C_t^Y + \frac{C_t^O}{1+n} + (1+n)k_{t+1} = F(k_t, H_t) + (1-\delta)k_t, \tag{A5.135}$$

where $k_t \equiv K_t/L_t$ is capital per worker.

The Lagrangian associated with the social optimization problem is given by:

$$\begin{aligned}\mathcal{L}_0 \equiv & \sum_{t=0}^{\infty} \lambda_t \Lambda^Y(C_t^Y, C_{t+1}^O, Z_t, (1+n)H_{t+1}) \\ & - \sum_{t=0}^{\infty} \mu_t^R \left[C_t^Y + \frac{C_t^O}{1+n} + (1+n)k_{t+1} - F(k_t, H_t) - (1-\delta)k_t \right] \\ & - \sum_{t=0}^{\infty} \mu_t^T \left[Z_t + (1+n)E_t - 1 \right] - \sum_{t=0}^{\infty} \mu_t^H \left[H_{t+1} - G(E_t) H_t^{\beta} \right],\end{aligned} \tag{A5.136}$$

where μ_t^R, μ_t^T, and μ_t^H are the Lagrange multipliers associated with, respectively, the resource constraint, the time constraint, and the training technology.

After some manipulation we find the following first-order conditions for the social optimum for $t = 0, ..., \infty$:

$$\frac{\partial \mathcal{L}_0}{\partial C_t^Y} = \lambda_t \frac{\partial \Lambda^Y}{\partial C_t^Y} - \mu_t^R = 0, \tag{A5.137}$$

$$\frac{\partial \mathcal{L}_0}{\partial C_{t+1}^O} = \lambda_t \frac{\partial \Lambda^Y}{\partial C_{t+1}^O} - \frac{\mu_{t+1}^R}{1+n} = 0, \tag{A5.138}$$

$$\frac{\partial \mathcal{L}_0}{\partial Z_t} = \lambda_t \frac{\partial \Lambda^Y}{\partial Z_t} - \mu_t^T = 0, \tag{A5.139}$$

$$\frac{\partial \mathcal{L}_0}{\partial E_t} = -(1+n)\mu_t^T + \mu_t^H G'(E_t) H_t^{\beta} = 0, \tag{A5.140}$$

$$\frac{\partial \mathcal{L}_0}{\partial H_{t+1}} = (1+n)\lambda_t \frac{\partial \Lambda^Y}{\partial O_t} - \mu_t^H + \beta \mu_{t+1}^H G(E_{t+1}) H_{t+1}^{\beta-1} + \mu_{t+1}^R F_N(k_{t+1}, H_{t+1}) = 0, \tag{A5.141}$$

$$\frac{\partial \mathcal{L}_0}{\partial k_{t+1}} = -(1+n)\mu_t^R + \mu_{t+1}^R \left[F_K(k_{t+1}, H_{t+1}) + (1-\delta)\right] = 0. \tag{A5.142}$$

By combining (A5.137)-(A5.138) and (A5.142) we obtain the socially optimal consumption Euler equation:

$$(1+n)\frac{\mu_t^R}{\mu_{t+1}^R} = \frac{\partial \Lambda^Y(\hat{x}_t)/\partial C_t^Y}{\partial \Lambda^Y(\hat{x}_t)/\partial C_{t+1}^O} = F_K(\hat{k}_{t+1}, \hat{H}_{t+1}) + (1-\delta)\left[\equiv 1 + \hat{r}_{t+1}\right], \tag{A5.143}$$

where $x_t \equiv (C_t^Y, C_{t+1}^O, Z_t, O_{t+1})$, hats ("$\hat{\ }$") denote socially optimal values, and $\hat{r}_{t+1}$ thus represents the socially optimal interest rate. Similarly, by using (A5.137) for period $t+1$ and (A5.138) we obtain an expression determining the socially optimal division of consumption between old and young agents living at the same time:

$$\frac{\lambda_{t+1}}{\lambda_t} = (1+n)\frac{\partial \Lambda^Y(\hat{x}_t)/\partial C_{t+1}^O}{\partial \Lambda^Y(\hat{x}_{t+1})/\partial C_{t+1}^Y}. \tag{A5.144}$$

This expression shows that, by adopting a particular sequence of generational weights $\{\lambda_t\}_{t=0}^{\infty}$, the social planner in fact chooses the generational consumption profile between the young and the old (see Calvo and Obstfeld, 1988, p. 417).

Intermezzo 12.1

Dynamic consistency. There are some subtle issues that must be confronted when using a social welfare function like (A5.134). If we are to attach any importance to the social planning exercise we must assume that either one of the following two situations holds:

Commitment The policy maker only performs the social planning exercise once and can credibly commit never to re-optimize. Economic policy is a one-shot event and no further restrictions on the generational weights are needed.

Consistency The policy maker can re-optimize at any time but the generational weights are such that the socially optimal plan is dynamically consistent, i.e. the mere evolution of time itself does not make the planner change his mind.

This intermezzo shows how dynamic consistency can be guaranteed in the absence of credible commitment. We study dynamic consistency in the context of the standard

Diamond-Samuelson model. The social welfare function in the planning period 0 is given in general terms by:

$$SW_0 \equiv \lambda_{0,-1}\Lambda^Y(C_{-1}^Y, C_0^O) + \sum_{\tau=0}^{\infty} \lambda_{0,\tau}\Lambda^Y(C_\tau^Y, C_{\tau+1}^O), \tag{A}$$

where $\lambda_{0,\tau}$ is the weight that the planner in time 0 attaches to the lifetime utility of the generation born in period τ (for $\tau = -1, 0, 1, 2, ...$). The social planner chooses sequences for consumption during youth and old age ($\{C_\tau^Y\}_{\tau=0}^{\infty}$ and $\{C_\tau^O\}_{\tau=0}^{\infty}$) and the capital stock ($\{k_{\tau+1}\}_{\tau=0}^{\infty}$) in order to maximize social welfare (A) subject to the resource constraint:

$$C_\tau^Y + \frac{C_\tau^O}{1+n} + (1+n)k_{\tau+1} = f(k_{\tau+1}) + (1-\delta)k_\tau, \tag{B}$$

and taking the initial capital stock, k_0, as given. Obviously, since the past cannot be undone, consumption during youth of the initially old generation (C_{-1}^Y) is also taken as given. After some straightforward computations we find the following first-order conditions characterizing the social optimum:

$$\frac{\partial\Lambda^Y(\hat{x}_\tau)/\partial C_\tau^Y}{\partial\Lambda^Y(\hat{x}_\tau)/\partial C_{\tau+1}^O} = f'(\hat{k}_{\tau+1}) + 1 - \delta, \tag{C}$$

$$\frac{\partial\Lambda^Y(\hat{x}_\tau)/\partial C_\tau^Y}{\partial\Lambda^Y(\hat{x}_{\tau-1})/\partial C_\tau^O} = \frac{(1+n)\lambda_{0,\tau-1}}{\lambda_{0,\tau}}, \quad \tau = 0, 1, 2, ... \tag{D}$$

where $x_\tau \equiv (C_\tau^Y, C_{\tau+1}^O)$ and hats denote socially optimal values.

Now consider a planner who performs the social planning exercise at some later planning period $t > 0$. The social welfare function in planning period t is:

$$SW_t \equiv \lambda_{t,t-1}\Lambda^Y(C_{t-1}^Y, C_t^O) + \sum_{\tau=t}^{\infty} \lambda_{t,\tau}\Lambda^Y(C_\tau^Y, C_{\tau+1}^O), \tag{E}$$

where $\lambda_{t,\tau}$ is the weight that the planner in time t attaches to the lifetime utility of the generation born in period τ (for $\tau = t-1, t, t+1, t+2, ...$). The social planner chooses sequences for consumption during youth and old age ($\{C_\tau^Y\}_{\tau=t}^{\infty}$ and $\{C_\tau^O\}_{\tau=t}^{\infty}$) and the capital stock ($\{k_{\tau+1}\}_{\tau=t}^{\infty}$) in order to maximize social welfare (E) subject to the resource constraint (B). The (interesting) first-order conditions consist of (C) and:

$$\frac{\partial\Lambda^Y(\hat{x}_\tau)/\partial C_\tau^Y}{\partial\Lambda^Y(\hat{x}_{\tau-1})/\partial C_\tau^O} = \frac{(1+n)\lambda_{t,\tau-1}}{\lambda_{t,\tau}}, \quad \tau = t, t+1, t+2, ... \tag{F}$$

The crucial thing to note is that conditions (D) and (F) overlap for the time interval $\tau = t, t+1, t+2, ...$ The sequences $\{C_\tau^Y\}_{\tau=0}^{t-1}$, $\{C_\tau^O\}_{\tau=0}^{t-1}$, and $\{k_{\tau+1}\}_{\tau=0}^{t-1}$ are chosen by the planner at time 0 but taken as given ("water under the bridge") by the planner at time t. But the sequences $\{C^Y_\tau\}_{\tau=t}^{\infty}$, $\{C^O_\tau\}_{\tau=t}^{\infty}$, and $\{k_{\tau+1}\}_{\tau=t}^{\infty}$ are chosen by both planners. Unless the planner at time 0 can commit to his plan (and thus can stop any future planner from re-optimizing the then relevant social welfare function), the sequences chosen by the planners at time 0 and at time t will not necessarily be the same. If they are not the same we call the social plan dynamically inconsistent.

Following the insights of Strotz (1956), Burness (1976) has derived conditions on the admissible pattern of generational weights, $\lambda_{t,\tau}$, that ensure that the optimal social plan is dynamically consistent. Comparing (D) and (F) reveals that dynamic consistency requires the following condition to hold for any planning period t:

$$\frac{\lambda_{t,\tau-1}}{\lambda_{t,\tau}} = \frac{\lambda_{0,\tau-1}}{\lambda_{0,\tau}}, \quad \tau = t, t+1, t+2, \ldots \tag{G}$$

Condition (G) means that $\lambda_{t,\tau}$ must be multiplicatively separable in time (τ) and the planning date (t), i.e. it must be possible to write $\lambda_{t,\tau} = g(t)\lambda_\tau$, where g is some function of t. A simple example of such a multiplicatively separable function is:

$$\lambda_{t,\tau} = \left(\frac{1}{1+\lambda}\right)^{\tau-t}, \tag{H}$$

where $\lambda > 0$ is the planner's constant discount rate. By using (H) we normalize the weight attached to the young in the planning period to unity ($\lambda_{t,t} = 1$). It follows necessarily, that in order to preserve dynamic consistency, there must be reverse discounting applied to the old generation in the planning period. Indeed, the dynamic consistency requirement (G) combined with (H) implies $\lambda_{t,\tau-1}/\lambda_{t,\tau} = 1+\lambda$ so that $\lambda_{t,t-1} = (1+\lambda)\lambda_{t,t} = 1+\lambda$. Calvo and Obstfeld (1988) apply this notion of reverse discounting in the context of the continuous-time Blanchard (1985) model of overlapping generations.

****

Finally, by using (A5.138)-(A5.140), and (A5.144) in (A5.141) we can derive the following expression:

$$\begin{aligned}\frac{\partial \Lambda^Y(\hat{x}_t)}{\partial Z_t} = G'(\hat{E}_t)\hat{H}_t^{\beta}\Bigg[&\frac{\partial \Lambda^Y(\hat{x}_t)}{\partial O_t} \\ &+ \frac{\partial \Lambda^Y(\hat{x}_t)}{\partial C^O_{t+1}} F_N(\hat{k}_{t+1}, \hat{H}_{t+1}) \\ &+ \frac{\beta(1+n)\hat{H}_{t+2}}{G'(\hat{E}_{t+1})\hat{H}_{t+1}^{1+\beta}} \frac{\partial \Lambda^Y(\hat{x}_t)}{\partial C^O_{t+1}} \frac{\partial \Lambda^Y(\hat{x}_{t+1})/\partial Z_{t+1}}{\partial \Lambda^Y(\hat{x}_{t+1})/\partial C^Y_{t+1}}\Bigg].\end{aligned} \tag{A5.145}$$

In the social optimum the marginal social cost of educational activities (left-hand side of (A5.145)) should be equated to the marginal social benefits of these activities (right-hand side of (A5.145)). The marginal social costs are just the value of leisure time of the parent, but the marginal social benefits consist of three terms. All three terms on the right-hand side of (A5.145) contain the expression $G'(\hat{E}_t)\hat{H}_t^{\beta}$, which represents the marginal product of time spent on educational activities in the production of human capital (see (A5.128)). The first line on the right-hand side of (A5.145) is the "own" effect of educational activities on the parent's utility. This term also features in the first-order condition for the privately optimal (internal) child-education decision, namely (A5.133). The second and third lines show the additional effects that the social planner takes into account in determining the optimal level of child education. The second line represents the effect of the parent's decision on the children's earnings: by endowing each child with more human capital they will have a higher skill level and thus command a higher wage. The third line represents the impact of the parent's investment on the children's incentives to provide education for their own children (i.e. the parent's grandchildren).

Eckstein and Zilcha are able to prove that (a) the competitive allocation is suboptimal, and (b) that under certain reasonable assumptions regarding the lifetime utility function there is underinvestment of human capital. Intuitively, this result obtains because the parents ignore some of the benefits of educating their children (1994, pp. 345-346). To internalize the externality in the human capital investment process, the policy maker would need to construct a rule such that the parent's decision regarding educational activities would take account of the effect on the children's wages and education efforts. As Eckstein and Zilcha argue, it is not likely that such a complex rule can actually be instituted in the real world. For that reason, the institution of compulsory education, which is practicable, may well achieve a welfare improvement over the competitive allocation because it imposes a minimal level of educational activities on parents (1994, pp. 341, 346).

12.3.2 Public investment

At least since the seminal work by Arrow and Kurz (1970), macroeconomists have known that the stock of public infrastructure is an important factor determining the productive capacity of an economy. Somewhat surprisingly, however, the public capital stock has played only a relatively minor role in the literature up until recently. This unfortunate state of affairs changed dramatically a decade ago when the pathbreaking and provocative empirical research of Aschauer (1989, 1990b) triggered a veritable boom in the econometric literature on public investment (see Gramlich, 1994 for an excellent survey of this literature). Aschauer (1989) showed that public capital exerts a strong positive effect on the productivity of private capital and argued that the slowdown in productivity growth in the US since the early 1970s is due to a shortage of investment in public infrastructure. Indeed, his estimates suggest implicit rates of return on government capital of 100% or more, values which are seen as highly implausible by many commentators (see e.g. Gramlich, 1994, p. 1186). Although Aschauer's results were controversial and many subsequent studies have questioned their robustness, it is nevertheless fair to conclude that economists generally support the notion that public capital is indeed productive.

In this subsection we show how productive public capital can be introduced into the Diamond-Samuelson model. We show how the dynamic behaviour of the economy is affected if the government adopts a constant infrastructural investment policy. Finally, we study how the socially optimal capital stock can be determined. To keep things simple we assume that labour supply is exogenous, and that the government has access to lump-sum taxes. We base our discussion in part on Azariadis (1993, pp. 336-340).

Prototypical examples of government capital are things like roads, bridges, airports, hospitals, etc., which all have the stock dimension. Just as with the private capital stock, the public capital stock is gradually built up by means of infrastructural investment and gradually wears down because depreciation takes place. Denoting the stock of government capital by G_t we have:

$$G_{t+1} - G_t = I_t^G - \delta_G G_t, \tag{A5.146}$$

where I_t^G is infrastructural investment and $0 < \delta_G < 1$ is the depreciation rate of public capital. Assuming that the population grows at a constant rate as in (A5.21), per capita public capital evolves according to:

$$(1+n)g_{t+1} = i_t^G + (1-\delta_G)g_t, \tag{A5.147}$$

where $g_t \equiv G_t/L_t$ and $i_t^G \equiv I_t/L_t$.

