CHRISTOPH BRAUN

OPTIMAL TAXATION OF HUMAN CAPITAL

Optimal Taxation of Human Capital

Inauguraldissertation
zur Erlangung des akademischen Grades
Doctor rerum politicarum
an der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Technischen Universität Dortmund

vorgelegt von
Diplom-Volkswirt Christoph Braun
aus Siegen

Tag der Einreichung: 28.01.2011
Tag der Disputation: 04.05.2011

DISSERTATION COMMITTEE:

Prof. Dr. Wolfram F. Richter (Chairman)

Prof. Dr. Andreas Schabert

Prof. Dr. Ludger Linnemann

LOCATION:

Dortmund

FACULTY:

Technische Univeristät Dortmund

Wirtschafts- und Sozialwissenschaftliche Fakultät

Dean: Prof. Dr. Wolfgang B. Schünemann

Für meine Tochter Eleonore.

ACKNOWLEDGMENTS

This dissertation was written while I was a Ph.D student at the Ruhr Graduate School in Economics. The whole enterprise benefited from the help of Professor Wolfram F. Richter. His support, guidance, and patience cannot be overestimated. I owe the greatest intellectual debt to him.

Over the course of time, I have discussed this work with many people at home, at work, at conferences, and at many other occasions. For some, it has hopefully been interesting and maybe even stimulating; for others, it may have sometimes been hard to follow my ideas. I would like to express my gratitude to all people who listened to me, who supported me, and who talked with me.

Financial support by the Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz is gratefully acknowledged.

Last but not least, I thank my beloved wife, Judith, for her continuous support and encouragement. Finishing this enterprise would not have been possible without her.

I dedicate this dissertation to my daughter, Eleonore. She is my fountain of joy who always reminds me that seeing her laughing and growing up is even more fascinating and mysterious than thinking about economics.

CONTENTS

LIST OF TABLES

1

INTRODUCTION

This dissertation is on optimal taxation of human capital. Why it is intriguing to examine this topic shall be the subject of what follows. I will first introduce and briefly discuss the notion of human capital. Then I will give a summary of the empirical and theoretical findings on (i) the link between human capital and economic well-being and (ii) how tax policies affect the accumulation of human capital. By then, I hope, the reader will be convinced that taxation matters for human capital and hence it matters for economic well-being. Taking this as given, it is fruitful to further pursue and to tackle the leading question of this dissertation:

> How should the tax system be optimally designed to promote the accumulation of human capital to maximize economic well-being?

The basic approach to this question will then be presented along with a discussion of related literature. This introduction concludes with a summary of my contributions to the literature on optimal taxation of human capital.

1.1 HUMAN CAPITAL AND ECONOMIC WELL-BEING

Education policy ranks high on the political agenda. The following statements are taken from German and US American politics:

- “Education is a key to personal prosperity, social justice, and wealth.” (Coalition agreement of the CDU, CSU and FDP, p. 6, 2009)

- “Growth. Education. Cohesion. Leading Germany to new Strength.” (Title of Federal Chancellor Angela Merkel’s inaugural policy statement, 2009)

- “The [] challenge we must address is the urgent need to expand the promise of education in America. In a global economy where the most valuable skill you can sell is your knowledge, a good education is no longer just a pathway to opportunity - it is a pre-requisite.” (President Barack Obama’s speech at a joint session of the United States Congress, 2009)

Put a little bit less flowery, one constituent of economic well-being is human capital. A deep and sound understanding of what human capital is and how it affects economic well being is therefore indispensable. In this section I would like to first discuss the notion of human capital and then survey literature that attempts at identifying the links between human capital and economic well being.

1.1.1 *Notion of Human Capital*

The term human capital was introduced into economics by Theodore W. Schultz, who was awarded the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel joint with Sir Arthur Lewis in 1979 “for their pioneering research into economic development research with particular consideration of the problems of developing countries”. Schultz’s seminal contribution does not give a formal definition, the usage of the term human capital is motivated by the following example:

> Much of what we call consumption constitutes investment in human capital. Direct expenditures on education, health, and internal migration to take advantage of better job opportunities are clear examples. Earnings foregone by mature students attending school and by workers acquiring on-the-job training are equally clear examples. Yet, nowhere do these enter our national accounts. The use of leisure time to improve skills and knowledge is widespread and it too is unrecorded. In these and similar ways the quality of human effort can be greatly improved and its productivity enhanced. I shall contend that such investments in human capital accounts for most of the impressive rise in real earnings per worker. (Schultz, 1961a, p. 1)

Gary S. Becker, who was awarded the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel "for having extended the domain of microeconomic analysis to a wide range of human behaviour and interaction, including nonmarket behaviour" in 1992, also did not give a formal definition of what he meant by human capital. He writes that

> expenditures on education, training, medical care, etc., are investments in capital. However, these produce human, not physical or financial, capital because you cannot separate a person from his or her knowledge, skills, health, or

> values the way it is possible to move financial and physical assets while the owner stays put. (Becker, 1993, p. 16)

Both Schultz and Becker have an analogy to the stock of physical capital in mind when they speak of human capital. The way savings act as a flow that increases the stock of physical capital, which then can be used in production, it is only natural also to speak of human capital as a stock variable in which various types of investments are made, of which education is only one example next to many others. Although the other types of investments in human capital are also important, education is the type of investment that will be further studied below. Thus, in a narrower sense, this dissertation could be equally titled as "Optimal Taxation of Education".

From the great many ways how the stock of human capital can be increased, it becomes only evident that is rather difficult to precisely determine its size. Whereas it is relatively easy to determine the value of tangible assets, which is the stock of physical capital holdings, it is relatively complicated to determine the value of the stock of human capital, which is part of the intangible assets. Empirical studies therefore differ with respect to how the stock of human capital is measured.

The rather abstract term "well-being" in the beginning is used to indicate that it is not clear how and to which kind of well-being human capital contributes. President Obama's quote rather refers to the individual, microeconomic view, whereas the two quotes from German politics rather refer to the societal, macroeconomic view. At the macroeconomic level well-being refers to the rate of growth in national income or the level of national income, which may also be expressed in per capita terms. The social rate of return to investments in human capital is another measure. At the microeconomic level one studies the private rate of return to investments in human capital.

There is a vast literature that examines the links between human capital and economic well-being at the macro and micro level. Many excellent surveys of each strand of the literature are already available. The macro literature is reviewed by Topel (1999) and Sianesi and Reenen (2003). Card (1999) and Harmon, Oosterbeek, and Walker (2003) review the micro literature. Finally, Krueger and Lindahl (2001) offer a discussion that tries to bridge the gap between the macro and micro approaches. In what follows, I will draw on these surveys, only sketch the most important contributions and report major results. This is meant only to show that the quoted politicians are right about that human capital is one of the most important constituents of economic well-being.

1.1.2 *Macro Growth Literature*

Very early attempts have been made by Schultz (1961b), Denison (1962, 1967). They differ with respect to the methodology applied to measure what constitutes human capital.[1] The starting point of both Denison's and Schultz's work is the growth accounting exercise pioneered by Solow (1957). In Denison's study, the production factor labor is adjusted to account for different schooling levels to have a mea-

1 Bowman (1964) and Psacharopoulos (1973, 111-118) provide surveys and discussions of their approaches.

sure of quality. He simply asks: "What was the division of growth among the sources?" (Denison, 1967, p. 296) For instance, the United States grew by 3.36% between 1950 and 1962. Education's contribution to growth was 0.49%. As a result, education's contribution to growth amounted to 15%. Further results can be found in table 1.1 in appendix 1.A.

Other contributions in the tradition of growth accounting, but using other measures of the stock of human capital, include Jorgenson and Fraumeni (1992), Mankiw, Romer, and Weil (1992), Hall and Jones (1999) and Klenow and Rodríguez-Clare (1997).

Work devoted to actually identifying the correlations between different regressors of human capital and measures of economic well being has been initiated by Barro (1991). To honor his seminal contribution, these cross-country growth regressions are sometimes referred to as Barro-regressions. Sianesi and Reenen (2003) have nicely compiled a table that tries to ease the interpretation and comparability of studies done by Barro (1991), Levine and Renelt (1992), Murphy, Shleifer, and Vishny (1991), Barro (1998), Hanushek and Kim (1995), Gemmell (1996), Judson (1998), Englander and Gurney (1994), Barro and Lee (1994), Benhabib and Spiegel (1994), Mankiw, Romer, and Weil (1992), de la Fuente and Doménech (2006) and Bassanini and Scarpetta (2001). Table 1.2 shows the results. The overall message from these studies is that a 1 percentage increase in human capital is significantly correlated to an increase in GDP. For instance, Barro (1991) finds that 1 percentage point increase in primary (secondary) school enrolment rates is significantly correlated to a 2.5 (3.0) percentage points increase in per capita GDP growth rates.

Another macro related approach is pursued by the OECD in its 'Education at a Glance' series, which is based on investment theory. The discount rate is the basis that allows to compare cash flows over time. The internal rate of return is the discount rate at which the discounted streams of benefits and costs are equal. The net present value is calculated by setting the discount rate at some required rate (OECD, 2010, p. 137). Both the internal rate of return and the net present values are thus measures of the profitability of an investment, which then can be easily compared to other investments. The data provided by the OECD can be used to compute what will be called the effective rate of subsidization of education in chapter 4. It measures the extent to which the private rate of return exceeds the social rate of return. Table 1.3 reports the results obtained from the most recent 2010 issue of 'Education at a Glance'. The evidence whether the private rate of return exceeds the social rate of return is mixed. In a slight majority of countries education is effectively subsidized. Interesting are the cases of Finland, Hungary, Italy, Korea, New Zealand, and the United Kingdom where males and females are treated differently. Belgium and Germany are clear cases in which the social rate of return is larger than the private rate of return, which potentially indicates that tertiary education may not be sufficiently subsidized.

1.1.3 *Micro Labor Literature*

Another measure of economic well being is the wage rate earned by an individual. Beginning with the seminal study by Mincer (1974) a vast literature has emerged that provides estimates of the monetary return to education. He found that the returns to schooling and experience were around 10% and 8%, respectively. See table 1.4 in appendix 1.A for the details. Willis (1986) and Psacharopoulos (1994) provide more examples. Ammermüller and Weber (2005) report results for Germany between 1985 and 2002 and find that the returns to education are in the range from 8% to 10% in West Germany and between 7% and 8% in East Germany.

To sum up, the positive impact of human capital on economic well being is a well-established fact although estimates of its magnitude differ. A deeper understanding of how the accumulation of human capital is affected is therefore necessary.

1.2 Human Capital and Fiscal Policies

Easily one can think of many different things that matter for human capital: Families, nutrition, health, the school and university system, taxes and subsidies, and many other phenomena and institutions shape the environment in which the accumulation of human capital takes place. Exploring each single issue is undoubtedly an interesting task. What I would like to focus on in the following is the role of fiscal policies and its impact on the accumulation of human capital. Schultz (1961a) has a rather pessimistic view:

> Our tax laws everywhere discriminate against human capital. Although the stock of such capital has become large and even though it is obvious that human capital, like other forms of reproducible capital, depreciates, becomes obsolete, and entails maintenance, our tax laws are all but blind on these matters.

Whether Schultz is right or wrong has been the subject of numerous theoretical and empirical studies. Myles (2007) provides a comprehensive review of both strands of the literature.[2] A priori it is not clear how fiscal policies affect the accumulation of human capital. On the one hand, individual decisions are affected by taxes. Returns to certain economic decisions such as saving, education and labor are reduced, which may be detrimental to individual prosperity and growth. But on the other hand, tax revenue may be spent on institutions, e.g. school quality, that partly offset the negative individual incentives effect and furthermore provide for a conducive and stimulating environment.

The studies by Lucas (1990), King and Rebelo (1990), Jones, Manuelli, and Rossi (1993) and Pecorino (1993) do not take into account how tax revenue is spent and only analyze the effects a tax reform has on growth, consumption and the size of the stock of physical capital. These simulation studies demonstrate a clear negative relationship between rising taxes on labor and capital income and growth. See table 1.5 in appendix 1.A for the details. For instance, the model by Lucas (1990) is most closely related to the models studied in the present dissertation. Lucas predicts that the tax rates on capital and labor income change from 36% and

2 Follow-ups are Myles (2009a, 2009b,2009c).

40% to 0% and 46%, respectively. The growth rate slightly decreases from 1.50% to 1.47%. But even more important is that these changes bring about increases of the capital stock and consumption of about 33% and 6%, indicating the large positive welfare effects of such a tax reform.

The model by Trostel (1993) is closely related to the present dissertation as it explicitly models the accumulation of human capital. Time spent on education together with goods are the investments into human capital. More precisely, the accumulation of human capital is described by a Ben-Porath (1967) type production function. For the baseline calibration the result is that a one percent increase in the labor income tax rate causes the stock of human capital to decrease by about 0.39%. Thus, theory clearly shows the negative impact of taxation on the accumulation of human capital. Heckman, Lochner, and Taber (1998) further pursue Trostel's idea and study a model similar to that by Auerbach and Kotlikoff (1987). Two policy experiments are conducted. First, the experiment is a move from a tax system with a progressive tax on labor income and a 15-percent flat tax on capital income towards a system in which the tax rate on capital income remains at the initial level and a flat tax on labor income is set equal to 7.7% such that it balances the budget. In the second experiment only consumption is taxed at a flat rate of 10%. The tax reforms entail significant increases in the stock of human capital. For instance, the stock of college human capital increases at 2.82% and 1.85% in the flat tax and flat consumption tax experiment, respectively, compared to the benchmark case. One can see that the flat tax is more favorable to human capital accumulation than the flat consumption tax, which is more pro-capital.

Table 1.6 in 1.A provides further details. Other contributions that empirically study the effect of progressive taxation on schooling and on-the-job training include Dupor, Lochner, Taber, and Wittekind (1996), Heckman, Lochner, and Taber (1999) and Taber (2002).

Work done by Blankenau and Simpson (2004) and Ciriani (2007) do take into account how the government spends the tax revenue. The evidence provided is mixed. One has to closely bear in mind on which activities resources are spent. The government may directly provide education or it may subsidize privately provided education. Distortionary taxation may even nullify the positive effect of financing education.

To conclude this section, theoretical and empirical work indicates that the way how the returns to saving, education and labor are taxed and subsidized has effects on the accumulation of human capital in the short and long run.

1.3 STARTING POINT OF THIS DISSERTATION

The discussion so far has shown that (i) human capital is an important ingredient of economic well-being and (ii) taxation may have adverse effects on the accumulation of human capital. The objective of this dissertation is to provide a normative analysis of these relationships and attempts to provide answers to the question of how the tax system should be optimally designed to promote the accumulation of human capital to maximize economic well-being. The following chapters present several models of optimal taxation that approach this question from different perspectives. The basic economic problem in each model is the same: There is a single individual, endowed with perfect foresight, that lives for a given number of periods. In each period of time it has to make a consumption-leisure decision. It may choose to consume today or tomorrow and thus to save in form of physical capital, which is used as a means of production. Total time available may be spent on working, education and leisure. Working today increases income from labor today whereas time spent on education is the means to accumulate human capital which in turn increases productivity tomorrow and thus labor earnings. Besides foregone earnings, education involves some direct cost, which the government may choose to subsidize. Hence, saving and education are two possibilities to smooth consumption over time. Because the government levies linear taxes on the returns to capital and labor to finance its expenditures, the individual's decisions how much to save and how much time to spend on working and education are distorted. How theses taxes and the subsidy should be set to meet the government's revenue needs is the chief concern of this dissertation. The results depend on the time horizon of the individual, the way human capital is accumulated, and how the individual internalizes the effects of his own education decisions. Particular attention will be paid to how the optimal tax system affects the education decision. The benchmark case is that the individual devotes time to education until the point where the social marginal benefit to education equals the social marginal cost of education. When this point is achieved, education efficiency is said to prevail.

All models are set up in the spirit of Ramsey (1927). The basic methodological approach is the same in all models.

In what follows, I will lay out the approach and draw on the expositions in Chari and Kehoe (1999), Ljungqvist and Sargent (2004, pp. 490) and Christiano (2010). First, the individual's and firm's problem is set up. How the individual and the firms behave given prices and a tax system is described by a set of first-order conditions, which are the best-response functions. A competitive equilibrium consists of a feasible allocation, a price system, and a government policy, such that given a price system and the government policy, the allocation solves the individual's and firm's problem, and the government policy satisfies the government's budget constraint given the allocation and the price system. The set $\mathcal{C}$ collects all competitive equilibria resulting from different government policies. The government then aims at designing a linear tax system that maximizes the individual's utility while taking the feasibility constraint and the individual's and firm's competitive equilibrium behavior into account. Put differently, as each government policy gives rise to a different competitive equilibrium, the Ramsey problem is to choose the one that yields the highest utility.

There are two major approaches how to incorporate the individual's competitive equilibrium behavior: The primal and the dual approach. I first would like to shortly describe the dual approach before I turn to the primal approach, which will be used later on. The dual approach takes all equations describing the competitive equilibrium into account and solves the maximization problem by choosing the allocation and prices. This approach may be intuitively more appealing. But the major disadvantage is that the optimization problem involves solving for a large number of variables, which may be a cumbersome task. Instead of following this direct way, the primal approach asks to choose the optimal allocation that is consistent with competitive equilibrium behavior and then to solve for the government policy and prices that support this optimal allocation as a competitive equilibrium. The key to solving this problem is to use the so-called implementability constraint that summarizes the individual's competitive equilibrium behavior.

The implementability constraint is the individual's budget constraint after having substituted out all after-tax prices using the sufficient and necessary individual's first-order conditions. It is also possible not to use all first-order conditions and then to add these conditions separately. This then mixes the dual and the primal approach. The other constraint is the economy's resource constraint. Let the set $\mathcal{R}$ consist of all allocations that satisfy the aforementioned constraints. The Ramsey problem is to choose an allocation, which will be called Ramsey allocation, from the set $\mathcal{R}$ that yields the highest utility.