We assume that public capital enters the production function of the private sector, i.e. instead of (A5.10) we have:

$$Y_t = F(K_t, L_t, g_t), \tag{A5.148}$$

where we assume that $F(\cdot)$ is linearly homogeneous in the *private* production factors, K_t and L_t. This means that we can express per capita output ($y_t \equiv Y_t/L_t$) as follows:

$$y_t = f(k_t, g_t), \tag{A5.149}$$

where $k_t \equiv K_t/L_t$ and $f(k_t, g_t) \equiv F(K_t/L_t, 1, g_t)$. We make the following set of assumptions regarding technology:

$$f_k \equiv \frac{\partial f}{\partial k_t} > 0, \quad f_g \equiv \frac{\partial f}{\partial g_t} > 0, \tag{P1}$$

$$f_{kk} \equiv \frac{\partial^2 f}{\partial k_t^2} < 0, \quad f_{gg} \equiv \frac{\partial^2 f}{\partial g_t^2} < 0, \tag{P2}$$

$$f(0, g_t) = f(k_t, 0) = 0, \tag{P3}$$

$$f_{kg} \equiv \frac{\partial^2 f}{\partial k_t \partial g_t} > 0, \tag{P4}$$

$$f_g - k f_{kg} > 0. \tag{P5}$$

Private and public capital both feature positive (property (P1)) but diminishing marginal productivity (property (P2)). Both types of capital are essential in production, i.e. output is zero if either input is zero (property (P3)). Finally, properties (P4)-(P5) ensure that public capital is complementary with both private capital and labour. This last implication can be seen by noting that perfectly competitive firms hire capital and labour according to the usual rental expressions $r_t + \delta = F_K(K_t, L_t, g_t)$ and $W_t = F_L(K_t, L_t, g_t)$. These can be expressed in per capita form as:

$$r_t = r(k_t, g_t) \equiv f_k(k_t, g_t) - \delta, \tag{A5.150}$$

$$W_t = W(k_t, g_t) \equiv f(k_t, g_t) - k_t f_k(k_t, g_t), \tag{A5.151}$$

where $0 < \delta < 1$ is the depreciation rate of the private capital stock. We can deduce from Properties (P4)-(P5) that $r_k \equiv \partial r/\partial k_t < 0$ and $W_k \equiv \partial W/\partial k_t > 0$ (as in the standard model) and $r_g \equiv \partial r/\partial g_t > 0$ and $W_g \equiv \partial W/\partial g_t > 0$ (public capital positively affects both the interest rate and the wage rate). To illustrate the key properties of the model we shall employ a simple Cobb-Douglas production function below of the form $Y_t = K_t^{1-\epsilon_L} L_t^{\epsilon_L} g_t^{\eta}$, with $0 < \eta < \epsilon_L < 1$. This function satisfies properties (P1)-(P5) and implies $W(k_t, g_t) = \epsilon_L k_t^{1-\epsilon_L} g_t^{\eta}$ and $r(k_t, g_t) = (1-\epsilon_L) k_t^{-\epsilon_L} g_t^{\eta} - \delta$.

To keep things simple, we assume that the representative young agent has the following lifetime utility function:

$$\Lambda_t^Y = \ln C_t^Y + \frac{1}{1+\rho} \ln C_{t+1}^O. \tag{A5.152}$$

The budget identities facing the household are:

$$C_t^Y + S_t = W_t - T_t^Y, \tag{A5.153}$$

$$C_{t+1}^O = (1 + r_{t+1}) S_t - T_{t+1}^O, \tag{A5.154}$$

where T_t^Y and T_{t+1}^O are lump-sum taxes paid by the agent during youth and old age respectively. The consolidated budget constraint is:

$$\hat{W}_t \equiv W_t - T_t^Y - \frac{T_{t+1}^O}{1 + r_{t+1}} = C_t^Y + \frac{C_{t+1}^O}{1 + r_{t+1}}, \tag{A5.155}$$

where $\hat{W}_t$ is after-tax non-interest lifetime income. The optimal household choices are $C_t^Y = c\hat{W}_t$ and $C_{t+1}^O/(1 + r_{t+1}) = (1-c)\hat{W}_t$, where $c \equiv (1+\rho)/(2+\rho)$. The savings function can then be written as follows:

$$S_t \equiv S(W_t, r_{t+1}, T_t^Y, T_{t+1}^O) = (1-c)\left(W_t - T_t^Y\right) + c\frac{T_{t+1}^O}{1+r_{t+1}}. \quad \text{(A5.156)}$$

It follows that, ceteris paribus, lump-sum taxes during youth reduce private saving whilst taxes during old age increase saving. As before, private saving by the young is next period's stock of private capital, i.e. $L_t S_t = K_{t+1}$. In per capita form we have:

$$S_t = (1+n)k_{t+1}. \quad \text{(A5.157)}$$

The government budget constraint is very simple and states that government infrastructural investment (I_t^G) is financed by tax receipts from the young and the old, i.e. $I_t^G = L_t T_t^Y + L_{t-1} T_t^O$ which can be written in per capita form as:

$$i_t^G = T_t^Y + \frac{T_t^O}{1+n}. \quad \text{(A5.158)}$$

We now have a complete description of the economy. The key expressions are the accumulation identity for the public capital stock (A5.147), the government budget constraint (A5.158), and the accumulation expression for private capital. The latter can be written in the following format by using (A5.150), (A5.151), and (A5.156) in (A5.157):

$$(1+n)k_{t+1} = (1-c)\left[W(k_t, g_t) - T_t^Y\right] + \frac{cT_{t+1}^O}{1+r(k_{t+1}, g_{t+1})}. \quad \text{(A5.159)}$$

Once a path for public investment and a particular financing method are chosen, (A5.158) and (A5.159) describe the dynamical evolution of the public and private capital stocks. We derive the phase diagram for the case of Cobb-Douglas technology and a constant public investment policy (so that $i_t^G = i^G$ for all t) financed by taxes on only the young generations (so that $T_t^Y = i^G$ and $T_t^O = 0$ for all t). The consequences of alternative assumptions regarding financing are left as an exercise for the reader.

The phase diagram has been drawn in Figure 12.6. The GE line is the graphical representation of (A5.147) for the constant public investment policy $i_t^G = i^G$, i.e. along the line we have $g_{t+1} = g_t$. The GE line is horizontal and defines a unique steady-state equilibrium value for the stock of public capital equal to $g = i^G/(n+\delta_G)$. The dynamics for public capital are derived from the rewritten version of (A5.147):

$$g_{t+1} - g_t = \frac{i^G - (n+\delta_G)g_t}{1+n} = -\frac{n+\delta_G}{1+n}\left[g_t - g\right], \quad \text{(A5.160)}$$

from which we conclude that for points above (below) the GE line, $g_t > g$ ($< g$) and the public capital stock falls (rises) over time, $g_{t+1} < g_t$ ($> g_t$) . This (stable) dynamic pattern has been illustrated with

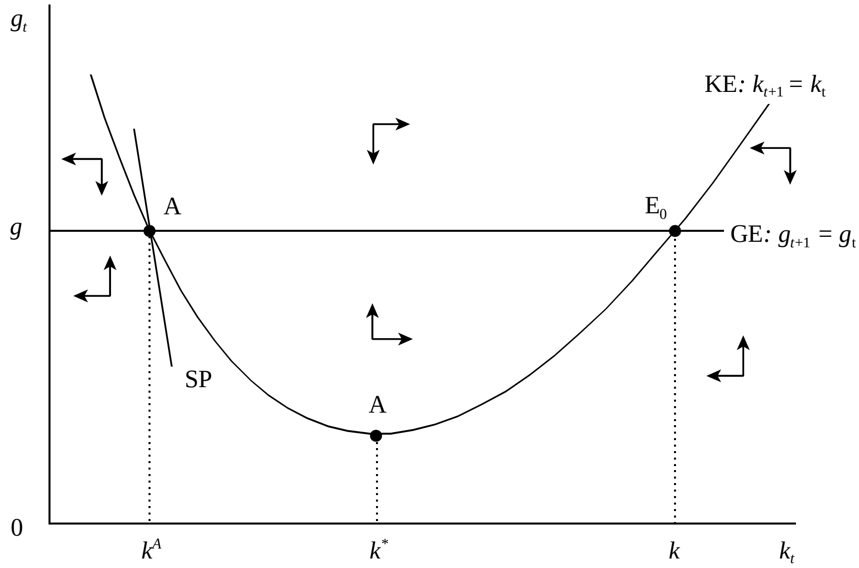

Figure 12.6: Public and private capital

vertical arrows in Figure 12.6.

The KE line in Figure 12.6 is the graphical representation of (A5.159), with the constant investment policy and the financing assumption both substituted in and imposing the steady state, $k_{t+1} = k_t$. For the Cobb-Douglas technology, the KE line has the following form:

$$g_t = \left(\frac{1+n}{\epsilon_L(1-c)}\right)^{1/\eta} \left[k_t^{\epsilon_L} + i^G \frac{1-c}{1+n} k_t^{\epsilon_L - 1}\right]^{1/\eta}, \tag{A5.161}$$

from which we derive that $\lim_{k_t \to 0} g_t = \lim_{k_t \to \infty} g_t = \infty$ and that g_t reaches its minimum value along the KE curve for $k_t = k^*$, where k^* is defined as:

$$k^* \equiv i^G \frac{1-c}{1+n} \frac{1-\epsilon_L}{\epsilon_L}. \tag{A5.162}$$

Hence, the KE line is as drawn in Figure 12.6. There are two steady-state equilibria (at A and E_0, respectively). The dynamics of the private capital stock are obtained by rewriting (A5.159) as:

$$k_{t+1} - k_t = \frac{1-c}{1+n} \left(\epsilon_L k_t^{1-\epsilon_L} g_t^{\eta} - i^G\right) - k_t, \tag{A5.163}$$

and noting that $\partial [k_{t+1} - k_t] / \partial g_t > 0$. Hence, since the wage rate increases with public capital and future consumption is a normal good, private saving increases with g_t. Hence, the capital stock is increasing (decreasing) over time for points above (below) the KE line. These dynamic forces have been illustrated with horizontal arrows in Figure 12.6.

It follows from the configuration of arrows (and from a formal local stability analysis of the linearized model) that the low-private-capital equilibrium at A is a saddle point whereas the high-private-capital equilibrium at E_0 is a stable node. For the latter equilibrium it holds that, regardless of the initial stocks of private and public capital, provided the economy is close enough to E_0 it will automatically return to E_0.

What about the steady-state equilibrium at A? Is it stable or unstable? In the typical encounters that we have had throughout this book with two-dimensional saddle-point equilibria, we called such equilibria stable because there always was one predetermined and one non-predetermined variable. By letting the non-predetermined variable jump onto the saddle path, stability was ensured. For example, in Chapter 4 we studied the Ramsey growth model and showed that the capital stock and the consumption are, respectively, the predetermined and jumping variable. In the present application, however, both K and G are predetermined variables so neither can jump. Only if the initial stocks of private and public capital by pure coincidence happen to lie on the saddle path (SP in Figure 12.6), will the equilibrium at A eventually be reached given the constant investment policy employed by the government. Appealing to the Samuelsonian correspondence principle we focus attention in the remainder of this subsection on the truly stable equilibrium at E_0.

Now consider what happens if the government increases its public investment. It follows from, respectively (A5.160) and (A5.161), that both the GE and KE lines shift up. Clearly, the higher public investment level will lead to a higher long-run stock of public capital, i.e. $dg/di^G = 1/(n+\delta_G) > 0$. The long-run effect on the private capital stock is ambiguous and depends on the relative scarcity of public capital. By imposing the steady state in (A5.163) and differentiating we obtain:

$$\left[1 - \frac{1-c}{1+n} W_k\right] \frac{dk}{di^G} = \frac{1-c}{1+n} \left[W_g \frac{dg}{di^G} - 1\right], \tag{A5.164}$$

where the term in square brackets on the left-hand side is positive because the model is outright stable around the initial steady-state equilibrium E_0.[17] The first term in square brackets on the right-hand side represents the positive effect on the pre-tax wage of the young households whilst the second term is the negative tax effect. Since $W_g = \eta W/g$, $W = \epsilon_L y$, and $g = i^G/(n+\delta_G)$, it follows from (A5.164) that the steady-state private capital stock rises (falls) as a result of the shock if $i^G/y < \eta\epsilon_L$ ($> \eta\epsilon_L$), i.e. if public capital is initially relatively scarce (abundant).

12.3.2.1 Modified golden rules

Now that we have established the macroeconomic effects of public capital, we can confront the equally important question regarding the socially optimal amount of public infrastructure. Just as in the previous subsection on education, we study this issue by computing the public investment plan that a social

[17]Recall that for a constant level of public capital, the model is stable provided the following stability condition is satisfied around the initial steady state, E_0:

$$0 < \frac{\partial k_{t+1}}{\partial k_t} \equiv \frac{1-c}{1+n} W_k < 1.$$

planner would choose. Following Calvo and Obstfeld (1988, p. 414) and Diamond (1973, p. 219) we assume that the social welfare function takes the following Benthamite form (see also Chapter 9 above):

$$SW_0 \equiv \left(\frac{1+n}{1+\rho_G}\right)^{-1} \Lambda^Y(C^Y_{-1}, C^O_0) + \sum_{t=0}^{\infty} \left(\frac{1+n}{1+\rho_G}\right)^t \Lambda^Y(C^Y_t, C^O_{t+1}), \tag{A5.165}$$

where we assume that $\rho_G < n$. Equation (A5.165) is a special case of (A5.134) with the generational weight set equal to $\lambda_t \equiv [(1+n)/(1+\rho_G)]^t$. This means that the social planner discounts the lifetime utility of generations at a constant rate ρ_G which may or may not be equal to the rate employed by the agents to discount their own periodic utility (namely ρ). The social planner chooses sequences for consumption for young and old ($\{C^Y_t\}_{t=0}^{\infty}$ and $\{C^O_t\}_{t=0}^{\infty}$), the per capita stocks of public and private capital ($\{g_{t+1}\}_{t=0}^{\infty}$ and $\{k_{t+1}\}_{t=0}^{\infty}$), in order to maximize (A5.165) subject to the following resource constraint:

$$C^Y_t + \frac{C^O_t}{1+n} + (1+n)\left[k_{t+1} + g_{t+1}\right] = f(k_t, g_t) + (1-\delta)k_t + (1-\delta_G)g_t, \tag{A5.166}$$

and taking as given k_0 and g_0. The Lagrangian associated with the social optimization problem is given by:

$$\begin{aligned}\mathcal{L}_0 \equiv & \left(\frac{1+n}{1+\rho_G}\right)^{-1} \Lambda^Y(C^Y_{-1}, C^O_0) + \sum_{t=0}^{\infty} \left(\frac{1+n}{1+\rho_G}\right)^t \Lambda^Y(C^Y_t, C^O_{t+1}) \\ & - \sum_{t=0}^{\infty} \mu^R_t \Big[C^Y_t + \frac{C^O_t}{1+n} + (1+n)\left[k_{t+1} + g_{t+1}\right] - f(k_t, g_t) \\ & - (1-\delta)k_t - (1-\delta_G)g_t \Big],\end{aligned} \tag{A5.167}$$

where μ^R_t is the Lagrange multiplier associated with the resource constraint.

After some manipulation we find the following first-order conditions for the social optimum for $t = 0, ..., \infty$:

$$\frac{\partial \mathcal{L}_0}{\partial C^Y_t} = \left(\frac{1+n}{1+\rho_G}\right)^t \frac{\partial \Lambda^Y(x_t)}{\partial C^Y_t} - \mu^R_t = 0, \tag{A5.168}$$

$$\frac{\partial \mathcal{L}_0}{\partial C^O_t} = \left(\frac{1+n}{1+\rho_G}\right)^{t-1} \frac{\partial \Lambda^Y(x_{t-1})}{\partial C^O_t} - \frac{\mu^R_t}{1+n} = 0, \tag{A5.169}$$

$$\frac{\partial \mathcal{L}_0}{\partial g_{t+1}} = -(1+n)\mu^R_t + \mu^R_{t+1}\left[f_g(k_{t+1}, g_{t+1}) + 1 - \delta_G\right] = 0, \tag{A5.170}$$

$$\frac{\partial \mathcal{L}_0}{\partial k_{t+1}} = -(1+n)\mu^R_t + \mu^R_{t+1}\left[f_k(k_{t+1}, g_{t+1}) + 1 - \delta\right] = 0, \tag{A5.171}$$

where $x_t \equiv (C^Y_t, C^O_{t+1})$. By combining (A5.168)-(A5.171) to eliminate the Lagrange multipliers we find some intuitive expressions characterizing the social optimum:

$$\frac{\partial \Lambda^Y(\hat{x}_t)/\partial C_t^Y}{\partial \Lambda^Y(\hat{x}_t)/\partial C_{t+1}^O} = f_k(\hat{k}_{t+1}, \hat{g}_{t+1}) + 1 - \delta = f_g(\hat{k}_{t+1}, \hat{g}_{t+1}) + 1 - \delta_G, \tag{A5.172}$$

$$\frac{\partial \Lambda^Y(\hat{x}_t)/\partial C_t^Y}{\partial \Lambda^Y(\hat{x}_{t-1})/\partial C_t^O} = 1 + \rho_G, \tag{A5.173}$$

where hatted variables once again denote socially optimal values. The first equality in (A5.172) is the socially optimal consumption Euler equation calling for an equalization of, on the one hand, the marginal rate of substitution between present and future consumption and, on the other hand, the socially optimal gross interest factor, $1 + \hat{r}_{t+1}$, where $\hat{r}_{t+1} \equiv f_k(\hat{k}_{t+1}, \hat{g}_{t+1}) - \delta$. The second equality in (A5.172) says that the socially optimal stock of public capital per worker should be such that the yields on private and public capital are equalized, i.e. $\hat{g}_{t+1}$ should be set in such a way that $\hat{r}_{t+1}^G = \hat{r}_{t+1}$, where $\hat{r}_{t+1}^G \equiv f_g(\hat{k}_{t+1}, \hat{g}_{t+1}) - \delta_G$. Finally, equation (A5.173) determines the socially optimal *intra*temporal division of consumption. Its intuitive meaning, and especially the interplay between the agent's and the planner's discount rate, can best be understood by considering the case of intertemporally separable preferences (which has been used throughout this chapter). By using $\Lambda_t^Y(x_t) \equiv U(C_t^Y) + (1+\rho)^{-1} U(C_{t+1}^O)$ we can rewrite (A5.173) in terms of the agent's felicity function ($U(\cdot)$) and the pure rate of time preference (ρ):

$$\frac{U'(\hat{C}_t^Y)}{U'(\hat{C}_t^O)} = \frac{1+\rho_G}{1+\rho}. \tag{A5.174}$$

It follows from (A5.174) that if the planner's discount rate exceeds (falls short of) the agent's rate of time preference, $\rho_G > \rho$ ($< \rho$), then the social planner ensures that $U'(\hat{C}_t^Y)$ exceeds (falls short of) $U'(\hat{C}_t^O)$, and thus (since $U'' < 0$) that $\hat{C}_t^Y$ falls short of (exceeds) $\hat{C}_t^O$. If $\rho_G = \rho$, the planner chooses the egalitarian solution ($\hat{C}_t^O = \hat{C}_t^Y$).

Intermezzo 12.2

Calvo-Obstfeld two-step procedure. Calvo and Obstfeld (1988) have shown that with intertemporally separable preferences, the social planning problem can be solved in two stages. In the first stage, the planner solves a static problem and in the second stage a dynamic problem is solved. Their procedure works as follows. Aggregate consumption at time τ, expressed per worker, is defined as:

$$C_\tau \equiv C_\tau^Y + \frac{1}{1+n} C_\tau^O \tag{A}$$

With intertemporally separable preferences (and ignoring a constant like $U(C_{-1}^Y)$) the social welfare function in period t can be rewritten as:

$$\begin{aligned} SW_t &\equiv \frac{1+\rho_G}{(1+n)(1+\rho)} U\left(C_t^O\right) \\ &\quad + \sum_{\tau=t}^{\infty} \left(\frac{1+n}{1+\rho_G}\right)^{\tau-t} \left[U\left(C_\tau^Y\right) + \left(\frac{1}{1+\rho}\right) U\left(C_{\tau+1}^O\right)\right] \\ &= \sum_{\tau=t}^{\infty} \left(\frac{1+n}{1+\rho_G}\right)^{\tau-t} \left[U\left(C_\tau^Y\right) + \frac{1+\rho_G}{(1+n)(1+\rho)} U\left(C_\tau^O\right)\right], \end{aligned} \tag{B}$$

where the term in square brackets in (B) now contains the weighted felicity levels of old and young agents living in the same time period. The special treatment of period-t felicity of the old is to preserve dynamic consistency (see the Intermezzo above). We can now demonstrate the two-step procedure.

In the first step, the social planner solves the static problem of dividing a given level of aggregate consumption, C_τ, over the generations that are alive at that time:

$$\bar{U}(C_\tau) \equiv \max_{\{C_\tau^Y, C_\tau^O\}} \left[U\left(C_\tau^Y\right) + \frac{1+\rho_G}{(1+n)(1+\rho)} U\left(C_\tau^O\right)\right], \quad \text{s.t. (A)}, \tag{C}$$

where $\bar{U}(C_\tau)$ is the (indirect) social felicity function. The first-order condition associated with this optimization problem is:

$$\frac{U'(C_\tau^Y)}{U'(C_\tau^O)} = \frac{1+\rho_G}{1+\rho}, \tag{D}$$

which is the same as (A5.174). Furthermore, by differentiating (C) and using (A) and (D) we find the familiar envelope property:

$$\bar{U}'(C_\tau) \equiv \frac{d\bar{U}(C_\tau)}{dC_\tau} = U'(C_\tau^Y). \tag{E}$$

For the special case of logarithmic preferences, for example, individual felicity is $U(x) \equiv \ln x$ and the social felicity function would take the following form:

$$\begin{aligned} \bar{U}(C_\tau) &= \ln\left(\frac{(1+n)(1+\rho)C_\tau}{(1+n)(1+\rho)+1+\rho_G}\right) \\ &\quad + \frac{1+\rho_G}{(1+n)(1+\rho)} \ln\left(\frac{(1+n)(1+\rho_G)C_\tau}{(1+n)(1+\rho)+1+\rho_G}\right) \\ &\equiv \omega_0 + \frac{(1+n)(1+\rho)+1+\rho_G}{(1+n)(1+\rho)} \ln C_\tau. \end{aligned} \tag{F}$$

In the second step the social planner chooses sequences of aggregate consumption and the two types of capital in order to maximize social welfare:

$$SW_t = \sum_{\tau=t}^{\infty} \left(\frac{1+n}{1+\rho_G} \right)^{\tau-t} \bar{U}(C_\tau), \tag{G}$$

subject to the initial conditions (k_t and g_t given) and the resource constraint:

$$C_\tau + (1+n) \left[k_{\tau+1} + g_{\tau+1} \right] = f(k_\tau, g_\tau) + (1-\delta)k_\tau + (1-\delta_G) g_\tau, \tag{H}$$

where we have used (A) in (A5.166) to get (H). Letting μ_τ^R denote the Lagrange multiplier for the resource constraint in period τ we obtain the following first-order conditions:

$$\frac{(1+n)\mu_\tau^R}{\mu_{\tau+1}^R} = f_k(k_{t+1}, g_{t+1}) + 1 - \delta = f_g(k_{t+1}, g_{t+1}) + 1 - \delta_G, \tag{I}$$

$$\mu_\tau^R = \left(\frac{1+n}{1+\rho_G} \right)^{\tau-t} \bar{U}'(C_\tau). \tag{J}$$

By using (J) for period $t+1$ and noting (D) and (E) we find that (I) coincides with (A5.172).

****

We now return to the general first-order conditions (A5.172)-(A5.173) and study the steady state. In the steady state we have $C_t^Y = C^Y$, $C_t^O = C^O$, $k_t = k$, $g_t = g$, and $\hat{x}_t = \hat{x}$ for all t so that (A5.172)-(A5.173) simplify to:

$$\frac{\partial \Lambda^Y(\hat{x})/\partial C_t^Y}{\partial \Lambda^Y(\hat{x})/\partial C_t^O} = 1 + \rho_G, \tag{A5.175}$$

$$[\hat{r} \equiv] \, f_k(k,g) - \delta = \rho_G = f_g(k,g) - \delta_G \, [\equiv \hat{r}_G] \, . \tag{A5.176}$$

Equation (A5.175) calls for an optimal division of consumption over the young and the old. The first equality in (A5.176) is the *modified golden rule* (MGR) equating the steady-state yield on the private capital stock (the steady-state rate of interest) to the rate of time preference of the social planner. There is an important difference between this version of the MGR and the one encountered in Chapter 8 in the context of the Ramsey representative-agent model. In the OLG setting, the *planner's* rate of time preference features in the MGR whereas in the Ramsey model the representative *agent's own* rate of time preference is relevant.

The second equality in (A5.176) is a modified golden rule for public capital that was initially derived by Pestieau (1974). It calls for an equalization of the public rate of return and the planner's rate of time preference. The two equalities in (A5.176) together determine the optimal per worker stocks of public and private capital. For example, for Cobb-Douglas technology we have $y_t = k_t^{1-\epsilon_L} g_t^{\eta}$ (with $\eta < \epsilon_L$) so that $k/y = (1-\epsilon_L)/(\rho_G+\delta)$, $g/y = \eta/(\rho_G+\delta_G)$. It follows from these results that output per worker is:

$$y = \left[\left(\frac{k}{y}\right)^{1-\epsilon_L}\left(\frac{g}{y}\right)^{\eta}\right]^{1/(\epsilon_L-\eta)} = \left[\left(\frac{1-\epsilon_L}{\rho_G+\delta}\right)^{1-\epsilon_L}\left(\frac{\eta}{\rho_G+\delta_G}\right)^{\eta}\right]^{1/(\epsilon_L-\eta)} \tag{A5.177}$$

Now that we have characterized the necessary conditions for the steady-state social optimum, a relevant question concerns the *decentralization* of this optimum. Can the policy maker devise a set of policy tools in such a way that the private sector choices concerning consumption and private capital accumulation coincide exactly with their respective values in the social optimum? The answer is affirmative provided the policy maker has access to the right kind of policy instruments. In the present context, for example, the first-best social optimum can be mimicked in the market place if (i) the level of public investment (and thus the public capital stock) is chosen to be consistent with (A5.176), and (ii) there are age-specific lump-sum taxes available (see Pestieau, 1974 and Ihori, 1996, p. 114). The latter instrument is needed to ensure that the market replicates the socially optimal mix of consumption by the young and the old (cf. (A5.175)).

12.3.3 Intergenerational accounting

One of the most hotly debated concepts in policy circles has been the correct definition and measurements of the government budget deficit. Simply put, there exists a fundamental ambiguity in the concept of the deficit. Auerbach, Gokhale, and Kotlikoff (1991, p. 57) give the most radical statement of the problem to date by arguing that "...every dollar the government takes in or pays out is labeled in a manner that is economically arbitrary". They suggest doing away with the concept of the government deficit altogether and to focus instead on what they label the *generational accounts*. The background to their proposal is the notion that "[t]he conceptual issue associated with the word 'deficit' is the intergenerational distribution of welfare" (Auerbach et al., 1991, p. 57) and that the intertemporal budget constraint of the government should be the focus of attention. In words, this constraint says that "the government's current net wealth plus the present value of the government's net receipts from all current and future generations (the generational accounts) must be sufficient to pay for the present value of the government's current and future consumption" (1991, p. 58).

Auerbach et al. (1991, 1994) claim a number of advantages that a system of generational accounts has over the traditional government budget deficit: (i) generational accounts are invariant to changes in accounting labels, (ii) they bring out the zero-sum feature of the intertemporal government budget constraint (what some generation gets will have to be paid for by some other generation), and (iii) they can be used to study the fiscal and intergenerational consequences of alternative policies

In this subsection we follow Buiter (1997) by illustrating the system of generational accounts in a simple version of the Diamond-Samuelson model. Assuming that the population is constant and normalized to unity (so that $L_{t-1} = L_t = 1$), the government (flow) budget identity is given by:

$$B_{t+1} = (1+r_t)B_t + G_t^O + G_t^Y - T_t^O - T_t^Y, \tag{A5.178}$$

where T_t^O and T_t^Y are the taxes paid by the old and young respectively, and G_t^O and G_t^Y are pure public consumption goods that the government provides free of charge to, respectively, the old and the young. Following Buiter (1997, p. 607) we assume that these public goods are non-rival and non-excludable. Iterating (A5.178) forward in time yields the following expression:

$$\begin{aligned} B_t &= \prod_{s=0}^{T} \frac{1}{1+r_{t+s}} B_{t+T+1} + \sum_{\tau=0}^{T} \prod_{s=0}^{\tau} \frac{1}{1+r_{t+s}} \left[T_{t+\tau}^O + T_{t+\tau}^Y - \left(G_{t+\tau}^O + G_{t+\tau}^Y\right)\right] \\ &\equiv R_{t-1,T} B_{t+T+1} + \sum_{\tau=0}^{T} R_{t-1,\tau} \left[T_{t+\tau}^O + T_{t+\tau}^Y - \left(G_{t+\tau}^O + G_{t+\tau}^Y\right)\right], \end{aligned} \tag{A5.179}$$

where $R_{t-1,\tau}$ is a discounting factor:

$$R_{t-1,\tau} = \prod_{s=0}^{\tau} \frac{1}{1+r_{t+s}}. \tag{A5.180}$$

By letting $T \to \infty$ in (A5.179) we find that the government NPG condition is:

$$\lim_{T\to\infty} R_{t-1,T} B_{t+T+1} = 0, \tag{A5.181}$$

so that the government budget constraint is:

$$B_t = \sum_{\tau=0}^{\infty} R_{t-1,\tau} \left[T_{t+\tau}^O + T_{t+\tau}^Y - \left(G_{t+\tau}^O + G_{t+\tau}^Y\right)\right]. \tag{A5.182}$$

If there is government debt outstanding at time t ($B_t > 0$ on the left-hand side of (A5.182)), then the solvent government must ultimately run primary surpluses. Note that (A5.181) does not require the government to pay off its debt eventually. All that solvency requires is that government debt must not grow faster in the long run than the rate of interest.

The household sector is standard. Households consume during youth and old age (C_t^Y and C_{t+1}^O, respectively), practise consumption smoothing by saving (S_t) which can be in the form of physical capital or government bonds. The relevant expressions characterizing the household sector are:

$$C_t^Y + S_t = W_t - T_t^Y, \tag{A5.183}$$

$$C_{t+1}^Y = (1+r_{t+1})S_t - T_{t+1}^O, \tag{A5.184}$$

$$S_t = B_{t+1} + K_{t+1}. \tag{A5.185}$$

Equations (A5.183)-(A5.185) are the same as (A5.153)-(A5.154) and (A5.185) is the same as (A5.61) but with the size of the (young) population set equal to unity ($L_t = 1$). The consolidated budget constraint facing households is obtained in the usual manner by combining (A5.183) and (A5.184):

$$C_t^Y + \frac{C_{t+1}^O}{1+r_{t+1}} = W_t - T_{t,t}, \tag{A5.186}$$

where $T_{t,t}$ is the present value of (lump-sum) taxes that a generation born in period t (second subscript) must pay over the course of its life seen from the perspective of period t (first subscript):

$$T_{t,t} \equiv T_t^Y + \frac{T_{t+1}^O}{1+r_{t+1}}. \tag{A5.187}$$

We can now develop the generational accounts for existing and future generations by decomposing the government budget constraint (A5.182). Because it is very easy indeed to get tangled up in the different subscripts identifying time and generations we show some of the details of the derivation.[18] First we note that by using (A5.180) equation (A5.182) can be written as follows:

$$\begin{aligned} B_t = {} & \frac{1}{1+r_t}\left[T_t^O + T_t^Y\right] + \frac{1}{1+r_t}\cdot\frac{1}{1+r_{t+1}}\left[T_{t+1}^O + T_{t+1}^Y\right] \\ & + \frac{1}{1+r_t}\cdot\frac{1}{1+r_{t+1}}\cdot\frac{1}{1+r_{t+2}}\left[T_{t+2}^O + T_{t+2}^Y\right] \\ & + \frac{1}{1+r_t}\cdot\frac{1}{1+r_{t+1}}\cdot\frac{1}{1+r_{t+2}}\cdot\frac{1}{1+r_{t+3}}\left[T_{t+3}^O + T_{t+3}^Y\right] + ... \\ & - \sum_{\tau=0}^{\infty} R_{t-1,\tau}\left[G_{t+\tau}^O + G_{t+\tau}^Y\right]. \end{aligned} \tag{A5.188}$$

Next we look for terms pertaining to the same generation and group these together:

$$\begin{aligned} B_t = {} & \frac{1}{1+r_t}T_t^O + \frac{1}{1+r_t}\left[T_t^Y + \frac{T_{t+1}^O}{1+r_{t+1}}\right] \\ & + \frac{1}{1+r_t}\cdot\frac{1}{1+r_{t+1}}\left[T_{t+1}^Y + \frac{T_{t+2}^O}{1+r_{t+2}}\right] \\ & + \frac{1}{1+r_t}\cdot\frac{1}{1+r_{t+1}}\cdot\frac{1}{1+r_{t+2}}\left[T_{t+2}^Y + \frac{T_{t+3}^O}{1+r_{t+3}}\right] + ... \\ & - \sum_{\tau=0}^{\infty} R_{t-1,\tau}\left[G_{t+\tau}^O + G_{t+\tau}^Y\right]. \end{aligned} \tag{A5.189}$$

In the first line of (A5.189) we find the remaining taxes to be paid by the old at time t and the lifetime taxes, $T_{t,t}$, of the young at time t. Both these terms are, however, expressed in present-value terms, i.e. they are discounted back to the end of period $t-1$. The same holds for all the other terms pertaining

[18] A more direct derivation makes use of the fact that the discount factor in (A5.180) satisfies the following property:

$$R_{t-1,\tau+1} = \frac{1}{1+r_{t+\tau+1}}R_{t-1,\tau}.$$

Using this property in (A5.182) yields (A5.189) in a single step.

to future generations (namely the second and third lines in (A5.189)). The reason for this discounting is that B_t is debt at the beginning of period t (which was accumulated at the end of period $t-1$), over which interest must be paid at the beginning of period t.