The key result is that one can find a price system and a government policy that implement the Ramsey allocation as a competitive equilibrium, or put shorter, that the two sets $\mathcal{C}$ and $\mathcal{R}$ are equal. First, one has to show that any allocation that is in the set $\mathcal{C}$ is also in the set $\mathcal{R}$. This is fairly intuitive because the implementability constraint is directly derived from the individual's competitive equilibrium conditions. Because the individual's and government's budget constraints are satisfied the resource constraint is satisfied by Walras's law. This proves the first inclusion.

Then, second, one needs to show that any allocation satisfying the implementability and resource constraint satisfies competitive equilibrium behavior. This means one has to find prices and a government policy, namely the tax rates, such that the allocation that is in $\mathcal{R}$ is also in $\mathcal{C}$. To do this, use the individual's and firm's first-order conditions, evaluate them at the Ramsey allocation, and determine prices and tax rates such that the allocation is in line with the individual's and firm's competitive equilibrium behavior as described by the first-order conditions. By construction, the Ramsey allocation satisfies the individual's budget constraint and the economy's resource constraint. By Walras' law, the government's budget constraint is satisfied as well.

The preceding discussion also helps to clarify the notion of the implementability constraint. The aim is to look for prices and tax rates that *implement* the Ramsey allocation as a competitive equilibrium. Any Ramsey allocation, which by construction satisfies the implementability constraint, can be implemented as a competitive equilibrium via an appropriate choice of tax rates.

To sum up, Ljungqvist and Sargent (2004, p. 491) provide the following recipe:

1. Derive the first-order conditions of the individual's and firm's problem. These are the best-response functions given prices and tax-rates. Solve for these prices and tax-rates as functions of the allocation.

2. Substitute out any prices and tax rates in the individual's intertemporal budget constraint. This leaves one with a constraint only taking the allocation as its argument.

3. Solve the Ramsey problem by maximizing the individual's utility subject to the implementability and resource constraint.

4. Use the Ramsey allocation to find the tax rates and prices.

Given the optimal tax rates and prices, the Ramsey allocation is feasible and simultaneously solves the individual's and firm's problem. As a result, one can state the following proposition, which fully characterizes the methodological foundation of this dissertation:

Proposition 1.1 (taken from Chari and Kehoe (1999), Proposition 1)**.** *The allocations in a competitive equilibrium satisfy the resource constraints and the implementability constraint. Furthermore, given allocations which satisfy these constraints, we can construct policies and prices which, together with the given allocations, constitute a competitive equilibrium.*

To get a deeper understanding of this rather abstract description of the primal approach, appendix 1.B provides a simple and fully fleshed out example taken from Christiano (2010).

1.4 RELATED LITERATURE AND SHORTCOMINGS

The models presented in this dissertation are simple. Many important issues have not been considered, which I would like to comment on now.

It will be assumed that the government commits to a policy chosen at the outset of time. Reoptimization during the course of time is ruled out. The well-known time inconsistency (Kydland and Prescott (1977), Fischer (1980)) problems are assumed away.[3] Boadway, Marceau, and Marchand (1996) and Andersson and Konrad (2003) address these issues.

Two forms of market failure are not discussed: (i) Borrowing constraints and (ii) incomplete insurance markets. The presence of market failure is generally taken as an argument favoring government intervention if the government has the means at its disposal to yield a better outcome than the market solution.

If the individual when young is borrowing constrained, he is not able to finance consumption and the direct cost of education by borrowing against his future income when old. Credit markets for young individuals are thus imperfect. The assumption that individuals are borrowing constrained has been reassessed in several models by Kane (1994), Card (1999) and Ellwood and Kane (2000) who argue that low-income earners are indeed borrowing constrained which then explains their lower participation rates in tertiary education. But the more important question actually is in which way individuals are constrained. Certainly, in the short run individuals may be borrowing constrained which is why the government must intervene. Cameron and Heckman (2001) and Carneiro and Heckman (2002) argue that long-term constraints serve better to explain the relative absence of children of low-income earners in tertiary education.

Incomplete insurance markets are an issue if the individual is subject to idiosyncratic shocks that are uninsured. Instead it is assumed that the individual is endowed with perfect foresight. This means there is no uncertainty about the state of the world tomorrow. The individual knows

3 See also Golosov and Tsyvinski (2008) for a short discussion.

today what will be the output of the production function tomorrow given his savings and time devoted to working. He also knows what will be the size of the stock of human capital tomorrow given the time spent on education today. The chance of failing or not being able to fully use the stock of human capital is absent. Put shorter, human capital is not risky. Contributions tackling this issue include Wigger and von Weizsäcker (2001), Krebs (2003), da Costa and Maestri (2007), Anderberg (2009) and Jacobs, Schindler, and Yank (2010).

There are two approaches to study questions of optimal taxation, which have been pioneered by Ramsey (1927) and Mirrlees (1971). Both approaches derive different results regarding the optimal taxation of human capital. The Mirrlees approach[4] rests on the assumption that the government is not able to observe the individuals' skills. Hence, it has to take an informational asymmetry into account and the aim then is to set up a tax system that is incentive-compatible such that the high-skill individual does not mimic the low-skill individual. Contributions by Bovenberg and Jacobs (2005), Jacobs (2005), Jacobs and Bovenberg (2010a, 2010b) and Grochulski and Piskorski (2010) put emphasis on endogenous human capital accumulation. A positive tax wedge on saving is used to set correct incentives.[5] Ramsey models have been studied by Lucas (1990), Yuen (1991), Bull (1993), Jones, Manuelli, and Rossi (1997), Chari and Kehoe (1999), Erosa and Gervais (2002), Barbie

4 Useful surveys include Mirrlees (1986) and Golosov, Tszvinski, and Werning (2007). The latter survey also contains a discussion of the pros and cons of the Ramsey and Mirrlees approach on which Diamond (2007) and Judd (2007) comment.

5 A wedge is the difference between an individual marginal rates of substitution and marginal rates of transformation. These wedges determine optimal taxes. See Kocherlakota (2005) for a discussion.

and Hermeling (2006) and Richter (2009). Both strands of the literature provide results on education efficiency. But results differ with respect to the role of capital income taxation. Chamley (1986) and Judd (1985) show that within Ramsey models there should be no wedge between the intertemporal rate of substitution and the marginal rate of transformation, or, alternatively, that the capital income tax rate is zero in the long run.[6] The converse is true in Mirrlees models if skills evolve stochastically over time.[7] The difference is that within Ramsey models the prime objective is achieving allocative efficiency, that is, not to distort the intertemporal allocation of consumption. This objective is pursued independent of whether the models feature human capital or not (see Jones, Manuelli, and Rossi (1997)). Within Mirrlees models, capital income is taxed to discourage savings and therefore to make invests in human capital more attractive. This result is driven by the informational asymmetry between the government and the individuals.

1.5 SUMMARY OF THE DISSERTATION

Chapter 2 discusses contributions by Bovenberg and Jacobs (2005) and Richter (2009), who set up two-period models of the Mirrlees and of the Ramsey type and derive the so-called *education efficiency theorem*: In a second-best optimum, the education decision is undistorted. The before- and after-tax rates of return to education are equal. This

6 When one imposes certain restrictions on the tax rates, capital income may be taxed as Chari and Kehoe (1999) show.

7 Otherwise, if the distribution of skills over time is constant, the intertemporal margin is not distorted as Golosov, Kocherlakota, and Tsyvinski (2003) show.

result crucially depends on the way the accumulation of human capital is modeled, which is as follows. In the first period, the individual spends time on education, which enters a human capital production function as the only input. The output increases the stock of human capital in the second period. Bovenberg and Jacobs (2005) and Richter (2009) assume that the stock of human capital in the second period equals only the output of the human capital production function, which is assumed to be isoelastic with respect to education. This means that the function expressing the stock of human capital in the second period and the human capital production function are the same. By contrast, I assume that in the first period the individual is endowed with an initial stock of human capital. The chapter then studies the effect that the initial stock of human capital has on optimal taxation. The stock of human capital in the second period is assumed to be the sum of the output of the isoelastic human capital production function, which takes education as the only input, and the initial stock of human capital net of depreciation. This implies that the elasticity of the function expressing the stock of human capital in the second period is increasing. Then the education efficiency theorem no longer holds. In a second-best optimum, the discounted marginal social return to education is smaller than the marginal social cost. The individual overinvests in human capital relative to the first best. As a result, the government effectively subsidizes the return to education.

Chapter 3 tackles the following question asked by Jones, Manuelli, and Rossi (1997): "Is physical capital special?" Using the Ramsey approach, they add human capital to an optimal taxation model with physical capital. By modeling human capital almost symmetrically to physical capital they show that in a stationary state all taxes are zero. What drives this zero-tax result is that the human capital production function features constant returns to scale with respect to the stock of human capital. But they acknowledge that if the human capital production function violates this assumption, the stationary-state labor tax will not be zero. This chapter takes up the issue of modeling human capital almost symmetrically to physical capital. I drop the constant-returns-to-scale assumption. The human capital production function does not include the current stock of human capital and only includes the individual's time devoted to education. The individual has to pay for verifiable direct costs, e.g., tuition fees, that depend on the amount of education. The government may choose to subsidize this cost. I derive two results: First, optimal taxation in the stationary state prescribes not taxing capital income. The zero-capital-tax result holds despite the presence of human capital. The education decision depends only on how the labor tax and the education subsidy interact with each other. This relates to the second result, stating that in the optimum the marginal social return to education is larger than the marginal social cost. The so-called *Education Efficiency Theorem* (Richter, 2009) does not hold. From the inequality between the marginal social return and the marginal social cost it follows that education is effectively taxed. Turning to the underlying tax rates, it results that the cost of education is not fully tax-deductible, the labor income tax rate is higher than the rate of subsidization. As a consequence, the individual underinvests in human capital relative to the first best.

Chapter 4, which is joint work with Wolfram F. Richter, studies a model with overlapping generations and endogenous growth. Individuals live for two periods. They decide on education, saving, and nonqualified labor in their youth. They supply qualified labor when old. The productivity of qualified labour increases in the stock of human capital inherited from preceding generations, and it also increases in own educational investments. Individuals either may be perfect altruists with respect to descendent generations or may behave selfishly. Assuming selfish individuals yields the result that it is second best not to distort education if the human capital investment function is isoelastic in education. If, however, the elasticity of the investment function is increasing, which happens when the human capital stock does not depreciate completely if just one generation fails to invest, it is second best at balanced growth to subsidize education even relative to the first best. Assuming altruistic individuals changes some conclusions, but not all. Altruists internalize the positive effect that education has on descendents' productivity. For all generations except the first one the accumulation of human capital should not be distorted, and this result is obtained for arbitrary utility and human capital investment functions. The accumulation of physical capital should not be distorted if the taxpayer's utility is weakly separable between consumption and non-leisure and homothetic in consumption. Furthermore, qualified and nonqualified labour should be taxed uniformly across the life cycle when utility exhibits balanced growth path properties.

Chapter 5 provides a numerical analysis of an important result derived in chapter 4. Two effects have been identified that serve to justify the subsidization of education if individuals are selfish. First, the most well-known justification is to internalize intergenerational external effects of education. Second, distortionary labor taxation has a negative effect on education and thereby on growth. For this reason, education should be subsidized relative to the first best if the elasticity of the human capital investment function is increasing. The chapter numerically analyzes the impact an increasing elasticity has on optimal taxation and studies the importance of the intergenerational external effect and of the distortionary labor taxation effect as reasons to justify the subsidization of education. As it turns out, the case for subsidizing education to account for distortionary labor taxation is rather weak. The still dominant justification for subsidizing education is to internalize intergenerational externalities.

1.A TABLES

Table 1.1: The contribution of education to economic growth (percentage), 1950-1962

Country	Growth rate		Growth rate explained by education
	National income	Education	
United States	3.36	0.49	15
Belgium	3.03	0.43	14
Denmark	3.36	0.14	4
France	4.70	0.29	6
Germany	7.26	0.11	2
Italy	5.95	0.40	7
Netherlands	4.52	0.24	5
Norway	3.47	0.24	7
United Kingdom	2.38	0.29	12

Source: Denison (1967, tables 21-1–21-20)

Table 1.2: Cross-country growth regressions

Study	Dependent Variable	Human Capital Proxy	Flow/Stock	Estimated Coefficient	Illustration of Impact
Barro (1991)	growth rate of real per capita GDP annual 1960–85	school enrolment rate: number of students enrolled in the designated grade levels (primary and secondary respectively) relative to the total population of the corresponding age group in 1960	initial flow mean: prim60: 0.78 sec60: 0.23	prim = 0.025 sec = 0.030	A 1 percentage point increase in primary (secondary) school enrolment rates is associated with a 2.5 (3.0) percentage points increase in per capita GDP growth rate.
Levine and Renelt (1992)	growth rate of real per capita GDP annual 1960–89	secondary school enrolment rate in 1960	initial flow	high = 3.71 base = 3.17 low = 2.5	A 1 percentage point increase in secondary school enrolment rate is associated with a between 2.5 and 3.7 percentage points increase in per capita GDP growth rate.
Murphy, Schleifer and Vishny (1991)	growth rate of real per capita GDP between 1970–85	primary school enrolment rate in 1960	initial flow	full sample: 0.022 (OECD: not significant)	A 1 percentage point increase in primary school enrolment rate is associated with a 2.2 percentage points increase in per capita GDP growth rate.
Barro (1997)	growth rate of real per capita GDP over 1965–75, 1975–85, 1985–90	average years of attainment for males aged 25 and over in secondary and higher schools at the start of each period	initial stocks in 1965, 75 and 85 mean in 1990 = 1.9 years	0.012	An extra year of male upper-level schooling is associated with a 1.2 percentage point increase in per capita GDP growth rate.

(continued)

Study	Dependent Variable	Human Capital Proxy	Flow/Stock	Estimated Coefficient	Illustration of Impact
Hanushek and Kim (1995)	growth rate of real per capita GDP between 60–90	average years of secondary schooling of adult male population at beginning of period	initial stock	0.36	An extra year of male secondary schooling is associated with a 0.36 percentage point increase in per capita GDP growth rate.
Gemmel (1996)	growth rate of real per capita GDP annual 60–85	constructed human capital stock in 1960 and human capital annual average growth rates at primary, secondary and tertiary levels. These measures are both entered in the equation simultaneously.	initial stock mean: prim = 72.8 sec = 19.5 tert = 4.0 annual flows mean: prim = 2.5 sec = 3.7 tert = 2.7	Full sample: prim stock = 0.81 prim flow = 2.68 Poorest LDCs: prim stock = 0.91 prim flow = 4.19 Intermediate LDCs: sec stock = 1.09 OECD: tert stock = 1.10 tert flow = 5.89	For OECD: A 1 percent increase in tertiary human capital stock is associated with a 1.1 percentage point increase in per capita GDP growth rate. A 1 percentage point increase in tertiary human capital growth is associated with a 5.9 percentage points increase in per capita GDP growth rate.
Judson (1998)	growth rate of real GDP 5-years averages, 1960–90	growth of her constructed measure of human capital stock	period flows	10.8 low-efficiency countries = 3.0 high-efficiency = 12.9	A 1 percentage point increase in human capital growth is associated with an 11 percentage points increase in GDP growth rate.

(continued)

Englander and Gurney (1994)	growth of labour productivity (and total factor productivity) over four time periods	school enrolment rates: number of students enrolled in secondary school relative to the total population of the corresponding age group in beginning of period	initial flow	OECD: 1.45–1.78	A 1 percentage point increase in secondary school enrolment rate is associated with around 1.5 percentage point increase in productivity growth.
Barro and Lee (1994)	ΔlnGDP per worker	average years of secondary schooling of adult male population at beginning of period	initial stock	0.014	An extra year of male secondary schooling is associated with a 1.4 percent increase in per worker GDP growth.
Benhabib and Spiegel (1994)	ΔlnGDP per capita	human capital stock estimates from Kyriacou: average level of log human capital over the period (log of average level of human capital; log of average levels)	average stock	0.12–0.17	A 1 percent increase in the stock of human capital is associated with a 12 to 17 percent increase in per capita GDP growth.
Mankiw, Romer and Weil (1992)	lnGDP per working-age person	average percentage of working-age population in secondary school, 1960–85	period flow	0.66	A 1 percent increase in the average percentage of working-age population in secondary school is associated with a 0.66 percent increase in GDP per working-age person.
				implied output elasticity with respect to human capital stock = 0.28	A 1 percent increase in human capital stock is associated with a 0.28 percent increase in GDP

(continued)

Study	Dependent Variable	Human Capital Proxy	Flow/Stock	Estimated Coefficient	Illustration of Impact
de la Fuente and Doménech (2000)	GDP per worker Annual, 1960–90 OECD	average number of years of schooling of the adult population	stock mean '90 = 10.49 (UK'90 = 10.94)	implied output elasticity with respect to human capital stock = 0.27	A 1 percent increase in human capital stock is associated with a 0.27 percent increase in GDP. At the sample mean, an increase in average education by one year would raise output per capita by ca. 3 percent.
Bassanini and Scarpetta (2001)	GDP per working-age person Annual, 1971–98 OECD	average number of years of schooling of the adult population	stock mean = 10.15	implied output elasticity with respect to human capital stock = 0.57	A 1 percent increase in human capital stock is associated with a 0.57 percent increase in GDP. At the sample mean, an increase in average education by one year would raise output per capita by ca. 6 percent.