Equation (A5.189) gives the generational accounts for the different generations. The first line contains the accounts for the two existing generations at time t, whilst lines two and three contain the generational accounts for future generations. Kotlikoff and co-authors often write the generational accounts in a more compact format as:

$$B_t + \sum_{\tau=0}^{\infty} R_{t-1,\tau} \left[G^O_{t+\tau} + G^Y_{t+\tau} \right] = \sum_{k=t-1}^{\infty} T_{t-1,k}, \tag{A5.190}$$

where the $T_{t-1,k}$ terms are defined as follows:

$$T_{t-1,t-1} \equiv \frac{1}{1+r_t} T^O_t, \qquad \text{(existing old)}$$

$$T_{t-1,t} \equiv \frac{1}{1+r_t} T_{t,t} = \frac{1}{1+r_t} \left[T^Y_t + \frac{T^O_{t+1}}{1+r_{t+1}} \right], \qquad \text{(existing young)}$$

$$T_{t-1,k} \equiv R_{t-1,k-1} T_{k,k} = R_{t-1,k-t} \left[T^Y_k + \frac{T^O_{k+1}}{1+r_{k+1}} \right], \qquad \text{(future generations)}$$

where $k = t+1, t+2, \ldots$. Equation (A5.190) says that the sum of outstanding government debt plus the present value of government consumption (left-hand side) must equal the sum of the generational accounts of existing and future generations (right-hand side).

Having completed our description of the generational accounting system in the context of the Diamond-Samuelson model we can now turn to an actual empirical implementation of the method. Auerbach et al. (1991, pp. 65-75) explain in detail how the method of generational accounting can be applied to actual economies. Table 12.2 contains the 1991 generational accounts for US males. (This table is an abbreviated version of Table 1 of Auerbach et al., 1994, p. 80.) Of course, for the method to have any practical use, an actual implementation must contain much more detail than is contained in our stylized model. Table 12.2 therefore distinguishes ten (rather than just two) existing generations and gives the accounts for males only because females are different in labour force participation, family structure, and mortality. (Auerbach et al., 1994, give figures for nineteen five-year cohorts and also present generational accounts for females.) Furthermore, Auerbach et al. allow for transfers, distorting taxes etc. that were abstracted from in the stylized model.

In Table 12.2 the first column gives the age of the particular generation of US males in 1991, e.g. the row marked '0' pertains to agents born in 1991 whereas the row marked '40' gives the data for agents who were 40 years of age in 1991 (who were thus born in 1951). The second column gives the net generational accounts for the different generations whilst the third and fourth columns distinguishe, respectively, the underlying tax payments and transfer receipts. A positive entry in the second column

Table 12.2: Male generational accounts

Generation's age in 1991	*Net payments*	*Tax payments*	*Transfer receipts*
	×$1000	×$1000	×$1000
0	78.9	99.3	20.4
10	125.0	155.3	30.3
20	187.1	229.6	42.5
30	205.5	258.5	53.0
40	180.1	250.0	69.9
50	97.2	193.8	96.6
60	-23.0	112.1	135.1
70	-80.7	56.3	137.0
80	-61.1	30.2	91.3
90	-3.5	8.8	12.3
Future	166.5		

means that the particular generation will pay more in present value to the government than it will receive. For example, for a 40–year old male in 1991, the present value of taxes to be paid during his remaining lifetime amount to $250,000 whilst the present value of transfers is $69,900. In contrast, a 70-year old in 1991 has a negative generational account of $80,700 because the present value of transfers (on disability, health, and welfare transfers) far exceeds the present value of taxes.

The final row labelled 'Future' in Table 12.2 gives the generational account for the typical future generation. For future generations, the generational account measures the present value of net payments over their entire lives. Since the same holds for newborns in 1991, the figures for newborns and future generations can be meaningfully compared. As Auerbach et al. (1994, p. 82) point out, there is a striking generational imbalance in US fiscal policy in the sense that future newborns have a generational account of $166,500, which is a whopping $87,600 more than newborns in 1991 have to pay.

12.3.3.1 Discussion

Buiter (1997) agrees with the proponents of the generational accounting method that the traditional measure of the government deficit is a meaningless indicator for the effects of fiscal policy not only on aggregate demand and private saving but also on the intergenerational distribution of resources. He is nevertheless quite critical of the method of generational accounting. Buiter's objections centre on the following three issues. First, the usefulness of generational accounts "lives or dies with the validity of the life-cycle model" (1997, p. 606), of which the Diamond model is a simple representation. The validity of the life-cycle model depends critically on the following assumptions (which must all hold): (i) households have finite lives, (ii) generations are not linked via operative bequests, and (iii) markets are complete (no borrowing constraints). If condition (ii) is violated and Ricardian equivalence is valid (see Chapter 8), then the generational accounts are completely uninformative about the effect of the government budget on both the intergenerational distribution of resources and on saving (see Buiter, 1997, p. 612 and Bohn, 1992). A similar conclusion follows if condition (iii) is violated and households face binding liquidity constraints because in that case the timing of tax payments over the life cycle matters (in addition to the present value of these taxes).

Second, even if the strict life-cycle model is valid, generational accounts should be interpreted quite carefully. Indeed, existing applications of the generational accounting method say nothing about the intergenerational distribution of benefits from government spending on public goods. Take, for example, the case of a government abatement programme aimed at cleaning up the natural environment. If the environment improves only slowly over time, future generations may be the principal beneficiaries of the policy measure even though the current generations have paid for it. In generational accounts, the tax payments associated with the programme feature prominently but the benefits to future generations are not included.

Third, the method of generational accounting does not take into account the general equilibrium repercussions of alternative budgetary policies. In particular, the method ignores (i) the endogeneity of the various tax (and subsidy) bases and (ii) the endogeneity of pre-tax factor prices and incomes. Buiter gives several examples for which the general equilibrium effects turn out to be quite important (1997, pp. 616-622).[19]

In principle all the issues raised above can be studied with the aid of a computable dynamic general equilibrium model although the construction of such a model is clearly not a trivial task. On the one hand, such models can readily deal with the general equilibrium repercussions of alternative budgetary policies (see Auerbach and Kotlikoff, 1987) and can be extended to include all kinds of market imperfections and alternative intergenerational linkages. One the other hand, there are huge practical difficulties in quantifying the (intergenerational) welfare effects of public spending. In this context, the method of generational accounting is valuable because its data can provide some of the inputs needed for a realistic simulation model.

Key literature

- Heijdra & Van der Ploeg (2002, ch. 17), Jha (1998, ch. 5), or Myles (1995, ch. 14) on theory.
- Feldstein and Liebman (2002) on theory and empirics.
- Lindbeck and Persson (2003) on pension reform.
- Renstrom (1998) on optimal taxation in OLG setting.

[19] Fehr and Kotlikoff (1995), on the other hand, present a number of general equilibrium examples where the generational accounting method appears to work quite well.

- From FoMM: Samuelson (1975a, 1975b) and Feldstein (1974, 1976, 1985, 1987). In recent years a large literature has been developed on the issue of pension system reform. See Diamond (1997, 1999), Feldstein (1997, 1998), and Sinn (2000). For a recent survey on the economic effects of ageing, see Bosworth and Burtless (1998). The Diamond-Samuelson model has been generalized in a number of directions. Barro (1974) studies intergenerational linkages. Jones and Manuelli (1992) consider the growth effects of finite lives. Tirole (1985) and O'Connell and Zeldes (1988) consider the possibility of asset bubbles. Grandmont (1985) presents a model exhibiting endogenous business cycles. Michel and de la Croix (2000) study the model properties under both myopic foresight and perfect foresight. Bierwag, Grove, and Khang (1969) show that a full set of age-specific taxes renders debt policy redundant. Abel (1986) and Zilcha (1990, 1991) introduce uncertainty into the model. On intergenerational risk sharing, see Gordon and Varian (1988). Barro and Becker (1989) present a model of endogenous fertility. For applications of endogenous fertility models, see Wildasin (1990), Zhang (1995), Robinson and Srinivasan (1997), and Nerlove and Raut (1997). Galor (1992) and Nourry (2001) study a two-sector version of the Diamond-Samuelson model.

The Diamond-Samuelson model has been applied in a large number of fields. For public finance applications, see Auerbach (1979a), Kotlikoff and Summers (1979), and Ihori (1996). On the economics of education, see Loury (1981), Glomm and Ravikumar (1992), Zhang (1996), Buiter and Kletzer (1993), and Kaganovich and Zilcha (1999). Environmental policy applications include Howarth (1991, 1998), Howarth and Norgaard (1990, 1992), John and Pecchenino (1994), John et al. (1995), and Mourmouras (1993).

There is a large literature on generational accounting. Some key references are Auerbach, Gokhale, and Kotlikoff (1991, 1994), Kotlikoff (1993a, 1993b), and Fehr and Kotlikoff (1995). For critical papers on the topic, see Bohn (1992), Haveman (1994), and Buiter (1997). International applications of the method are collected in Auerbach, Kotlikoff, and Leibfritz (1999).

12.4 Punchlines

??? Key points restated simply

Further Reading

??? What should interested students read?

Chapter 13

Chapter 13: Rent seeking

The purpose of this chapter is to discuss the following topics:

- Theory of rent-seeking
- Applications: pollution taxation, privatisation, growth

13.1 First section

??? Bla Bla

Intermezzo 13.1

Title of the intermezzo. ??? bla bla bla ???

****

13.2 Second section

??? Bla Bla

13.3 Punchlines

??? Key points restated simply

Further Reading

??? What should interested students read?

Key Literature

- The collected works by Brooks and Heijdra (1987, 1988, 1989, 1990) on theory and applications.

Outline of the Chapter

??? Ideas about the outline of the chapter here

Chapter 14

Social Insurance

The purpose of this chapter is to discuss the following topics:

- Systems of social insurance
- Incentive effects.
- Labour supply.
- Empirics.

14.1 First section

??? Bla Bla

Intermezzo 14.1

Title of the intermezzo. ??? bla bla bla ???

****

14.2 Second section

??? Bla Bla

14.3 Punchlines

??? Key points restated simply

Further Reading

??? What should interested students read?

Key Literature

- Krueger and Meyer (2002) on theory and empirics.
- Moffitt (2002) on theory and empirics.

Outline of the Chapter

??? Ideas about the outline of the chapter here

Chapter 15

Redistribution

The purpose of this chapter is to discuss the following topics:

- Demand for redistribution.
- Supply of redistribution.
- Reforming the welfare state.
- Empirics.

15.1 First section

??? Bla Bla

Intermezzo 15.1

Title of the intermezzo. ??? bla bla bla ???

****

15.2 Second section

??? Bla Bla

15.3 Punchlines

??? Key points restated simply

Further Reading

??? What should interested students read?

Key Literature

- Boadway and Keen (2002) on theory and empirics.

Outline of the Chapter

??? Ideas about the outline of the chapter here

Bibliography

Aaron, H. J. (1966). The social insurance paradox. *Canadian Journal of Economics and Political Science*, 32:371–374.

Abel, A. B. (1986). Capital accumulation and uncertain lifetimes with adverse selection. *Econometrica*, 54:1079–1097.

Abel, A. B., Mankiw, N. G., Summers, L. H., and Zeckhauser, R. (1989). Assessing dynamic efficiency: Theory and evidence. *Review of Economic Studies*, 56:1–19.

Aghion, P. and Howitt, P. (1998). *Endogenous Growth Theory*. MIT Press, Cambridge, MA.

Ahsan, S. M. (1976). Taxation in a two-period temporal model of consumption and portfolio allocation. *Journal of Public Economics*, 5(3-4):337–352.

Akerlof, G. (1978). The economics of tagging as applied to the optimal income tax. *American Economic Review*, 68(1):8–19.

Allais, M. (1947). *Économie et Intérêt*. Imprimerie Nationale, Paris. Second edition with a new introduction, 1998, Paris: Clément Juglar.

Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1(3-4):323–338.

Anderberg, D. and Andersson, F. (2003). Investments in human capital, wage uncertainty, and public policy. *Journal of Public Economics*, 87(7-8):1521–1537.

Anderson, S. P., de Palma, A., and Kreidler, B. (2001). Tax incidence in differentiated product oligopoly. *Journal of Public Economics*, 81(2):173–192.

Apps, P. F. and Rees, R. (1988). Taxation and the household. *Journal of Public Economics*, 35:355–369.

Apps, P. F. and Rees, R. (1997). Collective labor supply and household production. *Journal of Political Economy*, 105(1):178–190.

Apps, P. F. and Rees, R. (1999a). Individual versus joint taxation in models with household production. *Journal of Political Economy*, 107(2):393–403.

Apps, P. F. and Rees, R. (1999b). On the taxation of trade within and between households. *Journal of Public Economics*, 73:241–263.

Archibald, G. C. (1987). Firm, theory of the. In Eatwell, J., Milgate, M., and Newman, P., editors, *The New Palgrave: A Dictionary of Economics*. Macmillan, London.

Arrow, K. J. (1962). The economic implications of learning by doing. *Review of Economic Studies*, 29:155–173.

Arrow, K. J. (1965). *Aspects of the Theory of Risk-Bearing*. Yrjö Jahnssonin Säätio, Helsinki.

Arrow, K. J. (1971a). *Essays in the Theory of Risk-Bearing*. North-Holland, Amsterdam.

Arrow, K. J. (1971b). A utilitarian approach to the concept of equality in public expenditures. *Quarterly Journal of Economics*, 85(3):409–415.

Arrow, K. J. and Kurz, M. (1970). *Public Investment, the Rate of Return, and Optimal Fiscal Policy*. Johns Hopkins Press, Baltimore, MD.

Aschauer, D. A. (1989). Is public expenditure productive? *Journal of Monetary Economics*, 23:177–200.

Aschauer, D. A. (1990). Why is infrastructure important? In Munnell, A. H., editor, *Is There a Shortfall in Public Investment?*, pages 21–68. Federal Reserve Bank of Boston, Boston.

Ashenfelter, O. and Heckman, J. (1974). The estimation of income and substitution effects in a model of family labor supply. *Econometrica*, 42(1):73–86.

Atkinson, A. B. (1973). How progressive should income tax be? In Parkin, M. and Nobay, A. R., editors, *Essays in Modern Economics*, pages 90–109. Longman, London.

Atkinson, A. B. (1990). Public economics and the economic public. *European Economic Review*, 34(2-3):225–248.

Atkinson, A. B. (1994). The distribution of the tax burden. In Quigley, J. M. and Smolensky, E., editors, *Modern Public Finance*, pages 13–57. Harvard University Press, Cambridge, MA.

Atkinson, A. B. (1995). *Public Economics in Action: The Basic Income / Flat Tax Proposal*. Oxford University Press.

Atkinson, A. B. and Stern, N. H. (1974). Pigou, taxation, and public goods. *Review of Economic Studies*, 41(1):119–128.

Atkinson, A. B. and Stern, N. H. (1979). A note on the allocation of time. *Economics Letters*, 3:119–123.

Atkinson, A. B. and Stern, N. H. (1980). On the switch from direct to indirect taxation. *Journal of Public Economics*, 14:195–224.