Source: Sianesi and Reenen (2003, p. 177-180, table 2)

Table 1.3: Effective rate of subsidization for an individual obtaining tertiary education as part of initial education, ISCED 5/6 (2006), in OECD countries

Country	Male	Female
Australia	$-0,32$	$-0,05$
Austria	$0,07$	$0,04$
Belgium	$-0,32$	$-0,23$
Canada	$0,14$	$0,17$
Czech Republic	$0,23$	$0,19$
Denmark	$-1,46$	$-0,22$
Finland	$-0,03$	$0,07$
France	na	na
Germany	$-0,34$	$-0,32$
Hungary	$0,01$	$-0,31$
Italy	$0,17$	$-0,26$
Korea	$-0,03$	$0,07$
Netherlands	$-0,16$	$-0,15$
New Zealand	$-0,14$	$0,10$
Norway	$0,11$	$0,34$
Poland	$0,20$	$0,22$
Portugal	$0,18$	$0,01$
Spain	$0,26$	$0,31$
Sweden	$0,07$	$0,62$
Turkey	$0,41$	$0,44$
United Kingdom	$0,05$	$-0,11$
United States	na	na
OECD average	$0,06$	$0,07$

Source: OECD (2010, chapter A8), own calculations

The OECD provides the following pieces of information that are used to compute the effective rate of subsidization:

- Private net present value, PrivNPV
- Public net present value, PubNPV
- Private direct costs, PrivDC

- Private foregone earnings, PrivFE
- Public direct costs, PubDC
- Public foregone earnings, PubFE

The private rate of return is defined as follows:

$$\text{PRR} = \frac{\text{PrivNPV}}{\text{PrivDC} + \text{PrivFE}}$$

The social rate of return is defined as follows:

$$\text{SRR} = \frac{\text{PrivNVP} + \text{PubNPV}}{\text{PrivDC} + \text{PrivFE} + \text{PubDC} + \text{PubFE}}$$

The effective rate of subsidization s_{eff} then is:

$$s_{\text{eff}} = \frac{\text{PRR} - \text{SRR}}{\text{PRR}}$$

Table 1.4: Regressions of individual earnings on schooling s and experience x (1959 annual earnings of white, nonfarm men)

Equation forms	R^2
1. $\ln y = 7.58 + 0.070s$	0.067
(43.8)	
2. $\ln y = 6.20 + 0.107s + 0.081x - 0.0012x^2$	0.285
(72.3) (75.5) (-55.8)	

Source: Mincer (1974, p. 92, table 5.1)

Table 1.5: Growth effects of tax reform

Author	Features	Utility Parameters	Initial Tax Rates and Growth Rate	Final Position	Additional Observations
Lucas (1990)	Production of human capital did not require physical capital	$\sigma = 2$ $\alpha = 0.5$	Capital 36% Labor 40% Growth 1.50%	Capital 0% Labor 46% Growth 1.47%	33% increase in capital stock 6% increase in consumption
King and Rebelo (1990)	Production of human capital requires physical capital (proportion = 1/3)	$\sigma = 2$ $\alpha = 0$	Capital 20% Labor 20% Growth 1.02%	Capital 30% Labor 20% Growth 0.50%	Labor supply is inelastic
Jones, Manuelli and Rossi (1993)	Time and physical capital produce human capital	$\sigma = 2$ $\alpha = 4.99$ α calibrated given σ	Capital 21% Labor 31% Growth 2.00%	Capital 0% Labor 0% Growth 4.00%	10% increase in capital stock 29% increase in consumption
Pecorino (1993)	Production of human capital requires physical capital	$\sigma = 2$ $\alpha = 0.5$	Capital 42% Labor 20% Growth 1.51%	Capital 0% Labor 0% Growth 2.74%	Capital and consumption different goods, consumption tax replaces income taxes

Source: Myles (2009a, p. 28, table 2)

Table 1.6: Comparison of steady states under alternative tax regimes

	Percentage difference from benchmark progressive case	
	Flat tax	Flat consumption tax
Stock of physical capital	−0.79	19.55
Stock of college human capital	2.82	1.85
Stock of highschool human capital	0.90	0.08
Aggregate output	1.15	4.98

Source: Heckman, Lochner, and Taber (1998, table 1)

The individual's problem is to maximize utility $u(c, l)$ over the choice of consumption, c, and labor supply, l, subject to the budget constraint $c \leqslant w(1-\tau)l$ where w is the wage rate and τ is the labor tax rate. The first-order condition

$$u_c w(1-\tau) + u_l = 0$$

along with the budget constraint define the optimal choices of c and l as functions of the tax rate τ, that is, $\tilde{c} = c(\tau)$ and $\tilde{l} = l(\tau)$.

The Ramsey problem is to find a government policy τ that maximizes the individual's utility $u(c(\tau), l(\tau))$ subject to the budget constraint $g \leqslant wl(\tau)\tau$. The Ramsey equilibrium consists of an optimal government policy τ^* and an optimal allocation $\{c^*, l^*\}$ such that the allocation is an optimal choice of the individual given τ^*, that is, $c^* = c(\tau^*)$ and $l^* = l(\tau^*)$, and τ^* solves the Ramsey problem.

Given a government policy τ, let $\mathcal{C}$ denote the set of allocations that constitute a competitive equilibrium. Put formally:

$$\begin{aligned}\mathcal{C} = \big\{(c, l) : \exists\tau \text{ s.t. } u_c w(1-\tau) + u_l = 0,& \\ c = w(1-\tau)l, g \leqslant wl(\tau)\tau\big\}&\end{aligned}$$

To employ the Ramsey approach, one has to derive the implementability constraint. Replace $(1-\tau)w$ in the individual's budget constraint using the first-order condition

$$(1-\tau)w = -\frac{u_l}{u_c}. \tag{1.1}$$

Then the implementability constraint reads

$$u_c c + u_l l = 0. \tag{1.2}$$

The set $\mathcal{R}$ consists of all allocations that satisfy the economy's resource constraint, $c + g \leqslant wl$, and the implementability constraint:

$$\mathcal{R} := \{(c, l) : c + g \leqslant wl, u_c c + u_l l = 0\}$$

Proposition 1.2. $\mathcal{C} = \mathcal{R}$

Proof. First, prove that $\mathcal{C} \subseteqq \mathcal{R}$. $(c, l) \in \mathcal{C}$ satisfies the implementability constraint (1.2) by construction. If one combines the individual's and government's budget constraint, one derives the resource constraint.

Second, show that $\mathcal{R} \subseteqq \mathcal{C}$. Choose τ such that (1.1) is satisfied. Then multiply (1.1) by l and use (1.2):

$$(1-\tau)wl = -\frac{u_l l}{u_c} = c$$

The individual's budget constraint is satisfied. If one combines the individual's budget constraint and the resource constraint, one ends up with the government's budget constraint which is also a direct consequence of Walras's law. $\square$

The proposition shows that it is possible to find a tax rate τ that implements the optimal allocation (c, l) as a competitive equilibrium.

2

OPTIMAL TAXATION OF EDUCATION WITH AN INITIAL ENDOWMENT OF HUMAN CAPITAL

2.1 INTRODUCTION

Bovenberg and Jacobs (2005) and Richter (2009) set up two-period models of the Mirrlees and of the Ramsey type and derive the so-called *education efficiency theorem*: In a second-best optimum, the education decision is undistorted. The before- and after-tax rates of return to education are equal. This result crucially depends on the way the accumulation of human capital is modeled, which is as follows. In the first period, the individual spends time on education, which enters a human capital production function as the only input. The output increases the stock of human capital in the second period. Bovenberg and Jacobs (2005) and Richter (2009) assume that the stock of human capital in the second period equals only the output of the human capital production function, which is assumed to be isoelastic with respect to education.[1] This means that the function expressing the stock of human capital in the second period and the human capital production function are the same. A debatable im-

1 Jacobs and Bovenberg (2010b) further analyze the human capital production function's properties and find that a constant elasticity is crucial for their results in Bovenberg and Jacobs (2005). Richter (2009) refers to the so-called power law of learning, a result from cognitive psychology that provides evidence in favor of a constant elasticity. See Ritter and Schooler (2001) for more details.

plication of modeling the law of motion for human capital this way is that the stock of human capital in the second period is zero if the individual does not spend any time on education in the first period. Then effective labor supply is zero, because it is modeled as the product of raw labor supply and the then existing stock of human capital. Consequently, the individual does not earn any labor income. Put more briefly, if the individual wants to reap benefits of human capital, it first has to spend time on education. Or, as an alternative interpretation of this implication, consider a two-period overlapping-generations (OLG) model. When young, the individual accumulates human capital, which it uses when old. When old, it passes on the then existing stock of human capital to the newly born young individual so that it can further increase the stock by spending time on education. But when a young individual stops devoting time to education, the stock of human capital in the second period is zero. Consequently, the old individual could not pass on human capital to the newly born young individual.

By contrast, I assume that in the first period the individual is endowed with an initial stock of human capital. The present paper then studies the effect that the initial stock of human capital has on optimal taxation. The stock of human capital in the second period is assumed to be the sum of the output of the isoelastic human capital production function, which takes education as the only input, and the initial stock of human capital net of depreciation. This implies that the elasticity of the function expressing the stock of human capital in the second period is increasing. With this specification, the stock of human capital in the second period does not drop to zero even when the individual stops spending time on education in the first period, or, referring back to the OLG interpretation, the old individual can then pass on human capital even when it may not have spent time on education when young. Then the education efficiency theorem no longer holds. In a second-best optimum, the discounted marginal social return to education is smaller than the marginal social cost. The individual overinvests in human capital relative to the first best. As a result, the government effectively subsidizes the return to education.

The general-equilibrium model used here comprises a single individual, a firm, and a government. The individual lives for two periods, in which it faces a consumption-labor-leisure choice. In the first period, it chooses how much time to devote to work and education. In the second period, it only decides how much to work. Time spent on education is transformed into human capital by means of a human capital production function. The individual combines its raw labor supply with the then existing stock of human capital, giving the effective labor supply. It chooses to lend capital to a firm, which takes it as an input, jointly with the effective labor supply, and pays a return. The firm produces a single consumption good. Time spent on education brings about disutility and comes at the cost of forgone earnings and some direct costs such as tuition fees. All actions of the individual are assumed to be fully observable. The government levies linear taxes on the individual's income from work and saving to finance an exogenously given stream of expenditures. Furthermore, it may choose to subsidize the direct cost of education. The question then is how to optimally choose linear taxes and the subsidy to maximize the individual's utility given exogenous government expenditures and subject to the individual's competitive equilibrium behavior.

2.2 THE MODEL

2.2.1 *Individual's Problem*

The individual solves the following maximization problem:

$$\begin{aligned}\mathcal{L} &= U(C_0, L_0 + E) + \beta U(C_1, L_1) \\ &+ \lambda_0 \Big(\omega_0 L_0 H_0 + R_0^\tau K_0 - C_0 - K_1 - \varphi E \Big) \\ &+ \beta\lambda_1 \Big(\omega_1 L_1 H_1 + R_1^\tau K_1 - C_1 \Big) \\ &\qquad + \mu \Big(G(E) + (1 - \delta_H) H_0 - H_1 \Big). \end{aligned} \quad (2.1)$$

The individual's utility function is strictly increasing in consumption, C_t, and strictly decreasing in the nonleisure times $L_0 + E$ and L_1. It is strictly concave in both arguments and time-separable.

Savings serve as a means to smooth consumption over time. They pay the net rate of return $R_t^\tau \equiv (1 - \tau_t^K) r_t + 1 - \delta_K$, where τ_t^K is a linear tax on the gross rate of return r_t, and δ_K is the rate at which the stock of capital K_t depreciates. Raw labor supply L_t is combined with the stock of human capital H_t accumulated so far. Effective labor supply $L_t H_t$ earns the net wage rate $\omega_t \equiv (1 - \tau_t^L) w_t$, where τ_t^L is a linear tax on the gross wage rate w_t. Let $\varphi \equiv (1 - \tau^H) f$ be the direct cost of education net of the subsidy τ^H, where f is an exogenous (fee) parameter. The endowments of the initial stocks of human capital, H_0, and capital, K_0, are given. β is the private discount factor.

The law of motion

$$H_1(E) = G(E) + (1 - \delta_H)H_0 \tag{2.2}$$

governs the evolution of the stock of human capital, which depreciates at the rate $\delta_H \leqslant 1$. Here G is the human capital production function, which takes time E as its only input factor. It is isoelastic:

$$G(E) = aE^{\gamma} \tag{2.3}$$

with $0 < \gamma < 1$. The coefficient $a > 0$ is a shift parameter. (2.2) and (2.3) imply that the elasticity η of the function of the stock of human capital $H_1(E)$ is strictly increasing as long as the initial stock of human capital does not fully depreciate.[2] By setting $H_0 = 0$ or $\delta_H = 1$, one obtains the model underlying the analysis in Bovenberg and Jacobs (2005) or Richter (2009). Then the elasticity of $H_1(E)$ equals γ.

To have a well-behaved problem, it does not suffice to apply the Inada conditions. An analysis of the second-order conditions, which is done in Appendix 2.A, reveals that moreover one has to assume that the utility function is sufficiently concave to compensate for the lack of concavity of the law of motion (2.2) for human capital. Put formally, the requirement says that $\gamma < \upsilon_1/(1+\upsilon_1)$, where γ is the elasticity of the human capital production function (2.3), and $\upsilon_1 = L_1 U_{L_1 L_1}/U_{L_1}$, which captures the concavity of the utility function with respect to second-period labor supply.

2 *Proof.* $\frac{d}{dE}\eta \equiv \frac{d}{dE}\frac{EH_1'(E)}{H_1(E)} = \gamma G'(E)\frac{H_1(E)-G(E)}{H_1(E)^2} > 0$ for $\delta_H < 1$. $\square$

This condition will show up again when labor taxation is analyzed in section 2.2.7.3.

Let U_{C_t} and U_{L_t} denote the partial derivatives with respect to consumption and nonleisure time, taking the corresponding period t variables as arguments. Maximization over consumption, time spent on working, and investments in human and physical capital yields the following first-order conditions:

$$C_t : U_{C_t} = \lambda_t, \; t = 0, 1, \tag{2.4}$$

$$L_t : -U_{L_t} = \omega_t H_t \lambda_t, \; t = 0, 1, \tag{2.5}$$

$$E : -U_{L_0} + \lambda_0 \varphi = \mu G'(E), \tag{2.6}$$

$$H_1 : \lambda_1 \beta \omega_1 L_1 = \mu, \tag{2.7}$$

$$K_1 : \lambda_0 = \lambda_1 \beta R_1^\tau. \tag{2.8}$$

By eliminating μ and using all first-order conditions, the following optimality condition results:

$$\frac{\omega_1 L_1 G'(E)}{R_1^\tau} = \varphi + \omega_0 H_0. \tag{2.9}$$

The individual chooses education up to the point where the discounted marginal (private) return $\omega_1 L_1 G'(E)/R_1^\tau$ equals the marginal (private) cost $\varphi + \omega_0 H_0$, which is sum of the direct cost and the forgone earnings.

2.2.2 *The Government*

The government uses linear taxes to finance an exogenously given stream of government expenditures $\{g_t\}_{t=0}^1$. Its budget constraints are

$$g_0 + \tau^H fE = \tau_0^K r_0 K_0 + \tau_0^L w_0 L_0 H_0, \tag{2.10}$$

$$g_1 = \tau_1^K r_1 K_1 + \tau_1^L w_1 L_1 H_1. \tag{2.11}$$

2.2.3 *Firm's Problem*

The stock of physical capital K_t and the individual's effective labor supply, $Z_t \equiv L_t H_t$, enter the firm's constant-returns-to-scale production function $F(K_t, Z_t)$. Factors are paid their marginal products:

$$F_{K_t} \equiv \frac{\partial}{\partial K_t} F(K_t, Z_t) = r_t, \; t = 0, 1, \tag{2.12}$$

$$F_{Z_t} \equiv \frac{\partial}{\partial Z_t} F(K_t, Z_t) = w_t, \; t = 0, 1. \tag{2.13}$$

2.2.4 *Competitive Equilibrium*

A competitive equilibrium consists of a feasible allocation

$$\left\{ \{C_t, L_t, K_t, H_t\}_{t=0}^1, E \right\},$$

a price system

$$\{w_t, r_t\}_{t=0}^1,$$

a government policy

$$\left\{ \{g_t, \tau_t^K, \tau_t^L\}_{t=0}^1, \tau^H \right\},$$

an exogenously given direct cost of education f, and initial stocks of human and physical capital, H_0 and K_0, respectively. The feasible allocation and the price system solve the individual's and firm's problems. The government policy satisfies the budget constraints (2.10) and (2.11).

2.2.5 *First-Best Solution*

Studying the first-best problem serves to establish a benchmark case. The planner chooses consumption, investments in physical and human capital, and the allocation of time to solve the following maximization problem:

$$\begin{aligned}
\mathcal{L} = {} & U(C_0, L_0 + E) + \beta U(C_1, L_1) \\
& + \theta_0 \Big(F(K_0, L_0 H_0) + (1 - \delta_K) K_0 \\
& - C_0 - K_1 - fE - g_0 \Big) && (2.14) \\
& + \theta_1 \beta \Big(F(K_1, L_1 H_1) + (1 - \delta_K) K_1 - C_1 - g_1 \Big) && (2.15) \\
& + \mu \Big(G(E) + (1 - \delta_H) H_0 - H_1 \Big).
\end{aligned}$$

He maximizes the individual's discounted sum of utilities subject to the per-period resource constraints (2.14) and (2.15) and the law of motion for human capital.

The first-order conditions are

$$\begin{aligned}
C_t:\ & U_{C_t} = \theta_t,\ t = 0, 1, && (2.16) \\
L_t:\ & -U_{L_t} = \theta_t F_{Z_t} H_t,\ t = 0, 1, && (2.17) \\
E:\ & \mu G'(E) = -U_{L_0} + \theta_0 f, && (2.18) \\
K_1:\ & \theta_0 = \theta_1 \beta \big(F_{K_1} + 1 - \delta_K \big) \equiv \theta_1 \beta R_1^s, && (2.19) \\
H_1:\ & \theta_1 \beta F_{Z_1} L_1 = \mu. && (2.20)
\end{aligned}$$

Proposition 2.1.

The discounted marginal social return to education equals the marginal social cost:

$$\frac{F_{Z_1} L_1 G'(E)}{R_1^s} = f + F_{Z_0} H_0. \tag{2.21}$$

Proof.

Eliminate θ_0, μ, and U_{L_0} in the condition (2.18) using (2.17), (2.19), and (2.20). □

The social planner chooses education up to the point where the discounted marginal social return $F_{Z_1} L_1 G'(E)/R_1^s$ equals the marginal social cost $f + F_{Z_0} H_0$, which is sum of the direct cost and the loss in marginal productivity of first period's labor supply. Proposition 2.1 therefore suggests the following definition to gauge education efficiency.

Definition 2.1.

Education efficiency is achieved if the discounted marginal social return to education equals the marginal social cost, which is the direct cost of education plus the loss in marginal productivity of the first period's labor supply. In the first best, there is no wedge between the discounted marginal social return and the marginal social cost of education.

The efficiency condition (2.21) can be further used to assess under which circumstances a competitive equilibrium implies education efficiency in the sense of Definition 2.1. The wedge between the discounted marginal social return and the marginal social cost of education can be manipulated as follows:

$$\begin{aligned}\Delta &\equiv \frac{F_{Z_1} L_1 G'(E)}{R_1^s} - \left(f + F_{Z_0} H_0\right) \\ &= \frac{R_1^\tau\left(\varphi + \omega_0 H_0\right)}{R_1^s}\left[\frac{F_{Z_1}}{\omega_1} - \frac{R_1^s\left(f + F_{Z_0} H_0\right)}{R_1^\tau(\varphi + \omega_0 H_0)}\right].\end{aligned} \tag{2.22}$$

The last equality follows from the individual's optimality condition (2.9). Education efficiency holds if and only if the bracketed factor vanishes. Therefore,

$$\frac{F_{Z_1}}{R_1^s\left(f + F_{Z_0} H_0\right)} = \frac{\omega_1}{R_1^\tau\left(\varphi + \omega_0 H_0\right)}. \tag{2.23}$$

Put verbally, if and only if before- and after-tax rates of return are equal, education efficiency prevails in a competitive equilibrium. The wedge Δ is positive (negative) if and only if education is effectively taxed (subsidized). Richter (2009) uses the condition (2.23) to assess education efficiency.

2.2.6 *Second-Best Solution*

The Ramsey problem is to choose a government policy that maximizes the individual's utility subject to the individual's and the firm's competitive equilibrium behavior, given initial stocks of human and physical capital, H_0 and K_0, and direct cost of education f. The primal approach is adopted to study the Ramsey problem of optimal taxation (Atkinson and Stiglitz (1980), Chari and Kehoe (1999)). The difference to the dual approach is how it incorporates the individual's competitive equilibrium behavior. The individual's first-order conditions serve to eliminate all prices and taxes in the intertemporal budget constraint. As a result, this constraint then fully captures how the individual behaves in a competitive equilibrium. Given the allocation, the first-order conditions yield the prices and taxes that implement the second-best outcome as a competitive equilibrium. By contrast, the dual approach includes all constraints separately, which requires optimizing over the allocations and prices.

To derive the so-called implementability constraint, first the intertemporal budget results after combining the per-period budget constraints from the individual's problem (2.1) by eliminating K_1 and using (2.9) to eliminate direct cost φE:

$$R_0^\tau K_0 + \omega_0 H_0 (L_0 + E) + \frac{1}{R_1^\tau} \omega_1 L_1 H_1 \,(1 - \eta) = C_0 + \frac{1}{R_1^\tau} C_1 \quad (2.24)$$

with

$$\eta \equiv \eta(E, H_1) \equiv \frac{H_1'(E)E}{H_1(E)} = \frac{G'(E)E}{H_1(E)},$$

which is the nondecreasing[3] elasticity of the function expressing the stock of human capital in the second period. The LHS of (2.24) is the individual's income side. $R_0^\tau K_0$ is the value of the initial stock of physical capital. The RHS is the expenditure side.

3 See footnote 2.

Using the individual's first-order conditions (2.4), (2.5), and (2.8), the intertemporal budget constraint (2.24) can be written as

$$A = U_{C_0} C_0 + \beta U_{C_1} C_1 \\ + U_{L_0}(L_0 + E) + \beta U_{L_1} L_1 (1 - \eta) \quad (2.25)$$

with

$$A \equiv U_{C_0} R_0^{\tau} K_0, \quad (2.26)$$

which is a function of the endogenous variables C_0, L_0, E, and τ_0^K, and of the exogenous variables K_0 and H_0.

The allocations that the individual's problem imply for a given government policy satisfy the implementability constraint (2.25) and the per-period resource constraints (2.14) and (2.15) (see Proposition 1 in Chari and Kehoe (1999)).

The government commits to a specific policy chosen at the outset of period 0, meaning that it does not reoptimize during the course of time.

The Ramsey problem reads

$$\begin{aligned}\mathcal{L} &= U(C_0, L_0 + E) + \beta U(C_1, L_1) \\ &+ \theta_0 \Big(F(K_0, L_0 H_0) + (1 - \delta_K) K_0 - C_0 - K_1 - fE - g_0 \Big) \\ &+ \beta \theta_1 \Big(F(K_1, L_1 H_1) + (1 - \delta_K) K_1 - C_1 - g_1 \Big) \\ &+ \mu \Big(G(E) + (1 - \delta_H) H_0 - H_1 \Big) \\ &+ \varphi \Big(U_{C_0} C_0 + \beta U_{C_1} C_1 \\ &\qquad + U_{L_0}(L_0 + E) + \beta U_{L_1} L_1 (1 - \eta) - A \Big).\end{aligned}$$

The following assumption simplifies the derivation of the first-order conditions.

Assumption 2.1.
The utility function U *is additively separable in consumption and nonleisure, that is,* $U_{C_t L_t} = 0, \ t = 0, 1$.

The first-order conditions for the Ramsey problem are

$$C_0: \ U_{C_0} - \theta_0 + \phi\Big(U_{C_0 C_0} C_0 + U_{C_0} - A_{C_0}\Big) = 0, \tag{2.27}$$

$$C_1: \ U_{C_1} - \theta_1 + \phi\Big(U_{C_1 C_1} C_1 + U_{C_1}\Big) = 0, \tag{2.28}$$

$$\begin{aligned} L_0: \ & U_{L_0} + \theta_0 F_{Z_0} H_0 \\ & + \phi\Big(U_{L_0 L_0}(L_0 + E) + U_{L_0} - A_{L_0}\Big) = 0, \end{aligned} \tag{2.29}$$

$$L_1: \ U_{L_1} + \theta_1 F_{Z_1} H_1 + \phi\Big(U_{L_1 L_1} L_1 + U_{L_1}\Big)(1 - \eta) = 0, \tag{2.30}$$

$$\begin{aligned} E: \ & U_{L_0} - \theta_0 \varphi + \mu G'(E) \\ & + \phi\left(U_{L_0 L_0}(L_0 + E) + U_{L_0} - \beta U_{L_1} L_1 \frac{d\eta}{dE}\right) = 0, \end{aligned} \tag{2.31}$$

$$K_1: \ -\theta_0 + \beta\theta_1\big(F_{K_1} + 1 - \delta_K\big) = 0, \tag{2.32}$$

$$H_1: \ \beta\theta_1 F_{Z_1} L_1 - \mu - \phi\beta U_{L_1} L_1 \frac{d\eta}{dH_1} = 0. \tag{2.33}$$

Maximizing over τ_0^K would be the same as taxing away the return to the initial stock of capital, which essentially is a lump-sum tax.[4] Assuming $\tau_0^K = 0$ rules out this form of taxation.

4 To see this point, maximize the Lagrangian over τ_0^K:

$$\frac{\partial \mathcal{L}}{\partial \tau_0^K} = \phi U_{C_0} F_{K_0} K_0$$

Introducing a lump-sum tax, namely τ_0^K, enhances welfare, as less distortionary taxation is necessary. ϕ measures the cost of using distortionary taxation. The other three factors are positive. Therefore, $\phi > 0$. Optimally, τ_0^K should be chosen such that all government expenditures could be financed. Then we would have $\phi = 0$ and the present problem

2.2.7 *Results*

2.2.7.1 *Taxation of Physical Capital*

To study the case of taxation of physical capital, the following assumption limits the analysis to a specific type of utility functions.

Assumption 2.2.
The instantaneous utility function shall have the following form:

$$U(C_t, \cdot) = \begin{cases} \frac{C_t^{1-\sigma}-1}{1-\sigma} - V(\cdot),\ t = 0, 1,\ 0 \leqslant \sigma \neq 1. \\ \\ \ln C_t - V(\cdot),\ t = 0, 1,\ \sigma = 1. \end{cases}$$

V *is strictly increasing and strictly convex. It is a function of* $L_0 + E$ *and* L_1*, respectively.* $1/\sigma$ *is the intertemporal elasticity of substitution in consumption.*

Proposition 2.2.
Given Assumption 2.2 with $\sigma > 0$*, if* $R_0^\tau K_0 > 0$*, then* $\tau_1^K > 0$*.*

Proof.
Combine the conditions (2.27), (2.28), and (2.32):

$$\frac{U_{C_0}}{\beta U_{C_1}} \frac{1 + \phi\,(1-\sigma) - \phi \frac{U_{C_0 C_0}}{U_{C_0}} R_0^\tau K_0}{1 + \phi(1-\sigma)} = R_1^s.$$

To determine τ_1^K, use the individual's conditions (2.4) and (2.8):

$$1 < \frac{1 + \phi\,(1-\sigma) - \phi \frac{U_{C_0 C_0}}{U_{C_0}} R_0^\tau K_0}{1 + \phi(1-\sigma)} = \frac{R_1^s}{R_1^\tau} = 1 + \frac{\tau_1^K r_1}{R_1^\tau}.$$

coincides with the first-best problem. This renders the whole analysis uninteresting. See also Jones, Manuelli, and Rossi (1997, p. 111).

$$(2.34)$$

The proposed result follows immediately. □

Assumption 2.2 is necessary because it allows one to compare the denominator and numerator in (2.34), which would not be possible if the coefficient σ were not constant.

Proposition 2.2 is a well-known result in macroeconomics.[5] Taxation of the return to physical capital in the first period was ruled out by assumption, and therefore the government was not able to extract the profit coming from the initial stock of physical capital. In period 1, the positive tax on capital income is due to this initial stock. One may view this capital tax as an attempt to take away part of the return to capital, which was ruled out in the first period.

The proof of Proposition 2.2 highlights also that taxation of physical capital income depends on the individual's preferences for consumption. To emphasize this point, suppose that the individual's utility from consumption is linear, $\sigma = 0$. Then $U_{C_0 C_0} = 0$, and $\tau_1^K = 0$ results.

The preceding analysis furthermore shows that the results are derived despite and not because of the presence of human capital in the model. This points out that taxes on the return to physical capital are not a vehicle to provide education incentives. As the next section will show, the wedge between the discounted marginal social return and the marginal social cost of education does not vanish with the optimal capital tax rate.

5 See, for instance, Proposition 7 in Chari and Kehoe (1999).

2.2.7.2 *Taxation of Human Capital*

Proposition 2.3.

The discounted marginal social return to education is smaller than the marginal social cost:

$$\frac{F_{Z_1} L_1 G'(E)}{R_1^s} < f + F_{Z_0} H_0. \tag{2.35}$$

Proof.

The first-order conditions (2.29) and (2.31) imply

$$\begin{aligned}
&\beta \frac{\theta_1}{\theta_0} F_{Z_1} L_1 G'(E) - \left(f + F_{Z_0} H_0\right) \\
&\quad = -\frac{\phi}{\theta_0} \left\{ \beta L_1 U_{L_1} \left(-G'(E) \frac{d\eta}{dH_1} - \frac{d\eta}{dE} \right) + A_{L_0} \right\}.
\end{aligned} \tag{2.36}$$

By (2.32), $\beta\theta_1/\theta_0$ equals the social discount factor $1/R_1^s$. As a result, the LHS of (2.36) is the wedge Δ as defined by (2.22).

To prove the inequality, one has to determine the sign of the factor in curly brackets. The first term in it can be rearranged using the law of motion (2.2) for human capital, the specific functional form (2.3) of G, and the individual's optimality condition (2.9). Therefore,

$$\Delta = -\frac{\phi}{\theta_0} \left\{ \frac{\gamma}{H_1} \underbrace{(1 - \delta_H) H_0}_{\#1} (\varphi + \omega_0) U_{C_0} + \underbrace{A_{L_0}}_{\#2} \right\} < 0.$$

From the definition (2.26), $A_{L_0} = U_{C_0} F_{K_0 Z_0} H_0 K_0$.[6] As long as $H_0 > 0$, the factor in the curly brackets is positive. It further increases as the human capital depreciation rate δ_H decreases. ϕ is positive for the reason explained above; see footnote 4, p. 49. □

6 Recall that $\tau_0^K = 0$ was assumed.

Corrolary 2.1.

In the second-best optimum, the individual overinvests in human capital relative to the first best.

Proof.

One can show that the discounted marginal return to education is a decreasing function of E, ceteris paribus. The marginal cost is constant. As a result, the individual overinvests in human capital relative to the first best. $\square$

The results qualify the education efficiency theorem and show under which circumstances it does not hold. To begin, Proposition 2.3 holds in any case if there is an initial stock of human capital, $H_0 > 0$. Then at least #2 does not vanish. The source of distortion is the term A_{L_0}, which is due to the endogeneity of the first-period interest rate F_{K_0}. It is an initial endowment effect similar to the one discussed in the context of physical capital taxation. In Richter (2009) and Bovenberg and Jacobs (2005) this effect is not present, because they use a partial equilibrium analysis in which the interest rate is fixed, which means $F_{K_0Z_0} = 0$ and $A_{L_0} = 0$ results.

Regarding #1, if $H_1 \equiv G(E) = aE^{\gamma}$, which follows from setting $\delta_H = 1$ or $H_0 = 0$ in the law of motion (2.2) for human capital, the elasticity of the function $H_1(E)$ is constant, that is, $\frac{d\eta}{dE} = \frac{d\eta}{dH_1} = 0$. This means the positive product $(1 - \delta_H)H_0$ is one source of distortion, because as long as the initial stock of human capital does not fully depreciate, the elasticity η is increasing. This is the essence of Remark 2 in Richter (2009).

Both effects #1 and #2 would vanish with $H_0 = 0$, and the education efficiency theorem would result. This case corresponds to ruling out labor supply in the first period, as is done, for instance, in Jacobs and Bovenberg (2010a).

The preceding results allow one to show that education is effectively subsidized. Use (2.35) and (2.9) to derive

$$\frac{F_{Z_1}}{R_1^s\left(f + w_0 H_0\right)} < \frac{\omega_1}{R_1^\tau\left(\varphi + \omega_0 H_0\right)} \Leftrightarrow \Delta < 0. \tag{2.37}$$

The before-tax rate of return to education is smaller than the after-tax rate of return, which means that the wedge Δ between the discounted marginal social return and the the marginal social cost is negative. Therefore, the following proposition results:

Proposition 2.4.
Education is effectively subsidized relative to the first best. The private rate of return to education is larger than the social rate of return.

Physical and human capital are two assets, which the individual can hold to smooth consumption over time. Above, on page 51, it is explained that it is optimal to tax the return to physical capital in the second period. Turning to human capital, the individual disposes of an initial stock of human capital H_0. The return to it can only be taxed in a distorting way, because the tax rate τ_0^L does not have the characteristic of a lump-sum tax. Consequently, this tax is an imperfect instrument, in the sense that it distorts the labor decision, to extract the return to the initial stock of human capital. For this reason, in the second period, when the individual reaps the fruits of education, I therefore would have expected the government to at least partly skim off the additional return that could be attributed to the initial stock of human capital. The striking result, however, is that the contrary is true. The government should subsidize the accumulation of human capital. To sum up this point, an initial stock of physical capital implies that it is optimal to tax physical capital, whereas an initial stock of human capital calls for a subsidization of human capital. The intuition for why it is optimal to subsidize human capital may be the following. Labor taxation exerts a depressing effect on the accumulation of human capital. To counter this, a subsidy is helpful.

One may view the above result in a different light and interpret the model as the steady state of an OLG model as in Nielsen and Sørensen (1997). Then the OLG interpretation of the present model is in line with Propositions 4.2 and 4.3 in chapter 4. The first result states that if the function $H_1(E)$ is isoelastic, education will remain undistorted. This case corresponds to setting $\delta_H = 1$ and thereby implicitly assuming that the young individual does not inherit any stock of human capital from the old individual. Term #1 vanishes. Proposition 3 states that if the elasticity of $H_1(E)$ is increasing, which is the case when $\delta_H < 1$, education will be subsidized relative to the first best. Also, the strength of the positive distortion depends on the cost resulting from the unavailability of lump-sum taxes, as captured by the Lagrange multiplier ϕ. Term #2 is not present in a steady state, because the initial endowment effect A_{L_0} only occurs in the first period and not later on.

2.2.7.3 *Taxation of Labor*

Proposition 2.5.

Labor tax rates are given by

$$\frac{\tau_0^L}{1-\tau_0^L} = -\frac{\phi}{1+\phi}\left(\upsilon_0 + \frac{A_{L_0} + A_{C_0}F_{Z_0}H_0 + F_{Z_0}H_0 U_{C_0C_0}C_0}{-U_{L_0}}\right) \tag{2.38}$$

and

$$\frac{\tau_1^L}{1-\tau_1^L} = -\frac{\phi}{1+\phi}\left((1-\eta)\,\upsilon_1 - \eta + \frac{U_{C_1C_1}}{-U_{L_1}}C_1F_{Z_1}H_1\right) \tag{2.39}$$

with

$$\upsilon_0 = \frac{(L_0+E)U_{L_0L_0}}{U_{L_0}} \text{ and } \upsilon_1 = \frac{L_1U_{L_1L_1}}{U_{L_1}}$$

denoting the reciprocals of the elasticities of nonleisure in periods 0 *and* 1 *in Frisch's sense.*[7]

Proof.

Combine the first-order conditions (2.27) and (2.29), and (2.28) and (2.30). □

τ_0^L depends on initial endowment effects and the individual's preferences for consumption. τ_1^L is affected by the effect of human capital, which is captured by the elasticity η.

The following assumption helps to gain further insight into (2.38) and (2.39).

Assumption 2.3.

7 L_t is implicitly defined by (2.5). Differentiating this condition with respect to, say, ω_0, holding λ_0 constant, yields $1/\upsilon_0 = U_{L_0}/[(L_0+E)U_{L_0L_0}]$. See Cahuc and Zylbergerg (2004, p. 20) for further details.

1. *Interpret the above model as the steady state of an OLG model.*

2. *The utility function* U *shall be linear in consumption.*

3. *Human capital fully depreciates* ($\delta_H = 1$)*, or the initial stock of human capital is zero* ($H_0 = 0$)*.*

The first assumption implies that the initial endowment effects A_{L_0} and A_{C_0} are not present. $U_{C_t C_t} = 0$ follows from the second assumption, which implies that savings are not taxed in the second period: $\tau_1^K = 0$. The third assumption implies that the elasticity of the function expressing human capital in the second period equals the human capital production function's elasticity: $\eta = \gamma$.