Atkinson, A. B. and Stern, N. H. (1981). On labour supply and commodity demands. In Deaton, A., editor, *Essays in the Theory and Measurement of Consumer Behaviour: Essays in Honour of Sir Richard Stone*, pages 265–296. Cambridge University Press, Cambridge.

Atkinson, A. B. and Stiglitz, J. E. (1972). The structure of indirect taxation and economic efficiency. *Journal of Public Economics*, 1(1):97–119.

Atkinson, A. B. and Stiglitz, J. E. (1976). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics*, 6(1-2):55–75.

Atkinson, A. B. and Stiglitz, J. E. (1980). *Lectures on Public Economics*. McGraw-Hill, New York.

Attanasio, O. P. (1999). Consumption. In Taylor, J. B. and Woodford, M., editors, *Handbook of Macroeconomics*, volume IB, pages 741–812. North-Holland, Amsterdam.

Auerbach, A. J. (1979a). Share valuation and corporate equity policy. *Journal of Public Economics*, 11(3):291–305.

Auerbach, A. J. (1979b). Wealth maximization and the cost of capital. *Quarterly Journal of Economics*, 93(3):433–446.

Auerbach, A. J. (1983). Taxation, corporate financial policy and the cost of capital. *Journal of Economic Literature*, 21(3):905–940.

Auerbach, A. J. (1984). Taxes, firm financial policy and the cost of capital: An empirical analysis. *Journal of Public Economics*, 23(1-2):27–57.

Auerbach, A. J. (1985). The theory of excess burden and optimal taxation. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume I, pages 61–127. North-Holland, Amsterdam.

Auerbach, A. J. (2002). Taxation and corporate financial policy. In Auerbach, A. and Feldstein, M., editors, *Handbook of Public Economics*, volume III, pages 1251–1292. North-Holland, Amsterdam.

Auerbach, A. J. and Feldstein, M., editors (1985). *Handbook of Public Economics*, volume I. North-Holland, Amsterdam.

Auerbach, A. J. and Feldstein, M., editors (1987). *Handbook of Public Economics*, volume II. North-Holland, Amsterdam.

Auerbach, A. J. and Feldstein, M., editors (2002a). *Handbook of Public Economics*, volume III. North-Holland, Amsterdam.

Auerbach, A. J. and Feldstein, M., editors (2002b). *Handbook of Public Economics*, volume IV. North-Holland, Amsterdam.

Auerbach, A. J., Gokhale, J., and Kotlikoff, L. J. (1991). Generational accounts: A meaningful alternative to deficit accounting. In Bradford, D., editor, *Tax Policy and the Economy*, volume 5, pages 55–110. MIT Press.

Auerbach, A. J., Gokhale, J., and Kotlikoff, L. J. (1994). Generational accounting: A meaningful way to evaluate fiscal policy. *Journal of Economic Perspectives*, 8:73–94.

Auerbach, A. J. and Hines, J. R. (2002). Taxation and economic efficiency. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume IV, pages 1347–1421. North-Holland, Amsterdam.

Auerbach, A. J. and Kotlikoff, L. J. (1987). *Dynamic Fiscal Policy*. Cambridge University Press, Cambridge.

Auerbach, A. J., Kotlikoff, L. J., and Leibfritz, W., editors (1999). *Generational Accounting around the World*. University of Chicago Press, Chicago.

Aura, S., Diamond, P., and Geanakoplos, J. (2002). Savings and portfolio choice in a two-period two-asset model. *American Economic Review*, 92(4):1185–1191.

Azariadis, C. (1993). *Intertemporal Macroeconomics*. Basil Blackwell, Oxford.

Azariadis, C. and Drazen, A. (1990). Threshold externalities in economic development. *Quarterly Journal of Economics*, 105:501–526.

Baldry, J. C. (1979). Tax evasion and labour supply. *Economics Letters*, 3:53–56.

Ballard, C. L., Shoven, J. B., and Whalley, J. (1985). General equilibrium computations of the marginal welfare costs of taxes in the United States. *American Economic Review*, 75(1):128–138.

Barr, N. (2004). *Economics of the Welfare State*. Cambridge University Press, Cambridge, fourth edition.

Barro, R. J. (1974). Are government bonds net wealth? *Journal of Political Economy*, 82:1095–1117.

Barro, R. J. (1990). Government spending in a simple model of endogenous growth. *Journal of Political Economy*, 98:S103–S125.

Barro, R. J. and Becker, G. S. (1989). Fertility choice in a model of economic growth. *Econometrica*, 57:481–501.

Barro, R. J. and Sala-i-Martin, X. (1992). Public finance in models of economic growth. *Review of Economic Studies*, 59(4):645–661.

Barro, R. J. and Sala-i-Martin, X. (1995). *Economic Growth*. McGraw-Hill, New York.

Barsky, R. B., Mankiw, N. G., and Zeldes, S. P. (1986). Ricardian consumers with Keynesian propensities. *American Economic Review*, 76(4):676–691.

Batra, R. N. and Ullah, A. (1974). Competitive firm and the theory of input demand under price uncertainty. *Journal of Political Economy*, 82(3):537–548.

Becker, G. S. (1965). A theory of the allocation of time. *Economic Journal*, 75:493–517.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217.

Belan, P., Michel, P., and Pestieau, P. (1998). Pareto improving social security reform. *Geneva Papers on Risk and Insurance Theory*, 23:119–125.

Belan, P. and Pestieau, P. (1999). Privatizing social security: A critical assessment. *Geneva Papers on Risk and Insurance Theory*, 24:114–130.

Bénassy, J.-P. (1998). Is there always too little research in endogenous growth with expanding product variety? *European Economic Review*, 42:61–69.

Benveniste, L. M. and Scheinkman, J. A. (1982). Duality theory for dynamic optimization models of economics: The continuous time case. *Journal of Economic Theory*, 27(1):1–19.

Bergson, A. (1938). A reformulation of certain aspects of welfare economics. *Quarterly Journal of Economics*, 52(2):310–334.

Bergstrom, T. C. (1996). Economics in a family way. *Journal of Economic Literature*, 34(4):1903–1934.

Bergstrom, T. C., Blume, L., and Varian, H. R. (1986). On the private provision of public goods. *Journal of Public Economics*, 29(1):25–49.

Bernheim, B. D. (2002). Taxation and saving. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume III, pages 1173–1249. North-Holland, Amsterdam.

Bernheim, B. D. and Bagwell, K. (1988). Is everything neutral? *Journal of Political Economy*, 96(2):308–338.

Besley, T. J. and Rosen, H. S. (1999). Sales taxes and prices: An empirical analysis. *National Tax Journal*, 52(2):157–178.

Bierwag, G. O., Grove, M. A., and Khang, C. (1969). National debt in a neoclassical growth model: Comment. *American Economic Review*, 59:205–210.

Binmore, K. and Dasgupta, P., editors (1987). *The Economics of Bargaining*. Basil Blackwell, Oxford.

Blanchard, O. J. (1985). Debts, deficits, and finite horizons. *Journal of Political Economy*, 93(2):223–247.

Blanchard, O. J. and Fischer, S. (1989). *Lectures on Macroeconomics*. MIT Press, Cambridge, MA.

Blinder, A. S. (1975). Distribution effects and the aggregate consumption function. *Journal of Political Economy*, 83(3):447–475.

Block, M. K. and Heineke, J. M. (1973). The allocation of effort under uncertainty: The case of risk-averse behavior. *Journal of Political Economy*, 81(2):376–385.

Block, M. K. and Heineke, J. M. (1975). Factor allocations under uncertainty: An extension. *Southern Economic Journal*, 41(3):526–530.

Blomquist, N. S. (1985). Labour supply in a two-period model: The effect of a nonlinear progressive income tax. *Review of Economic Studies*, 52(3):515–524.

Blundell, R. and MaCurdy, T. (1999). Labor supply: A review of alternative approaches. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume III, pages 1559–1695. North-Holland, Amsterdam.

Boadway, R. and Keen, M. (2000). Redistribution. In Atkinson, A. B. and Bourguignon, F., editors, *Handbook of Income Distibution*, volume I, pages 677–786. North-Holland, Amsterdam.

Boadway, R. W. and Bruce, N. (1984). *Welfare Economics*. Basil Blackwell, Oxford.

Bodie, Z., Merton, R. C., and Samuelson, W. F. (1992). Labor supply flexibility and portfolio choice in a life cycle model. *Journal of Economic Dynamics and Control*, 16(3-4):427–449.

Bohn, H. (1992). Budget deficits and government accounting. *Carnegie-Rochester Conference Series on Public Policy*, 37:1–84.

Boiteux, M. (1971). On the management of public monopolies subject to budgetary constraints. *Journal of Economic Theory*, 3(1):219–240. (Translation of paper originally published in *Econometrica* in 1956).

Bond, E. W., Wang, P., and Yip, C. K. (1996). A general two-sector model of endogenous growth with human and physical capital: Balanced growth and transitional dynamics. *Journal of Economic Theory*, 68:149–173.

Boone, J. and Bovenberg, A. L. (2002). Optimal labour taxation and search. *Journal of Public Economics*, 85(1):53–97.

Boone, J. and Bovenberg, A. L. (2004). The optimal taxation of unskilled labor with job search and social assistance. *Journal of Public Economics*, 88(11):2227–2258.

Boskin, M. J. (1978). Taxation, saving, and the rate of interest. *Journal of Political Economy*, 86(2):S3–S27.

Boskin, M. J. and Lau, L. J. (1982). Taxation, social security, and aggregate factor supply in the United States. In Kessler, D., Masson, A., and Strauss-Kahn, D., editors, *Accumulation et Réparation des Patrimoines*, pages 185–198. Economica, Paris.

Boskin, M. J. and Sheshinski, E. (1983). Optimal tax treatment of the family: Married couples. *Journal of Public Economics*, 20:281–297.

Bosworth, B. and Burtless, G., editors (1998). *Aging Societies: The Global Dimension*. Brookings Institution Press, Washington.

Bourguignon, F. and Chiappori, P.-A. (1994). The collective approach to household behaviour. In Blundell, R., Preston, I., and Walker, I., editors, *The Measurement of Household Welfare*, pages 70–85. Cambridge University Press, Cambridge.

Bovenberg, A. L. (1993). Investment-promoting policies in open economies: The importance of intergenerational and international distributional effects. *Journal of Public Economics*, 51:3–54.

Bovenberg, A. L. (1994). Capital taxation in the world economy. In van der Ploeg, F., editor, *Handbook of International Macroeconomics*. Basil Blackwell, Oxford.

Bovenberg, A. L. and Heijdra, B. J. (1998). Environmental tax policy and intergenerational distribution. *Journal of Public Economics*, 67:1–24.

Bovenberg, A. L. and van der Ploeg, F. (1994). Effects of the tax and benefit system on wage formation and unemployment. Mimeo.

Bradford, D. F. (1978). Factor prices may be constant but factor returns are not. *Economics Letters*, 1(3):199–203.

Bradford, D. F. (1981). The incidence and allocation effects of a tax on corporate distributions. *Journal of Public Economics*, 15(1):1–22.

Brakman, S. and Heijdra, B. J., editors (2004). *The Monopolistic Competition Revolution in Retrospect*. Cambridge University Press, Cambridge.

Breyer, F. (1989). On the intergenerational Pareto efficiency of pay-as-you-go financed pension systems. *Journal of Institutional and Theoretical Economics*, 145:643–658.

Breyer, F. and Straub, M. (1993). Welfare effects of unfunded pension systems when labor supply is endogenous. *Journal of Public Economics*, 50:77–91.

Brito, D. L., Hamilton, J. H., Slutsky, S. M., and Stiglitz, J. E. (1991a). Dynamic optimal income taxation with government commitment. *Journal of Public Economics*, 44(1):15–35.

Brito, D. L., Sheshinski, E., and Intriligator, M. D. (1991b). Externalities and compulsory vaccinations. *Journal of Public Economics*, 45(1):69–90.

Brock, W. A. and Turnovsky, S. J. (1981). The analysis of macroeconomic policies in perfect foresight equilibrium. *International Economic Review*, 22(1):179–209.

Brooks, M. A. and Heijdra, B. J. (1987). Rent-seeking and pollution taxation: An extension. *Southern Economic Journal*, 54:335–342.

Brooks, M. A. and Heijdra, B. J. (1988). In search of rent-seeking. In Rowley, C. K., Tollison, R. D., and Tullock, G., editors, *The Political Economiy of Rent-Seeking*, pages 27–49. Kluwer Academic Publishers, Boston.

Brooks, M. A. and Heijdra, B. J. (1989). An exploration of rent-seeking. *Economic Record*, 65:32–50.

Brooks, M. A. and Heijdra, B. J. (1990). Rent-seeking and the privatization of the commons. *European Journal of Political Economy*, 6:41–59.

Browning, M. and Chiappori, P.-A. (1998). Efficient intra-household allocations: A general characterization and empirical tests. *Econometrica*, 66(6):1241–1278.

Browning, M., Deaton, A., and Irish, M. (1985). A profitable approach to labor supply and commodity demands over the life-cycle. *Econometrica*, 53(3):503–543.

Buchanan, J. M. (1965). An economic theory of clubs. *Economica*, 32:1–14.

Buchanan, J. M. and Stubblebine, W. C. (1962). Externality. *Economica*, 29:371–384.

Buiter, W. H. (1997). Generational accounts, aggregate saving and intergenerational distribution. *Economica*, 64:605–626.

Buiter, W. H. and Kletzer, K. M. (1992). Government solvency, Ponzi finance and the redundancy and usefulness of public debt. NBER Working Paper 4076, National Bureau for Economic Research, Cambridge, MA.

Buiter, W. H. and Kletzer, K. M. (1993). Permanent international productivity growth differentials in an integrated global economy. *Scandinavian Journal of Economics*, 95:467–493.

Bulow, J. I. and Pfleiderer, P. (1983). A note on the effect of cost changes on prices. *Journal of Political Economy*, 91(1):182–185.

Bulow, J. I. and Summers, L. H. (1986). A theory of dual labor markets with applications to industrial policy, discrimination, and Keynesian unemployment. *Journal of Labor Economics*, 4(3):376–414.

Burmeister, E. and Dobell, A. R. (1970). *Mathematical Theories of Economic Growth*. Macmillan, London.

Burness, H. S. (1976). A note on consistent naive intertemporal decision making and an application to the case of uncertain lifetime. *Review of Economic Studies*, 43:547–549.

Calvo, G. A. and Obstfeld, M. (1988). Optimal time-consistent fiscal policy with finite lifetimes. *Econometrica*, 56:411–432.

Card, D. (1994). Intertemporal labour supply: An assessment. In Sims, C., editor, *Advances in econometrics: Sixth world congress*, volume II. Cambridge University Press, Cambridge.

Chambers, R. G. (1988). *Applied Production Analysis: A Dual Approach*. Cambridge University Press, Cambridge.

Chiappori, P.-A. (1988a). Nash-bargained household decisions: A comment. *International Econometic Review*, 29(4):791–796.

Chiappori, P.-A. (1988b). Rational household labor supply. *Econometrica*, 56(1):63–89.

Chiappori, P.-A. (1992). Collective labor supply and welfare. *Journal of Political Economy*, 100(3):437–467.

Chiappori, P.-A. (1997). Introducing household production in collective models of labor supply. *Journal of Political Economy*, 105(1):191–209.

Chipman, J. S. (1970). External economies of scale and competitive equilibrium. *Quarterly Journal of Economics*, 84(3):347–385.

Corlett, W. J. and Hague, D. C. (1953-1954). Complementarity and the excess burden of taxation. *Review of Economic Studies*, 21(1):21–30.

Corneo, G. and Marquardt, M. (2000). Public pensions, unemployment insurance, and growth. *Journal of Public Economics*, 75:293–311.

Cornes, R. and Sandler, T. (2000). Pareto-improving redistribution and pure public goods. *German Economic Review*, 1(2):169–186.

Corsetti, G. and Roubini, N. (1996). Optimal government spending and taxation in endogenous growth models. NBER Working Paper 5851, National Bureau for Economic Research, Cambridge, MA.

Coyte, P. C. (1986). The supply of individual hours and labor force participation under uncertainty. *Economic Inquiry*, 24(1):155–170.

Cremer, H., Gahvari, F., and Ladoux, N. (1998). Externalities and optimal taxation. *Journal of Public Economics*, 70(3):343–364.