Division of (2.39) by (2.38) then yields

$$\frac{\frac{\tau_1^L}{1-\tau_1^L}}{\frac{\tau_0^L}{1-\tau_0^L}} = \frac{(1-\gamma)\upsilon_1 - \gamma}{\upsilon_0}. \tag{2.40}$$

(2.40) is the analogue to equation 13 in Richter (2009), who terms it an extension of the inverse elasticity rule to cope with endogenous education.[8] He, however, assumed a utility function of the form $U = Z(C_0, C_1) - V(L_0 + E) - V(L_1)$, with the function Z being linear homogeneous. Because the utility function used here is time-separable in consumption, linear homogeneity means that utility is linear.

To have positive tax rates on labor income when young and old, the numerator in (2.40) has to be positive: $\gamma <$

8 Another insignificant difference is that Richter (2009) uses exclusive tax rates, whereas inclusive tax rates are used here.

$$1 - \tau = \frac{1}{1+\tau'}$$

is the formula for converting from a tax-inclusive basis to a tax-exclusive basis (Atkinson and Stiglitz, 1980, p. 70).

$\upsilon_1/(1+\upsilon_1)$. This inequality emerged from the analysis of the second-order conditions; see appendix 2.A. It is a sufficient condition that must hold to have a well-behaved problem and moreover ensures that the tax rates are positive.

2.3 CONCLUSION

This chapter has reassessed the models by Bovenberg and Jacobs (2005) and Richter (2009). Their papers and this chapter tackle the same set of questions, use different approaches, but in the end come to similar conclusions. First, I demonstrated that the question of how to tax the return to physical capital is not affected by the accumulation of human capital. Taxing the return to physical capital investments is not a means to yield efficient investments in human capital. Then I showed that an increasing elasticity of the function expressing the stock of human capital in the second period implies subsidizing the return to education. An increasing elasticity arises if the individual is endowed with an initial stock of human capital that does not fully depreciate.

The existing literature on Ramsey models of optimal taxation (see Atkeson, Chari, and Kehoe (1999) among others) and this chapter both come to the conclusion that the question of whether to tax the return to capital or not depends on individual consumption preferences. In the second period this result is special, because one has to allow for an initial endowment effect, which is due to the initial stock of capital. The issue of capital taxation is independent of whether the model features human capital accumulation or not. To sum up this point, taxing capital income in the second period is optimal, but for other reasons than achieving efficient investments in human capital.

Further research could discuss the present model in an infinite-horizon setup, as it is done by Jones, Manuelli, and Rossi (1997). The major difference is that their human capital production function exhibits constant returns to scale with respect to stock variables that enter as means of production. By this specification they model human capital very symmetrically to physical capital and show that the return to education should remain untaxed in a steady state. The natural question then arises what exactly is the difference between human and physical capital. This chapter works with a human capital production function that does not include the stock of human capital as a production factor. Then a model of optimal taxation could answer the question of how to tax the return to education if the human capital production function does not exhibit the restrictive properties as in Jones, Manuelli, and Rossi (1997). See also Ljungqvist and Sargent (2004, p. 534,) for a short discussion of this point.

To study the conditions that must hold to have a well-behaved problem with an interior solution, the original problem (2.1) is written as one with only a single constraint. This is done by deriving the intertemporal budget constraint by substituting out the capital K_1, and then replacing the stock of human capital H_1 with the law of motion (2.2) for human capital. The Lagrangian therefore reads

$$\begin{aligned}\mathcal{L} &= U(C_0, L_0 + E) + \beta U(C_1, L_1) \\ &+ \lambda \bigg(R_0^\tau K_0 + \omega_0 L_0 H_0 + \frac{1}{R_1^\tau} \omega_1 L_1 \big(G(E) + 1 - \delta_H \big) \\ &\qquad - C_0 - \frac{1}{R_1^\tau} C_1 - \varphi E \bigg).\end{aligned}$$

A sufficient condition for the solution to solve the constrained maximization problem is that the bordered Hessian of the Lagrangian satisfies the condition that the last four leading principal minors alternate in sign, the sign of the first one being positive.

Let ∇g denote the gradient of the constraint:

$$\nabla g = \left[-1, \omega_0 H_0, \frac{1}{R_1^\tau} \omega_1 L_1 G' - \varphi, -\frac{1}{R_1^\tau}, \frac{1}{R_1^\tau} \omega_1 G \right]. \quad (2.41)$$

The Hessian of the Lagrangian reads

$$H = \begin{pmatrix} U_{C_0 C_0} & 0 & 0 & 0 & 0 \\ 0 & U_{L_0 L_0} & U_{L_0 L_0} & 0 & 0 \\ 0 & U_{L_0 L_0} & U_{L_0 L_0} + \lambda \frac{1}{R_1^\tau} \omega_1 L_1 G'' & 0 & \lambda \frac{1}{R_1^\tau} \omega_1 G' \\ 0 & 0 & 0 & \beta U_{C_1 C_1} & 0 \\ 0 & 0 & \lambda \frac{1}{R_1^\tau} \omega_1 G' & 0 & \beta U_{L_1 L_1} \end{pmatrix}.$$

Then the bordered Hessian is

$$bH = \begin{pmatrix} 0 & \nabla g \\ \nabla g' & H \end{pmatrix}. \tag{2.42}$$

When deriving the Hessian, it was assumed that the cross derivatives of U are zero ($U_{CL} = 0$) and that human capital fully depreciates ($\delta_H = 1$). Dropping these simplifying assumptions does not change the following results.

Let D_k denote the kth leading principal minor of bH. Then straightforward but tedious calculations yield

$$\begin{aligned}
D_3 = & -(\omega_0 H_0)^2 U_{C_0C_0} - (-1)^2 U_{L_0L_0} > 0, \\
D_4 = & \lambda \frac{1}{R_1^\tau} \omega_1 L_1 G'' \big(-U_{L_0L_0} - U_{C_0C_0}(\omega_0 H_0)^2\big) < 0, \\
D_5 = & -\lambda \frac{1}{R_1^\tau} \omega_1 L_1 G'' \bigg(U_{C_0C_0} U_{L_0L_0} \left(-\frac{1}{R_1^\tau}\right)^2 \\
& + U_{L_0L_0} \beta U_{C_0C_0} (-1)^2 \\
& + U_{C_0C_0} \beta U_{C_1C_1} (\omega_0 H_0)^2 \bigg) > 0, \\
D_6 = & U_{C_0C_0} \beta U_{C_1C_1} U_{L_0L_0} \left(\frac{1}{R_1^\tau} \omega_1 G\right)^2 \left(-\lambda \frac{1}{R_1^\tau} \omega_1 L_1 G''\right) \\
& - \bigg(U_{L_0L_0} \beta U_{C_1C_1} (-1)^2 + U_{C_0C_0} U_{L_0L_0} \left(-\frac{1}{R_1^\tau}\right)^2 \\
& + U_{C_0C_0} \beta U_{C_1C_1} (\omega_0 H_0)^2 \bigg) \\
& \times \lambda \frac{1}{R_1^\tau} \omega_1 G'' U_{L_1} \frac{1}{1-\gamma} \big((1-\gamma)\upsilon_1 - \gamma\big) \overset{!}{<} 0 \\
& \text{with } \upsilon_1 = L_1 U_{L_1L_1} / U_{L_1} \text{ and } \gamma = \frac{G'}{G} E.
\end{aligned}$$

The sign of D_6 must be negative. The minuend is negative. The first factor of the subtrahend is positive. Hence, the second factor must be positive:

$$(1-\gamma)\upsilon_1 > \gamma \Leftrightarrow \gamma < \frac{\upsilon_1}{1+\upsilon_1}.$$

The requirement is that the concavity of the utility function, captured by υ_1, has to be sufficiently large to compensate for the lack of concavity of the law of motion (2.2) for human capital, measured by γ.

3

TAXING HUMAN CAPITAL: A GOOD IDEA

3.1 INTRODUCTION

"Is physical capital special?" Jones, Manuelli, and Rossi (1997) ask. Using the Ramsey approach (Ramsey, 1927), they add human capital to an optimal taxation model with physical capital similar to that of Chamley (1986) and Judd (1985). By modeling human capital almost symmetrically to physical capital they show that in a stationary state all taxes are zero. Chamely and Judd's result is thus shown to extend to human capital. What drives this zero-tax result is that the human capital production function features constant returns to scale with respect to the stock of human capital. Jones, Manuelli, and Rossi (1997) call this specification a zero-profit condition. As a consequence, human capital disappears as an object of taxation in a competitive equilibrium. But they acknowledge that if the human capital production function violates the assumption of constant returns to scale, the stationary-state labor tax will not be zero. Jones, Manuelli, and Rossi (1997) raise an intriguing question and provide useful insights into the nature of optimal taxation, but in the end, unfortunately, no answer is evident. The difference between physical and human capital still is not clear, because they have made it disappear by means of zero-profit conditions.[1]

1 See also Ljungqvist and Sargent (2004, pp. 534) for this line of argument.

This chapter takes up the issue of modeling human capital almost symmetrically to physical capital. I drop the constant-returns-to-scale assumption. The human capital production function does not include the current stock of human capital, which therefore is not self-productive. It does not raise the productivity of human capital investments, or interchangeably, education. The increasing and concave production function only includes the individual's time devoted to education. Time spent on education cannot be substituted by physical goods.[2] Instead, the individual has to pay for verifiable[3] direct costs, e.g., tuition fees, that depend on the amount of education. The government may choose to subsidize this cost. It therefore has two instruments at its disposal to guide education. Labor taxes and the subsidy both affect the opportunity cost of education. The next periods' labor tax rates affect the discounted stream of marginal earnings from education.

I derive two results: The first one is not surprising but nonetheless important, as it helps to clarify the role of zero capital taxation when the model features human capital. The other is new and shows how to deal with profits coming from education. First, optimal taxation in the stationary state prescribes not taxing capital income, as Chamley (1986) and Judd (1985) show. The zero-capital-tax result holds despite the presence of human capital. Lucas (1990), Jones, Manuelli, and Rossi (1997) and Chari and Kehoe (1999) also derive this result. The education decision depends only on

2 Allowing for physical goods as an additional production factor does not affect the results obtained by Jones, Manuelli, and Rossi (1997), as Chari and Kehoe (1999) show.

3 Reis (2007, chapter 4) assumes that the government cannot distinguish between consumption and expenditures on education and finds that it is optimal to tax human capital.

how the labor tax and the education subsidy interact with each other. This relates to the second result, stating that in the optimum the marginal social return to education is larger than the marginal social cost. The so-called *Education Efficiency Theorem* (Richter, 2009), which states that the education decision is undistorted given certain assumptions, does not hold. From the inequality between the marginal social return and the marginal social cost it follows that education is effectively taxed, i.e., the private rate of return to education is smaller than the social rate of return. Turning to the underlying tax rates, it results that the cost of education is not fully tax-deductible, the labor income tax rate is higher than the rate of subsidization. As a consequence, the individual underinvests in human capital relative to the first best.

The second result is striking. Since the individual is endowed with perfect foresight and therefore must be able to internalize the effects of its actions, one would have expected to derive an equality between the private and social rates of return to education, and the *Education Efficiency Theorem* to hold - a result that Jones, Manuelli, and Rossi (1997), among others,[4] also obtain. Their zero-tax results imply that all private and social rates of return from investments in physical and human capital are equal in the stationary state. The difference in results is due to how I model the accumulation of human capital. The specification used gives rise to profits in equilibrium. Profits from education are not pure in the strict sense, because they still depend on raw

4 For further reference, see Lucas (1990), Bull (1993), Milesi-Ferretti and Roubini (1998), Chari and Kehoe (1999), Barbie and Hermeling (2006), and Richter (2009).

labor supply. The government taxes away part of the return to education, thereby accepting the distortion of education.

To derive clear-cut results, the analysis is confined to an examination of the stationary state. In the stationary state, the individual's decision variables remain constant. As usual, it is assumed that a unique stationary state exists and that the economy converges to it. It would be straightforward to introduce exogenous growth. To allow for a setting in which the economy grows endogenously is however not possible. The reason for this limitation is the specification of the human capital production function. Lucas (1988) and Caballe and Santos (1993) provide a discussion of the existence and properties of a balanced growth path. They show that the human capital production function must feature constant returns to scale with respect to the stock of human capital.

3.2 THE MODEL

3.2.1 *Individual's Problem*

The individual solves the following maximization problem:

$$
\begin{aligned}
\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \Big\{ & u(c_t, 1 - n_t - e_t) \\
& - \lambda_t \Big(c_t + k_{t+1} + b_{t+1} + (1 - \tau_t^e) f e_t \\
& - (1 - \tau_t^n) w_t n_t h_t - R_t^k k_t - R_t^b b_t \Big) && (3.1) \\
& - \mu_t \Big(h_{t+1} - (1 - \delta_h) h_t - G(e_t) \Big) \Big\} && (3.2)
\end{aligned}
$$

The individual's utility function u is strictly increasing and concave in both arguments and continuously differentiable everywhere. The Inada conditions apply to ensure interior solutions. In each period t the individual faces a consumption-labor-leisure choice. It consumes c_t, which is not taxed,[5] and devotes n_t time units to work in the labor market and e_t time units to investment in human capital. The total time endowment is normalized to one, i.e., $n_t + e_t + \ell_t = 1$, where ℓ_t is the amount of leisure. The individual combines its raw labor supply n_t with the current stock of human capital h_t. The product $z_t \equiv n_t h_t$ is called the effective labor supply,[6] it earns the after-tax wage rate $(1 - \tau_t^n) w_t$ where w_t is the real wage rate. The individual must spend resources $(1 - \tau_t^e) f$ per time unit invested in human capital. One may think of f as tuition, books, and other related expenses. The government subsidizes this cost at rate τ_t^e. The individual lends capital k_{t+1} to the firm. The rate of return net of taxes and depreciation is $R_{t+1}^k \equiv (1 - \tau_t^k) r_{t+1} + 1 - \delta_k$, where r_t is the real interest rate and δ_k is the rate at which capital depreciates. The individual may lend b_{t+1} to the government which offers a rate of return of R_{t+1}^b in the next period. In period 0, the individual earns income from capital $R_0^k k_0$ and government debt $R_0^b b_0$. (3.1) is the individual's budget constraint in period t, which is associated with the Lagrange multipliers λ_t.

Associated with the Lagrange multiplier μ_t, the law of motion (3.2) describes the accumulation of human cap-

5 Taxing consumption only complicates the analysis without yielding further insights in the present context.

6 This specification is a special case of Jones, Manuelli, and Rossi (1997), who use the more general function $z = M(x, h, n)$ and assume that it exhibits constant returns to scale with respect to h and market goods x. Judd (1999) works out that this specification is not innocuous, as any deviation gives rise to positive taxation of human capital. This point Jones, Manuelli, and Rossi (1997, p. 103) acknowledge.

ital. Investments e_t enter the human capital production function G, which is strictly increasing and concave, i.e., $G'' < 0 < G'$. The crucial assumption is that the current stock of human capital does not enter the production function; it does not increase productivity. The output of G is added to the depreciated stock of human capital, the rate of depreciation being $0 < \delta_h \leqslant 1$.[7] Furthermore, the human capital production function G is assumed to be isoelastic, that is, $\gamma \equiv G'e/G < 1$.[8] Finally, β is the individual's discount factor, which, for simplicity, stays constant over time.

The first-order conditions are

$$\frac{\partial u}{\partial c_t} \equiv u_{c_t} = \lambda_t, \tag{3.3}$$

$$\frac{\partial u}{\partial \ell_t} \equiv u_{\ell_t} = (1 - \tau_t^n) w_t h_t \lambda_t, \tag{3.4}$$

$$u_{\ell_t} + \lambda_t (1 - \tau_t^e) f = \mu_t G'(e_t), \tag{3.5}$$

$$\lambda_{t+1} \beta (1 - \tau_{t+1}^n) w_{t+1} n_{t+1} + \mu_{t+1} \beta (1 - \delta_h) = \mu_t, \tag{3.6}$$

$$\lambda_t = \beta \lambda_{t+1} R_{t+1}^k, \tag{3.7}$$

$$\lambda_t = \beta \lambda_{t+1} R_{t+1}^b. \tag{3.8}$$

Combine (3.7) and (3.8) to derive

$$R_{t+1}^b = R_{t+1}^k. \tag{3.9}$$

7 $\delta_h = 1$ means the individual cannot use the stock of human capital accumulated so far in the next period.

8 This is an assumption that features prominently in the literature. See Jacobs and Bovenberg (2010b) for a discussion and their footnote 3 for more references.

(3.9) is a familiar condition that states that there is arbitrage-freeness between investments in physical capital and government bonds. Both investments promise the same rate of return in equilibrium.

Human capital can be regarded as an asset, similar to physical capital, that yields a rate of return, which in equilibrium must be equal to the other assets' rates of return. To see this, recursively eliminate μ_{t+1} in (3.6), and use (3.5) and (3.7)[9]:

$$R^k_{t+1} = \frac{\sum_{i=0}^{\infty} q^{t+1+i}_{t+1}(1-\tau^n_{t+1+i})w_{t+1+i}n_{t+1+i}G'(e_t)(1-\delta_h)^i}{(1-\tau^e_t)f + (1-\tau^n_t)w_t h_t}. \tag{3.10}$$

with

$$q^{t+1+i}_{t+1} = \prod_{j=1}^{i}\left(R^k_{t+1+j}\right)^{-1},$$

which denotes the private period $t+1$ price of a unit of the consumption good in period $t+i+1$.[10]The numerator in (3.10) summarizes the discounted sum of returns due to a marginal investment e_t, henceforth referred to as the marginal (private) return to education. The investment in

9 Then the transversality condition

$$\lim_{i\to\infty}\left(\prod_{j=0}^{i}\left(R^k_{t+j}\right)^{-1}\right)(1-\delta_h)^i\frac{G'(e_t)}{G'(e_{t+i})} \times \left[\left(1-\tau^n_{t+i}\right)w_{t+i}h_{t+i} + \left(1-\tau^h_{t+i}\right)f\right] = 0$$

also emerges, which holds as long as $0 < \delta_h \leqslant 1$.

10 This implies $q^{t+1}_{t+1} = 1$.

period t not only increases tomorrow's stock of human capital and thereby the wage earned, but also the stock afterwards at the decreasing rate $1 - \delta_h$.[11] The denominator summarizes the marginal (private) cost of education in period t, comprising direct cost and foregone earnings. The optimality condition (3.10) reveals arbitrage-freeness between investments in human and physical capital.

(3.10) also shows that the depreciation rate δ_h and the after-tax rate of return to physical capital investments R^k affect the discounted present value of a time unit e_t invested in human capital similarly. An increasing capital tax rate, which reduces R^k, and an increasing rate of depreciation both raise the marginal return to education (Davies and Whalley, 1991).

For further reference, the stationary state version of (3.10) is[12]

$$\frac{\beta}{1-\beta(1-\delta_h)}(1-\tau^n)wnG' = (1-\tau^n)wh + (1-\tau^e)f. \tag{3.11}$$

(3.11) can be interpreted in the same way as (3.10). The individual devotes time to education up to the point where the marginal cost equals the marginal return to education. One can also see that if the direct cost of education were 100% tax-deductible, i.e., $\tau^n = \tau^e$, the choice of education would be undistorted. Boskin (1975) was the first to state this insight.

(3.11) also reveals that capital taxation does not affect the marginal return to education, because only the individual's

11 To allow for $\delta_h = 1$, $0^0 = 1$ must hold.
12 Use $1 = \beta R^k$, which is the stationary state version of (3.7).

discount factor β matters. This means that only the labor tax rate τ^n and the rate of subsidization τ^e affect the wedge between the marginal return to and the marginal cost of education.

3.2.2 *Firm's problem*

The representative firm produces the single consumption good using a neoclassical constant-returns-to-scale production function. It maximizes profits

$$F(k_t, n_t h_t) - r_t k_t - w_t n_t h_t$$

in capital k_t and effective labor $z_t \equiv n_t h_t$, taking the capital rental rate r_t and the wage rate w_t as given . As a result,

$$F_{k_t} \equiv \frac{\partial F(k_t, z_t)}{\partial k_t} = r_t, \tag{3.12}$$

$$F_{z_t} \equiv \frac{\partial F(k_t, z_t)}{\partial z_t} = w_t. \tag{3.13}$$

The constant-returns-to-scale production technology implies that the firm makes zero profit in equilibrium.

3.2.3 *Government's problem*

The government finances an exogenously given stream of government expenditures $\{g_t\}_{t=0}^{\infty}$. Its per-period budget constraint is

$$g_t + R_t^b b_t = \tau_t^k r_t k_t + \tau_t^n w_t h_t n_t - \tau_t^e f e_t + b_{t+1}. \tag{3.14}$$

3.2.4 *Competitive Equilibrium*

A competitive equilibrium consists of a feasible allocation

$$\{c_t, n_t, e_t, k_t, h_t, g_t\}_{t=0}^{\infty},$$

a price system

$$\{w_t, r_t, R_t^b\}_{t=0}^{\infty},$$

and a government policy

$$\{\tau_t^n, \tau_t^k, \tau_t^e, b_t, g_t\}_{t=0}^{\infty},$$

such that, given the price system and the government policy, the allocation solves the individual's and firm's problems, and the government balances its budget. $\mathcal{C}$ is the set of the competitive equilibria that result from different government policies. Put formally:

$$\begin{aligned}
\mathcal{C} = \Big\{ & \{c_t, n_t, e_t, k_t, h_t, g_t\}_{t=0}^{\infty} : \\
& \exists \{\tau_t^n, \tau_t^k, \tau_t^e, b_t, g_t\}_{t=0}^{\infty}, \{w_t, r_t, R_t^b\}_{t=0}^{\infty} \\
& \text{s.t. } (3.3) - (3.8), (3.12) - (3.13), (3.1), (3.2) \text{ and } (3.14) \\
& \text{ hold for all } t = 0, 1, \dots, \\
& \qquad\qquad k_0, b_0 \text{ and } h_0 \text{ are given.} \Big\}.
\end{aligned}$$

3.2.5 *Social Planner's Problem – First-Best Analysis*

The social planner maximizes the individual's utility subject to the resource constraint and the law of motion for human capital. The Lagrangian reads

$$\begin{aligned}\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \Big\{ & u(c_t, 1 - n_t - e_t) \\ & + \theta_t \Big(F(k_t, z_t) + (1 - \delta_k) k_t - c_t - k_{t+1} - f e_t - g_t \Big) \\ & - \mu_t \Big(h_{t+1} - (1 - \delta_h) h_t - G(e_t) \Big) \Big\}.\end{aligned}$$

The first-order conditions for c_t, e_t, n_t, h_{t+1}, and k_{t+1} are

$$u_{c_t} = \theta_t, \tag{3.15}$$

$$\mu_t G'(e_t) = u_{\ell_t} + \theta_t f, \tag{3.16}$$

$$\theta_t F_{z_t} h_t = u_{\ell_t}, \tag{3.17}$$

$$\theta_{t+1} \beta F_{z_{t+1}} n_{t+1} - \mu_t + \beta \mu_{t+1} (1 - \delta_h) = 0, \tag{3.18}$$

$$\theta_t = \theta_{t+1} \beta (F_{k_t} + 1 - \delta_k). \tag{3.19}$$

Analogously to the individual's problem, the following condition shows how the social planner optimally chooses education[13]:

$$F_{k_{t+1}} + 1 - \delta_k = \frac{\sum_{i=0}^{\infty} \tilde{q}_{t+1}^{t+1+i} F_{z_{t+1+i}} n_{t+1+i} G'(e_t) (1 - \delta_h)^i}{f + F_{z_t} h_t}. \tag{3.20}$$

13 Recursively eliminate μ_{t+1} in (3.18), and use (3.16) and (3.19).

with

$$\tilde{q}_{t+1}^{t+1+i} = \prod_{j=1}^{i} \left(F_{k_{t+1+j}} \right)^{-1},$$

which denotes the social period $t+1$ price of a unit of the consumption good in period $t+i+1$. The numerator in (3.20) is the discounted sum of marginal returns to investment e_t, henceforth called marginal (social) return to education. The investment in period t increases not only tomorrow's stock of human capital and thereby the productivity but also the stock afterwards at the decreasing rate $1-\delta_h$. The denominator captures the marginal (social) cost in period t, comprising the direct cost of education and the loss of labor income. The optimality condition (3.20) reveals that the rates of return to physical and human capital accumulation are equal.

For further reference, the stationary-state version of (3.20) reads

$$\frac{\beta}{1-\beta(1-\delta_h)} F_z n G' = F_z h + f. \tag{3.21}$$

The efficiency condition (3.21) will serve as a benchmark when analyzing below how the education decision is affected by the use of distortionary taxation. The preceding discussion therefore suggests the following

Definition 3.1. *Education efficiency is achieved if the marginal social return to education equals the marginal social cost of education. In the first best, there is no wedge between the marginal social return to and the marginal social cost of education.*

Linear taxes are chosen to finance a given stream of government expenditures. The choice of taxes should maximize social welfare subject to resource and budget constraints and taking the individual's and firm's competitive equilibrium behavior into account. Each government policy gives rise to a different competitive equilibrium. The Ramsey problem is to choose the competitive equilibrium that yields the highest utility. To solve the problem, the primal approach (Lucas and Stokey (1983), Atkinson and Stiglitz (1980), Chari and Kehoe (1999)) is adopted.

This approach is one way to take into account the competitive equilibrium behavior. Instead of choosing the optimal policy directly, which yields the optimal allocation and prices, one chooses the optimal allocation that is consistent with competitive equilibrium behavior and then solves for the government policy and prices that support this outcome. The key to solving this problem is to use the so-called implementability constraint that summarizes the individual's competitive equilibrium behavior.

In the present model, three conditions on the Ramsey problem must hold. The first one, the implementability constraint, is the individual's budget constraint after having substituted for after-tax prices by means of the individual's first-order conditions.

Combining the per-period budget constraints (3.1) leads to the intertemporal budget constraint (using (3.8)):

$$R_0^k k_0 + R_0^b b_0 + \sum_{t=0}^{\infty} \left(\prod_{j=1}^{t} \left(R_j^k \right)^{-1} \right) (1 - \tau_t^n) w_t n_t h_t$$

$$= \sum_{t=0}^{\infty} \left(\prod_{j=1}^{t} \left(R_j^k \right)^{-1} \right) \left(c_t + (1 - \tau_t^e) f e_t \right). \quad (3.22)$$

The transversality conditions

$$\lim_{t \to \infty} \left(\prod_{j=0}^{t} \left(R_j^k \right)^{-1} \right) k_{t+1} = 0 \tag{3.23}$$

and

$$\lim_{t \to \infty} \left(\prod_{j=0}^{t} \left(R_j^b \right)^{-1} \right) b_{t+1} = 0 \tag{3.24}$$

must hold. If (3.23) and (3.24) were positive, then the individual could find an alternative allocation yielding a higher utility by simply consuming more in finite time. The reverse cannot hold either, because some other individual has to be on the lending side and could increase utility for the reason just explained.

Then, using the individual's first-order conditions (3.3) and (3.4) and thereby substituting out $(1 - \tau_t^n) w_t h_t$, the intertemporal budget constraint (3.22) can be written as

$$W_0 + \sum_{t=0}^{\infty} \beta^t u_{\ell_t} n_t = \sum_{t=0}^{\infty} \beta^t u_{c_t} \left(c_t + (1 - \tau_t^e) f e_t \right) \tag{3.25}$$

with $W_0 \equiv u_{c_0} \left(R_0^k k_0 + R_0^b b_0 \right)$, which is the value of the initial endowment of physical capital and government bonds. (3.25) is the first constraint in the planner's problem.

The first-order conditions for e_t and h_{t+1}, (3.5) and (3.6), which yield (3.10) and determine the dynamic choice of h_{t+1}, have not been used. Therefore, they give rise to a second constraint, which Jones, Manuelli, and Rossi (1997) and Atkeson, Chari, and Kehoe (1999) call an Euler equation for the accumulation of human capital:

$$\beta n_{t+1} u_{\ell_{t+1}} h_{t+1}^{-1} + \beta(1-\delta_h)\frac{u_{\ell_{t+1}} + u_{c_{t+1}}(1-\tau_{t+1}^e)f}{G'(e_{t+1})} \\ = \frac{u_{\ell_t} + u_{c_t}(1-\tau_t^e)f}{G'(e_t)}. \quad (3.26)$$

$(1-\tau_t^e)f$ could also be eliminated in (3.25) using (3.26). But dealing with the resulting double sum is cumbersome, which is why it is more convenient to work with two implementability constraints. In any case, either approach must yield the same solution. Pursuing the present way mixes the primal and the dual approach, as the planner has to optimize over the allocation and over the tax rate τ_t^e.

Third, the economy's resource constraint is

$$F(k_t, z_t) + (1-\delta_k)k_t - c_t - k_{t+1} - f e_t - g_t = 0. \quad (3.27)$$

The set $\mathcal{R}$ consists of all allocations that satisfy the three constraints above and the law of motion (3.2) for human capital. Put formally,

$$\mathcal{R} = \Big\{\{c_t, n_t, e_t, k_t, h_t, g_t\}_{t=0}^{\infty} : \\ (3.2), (3.25), (3.26) \text{ and } (3.27) \text{ hold for all } t = 0, 1, \ldots\Big\}.$$

The Ramsey problem is to choose a member belonging to the set $\mathcal{R}$ that yields the highest utility.

The key result to solving the Ramsey problem is the following

Proposition 3.1 (see Chari and Kehoe (1999), Proposition 1)**.** *The competitive equilibrium allocations satisfy the resource constraints and the implementability constraint. Furthermore, given* *allocations that satisfy these constraints, one can construct policies and prices that, together with the given allocations, constitute a competitive equilibrium. Put formally,* $\mathcal{C} = \mathcal{R}$.

Proof. See appendix 3.A. □

The Ramsey problem therefore reads

$$\begin{aligned}
\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \Bigg\{ & W(c_t, n_t, e_t, \tau_t^e, \phi) \\
& + \theta_t \Big(F(k_t, z_t) + (1-\delta_k)k_t - c_t - k_{t+1} - fe_t - g_t \Big) \\
& - \mu_t \Big(h_{t+1} - (1-\delta_h)h_t - G(e_t) \Big) \\
& - \eta_t \Bigg(\beta n_{t+1} u_{\ell_{t+1}} h_{t+1}^{-1} - \frac{u_{\ell_t} + u_{c_t}(1-\tau_t^e)f}{G'(e_t)} \\
& \qquad + \beta(1-\delta_h)\frac{u_{\ell_{t+1}} + u_{c_{t+1}}(1-\tau_{t+1}^e)f}{G'(e_{t+1})} \Bigg) \Bigg\} - \phi W_0,
\end{aligned}$$

with

$$\begin{aligned}
W(c_t, n_t, e_t, \tau_t^e, \phi) = {} & u(c_t, 1 - n_t - e_t) \\
& + \phi \Big(u_{c_t}\big(c_t + (1-\tau_t^e)fe_t\big) - u_{\ell_t} n_t \Big)
\end{aligned}$$

defining the so-called pseudo-welfare function, which includes the implementability constraint and also depends on the endogenous Lagrange multiplier ϕ. Jones, Manuelli, and Rossi (1997) follow the same approach. But their and the present setup differ substantially. First, due to the special assumptions made regarding the human capital production function, they show that human capital does not appear in the implementability constraint. Second, when solving the Ramsey problem they neglect the Euler equation for the accumulation of human capital (3.26). After having found the solution to this relaxed problem, they show that this equation is satisfied anyway. Similarly, they derive a stationary-state arbitrage-freeness condition for human capital and a corresponding Ramsey problem's condition. Because in their setup time devoted to education only gives rise to some cost in the form of forgone earnings and because the labor income tax is proportional, the tax cannot have an effect on education in a stationary state. Both the return and the cost are taxed at the same rate, and both are reduced in the same proportion.[14] It is the implementability constraint (3.26) that captures the transitional dynamics of the accumulation of human capital. Setting up the problem in a way that allows one to put this constraint aside and then to show that it is satisfied anyway does not, however, help to explore the special nature of human capital.

As k_0 is exogenous, τ_0^k works like a lump-sum tax.[15] To rule out this trivial form of taxation, it is common to assume $\tau_0^k = 0$.

14 Even more obviously, this is the case in Chari and Kehoe (1999), too.

15 To see this point, maximize the Lagrangian over τ_0^k:

$$\frac{\partial \mathcal{L}}{\partial \tau_0^k} = \phi u_{c_0} F_{k_0} k_0$$

ϕ measures the costs of using distortionary taxation. Optimally, τ_0^k should be chosen such that all government expenditures could be financed by taxing away the return to the initial stock of physical capital and thereby abstaining from levying distorting taxes on capital and labor. The other three factors are positive. Therefore, $\phi > 0$. It is then possible to increase τ_0^k until $\phi = 0$ and the present problem coincides with the first-best problem. This renders the whole analysis uninteresting. See also Jones, Manuelli, and Rossi (1997, p. 111).

Under the assumption that a unique stationary state exists,[16] the first-order conditions for c_t, e_t, n_t, h_{t+1}, k_{t+1} and τ_t^e, evaluated at the stationary state, are

$$W_c - \theta - \eta\left(nu_{c\ell}h^{-1} - \delta_h \frac{u_{\ell c} + u_{cc}(1-\tau^e)f}{G'}\right) = 0, \tag{3.28}$$

$$\begin{aligned} &W_e - \theta f + \mu G' - \eta\Bigg(-nu_{\ell\ell}h^{-1} \\ &- \delta_h \frac{\big(-u_{\ell\ell} - u_{c\ell}(1-\tau^e)f\big)G' - \big(u_\ell + u_c(1-\tau^e)f\big)G''}{G'^2}\Bigg) \\ &= 0, \end{aligned} \tag{3.29}$$

$$\begin{aligned} &W_n + \theta F_z h \\ &-\eta\left(h^{-1}(-u_{\ell\ell}n + u_\ell) - \delta_h \frac{-u_{\ell\ell} - u_{c\ell}(1-\tau^e)f}{G'}\right) = 0, \end{aligned} \tag{3.30}$$

$$\begin{aligned} \mu = {} & \theta \frac{\beta}{1-\beta(1-\delta_h)} F_z n \\ & + \eta \frac{\beta}{1-\beta(1-\delta_h)} nu_\ell h^{-2}, \end{aligned} \tag{3.31}$$

$$1 = \beta(F_k + 1 - \delta_k), \tag{3.32}$$

$$\eta \frac{\delta_h}{G'} = -\phi e. \tag{3.33}$$

The first-order conditions (3.28)-(3.33), the resource constraint (3.27), the implementability constraint (3.25), and the Euler equation (3.26) for the accumulation of human capital determine the Ramsey allocation $\{c, e, n, h, k, \tau^e\}$ along with the Lagrange multipliers θ, η, and ϕ.[17]

16 This is a common assumption frequently found in the literature. Judd (1999) uses a compactness assumption on the marginal social value of government wealth instead of the convergence assumption adopted here and shows that the average capital tax rate is zero for any long interval.

17 Chari, Christiano, and Kehoe (1994) explain how to compute ϕ. First, fix ϕ and solve for the entire allocation, using all first-order conditions

The following analysis is devoted to studying the tax rates τ^k, τ^n, and τ^e that implement the Ramsey allocation as a competitive equilibrium, $\mathcal{R} \subseteq \mathcal{C}$.[18]

Proposition 3.2.

Capital income is not taxed in the stationary state, i.e., $\tau^k = 0$.

Proof. Combine (3.32) with (3.7) evaluated at the stationary state. □

This is the seminal result by Chamley (1986) and Judd (1985). Evidently, the private and social rates of return to capital investments are equal. The zero-capital-tax result is independent of whether the model features human capital or not. From this follows that there is no trade-off between efficiency in physical and human capital formation.

Proposition 3.3.