Cremer, H., Pestieau, P., and Rochet, J.-C. (2001). Direct versus indirect taxation: The design of the tax structure revisited. *International Economic Review*, 42(3):781–799.

Cutler, D. (1988). Tax reform and the stock market: An asset price approach. *American Economic Review*, 78:1107–1117.

Dardanoni, V. (1988). Optimal choices under uncertainty: The case of two-argument utility functions. *Economic Journal*, 98(391):429–450.

Davis, G. K. (1989). Income and substitution effects for mean-preserving spreads. *International Economic Review*, 30(1):131–136.

de la Croix, D. and Michel, P. (2002). *A Theory of Economic Growth: Dynamics and Policy in Overlapping Generations*. Cambridge University Press, Cambridge.

Deaton, A. (1977). Equity, efficiency, and the structure of indirect taxation. *Journal of Public Economics*, 8:299–312.

Deaton, A. (1979). The distance function in consumer behaviour with applications to index numbers and optimal taxation. *Review of Economic Studies*, 46(3):391–405.

Deaton, A. (1981). Optimal taxes and the structure of preferences. *Econometrica*, 49(5):1245–1260.

Deaton, A. (1992). *Understanding Consumption*. Oxford University Press, Oxford.

Deaton, A. and Muellbauer, J. (1980). *Economics and Consumer Behavior*. Cambridge University Press, Cambridge.

Demmel, R. and Keuschnigg, C. (2000). Funded pensions and unemployment. *FinanzArchiv*, 57(1):22–38.

Diamond, P. A. (1965). National debt in a neoclassical growth model. *American Economic Review*, 55:1126–1150.

Diamond, P. A. (1970). Incidence of an interest income tax. *Journal of Economic Theory*, 2(3):211–224.

Diamond, P. A. (1973). Taxation and public production in a growth setting. In Mirrlees, J. A. and Stern, N. H., editors, *Models of Economic Growth*, pages 215–240. Macmillan, London.

Diamond, P. A. (1975). A many-person Ramsey tax rule. *Journal of Public Economics*, 4(4):335–342.

Diamond, P. A. (1997). Macroeconomic aspects of social security reform. *Brookings Papers on Economic Activity*, 2:1–87.

Diamond, P. A. (1998). Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review*, 88(1):83–95.

Diamond, P. A., editor (1999). *Issues in Privatizing Social Security*. MIT Press, Cambridge, MA.

Diamond, P. A., Helms, L. J., and Mirrlees, J. A. (1980). Optimal taxation in a stochastic economy. *Journal of Public Economics*, 14(1):1–29.

Diamond, P. A. and McFadden, D. L. (1974). Some uses of the expenditure function in public finance. *Journal of Public Economics*, 3(1):3–21.

Diamond, P. A. and Mirrlees, J. A. (1971a). Optimal taxation and public production: I: Production effiency. *American Economic Review*, 61(1):8–27.

Diamond, P. A. and Mirrlees, J. A. (1971b). Optimal taxation and public production: II: Tax rules. *American Economic Review*, 61(3):261–278.

Diamond, P. A. and Stiglitz, J. E. (1974). Increases in risk and in risk aversion. *Journal of Economic Theory*, 8(3):337–360.

Diamond, P. A. and Yaari, M. (1972). Implications of the theory of rationing for consumer choice under uncertainty. *American Economic Review*, 62(3):333–343.

Dixit, A. K. (1970). On the optimum structure of commodity taxes. *American Economic Review*, 60(3):295–301.

Dixit, A. K. (1975). Welfare effects of tax and price changes. *Journal of Public Economics*, 4(2):103–123.

Dixit, A. K. (1986). Comparative statics for oligopoly. *International Economic Review*, 27(1):107–122.

Dixit, A. K. and Munk, K. J. (1977). Welfare effects of tax and price changes: A correction. *Journal of Public Economics*, 8(1):103–107.

Dixit, A. K. and Sandmo, A. (1977). Some simplified formulae for optimal income taxation. *Scandinavian Journal of Economics*, 79(4):417–423.

Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *American Economic Review*, 67(3):297–308.

Doeringer, P. B. and Piore, M. J. (1971). *Internal Labor Markets and Manpower Analysis*. D. C. Heath, Lexington, MA.

Domar, E. D. (1946). Capital expansion, rate of growth and employment. *Econometrica*, 14(2):137–147.

Domar, E. D. (1947). Expansion and employment. *American Economic Review*, 37(1):34–55.

Domar, E. D. and Musgrave, R. A. (1944). Proportional income taxation and risk-taking. *Quarterly Journal of Economics*, 58(3):388–422.

Drazen, A. (2000). *Political Economy in Macroeconomics*. Princeton University Press, Princeton, NJ.

Drèze, J. and Modigliani, F. (1972). Consumption decisions under uncertainty. *Journal of Economic Theory*, 5(3):308–335.

Eaton, J. and Rosen, H. S. (1980a). Labor supply, uncertainty, and efficient taxation. *Journal of Public Economics*, 14(3):365–374.

Eaton, J. and Rosen, H. S. (1980b). Taxation, human capital, and uncertainty. *American Economic Review*, 70(4):715–715.

Ebert, U. (1992). A reexamination of the optimal nonlinear income tax. *Journal of Public Economics*, 49(1):47–73.

Eckstein, Z. and Zilcha, I. (1994). The effects of compulsory schooling on growth, income distribution and welfare. *Journal of Public Economics*, 54:339–359.

Edwards, J. S. and Keen, M. J. (984). Wealth maximization and the cost of capital: A comment. *Quarterly Journal of Economics*, 99(1):211–214.

Eeckhoudt, L. and Gollier, C. (1995). *Risk: Evaluation, Management and Sharing*. Harvester Wheatsheaf, Hertfordshire.

Eeckhoudt, L., Gollier, C., and Schlesinger, H. (2005). *Economic and Financial Decisions under Risk*. Princeton University Press, Princeton, NJ.

Eeckhoudt, L. and Hansen, P. (1980). Minimum and maximum prices, uncertainty, and the theory of the competitive firm. *American Economic Review*, 70(5):1064–1068.

Eissa, N. (1995). Taxation and labor supply of married women: The Tax Reform Act of 1986 as a natural experiment. NBER Working Paper 5023, National Bureau for Economic Research, Cambridge, MA.

Eissa, N. and Liebman, J. B. (1996). Labor supply response to the earned income tax credit. *Quarterly Journal of Economics*, 111(2):605–637.

Ekelund, R. B. and Hebert, R. F. (1990). *A History of Economic Theory and Method*. McGraw-Hill, New York.

Eldor, R. and Zilcha, I. (1990). Oligopoly, uncertain demand, and forward markets. *Journal of Economics and Business*, 42(1):17–26.

Elmendorf, D. W. (1996). The effect of interest-rate changes on household saving and consumption: A survey. Mimeo, Federal Reserve Board, Washington DC.

Elmendorf, D. W. and Kimball, M. S. (2000). Taxation of labor income and the demand for risky assets. *International Economic Review*, 41(3):801–832.

Evans, O. J. (1983). Tax policy, the interest elasticity of saving, and capital accumulation: Numerical analysis of theoretical models. *American Economic Review*, 73(3):398–410.

Fagan, E. D. (1938). Recent and contemporary theories of progressive taxation. *Journal of Political Economy*, 46(4):457–498.

Fair, R. C. (1971). The optimal distribution of income. *Quarterly Journal of Economics*, 85(4):551–579.

Feehan, J. P. (1998). Public investment: Optimal provision of Hicksian public inputs. *Canadian Journal of Economics*, 31(3):693–707.

Fehr, H. and Kotlikoff, L. (1995). Generational accounting in general equilibrium. NBER Working Paper 5090, National Bureau for Economic Research, Cambridge, MA.

Feldstein, M. and Liebman, J. B. (2002). Social security. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume IV, pages 2245–2324. North-Holland, Amsterdam.

Feldstein, M. S. (1972a). Distributional equity and the optimal structure of public prices. *American Economic Review*, 62(1/2):32–36.

Feldstein, M. S. (1972b). Equity and efficiency in public sector prising: The optimal two-part tariff. *Quarterly Journal of Economics*, 86(2):175–187.

Feldstein, M. S. (1974a). Incidence of a capital income tax in a growing economy with variable savings rates. *Review of Economic Studies*, 41(4):505–513.

Feldstein, M. S. (1974b). Social security, induced retirement, and aggregate capital accumulation. *Journal of Political Economy*, 82:905–926.

Feldstein, M. S. (1974c). Tax incidence in a growing economy with variable factor supply. *Quarterly Journal of Economics*, 88(4):551–573.

Feldstein, M. S. (1976). Social security and saving: The extended life cycle theory. *American Economic Review, Papers and Proceedings*, 66:77–86.

Feldstein, M. S. (1985). The optimal level of social security benefits. *Quarterly Journal of Economics*, 100:303–320.

Feldstein, M. S. (1987). Should social security benefits be means tested? *Journal of Political Economy*, 95:468–484.

Feldstein, M. S. (1997). Transition to a fully funded pension system: Five economic issues. NBER Working Paper 6149, National Bureau for Economic Research, Cambridge, MA.

Feldstein, M. S., editor (1998). *Privatizing Social Security*. University of Chicago Press, Chicago.

Feuerstein, S. (2002). Do coffee roasters benefit from high prices of green coffee? *International Journal of Industrial Organization*, 20(1):89–118.

Fischer, S. (1972). Assets, contingent commodities, and the Slutsky equations. *Econometrica*, 40(2):371–385.

Fisher, I. (1930). *The Theory of Interest*. Macmillan, New York. Reprinted by Augustus M. Kelley, Fairfield, NJ, 1986.

Friend, I. and Hasbrouck, J. (1983). Saving and after-tax rates of return. *Review of Economics and Statistics*, 65(4):537–543.

Frisch, R. (1932). *New Methods of Measuring Marginal Utility*. J. C. B. Mohr, Tübingen.

Frisch, R. (1959). A complete scheme for computing all direct and cross demand elasticities in a model with many sectors. *Econometrica*, 27(2):177–196.

Fullerton, D. and Metcalf, G. E. (2002). Tax incidence. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume IV, pages 1787–1872. North-Holland, Amsterdam.

Galor, O. and Ryder, H. E. (1989). Existence, uniqueness, and stability of equilibrium in an overlapping-generations model with productive capital. *Journal of Economic Theory*, 49:360–375.

Glomm, G. and Ravikumar, B. (1992). Public versus private investment in human capital: Endogenous growth and income inequality. *Journal of Political Economy*, 100:818–834.

Gollier, C. (2001). *The Economics of Risk and Time*. MIT Press, Cambridge, MA.

Gordon, R. H. and Varian, H. R. (1988). Intergenerational risk sharing. *Journal of Public Economics*, 37:185–202.

Gorman, W. M. (1953). Community preference fields. *Econometrica*, 21(1):63–80.

Gorman, W. M. (1959). Are social indifference curves convex? *Quarterly Journal of Economics*, 73(3):485–496.

Goulder, L. H. and Summers, L. H. (1989). Tax policy, asset prices, and growth: A general equilibrium analysis. *Journal of Public Economics*, 38(3):265–296.

Graham, J. W. and Green, C. A. (1984). Estimating the parameters of a household production function with joint products. *Review of Economics and Statistics*, 66(2):277–282.

Gramlich, E. M. (1994). Infrastructure investment: A review essay. *Journal of Economic Literature*, 32:1176–1196.

Grandmont, J.-M. (1985). On endogenous competitive business cycles. *Econometrica*, 53:995–1045.

Gravelle, J. G. and Kotlikoff, L. J. (1989). The incidence and efficiency costs of corporate taxation when corporate and noncorporate firms produce the same good. *Journal of Political Economy*, 97(4):749–780.

Gronau, R. (1977). Leisure, home production, and work – The theory of the allocation of time revisited. *Journal of Political Economy*, 85(6):1099–1123.

Gronau, R. (1980). Home production: A forgotten industry. *Review of Economics and Statistics*, 62(3):408–416.

Grossman, G. M. and Helpman, E. (1991). *Innovation and Growth in the Global Economy*. MIT Press, Cambridge, MA.

Gruber, J. and Saez, E. (2002). The elasticity of taxable income: Evidence and implications. *Journal of Public Economics*, 84(1):1–32.

Guesnerie, R. and Laffont, J.-J. (1984). A complete solution to a class of principal-agent poblems with an application to the control of a self-managed firm. *Journal of Public Economics*, 25(3):329–369.

Guiso, L., Jappelli, T., and Terlizzese, D. (1996). Income risk, borrowing constraints, and portfolio choice. *American Economic Review*, 86(1):158–172.

Hagen, K. P. (1971). Taxation and investment behaviour under uncertainty: A multiperiod portfolio analysis. *Theory and Decision*, 1(3):269–295.

Hall, R. E. (1978). The stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy*, 86(6):971–987.

Ham, J. C. and Reilly, K. T. (2002). Testing intertemporal substitution, implicit contracts, and hours restriction models of the labor market using micro-data. *American Economic Review*, 92(4):905–927.

Hamilton, J. H. (1987). Optimal wage and income taxation with wage uncertainty. *International Economic Review*, 28(2):373–388.

Hamilton, S. F. (1999). Tax incidence under oligopoly: A comparison of policy approaches. *Journal of Public Economics*, 71(2):233–245.

Harberger, A. C. (1962). The incidence of the corporation income tax. *Journal of Political Economy*, 70(3):215–240.

Harris, R. G. and MacKinnon, J. G. (1979). Computing optimal tax equilibria. *Journal of Public Economics*, 11:197–212.

Harrison, R. (1987). Bentham,Jeremy. In Eatwell, J., Milgate, M., and Newman, P., editors, *The New Palgrave: A Dictionary of Economics*. Macmillan, London.

Harrod, R. F. (1939). An essay in dynamic theory. *Economic Journal*, 49(193):14–33.

Harrod, R. F. (1948). *Towards a Dynamic Economics*. Macmillan, London.

Hartwick, J. M. (2000). Labor supply under wage uncertainty. *Economics Letters*, 68(3):319–325.

Hassett, K. A. and Hubbard, R. G. (2002). Tax policy and business investment. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume III, pages 1293–1343. North-Holland, Amsterdam.

Hausman, J. A. (1981a). Exact consumer's surplus and deadweight loss. *AER*, 71(4):662–676.

Hausman, J. A. (1981b). Labor supply. In Aaron, H. and Pechman, J., editors, *How Taxes Affect Economic Activity*, pages 27–83. Brookings Institution, Washington, D.C.

Hausman, J. A. (1985). Taxes and labor supply. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume I, pages 213–263. North-Holland, Amsterdam.

Haveman, R. (1994a). Should generational accounts replace public budgets and deficits? *Journal of Economic Perspectives*, 8:95–111.

Haveman, R. H. (1994b). Optimal taxation and public policy. In Quigley, J. M. and Smolensky, E., editors, *Modern Public Finance*, pages 247–256. Harvard University Press, Cambridge, MA.

Hayashi, F. (1982). Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica*, 50(1):213–224.

Hayashi, F. (1985). Corporate finance side of the Q theory of investment. *Journal of Public Economics*, 27(3):261–280.

Heckman, J. J. (1974). Life cycle consumption and labor supply: An explanation of the relationship between income and consumption over the life cycle. *American Economic Review*, 64(1):188–194.

Heckman, J. J. (1976). A life-cycle model of earnings, learning, and consumption. *Journal of Political Economy*, 84:S11–S44.

Heijdra, B. J. and Ligthart, J. E. (2000). The dynamic macroeconomic effects of tax policy in an overlapping generations model. *Oxford Economic Papers*, 52:677–701.

Heijdra, B. J. and Ligthart, J. E. (2002). Tax policy, the macroeconomy, and intergenerational distribution. *IMF Staff Papers*, 49:106–127.

Heijdra, B. J. and van der Ploeg, F. (2002). *Foundations of Modern Macroeconomics*. Oxford University Press, Oxford.

Hellwig, M. F. (1986). The optimal linear income tax revisited. *Journal of Public Economics*, 31(2):163–179.

Hirshleifer, J. and Riley, J. G. (1992). *The Analytics of Uncertainty and Information*. Cambridge University Press, Cambridge.

Hoel, M. (1990). Efficiency wages and income taxes. *Journal of Economics*, 51:89–99.

Homburg, S. (1990). The efficiency of unfunded pension systems. *Journal of Institutional and Theoretical Economics*, 146:640–647.

Houthakker, H. S. (1960). Additive preferences. *Econometrica*, 28(2):244–257.

Howarth, R. B. (1991). Intergenerational competitive equilibria under technological uncertainty and an exhaustible resource constraint. *Journal of Environmental Economics and Management*, 21:225–243.

Howarth, R. B. (1998). An overlapping generations model of climate-economy interactions. *Scandinavian Journal of Economics*, 100:575–591.

Howarth, R. B. and Norgaard, R. B. (1990). Intergenerational resource rights, efficiency, and social optimality. *Land Economics*, 66:1–11.

Howarth, R. B. and Norgaard, R. B. (1992). Environmental valuation under sustainable development. *American Economic Review, Papers and Proceedings*, 82:473–477.

Howitt, P. and Sinn, H.-W. (1989). Gradual reforms of capital income taxation. *American Economic Review*, 79(1):106–124.

Ihori, T. (1996). *Public Finance in an Overlapping Generations Economy*. Macmillan, London.

Ishii, Y. (1977). On the theory of the competitive firm under price uncertainty: Note. *American Economic Review*, 67(4):768–769.

Jehle, G. A. and Reny, P. J. (2001). *Advanced Microeconomic Theory*. Addison Wesley, Boston, second edition.

Jha, R. (1998). *Modern Public Economics*. Routledge, London and New York.

John, A., Pecchenino, R., Schimmelpfennig, D., and Schreft, S. (1995). Short-lived agents and the long-lived environment. *Journal of Public Economics*, 58:127–141.

John, A. and Pecchenino, R. A. (1994). An overlapping generations model of growth and the environment. *Economic Journal*, 104:1393–1410.

Jones, C. I. (1995). R & D-based models of economic growth. *Journal of Political Economy*, 103:759–784.

Jones, L. E. and Manuelli, R. E. (1992). Finite lifetimes and growth. *Journal of Economic Theory*, 58:171–197.

Jones, L. E., Manuelli, R. E., and Rossi, P. E. (1993). Optimal taxation in models of endogenous growth. *Journal of Political Economy*, 101(3):485–517.

Jones, L. E., Manuelli, R. E., and Rossi, P. E. (1997). On the optimal taxation of capital income. *Journal of Economic Theory*, 73:93–117.

Jones, R. W. (1965). The structure of simple general equilibrium models. *Journal of Political Economy*, 73(6):557–572.

Jones, R. W. (1971a). Distortions in factor markets and the general equilibrium model of production. *Journal of Political Economy*, 79(3):437–459.

Jones, R. W. (1971b). A three-factor model in theory, trade, and history. In Bhagwati, J. N., Jones, R. W., Mundell, R., and Vanek, J., editors, *Trade, Balance of Payments, and Growth*. North-Holland, Amsterdam.

Jones, R. W. (1974). Trade with non-traded goods: The anatomy of inter-connected markets. *Economica*, 41:121–138.

Judd, K. L. (1985). Short-run analysis of fiscal policy in a simple perfect foresight model. *Journal of Political Economy*, 93(2):298–319.

Judd, K. L. (1987a). Debt and distortionary taxation in a simple perfect foresight model. *Journal of Monetary Economics*, 20(1):51–72.

Judd, K. L. (1987b). The welfare cost of factor taxation in a perfect-foresight model. *Journal of Political Economy*, 95(4):675–709.

Judd, K. L. (1998). Taxes, uncertainty, and human capital. *American Economic Review, Papers and Proceedings*, 88(2):289–292.

Kaganovich, M. and Zilcha, I. (1999). Education, social security, and growth. *Journal of Public Economics*, 71:289–309.

Kaldor, N. (1961). Capital accumulation and economic growth. In Lutz, F. A. and Hague, D. C., editors, *The Theory of Capital*, pages 177–222. St Martin's Press, New York.

Kanbur, R. and Tuomala, M. (1994). Inherent inequality and the optimal graduation of marginal tax rates. *Scandinavian Journal of Economics*, 96(2):275–282.

Kanbur, S. M. (1981). Risk taking and taxation: An alternative perspective. *Journal of Public Economics*, 15(2):163–184.

Katz, M. L. and Rosen, H. S. (1985). Tax analysis in an oligopoly model. *Public Finance Quarterly*, 13(1):3–19.

Killingsworth, M. R. and Heckman, J. J. (1996). Female labor supply: A survey. In Ashenfelter, O. and Layard, R., editors, *Handbook of Labor Economics*, volume I, pages 103–204. North-Holland, Amsterdam.

Kimball, M. S. (1987). Making sense of two-sided altruism. *Journal of Monetary Economics*, 20(2):301–326.

Kimball, M. S. (1990). Precautionary saving in the small and in the large. *Econometrica*, 58(1):53–73.

King, M. (1974). Taxation and the cost of capital. *Review of Economic Studies*, 41(1):21–35.

Kleven, H. J. (2004). Optimum taxation and the allocation of time. *Journal of Public Economics*, 88:545–557.

Kleven, H. J. and Kreiner, C. T. (2004). Optimal taxation of married couples with household production. Mimeo, University of Copenhagen.

Kleven, H. J., Richter, W. F., and Sørenson, P. B. (2000). Optimal taxation with household production. *Oxford Economics Papers*, 52:584–594.

Konrad, K. A. (2001). Privacy and time-consistent optimal labor income taxation. *Journal of Public Economics*, 79(3):503–519.

Koskela, E. and Vilmunen, J. (1996). Tax progression is good for employment in popular models of trade union behaviour. *Labour Economics*, 3:65–80.

Kotlikoff, L. J. (1993a). From deficit delusion to the fiscal balance rule: Looking for an economically meaningful way to assess fiscal policy. In Felderer, B., editor, *Public Pension Economics*. Springer Verlag, New York.

Kotlikoff, L. J. (1993b). *Generational Accounting*. Free Press, New York.

Kotlikoff, L. J. and Summers, L. H. (1979). Tax incidence in a life cycle model with variable labor supply. *Quarterly Journal of Economics*, 94:705–718.

Kotlikoff, L. J. and Summers, L. H. (1987). Tax incidence. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume II, pages 1043–1092. North-Holland, Amsterdam.

Kreps, D. M. (1988). *Notes on the Theory of Choice*. Westview Press, Boulder, CO.

Krueger, A. B. and Meyer, B. D. (2002). Labor supply effects of social insurance. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume IV, pages 2327–2392. North-Holland, Amsterdam.

Kurz, M. (1985). Cooperative oligopoly equilibrium. *European Economic Review*, 27(1):3–24.

Laffont, J.-J. (1987). Incentives and the allocation of public goods. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume II, pages 537–569. North-Holland, Amsterdam.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, 74:132–157.

Leach, J. (2004). *A Course in Public Economics*. Cambridge University Press, Cambridge.

Lee, C.-I. (2001). Finite sample bias in IV estimation of intertemporal labor supply models: Is the intertemporal substitution elasticity really small? *Review of Economics and Statistics*, 83(4):638–646.

Leland, H. E. (1968). Saving and uncertainty: The precautionary demand for saving. *Quarterly Journal of Economics*, 82(3):465–473.

Leland, H. E. (1972). Theory of the firm facing uncertain demand. *American Economic Review*, 62(3):278–291.

Léonard, D. and Long, N. V. (1992). *Optimal Control Theory and Static Optimization in Economics*. Cambridge University Press, Cambridge.

Levhari, D. and Weiss, Y. (1974). The effect of risk on the investment in human capital. *American Economic Review*, 64(6):950–963.

Lindbeck, A. and Persson, M. (2003). The gains from pension reform. *Journal of Economic Literature*, 41:74–112.

Lipsey, R. G. and Lancaster, K. (1956). The general theory of second best. *Review of Economic Studies*, 24(1):11–32.

Little, I. M. D. (1957). *A Critique of Welfare Economics*. Oxford University Press, Oxford, second edition.

Lockwood, B. and Manning, A. (1993). Wage setting and the tax system: Theory and evidence for the United Kingdom. *Journal of Public Economics*, 52(1):1–29.

Lollivier, S. and Rochet, J.-C. (1982). Bunching and second-order conditions: A note on optimal tax theory. *Journal of Economic Theory*, 31(2):392–400.

Loury, G. C. (1981). Intergenerational transfers and the distribution of earnings. *Econometrica*, 49:843–867.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42.

Lucas, R. E. and Rapping, L. A. (1969). Real wages, employment, and inflation. *Journal of Political Economy*, 77:721–754.

Machina, M. J. (1987). Expected utility hypothesis. In Eatwell, J., Milgate, M., and Newman, P., editors, *The New Palgrave: A Dictionary of Economics*. Macmillan, London.

MaCurdy, T. E. (1981). An empirical model of labor supply in a life-cycle setting. *Journal of Political Economy*, 89(6):1059–1085.

MaCurdy, T. E. (1985). Interpreting empirical models of labor supply in an intertemporal framework with uncertainty. In Heckman, J. J. and Singer, B., editors, *Longitudinal Analysis of Labor Market Data*, pages 111–155. Cambridge University Press, Cambridge.

MaCurdy, T. E. (1992). Work disincentive effects of taxes: A reexamination of some evidence. *American Economic Review, Papers and Proceedings*, 82(2):243–249.

Malinvaud, E. (1972). *Lectures on Microeconomic Theory*. North-Holland, Amsterdam.

Manning, R., Markusen, J. R., and McMillan, J. (1985). Paying for public inputs. *American Economic Review*, 75(1):235–238.

Manser, M. and Brown, M. (1980). Marriage and household decision-making: A bargaining analysis. *International Economic Review*, 21(1):31–43.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1):77–91.

Marshall, A. (1920). *Principles of Economics*. Macmillan, London, eighth edition.

Martina, A. (2000). Could a basic result in optimal commodity taxation theory have been derived much earlier than it was? *History of Political Economy*, 32(1):139–157.

Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press, New York.

McDonald, I. M. and Solow, R. M. (1985). Wages and employment in a segmented labor market. *Quarterly Journal of Economics*, 100(4):1115–1141.

McElroy, M. and Horney, M. J. (1981). Nash-bargained household decisions: Toward a generalization of the theory of demand. *International Economic Review*, 22(2):333–349.

McFadden, D. L. (1978). Cost, revenue, and profit functions. In Fuss, M. A. and McFadden, D. L., editors, *Production Economics: A Dual Approach to Theory and Applications*, volume I, pages 3–109. North-Holland, Amsterdam.

McLure, C. E. (1975). General equilibrium incidence analysis: The Harberger model after ten years. *Journal of Public Economics*, 4(2):125–161.

McLure, C. E. (1979). *Must Corporate Income Be Taxed Twice?* Brookings Institution, Washington, DC.

McLure, C. E. and Thirsk, W. R. (1975). A simplified exposition of the Harberger model, I: Tax incidence. *National Tax Journal*, 28(1):1–27.

McLure, C. E., Thirsk, W. R., and Klein, R. (1975). A simplified exposition of the Harberger model, II: Expenditure incidence. *National Tax Journal*, 28(2):195–207.

McManus, M. (1959). Comments on the general theory of second best. *Review of Economic Studies*, 26(3):209–224.

Meade, J. E. (1952). External economies and diseconomies in a competitive situation. *Economic Journal*, 62(245):54–67.

Menezes, C. F. and Wang, X. H. (2005). Duality and Slutsky income and substitution effects of increases in wage rate uncertainty. *Oxford Economic Papers*, 57:545–557.

Menezes, C. F., Wang, X. H., and Bigelow, J. P. (2005). Duality and consumption decisions under income and price risk. *Journal of Mathematical Economics*, 41(3):387–405.

Michel, P. and de la Croix, D. (2000). Myopic and perfect foresight in the OLG model. *Economics Letters*, 67:53–60.

Mieszkowski, P. M. (1969). Tax incidence theory: The effects of taxes on the distribution of income. *Journal of Economic Literature*, 7(4):1103–1124.

Miller, M. H. and Scholes, M. S. (1978). Dividends and taxes. *Journal of Financial Economics*, 6(4):333–364.

Miller, M. H. and Scholes, M. S. (1982). Dividends and taxes: Some empirical evidence. *Journal of Political Economy*, 90(6):1118–1141.

Milleron, J.-C. (1972). Theory of value with public goods: A survey article. *Journal of Economic Theory*, 5(3):419–477.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38(2):175–208.

Mirrlees, J. A. (1972). On producer taxation. *Review of Economic Studies*, 39(1):105–111.

Mirrlees, J. A. (1975). Optimal commodity taxation in a two-class economy. *Journal of Public Economics*, 4(1):27–33.

Mirrlees, J. A. (1976). Optimal tax theory: A synthesis. *Journal of Public Economics*, 6(4):327–358.

Mirrlees, J. A. (1986). The theory of optimal taxation. In Arrow, K. J. and Intriligator, M. D., editors, *Handbook of Mathematical Economics*, volume III, pages 1197–1249. North-Holland, Amsterdam.

Mirrlees, J. A. (1990). Taxing uncertain incomes. *Oxford Economic Papers*, 42(1):34–45.

Moffitt, R. A. (2002). Welfare programs and labor supply. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume IV, pages 2393–2430. North-Holland, Amsterdam.

Mossin, J. (1968). Taxation and risk-taking: An expected utility approach. *Economica*, 35:74–82.

Mourmouras, A. (1993). Conservationist government policies and intergenerational equity in an overlapping generations model with renewable resources. *Journal of Public Economics*, 51:249–268.

Munro, A. (1991). The optimal provision of private goods. *Journal of Public Economics*, 44(2):239–261.

Musgrave, R. A. (1959). *The Theory of Public Finance*. McGraw-Hill, New York.

Musgrave, R. A. (1985). A brief history of fiscal doctrine. In Auerbach, A. and Feldstein, M., editors, *Handbook of Public Economics*, volume I, pages 1–59. North-Holland, Amsterdam.

Muth, R. F. (1966). Household production and consumere demand functions. *Econometrica*, 34(3):699–708.

Myles, G. D. (1995). *Public Economics*. Cambridge University Press, Cambridge.

Nath, S. K. (1969). *A Reappraisal of Welfare Economics*. Routledge & Kegan Paul, London.

Neary, J. P. (1978). Dynamic stability and the theory of factor market distortions. *American Economic Review*, 68(4):671–682.

Nerlove, M. and Raut, L. K. (1997). Growth models with endogenous population: A general framework. In Rosenzweig, M. R. and Stark, O., editors, *Handbook of Population and Family Economics*. North-Holland, Amsterdam.

Newbery, D. and Stern, N. (1987). Dynamic issues. In Newbery, D. and Stern, N., editors, *The Theory of Taxation for Developing Countries*, pages 114–140. Oxford University Press, New York.

Oakland, W. H. (1987). Theory of public goods. In Auerbach, A. and Feldstein, M., editors, *Handbook of Public Economics*, volume II, pages 485–535. North-Holland, Amsterdam.

O'Connell, S. A. and Zeldes, S. P. (1988). Rational Ponzi games. *International Economic Review*, 29:431–450.

Pencavel, J. (1986). Labor supply of men: A survey. In Ashenfelter, O. and Layard, R., editors, *Handbook of Labor Economics*, volume I, pages 3–102. North-Holland, Amsterdam.

Persson, T. and Tabellini, G. (2000). *Political Economics: Explaining Economic Policy*. MIT Press, Cambridge, MA.

Pestieau, P. M. (1974). Optimal taxation and discount rate for public investment in a growth setting. *Journal of Public Economics*, 3:217–235.

Phelps, E. S. (1973). Taxation of wage income for economic justice. *Quarterly Journal of Economics*, 87(3):331–354.

Piggott, J. and Whalley, J. (1996). The tax unit and household production. *Journal of Political Economy*, 104(2):398–418.

Pigou, A. C. (1947). *A Study in Public Finance*. Macmillan, London, third edition.

Piketty, T. (1997). La redistribution fiscale face au chômage. *Revue Française d'Economie*, 12(1):157–201.

Pissarides, C. A. (1998). The impact of employment tax cuts on unemployment and wages; The role of unemployment benefits and tax structure. *European Economic Review*, 42(1):155–183.

Plug, E. J. S., van Praag, B. M. S., and Hartog, J. (1999). If we knew ability, how would we tax individuals? *Journal of Public Economics*, 72(2):183–211.

Pollak, R. A. (2005). Bargaining power in marriage: Earnings, wage rates, and household production. NBER Working Paper 11239, National Bureau for Economic Research, Cambridge, MA.