1. *Labor income is taxed in the stationary state if the human capital production function's elasticity is sufficiently small.*
2. *The labor income tax rate is not higher than* 100% *if preferences satisfy the following condition:*

$$-\frac{u_{cc}c}{u_c} + \frac{u_{c\ell}}{u_c}\big((1-\gamma)n + e\big) < 1 + 1/\phi \tag{3.34}$$

Proof.

Combining (3.3), (3.4) and (3.28), (3.30) and rearranging yields

$$1 - \tau^n = \frac{1 + \phi\left(1 + \frac{u_{cc}}{u_c}c - \frac{u_{c\ell}}{u_c}\big((1-\gamma)n + e\big)\right)}{1 + \phi\left(1 - \gamma - \frac{u_{\ell\ell}}{u_\ell}\big((1-\gamma)n + e\big) + \frac{u_{c\ell}}{u_c}c\right)}.$$

and resource constraints. Then, check whether this allocation satisfies the implementability constraint. If not, iterate on ϕ until the constraint holds.

18 See Proposition 3.1 for the central argument.

1. $\tau^n > 0$ amounts to requiring

$$\gamma < \frac{\left(-\frac{u_{\ell\ell}}{u_\ell} + \frac{u_{c\ell}}{u_c}\right)(n+e) + \left(-\frac{u_{cc}}{u_c} + \frac{u_{c\ell}}{u_\ell}\right)c}{1 + \left(-\frac{u_{\ell\ell}}{u_\ell} + \frac{u_{c\ell}}{u_c}\right)n}.$$

2. $\tau^n < 1$ amounts to requiring the condition (3.34) to hold.

□

Restrictions are imposed on the individual's preferences and the properties of the human capital production function. Suppose the utility function reads $u(c, \ell) = \ln c + \kappa \ln \ell$. In this special case, conditions 1 and 2 then reduce to $\gamma < 1/(1 + n/\ell)$ and $1 < 1 + 1/\varphi$. Condition 1 says that the elasticity parameter has to be below unity, as has been assumed above on page 68. Condition 2 is always satisfied as long as distortionary taxes are used.

Proposition 3.4.
If the human capital production function is isoelastic, the education decision is distorted. In the second best, there is underinvestment in human capital relative to the first best.

Proof.
Combine (3.26), (3.29), (3.30), (3.31), and (3.33) to obtain[19]

$$\frac{\beta}{1 - \beta(1 - \delta_h)} F_z n G' = \frac{\varphi}{\theta} \gamma u_\ell + F_z h + f. \tag{3.35}$$

Equation (3.35) states that the discounted flow of marginal returns to education equals the marginal cost plus some distortion term,

$$\frac{\varphi}{\theta} \gamma u_\ell. \tag{3.36}$$

19 See appendix 3.B for the details.

The distortion term is positive, as each Lagrange multiplier is positive. Therefore, the marginal social return, which is decreasing in e, is larger than the marginal social cost, which is constant in e. The individual is required to underinvest in human capital relative to the first best. $\square$

Given that the education decision is distorted, the next question is what this means for the tax rates.

Corrolary 3.1.
In the stationary state, the direct cost of education is not fully tax-deductible, that is, $\tau^e < \tau^n$.

Proof.
Multiply (3.35) by $1 - \tau^n$, and combine the result with (3.11) using (3.13):

$$
\begin{aligned}
&(1-\tau^n)\frac{\phi}{\theta}\gamma u_\ell + (1-\tau^n)(F_z h + f) \\
&\quad = (1-\tau^n)F_z h + (1-\tau^e)f \\
\Leftrightarrow\ &(1-\tau^n)\frac{\phi}{\theta}\gamma u_\ell = (\tau^n - \tau^e)f.
\end{aligned}
$$

All the Lagrange multipliers are positive. Given that $\tau^n < 1$, the desired result follows. $\square$

The preceding results allow one to study how the social and the private return to education are related to each other.

Corrolary 3.2.

1. *The private rate of return to capital investments is equal to the social rate of return.*

2. *The private rate of return to capital investments is equal to the private rate of return to education.*

3. *The private rate of return to education is smaller than the social rate of return. Education is effectively taxed.*

Proof.
One has to show that

$$(1-\tau^k)r + 1 - \delta_k = F_k + 1 - \delta_k$$
$$= \frac{\frac{F_z(1-\tau^n)nG'}{1-\beta(1-\delta_h)}}{F_z(1-\tau^n)h + f(1-\tau^e)} < \frac{\frac{F_z nG'}{1-\beta(1-\delta_h)}}{F_z h + f}.$$

The first and second equalities follow from Proposition 3.2 and (3.11). (3.32) and Proposition 3.4 imply the inequality. □

The marginal social return is taxed at a higher rate than the marginal social cost. The result is that this tax scheme negatively distorts education incentives, as Proposition 3.4 clarifies.

To shed more light on the above results, consider the government's stationary-state budget constraint (3.14), which can be written as follows:

$$g + (R^B - 1)b = \tau^n(whn - fe) + (\tau^n - \tau^e)fe.$$

The direct cost of education is taxed at the rate $(\tau^n - \tau^e)$ as long as $\tau^n > \tau^e$. Suppose that the converse were true, and consider a marginal decrease of τ^e. Then τ^n has to decline as well if the government's budget constraint is to continue to hold. $\tau^n \leqslant \tau^e$ implies the private rate of return to education to be larger than the social rate of return. The considered tax reform has the effect that the marginal cost of education, consisting of the direct cost and forgone earnings, increases less than the marginal return. As a result, the private rate of return to education increases. Ceteris paribus, the individual earns more income and hence consumption rises which increases utility. An efficiency gain would result, which is not possible, given that the planner maximizes efficiency.

3.3 CONCLUSION AND DISCUSSION

This chapter explores the special nature of human capital compared to physical capital in an optimal taxation model. Capital income remains untaxed in the stationary state. The presence of human capital does not interfere with this result. This means taxing capital and human capital are two distinct issues and capital taxation is not a means to guide efficient education policy. This leaves labor taxation and the subsidization of the direct cost of education as the only instruments to set efficient education incentives.

As the human capital production function includes time as the only production factor and not the current stock of human capital, the analysis calls for effective taxation of education, thereby partly extracting the ability rent. To achieve this end, the cost of education is not fully tax-deductible. As a consequence, the subsidy is insufficient in encouraging education and to offset the distortions caused by the tax on labor.

Critical is the assumption that the cost of education is fully observable. This allows the government to use this piece of information to set an effective tax on education. Otherwise it has to resort to the labor tax alone to achieve this end, which would imply higher welfare costs. In reality it is not that easy to get exact data on the time spent on education. Likewise it is not possible to precisely estimate the stock of human capital.

1. $\mathcal{C} \subseteq \mathcal{R}$:

   The statement is true because the implementability constraint is the intertemporal budget constraint after having substituted out prices using the individual's first-order conditions. Derive (3.26) by combining (3.5) and (3.6) and substituting out prices again. Because the individual's and government's budget constraints are satisfied, the resource constraint is satisfied by Walras's law. This proves the first inclusion.

2. $\mathcal{R} \subseteq \mathcal{C}$:

   The converse, that any allocation satisfying the implementability and resource constraints satisfies competitive equilibrium behavior, is also true. This amounts to finding prices and a government policy, namely tax rates, such that the allocation that is in $\mathcal{R}$ is also in $\mathcal{C}$. To derive R^b_{t+1} use (3.3) and (3.8). Obtain r_t and w_t from (3.12) and (3.13). (3.3) and (3.4) yield τ^n_t. (3.3) and (3.7) determine τ^k_t. τ^e_t is defined recursively by (3.5) and (3.6).

   By construction, the Ramsey allocation satisfies the individual's budget constraint and the economy's resource constraint. By Walras' law, the government's budget constraint is satisfied as well.

(3.33) serves to eliminate η:

$$\eta = -\phi \frac{G'e}{\delta_h} \tag{3.37}$$

Equalize (3.29) and (3.30), and plug in (3.37):

$$\begin{aligned}
&-\theta f + \mu G' - \phi \frac{G'e}{\delta_h} \delta_h \frac{-u_{\ell\ell} G' - \big(u_\ell + u_c(1-\tau^e) f\big) G''}{G'^2} \\
&- \phi \frac{G'e}{\delta_h h} n u_{\ell\ell} + \phi\big(u_c(1-\tau^e) + u_\ell\big) \\
&= \theta F_z h + \phi \frac{G'e}{\delta_h h}(-u_\ell n + u_\ell) + \phi \frac{G'e}{\delta_h} \delta_h \frac{u_\ell \ell}{G'} \\
&\qquad + \phi u_{\ell\ell} n - \phi u_\ell
\end{aligned} \tag{3.38}$$

(3.2) yields $h = G + (1-\delta_h)h \Leftrightarrow 1/G = 1/(\delta_h h)$. Using the constant elasticity γ of G and substituting for μ by means of (3.31), one can manipulate (3.38) as follows:

$$\begin{aligned}
&\theta \frac{\beta}{1-\beta(1-\delta_h)} F_z n G' \\
&- \phi\gamma \left(\frac{\beta}{1-\beta(1-\delta_h)} G' n u_\ell h^{-1} - u_c(1-\tau^e) f \right) \\
&\qquad = \theta(F_z h + f)
\end{aligned} \tag{3.39}$$

The stationary-state version of (3.26) reads

$$\frac{\beta}{1-\beta(1-\delta_h)} G' n u_\ell h^{-1} = u_\ell + u_c(1-\tau^e) f. \tag{3.40}$$

Plug (3.40) into (3.39) to finally derive (3.35):

$$\frac{\beta}{1-\beta(1-\delta_h)} F_z n G' = \frac{\phi}{\theta} \gamma u_\ell + F_z h + f \tag{3.35}$$

4

EFFICIENT HUMAN CAPITAL POLICY WITH OVERLAPPING GENERATIONS AND ENDOGENOUS GROWTH

4.1 INTRODUCTION

Education is a field in which policies of OECD countries exhibit remarkable differences. This is borne out by the data published in 2009 (OECD, Tables A8.2 and A8.4). While various countries effectively subsidize education at the tertiary level, other countries effectively tax this activity. Such a finding does not only raise the question of which policy is superior, it also raises the question of whether and how an effective subsidization of education can be justified in terms of efficiency. This chapter studies this question in a framework of overlapping generations and endogenous growth. Two reasons of why it may be efficient to subsidize education are highlighted. The first one is well known from the literature. It is the potential need to internalize the positive effect that human capital investments of selfish individuals have on the productivity of descendent generations. Efficient internalization requires subsidizing investments up to the first best. This chapter stresses the second reason. This is the negative effect that distortionary taxation of labour has on education and growth. If the elasticity of the human capital investment function is strictly increasing, it is shown to be a second best policy to subsidize education even relative to the first best.

The traditional approach to optimal taxation follows Ramsey (1927) and takes the model of a representative taxpayer as a starting point. A critical feature of this literature is that the results characterizing optimal policy heavily depend on whether the representative taxpayer plans for finite or infinite periods. If the taxpayer's planning horizon is infinite, the rationale for employing distortionary linear taxes and subsidies turns out to be weak. This point was originally made by Chamley (1986) and Judd (1985) with respect to capital taxes. It extends, however, to the model with endogenous education, as has been demonstrated by Bull (1993), Jones, Manuelli, and Rossi (1993, 1997), and Atkeson, Chari, and Kehoe (1999). The question of whether human or non-human capital is accumulated is largely irrelevant. In the long run neither accumulation should be distorted.

The policy recommendations are less clear-cut if the taxpayer's planning horizon is finite. In the finite case it is primarily a matter of marginal rates of intertemporal substitution in consumption whether taxing saving is efficient or not. In particular, saving should be untaxed only if the taxpayer's utility is weakly separable between consumption and labour and homothetic in consumption (Atkinson and Stiglitz (1972), Sandmo (1974)). By contrast, the design of efficient education policy is more a reflection of the specific properties of the earnings function. This function has to be weakly separable in qualified labour supply and education and the elasticity with respect to the latter has to be constant if it shall be second best not to distort the choice of education (Jacobs and Bovenberg (2010b); Bovenberg and

Jacobs (2005)). If weak separability holds and if the elasticity is strictly increasing, it is second best to subsidize education (Richter (2009)). If the planner trades off efficiency and equity and if education and qualified labour are complementary, it is equally second best to subsidize education (Jacobs and Bovenberg, 2010b).

It somewhat discredits the Ramsey approach that the suggested policy recommendations so critically depend on the taxpayer's planning horizon. That is why the present chapter studies optimal taxation in a model with overlapping generations. Such a model stands between the static and dynamic Ramsey frameworks and it therefore promises less debatable policy recommendations. The broader objective of the present study is to characterize optimal policies for education, labour, and saving in a dynamic framework with overlapping generations. The narrower objective is to rationalize the effective subsidization of endogenous education. Such objectives may justify putting aside various shortcomings often turned against similar studies. In particular, we exclusively focus on efficiency and we stick to the representative taxpayer framework because one would not really be surprised to learn that subsidizing education can well be optimal when equity is traded off against efficiency. Furthermore, we rule out potential reasons of market failure because they may help to justify market intervention but certainly not the subsidization of education relative to the first best.

The model chosen is one with overlapping generations and endogenous growth. Individuals live for two periods. They decide on education, saving, and nonqualified labour in their youth. They supply qualified labour when old. The productivity of qualified labour increases in the stock of human capital inherited from preceding generations, and it also increases in own educational investments. Individuals either may be perfect altruists with respect to descendent generations or may behave selfishly. The implications of selfishness have been studied before by Wigger (2002, Sec. 3.4) and Docquier, Paddison, and Pestieau (2007) for a framework in which the government is not constrained in the use of policy instruments. It is shown that decentralizing the first best requires subsidizing education up to the first best. The present chapter goes beyond these earlier studies by endogenizing labour supply and by assuming that the government can only employ linear policy instruments. Most remarkably, major results characterizing efficient static policy extend to the dynamic framework. In particular, it is second best not to distort education if the human capital investment function is isoelastic in education. It is argued, however, that such constant elasticity has debatable implications in a dynamic framework. It implies that the human capital stock accumulated by preceding generations melts down to zero if just one generation stops investing. More appealing is the assumption that the elasticity of the investment function is increasing and that the human capital stock does not depreciate completely if just one generation fails to invest. If this is the case, it is second best at balanced growth to subsidize education even relative to the first best. This means that the marginal social cost of human capital should exceed the marginal social return in the long-run second-best optimum. This is a striking result. Not surprising is the need to subsidize education relative to *laissez faire*. This is so because the intergenerational externalities of human capital investments have to be internalized.[1] *A priori* it is not obvious, however, why investments should even exceed the first-best. Subsidizing education requires government revenue, which in the model has to be raised by distortionary taxes on labour and savings. With the intuition of Lipsey and Lancaster (1956 - 1957) in mind, one might hypothesize that it is second best to provide insufficient incentives for education if labour has to be taxed and if the level of comparison is the first best. The contrary, however, is true. The key assumption is the strictly increasing elasticity of the human capital investment function with respect to education. The effect is that it is second best to subsidize education in static analysis, and this effect is shown to extend to the dynamic framework. At balanced growth the need to subsidize increases in the derivative of the investment function's elasticity and in two further factors. One factor is the Lagrange multiplier on the planner's implementability constraint, and the other is the gap between the marginal return to capital and the rate of balanced growth. In other words, the more binding the non-availability of lump-sum taxes is and the more deficient the growth is, the more should human capital accumulation overshoot the first best.

Assuming altruistic individuals changes some conclusions, but not all. Altruists internalize the positive effect that education has on descendents' productivity. Hence the need for government intervention is reduced. However, the second source of inefficiency modelled in this chapter does not vanish. That second source is the need to employ

1 The need is highlighted by various earlier studies. An example is Rey and del Mar Racionero (2002).

distortionary taxes for financing government expenditures. The implications for second-best policy are shown to differ markedly between the first generation and all descendent generations. With respect to descendent generations the following results are obtained. The accumulation of human capital should not be distorted, and this result is obtained for arbitrary utility and human capital investment functions. The accumulation of physical capital should not be distorted if the taxpayer's utility is weakly separable between consumption and non-leisure and homothetic in consumption. Furthermore, qualified and nonqualified labour should be taxed uniformly across the life cycle when utility is compatible with balanced growth. Such results will not only give reason in Section 6 to qualify major results derived by Jones, Manuelli, and Rossi (1997), Atkeson, Chari, and Kehoe (1999), and Erosa and Gervais (2002). They strongly contrast with the results derived for the case of selfish individuals.

The results obtained for the first generation are less contrasting. In particular, it is second best not to distort the first generation's educational choice if the human capital investment function is isoelastic in education. If, however, this function fails to be isoelastic, the optimal education policy for the first generation depends on initial values. On neutralizing the effect of initialization by assuming balanced growth and assuming a strictly increasing elasticity of the human capital investment function, it turns out to be second best to subsidize education. The reason is the same encountered when individuals are selfish. Strictly increasing elasticity is the reason why it is second best to subsidize education in static analysis. This effect extends to the dynamic framework. The need to subsidize is the stronger the larger the derivative of the investment function's elasticity is, the more binding the non-availability of lump-sum taxes is, and the more deficient growth is.

The unifying bottom line for selfish and altruistic individuals is as follows. Altruism well reduces the need to subsidize education relative to *laissez faire*, and altruism also implies that descendent generations should have non-distorted incentives to invest in human capital. The short-run policy recommendations for altruism, however, agree with the long-run recommendations for selfishness. Labour has to be taxed, and – given that the elasticity of the human capital investment function is strictly increasing – education should be subsidized relative to the first best.

The chapter is structured as follows. Section 2 sets up the two-period overlapping-generations model with endogenous growth. The first-order conditions characterizing solutions of the planner's first-best maximization are derived. In Section 3 the utility functions are determined that are compatible with balanced growth in consumption and with constant use of labour and leisure. Section 4 studies the planner's problem when individuals behave selfishly and when no policy instruments but linear ones are available. Section 5 clarifies the relation between effective and efficient subsidization. Section 6 studies the planner's problem for individuals who are altruistic towards descendent generations. Section 7 summarizes.