Poterba, J. M. (1996). Retail price reactions to changes in state and local sales taxes. *National Tax Journal*, 49(2):165–176.

Poterba, J. M. and Summers, L. H. (1983). Dividend taxes, corporate investment, and Q. *Journal of Public Economics*, 22(2):135–167.

Poterba, J. M. and Summers, L. H. (1985). The economic effects of dividend taxation. In Altman, E. I. and Subrahmanyam, M. G., editors, *Recent Advances in Corporate Finance*, pages 227–284. Richard D. Irwin, Homewood, IL.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32(1-2):122–136.

Ramsey, F. P. (1927). A contribution to the theory of taxation. *Economic Journal*, 37(145):47–61.

Ramsey, F. P. (1928). A mathematical theory of savings. *Economic Journal*, 38:543–559.

Rawls, J. (1971). *A Theory of Justice*. Oxford University Press, Oxford.

Rebelo, S. (1991). Long-run policy analysis and long-run growth. *Journal of Political Economy*, 99(3):500–521.

Reichlin, R. (1986). Equilibrium cycles in an overlapping generations economy with production. *Journal of Economic Theory*, 40:89–102.

Renström, T. I. (1998). On optimal taxation in overlapping generations economies. Mimeo, University of Birmingham.

Robbins, L. (1998). *A History of Economic Thought*. Princeton University Press, Princeton, NJ. Edited by S. G. Medema and Warren J. Samuels.

Roberts, D. J. (2004). *The Modern Firm: Organizational Design for Performance and Growth*. Oxford University Press, Oxford.

Robinson, J. A. and Srinivasan, T. N. (1997). Long-term consequences of population growth: Technological change, natural resources, and the environment. In Rosenzweig, M. R. and Stark, O., editors, *Handbook of Population and Family Economics*. North-Holland, Amsterdam.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94:1002–1037.

Romer, P. M. (1987). Growth based on increasing returns due to specialization. *American Economic Review, Papers and Proceedings*, 77:56–62.

Romer, P. M. (1989). Capital accumulation in the theory of long-run growth. In Barro, R. J., editor, *Modern Business Cycle Theory*, pages 51–127. Basil Blackwell, Oxford.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98:S71–S101.

Rothschild, M. and Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, 2(3):225–243.

Rothschild, M. and Stiglitz, J. E. (1971). Increasing risk: II. Its economic consequences. *Journal of Economic Theory*, 3(1):66–84.

Sadka, E. (1976). On income distribution, incentive effects and optimal income taxation. *Review of Economic Studies*, 43(2):261–267.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68(1):205–229.

Saez, E. (2002a). The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics*, 83(2):217–230.

Saez, E. (2002b). Optimal income transfer programs: Intensive versus extensive labor supply responses. *Quarterly Journal of Economics*, 117(3):1039–1073.

Salanié, B. (2003). *The Economics of Taxation*. MIT Press, Cambridge, MA.

Samuelson, P. A. (1947). *Foundations of Economic Analysis*. Harvard University Press, Cambridge, MA.

Samuelson, P. A. (1954). The pure theory of public expenditure. *Review of Economics and Statistics*, 36(4):387–389.

Samuelson, P. A. (1955). Diagrammatic exposition of a theory of public expenditure. *Review of Economics and Statistics*, 37(4):350–356.

Samuelson, P. A. (1956). Social indifference curves. *Quarterly Journal of Economics*, 70(1):1–22.

Samuelson, P. A. (1958). Aspects of public expenditure theories. *Review of Economics and Statistics*, 40(4):332–338.

Samuelson, P. A. (1968). The two-part golden rule deduced as the asymptotic turnpike of catenary motions. *Western Economic Journal*, 6:85–89.

Samuelson, P. A. (1969a). Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, 51(3):239–246.

Samuelson, P. A. (1969b). Pure theory of public expenditure and taxation. In Margolis, J. and Guitton, H., editors, *Public Economics*, pages 98–123. Macmillan, London.

Samuelson, P. A. (1975a). The optimum growth rate for population. *International Economic Review*, 16(3):531–538.

Samuelson, P. A. (1975b). Optimum social security in a life-cycle growth model. *International Economic Review*, 16(3):539–544.

Samuelson, P. A. (1986). Theory of optimal taxation. *Journal of Public Economics*, 30(2):137–143. (Reprint of unpublished memorandum for the U.S. Treasury, 1951).

Sandmo, A. (1969). Capital risk, consumption, and portfolio choice. *Econometrica*, 37(4):586–599.

Sandmo, A. (1970). The effect of uncertainty on saving decisions. *Review of Economic Studies*, 37(3):353–360.

Sandmo, A. (1971). On the theory of the competitive firm under price uncertainty. *American Economic Review*, 61(1):65–73.

Sandmo, A. (1972). Optimality rules for the provision of collective factors of production. *Journal of Public Economics*, 1(1):149–158.

Sandmo, A. (1974a). A note on the structure of optimal taxation. *American Economic Review*, 64(4):701–706.

Sandmo, A. (1974b). Two-period models of consumption decisions under uncertainty: A survey. In Drèze, J. H., editor, *Allocation Under Uncertainty: Equilibrium and Optimality*, pages 24–35. Macmillan, London.

Sandmo, A. (1975). Optimal taxation in the presence of externalities. *Swedish Journal of Economics*, 77(1):86–98.

Sandmo, A. (1976). Optimal taxation: An introduction to the literature. *Journal of Public Economics*, 6(1-2):37–54.

Sandmo, A. (1977). Portfolio theory, asset demand and taxation: Comparative statics with many assets. *Review of Economic Studies*, 44(2):369–379.

Sandmo, A. (1983). Progressive taxation, redistribution, and labor supply. *Scandinavian Journal of Economics*, 85(3):311–323.

Sandmo, A. (1985). The effects of taxation on savings and risk taking. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume I, pages 265–311. North-Holland, Amsterdam.

Sandmo, A. (1990). Tax distortions and household production. *Oxford Economic Papers*, 42(1):78–90.

Sandmo, A. (1998). Redistribution and the marginal cost of public funds. *Journal of Public Economics*, 70(3):365–382.

Sargent, T. J. (1987). *Dynamic Macroeconomic Theory*. Harvard University Press, Cambridge, MA.

Schmeidler, D. and Wakker, P. (1987). Expected utility and mathematical expectation. In Eatwell, J., Milgate, M., and Newman, P., editors, *The New Palgrave: A Dictionary of Economics*. Macmillan, London.

Schumpeter, J. A. (1954). *History of Economic Analysis*. George Allen and & Unwin, London.

Seade, J. K. (1977). On the shape of optimal tax schedules. *Journal of Public Economics*, 7(2):203–235.

Seade, J. K. (1980a). On the effects of entry. *Econometrica*, 48(2):479–489.

Seade, J. K. (1980b). The stability of Cournot revisited. *Journal of Economic Theory*, 23(1):15–27.

Seade, J. K. (1982). On the sign of the optimal marginal income tax. *Review of Economic Studies*, 49(4):637–643.

Seade, J. K. (1985). Profitable cost increases and the shifting of taxation: Equilibrium responses of markets in oligopoly. Warwick Economic Research Paper 260, University of Warwick, Warwick, Coventry.

Shapiro, C. and Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review*, 74(3):433–444.

Sheshinski, E. (1972). The optimal linear income tax. *Review of Economic Studies*, 39(3):297–302.

Shoven, J. B. and Whalley, J. (1972). A general equilibrium calculation of the effects of differential taxation of income from capital in the U.S. *Journal of Public Economics*, 1(3-4):281–321.

Shoven, J. B. and Whalley, J. (1984). Applied general equilibrium models of taxation and international trade. *Journal of Economic Literature*, 22(3):1007–1051.

Shoven, J. B. and Whalley, J. (1992). *Applying General Equilibrium*. Cambridge University Press, Cambridge.

Silberberg, E. and Suen, W. (2001). *The Structure of Economics: A Mathematical Analysis*. McGraw-Hill, New York, third edition.

Sinn, H.-W. (1987). *Capital Income Taxation and Resource Allocation*. North-Holland, Amsterdam.

Sinn, H.-W. (1991). Taxation and the cost of capital: The old view, the new view, and another view. In Bradford, D., editor, *Tax Policy and the Economy*, volume 5, pages 25–54. MIT Press, Cambridge, MA.

Sinn, H.-W. (2000). Why a funded pension system is useful and why it is not useful. *International Tax and Public Finance*, 7:389–410.

Slemrod, J. (1998). Methodological issues in measuring and interpreting taxable income elasticities. *National Tax Journal*, 51(4):773–788.

Slemrod, J. and Yitzhaki, S. (2002). Tax avoidance, evasion, and administration. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume III, pages 1423–1470. North-Holland, Amsterdam.

Slemrod, J., Yitzhaki, S., Mayshar, J., and Lundholm, M. (1994). The optimal two-bracket linear income tax. *Journal of Public Economics*, 53(2):269–290.

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1):65–94.

Sørensen, P. B. (1994). Some old and new issues in the theory of corporate income taxation. *Finanzarchiv*, 51(4):425–456.

Starrett, D. A. (1988). Effects of taxes on saving. In Aaron, H. J., Galper, H., and Pechman, J. A., editors, *Uneasy Compromise: Problems of a Hybrid Income-Consumption Tax*, pages 237–259. Brookings Institution, Washington, DC.

Stern, N. (1987a). Aspects of the general theory of tax reform. In Newbery, D. and Stern, N., editors, *The Theory of Taxation for Developing Countries*, pages 60–91. Oxford University Press, New York.

Stern, N. (1987b). The theory of optimal commodity and income taxation: An introduction. In Newbery, D. and Stern, N., editors, *The Theory of Taxation for Developing Countries*, pages 22–59. Oxford University Press, New York.

Stern, N. H. (1976). On the specification of models of optimum income taxation. *Journal of Public Economics*, 6:123–162.

Stern, N. H. (1986). On the specification of labour supply functions. In Blundell, R. and Walker, I., editors, *Unemployment, Search and Labour Supply*, pages 143–189. Cambridge University Press, Cambridge.

Stern, N. H. (1987c). The effects of taxation, price control and government contracts in oligopoly and monpolistic competition. *Journal of Public Economics*, 32(2):133–158.

Stiglitz, J. E. (1969). The effects of income, wealth, and capital gains taxation on risk-taking. *Quarterly Journal of Economics*, 83(2):263–283.

Stiglitz, J. E. (1982). Self-selection and Pareto efficient taxation. *Journal of Public Economics*, 17(2):213–240.

Stiglitz, J. E. (1987). Pareto efficient and optimal taxation and the new new welfare economics. In Auerbach, A. and Feldstein, M., editors, *Handbook of Public Economics*, volume II, pages 991–1042. North-Holland, Amsterdam.

Stokey, N. L. (1980). A life-cycle model of the linear income tax. *Review of Economic Studies*, 47(4):777–785.

Stokey, N. L. and Rebelo, S. (1995). Growth effects of flat-tax rates. *Journal of Political Economy*, 103(3):519–550.

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, 23:165–180.

Sullivan, D. (1985). Testing hypotheses about firm behavior in the cigarette industry. *Journal of Political Economy*, 93(3):586–598.

Summers, L. H. (1981a). Capital taxation and accumulation in a life cycle growth model. *American Economic Review*, 71(4):533–544.

Summers, L. H. (1981b). Taxation and corporate investment: A q-theory approach. *Brookings Papers on Economic Activity*, 1:67–140.

Summers, L. H. (1984). The after-tax rate of return affects private savings. *American Economic Review, Papers and Proceedings*, 74(2):249–253.

Summers, L. H., Gruber, J., and Vergara, R. (1993). Taxation and the structure of labor markets: The case of corporatism. *Quarterly Journal of Economics*, 108(2):385–411.

Sumner, D. A. (1981). Measurement of monopoly behavior: An application to the cigarette industry. *Journal of Political Economy*, 89(5):1010–1019.

Swan, T. W. (1956). Economic growth and capital accumulation. *Economic Record*, 32:334–361.

Sydsæter, K., Strøm, A., and Berck, P. (2000). *Economists' Mathematical Manual*. Springer, Berlin, third edition.

Tawada, M. (1980). The production possibility set with public intermediate goods. *Econometrica*, 48(4):1005–1012.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64(5):416–424.

Tirole, J. (1985). Asset bubbles and overlapping generations. *Econometrica*, 53:1071–1100.

Tobin, J. (1958). Liquidity preference as behavior towards risk. *Review of Economic Studies*, 25(2):65–86.

Tresch, R. W. (2002). *Public Finance: A Normative Theory*. Academic Press, San Diego, CA, second edition.

Tuomala, M. (1990). *Optimal Income Tax and Redistribution*. Oxford University Press, Oxford.

Turnovsky, S. J. (1990). The effects of taxes and dividend policy on capital accumulation and macroeconomic behavior. *Journal of Economic Dynamics and Control*, 14(3-4):491–521.

Uzawa, H. (1961). On a two-sector model of economic growth, I. *Review of Economic Studies*, 29(1):40–47.

Uzawa, H. (1963). On a two-sector model of economic growth, II. *Review of Economic Studies*, 30(2):105–118.

Uzawa, H. (1964). Optimal growth in a two-sector model of capital accumulation. *Review of Economic Studies*, 31(1):1–24.

Uzawa, H. (1965). Optimum technical change in an aggregative model of economic growth. *International Economic Review*, 6:18–31.

Varian, H. R. (1980). Redistributive taxation as social insurance. *Journal of Public Economics*, 14(1):49–68.

Varian, H. R. (1992). *Microeconomic Analysis*. W. W. Norton, New York, third edition.

Varian, H. R. (1994a). Sequential contributions to public goods. *Journal of Public Economics*, 53(4):165–186.

Varian, H. R. (1994b). A solution to the problem of externalities when agents are well-informed. *American Economic Review*, 84(5):1278–1293.

Verbon, H. A. A. (1989). Conversion policies for public pension plans in a small open economy. In Gustafsson, B. A. and Klevmarken, N. A., editors, *The Political Economy of Social Security*. North-Holland, Amsterdam.

Vives, X. (1999). *Oligopoly Pricing: Old Ideas and New Tools*. MIT Press, Cambridge, MA.

Wales, T. J. and Woodland, A. D. (1977). Estimation of the allocation of time for work, leisure, and housework. *Econometrica*, 45(1):115–132.

Wales, T. J. and Woodland, A. D. (1979). Labour supply and progressive taxes. *Review of Economic Studies*, 46(1):83–95.

Wallace, N. (1981). A Modigliani-Miller theorem for open-market operations. *American Economic Review*, 71:267–274.

Weil, D. N. (1997). The economics of population aging. In Rosenzweig, M. and Stark, O., editors, *Handbook of Population and Family Economics*. North-Holland, Amsterdam.

Weiss, A. (1991). *Efficiency Wages: Models of Unemployment, Layoffs, and Wage Dispersion*. Oxford University Press, Oxford.

Wildasin, D. E. (1990). Non-neutrality of debt with endogenous fertility. *Oxford Economic Papers*, 42:414–428.

Wilson, J. D. (1991a). Optimal public good provision in the Ramsey tax model: A generalization. *Economics Letters*, 35(1):57–61.

Wilson, J. D. (1991b). Optimal public good provision with limited lump-sum taxation. *American Economic Review*, 81(1):153–166.

Wright, C. (1967). Some evidence on the interest elasticity of consumption. *American Economic Review*, 57(4):850–855.

Wright, C. (1969). Saving and the rate of interest. In Harberger, A. C. and Bailey, M. J., editors, *The Taxation of Income from Capital*, pages 275–300. Brookings Institution, Washingtom, DC.

Yitzhaki, S. (1974). A note on 'Income tax evasion: A theoretical analysis'. *Journal of Public Economics*, 3(2):201–202.

Zhang, J. (1995). Social security and endogenous growth. *Journal of Public Economics*, 58:185–213.

Zhang, J. (1996). Optimal public investments in education and endogenous growth. *Scandinavian Journal of Economics*, 98:387–404.

Zilcha, I. (1990). Dynamic efficiency in overlapping generations models with stochastic production. *Journal of Economic Theory*, 52:364–379.

Zilcha, I. (1991). Characterizing efficiency in stochastic overlapping generations models. *Journal of Economic Theory*, 55:1–16.

Zilcha, I. and Eldor, R. (2004). Firm's output under uncertainty and asymmetric taxation. *Economica*, 71:141–153.