4.2 THE MODEL AND THE PLANNER'S FIRST-BEST PROBLEM

Consider a sequence of overlapping generations with individuals living for two periods. The index t refers to the generation and to the period in which the representative individual of generation t is young and in her life period 0. Lifetime utility is given by $U_t \equiv U(C_{0t},, C_{1t}, L_{0t}, L_{1t})$ with the arguments C_{0t}, C_{1t}, L_{0t}, and L_{1t} denoting consumption and non-leisure in the life periods $i = 0, 1$. Utility is strictly increasing in consumption, is strictly decreasing in non-leisure, and is strictly concave. Additional restrictions on preferences required if the economy is to exhibit steady state growth are discussed in Section 3. Non-leisure in the second life period, L_{1t}, equals *qualified* labour supplied to the market in period $t+1$. By contrast, non-leisure in the first life period has to be divided between *nonqualified* labour supply $L_{0t} - E_t$ and education E_t. The effect of education is to increase human capital and labour productivity. H_{t-1} is the stock of human capital determining the productivity in period t. It is built up by generation $t-1$ and inherited by generation t. By spending time E_t on education, generation t determines the stock of human capital H_t effective in the second life period. The *human capital accumulation equation* is

$$G(E_t)H_{t-1} = H_t. \qquad (\mu_t \beta^t) \qquad (4.1)$$

$\mu_t \beta^t$ is a Lagrange multiplier associated with the planner's problem we are about to set up. The investment function $G_t \equiv G(E_t)$ is assumed to be non-negative and strictly monotone increasing with elasticity $\eta(E) \equiv EG'/G$ smaller than one. The case of constant elasticity η plays a prominent role in static models of endogenous education (Jacobs and Bovenberg (2010b); Richter (2009)) and equally in what follows. A critical implication is $G(0)=0$ so that the stock of human capital built up by generation t-1 melts down to zero, $H_t=0$, if generation t does not spend positive time on education. If one assumes instead $G(E_t) \equiv \tilde{G}(E_t) + 1 - \delta_H$ with $\delta_H < 1$ and some function $\tilde{G}(E)$ of constant elasticity $\tilde{\eta}$, then $H_t = (1 - \delta_H)H_{t-1}$ follows from $E_t = 0$ so that some human capital is passed on to the next generation even if there are no new investments. In this case, the elasticity of the investment function, $\eta(E) = \left[1 - \frac{1-\delta_H}{G(E)}\right]\tilde{\eta}$, is strictly increasing in E. To allow for both scenarios with constant and increasing elasticity of $G(E)$ we assume $\eta'(E) \geqslant 0$ in what follows.

The functional specification (4.1) is standard in the endogenous growth literature. It can be traced back to Uzawa (1965), and it has been used since by Lucas (1988), Atkeson, Chari, and Kehoe (1999), and others. A key feature is that H_t is linear homogenous in H_{t-1}. A notable implication of (4.1) is that time spent on education (*learning*) is the only variable input in the production of human capital. In particular, learning cannot be substituted by physical inputs or services supplied by instructors. There is however some cost of instruction which accrues in fixed proportion with education. For simplicity's sake, it is modelled as a linear function of inherited human capital and time spent on education, fE_tH_{t-1}. It is suggestive to interpret the exogenous parameter f as *tuition fee*.

There is a second stock variable, K_t, to be interpreted as (nonhuman) capital built up by generation t in their first life period. It is not productive before the second life period, and it depreciates at the rate δ_K. Production F is linear homogenous in capital and effective labour. The *resource constraint* is

$$\begin{aligned} F_t + (1-\delta_K)K_{t-1} &= C_{0t} + C_{1t-1} \\ &+ fE_tH_{t-1} + K_t + A_t \qquad (\alpha_t \beta^t) \end{aligned} \tag{4.2}$$

with $F_t \equiv F(K_{t-1}, (L_{0t} - E_t)H_{t-1}, L_{1t-1}H_{t-1})$.

The variable A_t denotes exogenous government spending. Such spending may be of consumptive and/or productive use. As A_t is exogenous, we refrain from making it an explicit argument of the utility and/or production functions. When taking partial derivatives use is made of the following short forms:

$$F_{Kt} \equiv \frac{\partial F}{\partial K_{t-1}}, \quad F_{L_0 t} \equiv \frac{\partial F}{\partial\big((L_{0t} - E_t)H_{t-1}\big)},$$
$$F_{L_1 t} \equiv \frac{\partial F}{\partial (L_{1t-1}H_{t-1})}.$$

Qualified and nonqualified labour may be perfect substitutes in production, but they need not be. Human capital is obviously labour augmenting. Note that education incurs two kinds of cost. There is the *cost of forgone earnings*, $F_{L_0 t}E_tH_{t-1}$, and the *cost of tuition*, fE_tH_{t-1}. It is not only realism suggesting an explicit modelling of both costs. We shall argue below that the explicit differentiation is the key to understanding why various results derived in the following sections deviate from related results derived in the literature.

The planner maximizes

$$\sum_{t=0}^{\infty} \beta^t U(C_{0t}, C_{1t}, L_{0t}, L_{1t}) \tag{4.3}$$

in C_{0t}, C_{1t}, L_{0t}, L_{1t}, E_t, H_t, and K_t $(t = 0, 1, \ldots)$ subject to the human capital accumulation equation (4.1) and the resource constraint (4.2). The parameters K_{-1}, H_{-1}, $L_{1-1} = L_{1,t=-1}$ are exogenously given. $0 < \beta < 1$ is a discount factor. Assume that this maximization – like all others still to follow – is well behaved and that it has an interior solution for which all choice variables are strictly positive. We abstain from stating all the assumptions needed to guarantee a well-behaved maximization with interior solutions. Identifying those assumptions must remain the object of independent research efforts. In the present chapter we just state those assumptions explicitly needed to derive meaningful first-order conditions of second-best policies. We study neither second-order conditions nor questions of existence. As argued in Richter (2009) and as will become clearer below, a well-behaved maximization requires a specification of $U_t = U(C_{0t}, C_{1t}, L_{0t}, L_{1t})$ which is sufficiently concave to compensate for the lack of concavity of the human capital accumulation equation (4.1). The first-order conditions of the planner's maximization are as follows:

$$U_{C_0t} = \alpha_t, \quad U_{C_1t} = \alpha_{t+1}\beta, \tag{4.4}$$

$$F_{L_0t}H_{t-1}U_{C_0t} = -U_{L_0t}, \quad F_{L_1t+1}H_tU_{C_1t} = -U_{L_1t}, \tag{4.5}$$

$$F_{Kt+1} + 1 - \delta_K = U_{C_0t}/U_{C_1t} = U_{C_0t}/\beta U_{C_0t+1}, \tag{4.6}$$

$$\mu_t G'_t = \alpha_t(f + F_{L_0t}), \tag{4.7}$$

$$\alpha_{t+1}\beta\left[F_{L_1t+1}L_{1t} + F_{L_0t+1}\cdot(L_{0t+1} - E_{t+1}) - fE_{t+1}\right] =$$

$$\mu_t - \beta G_{t+1}\mu_{t+1}. \tag{4.8}$$

The conditions (4.4) and (4.5) characterize efficient consumption and labour choices. The condition (4.6) characterizes efficient saving and efficient capital. The condition (4.7) characterizes the efficient choice of E_t, and (4.8) is the condition characterizing the efficient choice of H_t. Solving (4.7) for μ_t and inserting into (4.8) yields, after some straightforward manipulations, the condition characterizing the *efficient accumulation of human capital*,

$$\begin{aligned} F_{L_1 t+1}L_{1t} + F_{L_0 t+1}L_{0t+1} - \left(F_{L_0 t+1} + f\right) E_{t+1} \\ = [F_{Kt+1} + 1 - \delta_K]\frac{f + F_{L_0 t}}{G'_t} - G_{t+1}\frac{f + F_{L_0 t+1}}{G'_{t+1}}. \end{aligned} \tag{4.9}$$

For the sake of brevity we also speak of *efficient education* if (4.9) holds. The first term on the left-hand side, $F_{L_1 t+1}L_{1t}$, is the return to human capital accruing to generation t in the second life period, and the difference $F_{L_0 t+1}L_{0t+1} - (F_{L_0 t+1} + f)E_{t+1}$ is the return accruing to individuals of the next generation in their first life period. The factor

$$\frac{f + F_{L_0 t}}{G'_t} = \left(f + F_{L_0 t}\right) H_{t-1}\frac{dE_t}{dH_t} \tag{4.10}$$

is the marginal cost of human capital in period t, and $\frac{f+F_{L_0 t+1}}{G'_{t+1}}$ is the marginal cost of human capital one period later. Hence the right-hand side of (4.9) captures the cost resulting from investing in period t instead of postponing the investment to the next period. By separating terms referring to generation *t* from terms referring to generation $t+1$, (4.9) can be written as

$$
\begin{aligned}
&[F_{Kt+1} + 1 - \delta_K] \frac{f + F_{L_0 t}}{G'_t} - F_{L_1 t+1} L_{1t} \\
&= F_{L_0 t+1} L_{0t+1} + (F_{L_0 t+1} + f) E_{t+1} \left[\frac{1}{\eta_{t+1}} - 1 \right] \\
&\qquad\qquad\qquad\qquad \equiv MEB_{t,t+1}. \quad (4.11)
\end{aligned}
$$

Because $\eta_{t+1} < 1$ by assumption, $MEB_{t,t+1}$ is positive. It is the *marginal external benefit* enjoyed by generation *t*+1 and generated by the human capital investment of generation *t*. This excess benefit has to be internalized by first-best policy when individuals are selfish. As a result of internalization, generation *t*'s cost, $[F_{Kt+1} + 1 - \delta_K] \frac{f+F_{L_0 t}}{G'_t}$, exceeds generation t's return to human capital, $F_{L_1 t+1} L_{1t}$.

4.3 BALANCED GROWTH

We speak of *balanced growth* if the non-leisure choices $L_{0t} = L_0, L_{1t} = L_1$, and $E_t = E$ are constant across time while consumption, output, and both types of capital all grow at the common gross rate $G = G(E)$, so that we have $H_{t-1} = G^t H_{-1}$, $K_{t-1} = G^t K_{-1}$, $C_{it} = G^t C_{i0} \equiv G^t C_i$. At balanced growth, $F_{Kt+1} = F_K$ is constant in t. If an efficient allocation is to be compatible with balanced growth, then conditions (4.4) and (4.6) require the rates of substitution

$$
\frac{U_{C_0}(G^t C_0, G^t C_1, L_0, L_1)}{U_{C_0}(G^{t+1} C_0, G^{t+1} C_1, L_0, L_1)}, \frac{U_{C_0}(G^t C_0, G^t C_1, L_0, L_1)}{U_{C_1}(G^t C_0, G^t C_1, L_0, L_1)}
$$

to be both constant in t. Taking total derivatives with respect to *t* and setting the total derivatives equal to zero implies constancy of

$$
\Big[G^t C_0 \cdot U_{C_0 C_0}(G^t C_0, ..)
$$

$$+ G^t C_1 \cdot U_{C_1 C_0}(G^t C_0, ..)\Big] / U_{C_0}(G^t C_0, ..)$$

$$= \Big[G^t C_0 \cdot U_{C_0 C_1}(G^t C_0, ..)$$

$$+ G^t C_1 \cdot U_{C_1 C_1}(G^t C_0, ..)\Big] / U_{C_1}(G^t C_0, ..) \equiv d - 1 \quad (4.12)$$

in t. Upon substituting $\tilde{C}_i$ for $G^t C_i$ and integrating in $\tilde{C}_i$ one obtains

$$C_0 U_{C_0} + C_1 U_{C_1} = dU + cX \quad (4.13)$$

where d, c are constants and where X is a function of L_0, L_1. The following two types of utility specifications satisfy this condition:

(i)

$$U(C_0, C_1, L_0, L_1) = V(C_0, C_1) \cdot \Lambda(L_0, L_1) - D(L_0, L_1) \quad (4.14)$$

where $V(C_0, C_1)$ is homogeneous of degree $d \neq 0$;

(ii)

$$U(C_0, C_1, L_0, L_1) = \left[a_0 \ln C_0 + a_1 \ln C_1\right] \Lambda(L_0, L_1) - D(L_0, L_1) \quad (4.15)$$

Utility functions of type (4.14) satisfy condition (4.13) when setting $c \equiv d \neq 0$, $X \equiv D$ and utility functions of type (4.15) satisfy condition (4.13) when setting $c \equiv a_0 + a_1$, $d \equiv 0$, $X \equiv \Lambda$. Note that both specifications imply that U_{C_i} is homogenous of degree $d - 1$ in consumption.

Efficiency at balanced growth additionally requires

$$- U_{L_i}(G^t C_0, \ldots) = F_{L_i} G^{t+i} H_{-1} U_{C_i}(G^t C_0, \ldots) = \\ F_{L_i} G^{td+1} H_{-1}(C_0, \ldots).$$

In other words, $-U_{L_i}$ must be homogenous of degree d in consumption. This is clearly fulfilled only if $D \equiv$ constant in (4.14) and $\Lambda \equiv$ constant in (4.15). W.l.o.g. this means

(i)

$$U(C_0, C_1, L_0, L_1) = V(C_0, C_1)\Lambda(L_0, L_1), \tag{4.16}$$

where $V(C_0, C_1)$ is homogenous of degree $d \neq 0$;

(ii)

$$U(C_0, C_1, L_0, L_1) = a_0 \ln C_0 + a_1 \ln C_1 - D(L_0, L_1). \tag{4.17}$$

This may also be seen as follows. Start with equation (4.5) and take logs on both sides:

$$\log U_{C_0 t} + \log F_{L_0 t} H_{t-1} = \log(-U_{L_0 t})$$

The LHS grows at the common growth rate G. Hence, the RHS must grow at this rate as well. Differentiate both sides w.r.t. time t:

$$\frac{1}{U_{C_0 t}}\left(G^t C_0 U_{C_0 C_0 t} + G^t C_1 U_{C_0 C_1 t}\right) + \frac{F_{L_0} G^t H_{-1}}{F_{L_0} G^t H_{-1}} = \\ \frac{1}{-U_{L_0 t}}\left(-U_{L_0 C_0 t} G^t C_0 - U_{L_0 C_1 t} G^t C_1\right) \tag{4.18}$$

The factor $\log G$, which results from the differentiation w.r.t t, has already been canceled out. Furthermore $F_{L_0 t} =$

F_{L_0} for all t along the balanced growth path. The RHS is the elasticity of marginal utility of non-leisure time w.r.t. t, which will be denoted σ. From (4.12) follows that the LHS equals $d - 1 + 1 = d$. Thus, σ has to be constant along the balanced growth path with $\sigma = d$. When we calculate σ for the two candidate functions (4.14) and (4.15), then

$$\sigma = \begin{cases} \frac{V_{C_0 t} C_{0t} \Lambda_{L_0}}{V \Lambda_{L_0} - D_{L_0}} + \frac{V_{C_1 t} C_{1t} \Lambda_{L_0}}{V \Lambda_{L_0} - D_{L_0}} \stackrel{!}{=} d, \\ \frac{\frac{a_0}{C_{0t}} C_{0t} \Lambda_{L_0}}{V \Lambda_{L_0} - D_{L_0}} + \frac{\frac{a_1}{C_{1t}} C_{1t} \Lambda_{L_0}}{V \Lambda_{L_0} - D_{L_0}} \stackrel{!}{=} 0. \end{cases}$$

From these two conditions follows $D_{L_0} = 0$ in the non-log case and $\Lambda_{L_0} = 0$ in the log-case. Without loss of generality we set D and Λ equal to one in the non-log and log case, respectively.

An earlier characterization of utility functions compatible with growth in consumption and constancy in leisure is due to King, Plosser, and Rebelo (1988, 2002). These authors however restrict their study to dynamic equilibria and government policies in a Ramsey-type framework with exogenous growth. Furthermore, they work with utility functions $U(C, L)$ which have only two arguments. (4.16) and (4.17) extend their findings.

Assuming balanced growth and utility to be homogeneous of degree d in consumption, we obtain $U_{C_0 t} = G^{(d-1)t} U_{C_0 0}$. Hence $F_K + 1 - \delta_K = G^{1-d}/\beta$ by (4.6). Furthermore, the *condition of transversality*, $\beta^t U_{C_0 t} K_t \to 0$ for $t \to \infty$, implies $(\beta G^{d-1})^t U_{C_0 0} G^t K_{-1} \to 0$ for $t \to \infty$, i.e., $\beta G^d < 1$. As a result, the return to capital exceeds the growth rate:

$$F_K + 1 - \delta_K = G^{1-d}/\beta > G. \tag{4.19}$$

The following analysis studies second-best policy with regard to education, to saving, and also to labour. The focal question, however, is whether it is second best to provide or not to provide efficient incentives for education. As we shall see, much depends on the elasticity of the elasticity of the investment function $G(E)$ and on whether individuals are perfect altruists towards their children or not. In the altruistic model – also called the dynasty model – individuals are assumed to maximize (4.3). In the other case the representative individual is assumed to maximize own lifetime utility

$$U(C_{0t}, C_{1t}, L_{0t}, L_{1t}) \tag{4.20}$$

subject to the own lifetime budget constraint. We study both scenarios, and we start by analyzing efficient taxation in the standard OLG framework with selfish individuals. The approach taken is called the primal approach in optimal taxation.

4.4 Optimal Taxation in the Standard OLG Model with Selfish Individuals

The selfish individual representing generation t is assumed to maximize (4.20) in the five variables $C_{0t}, C_{1t}, L_{0t}, L_{1t}, E_t$, and savings S_t subject to the life-period budget constraints

$$\omega_{0t}(L_{0t} - E_t)H_{t-1} = C_{0t} + \varphi_t E_t H_{t-1} + S_t \quad (\lambda_{0t}) \tag{4.21a}$$

and

$$\omega_{1t}L_{1t}G(E_t)H_{t-1} + R_{t+1}S_t = C_{1t}. \qquad (\lambda_{1t}) \quad (4.21b)$$

In this optimization H_{t-1} is treated as an exogenous parameter. By assumption, any excess supply of savings, $S_t - K_t$, is invested in government bonds. ω_{0t} is the wage rate of nonqualified labour, ω_{1t} is the wage rate of qualified labour, φ_t is the tuition fee, and R_{t+1} is the return earned on savings. All these prices and costs are after tax and subsidy. For each t there are six first-order conditions

$$U_{C_0t} = \lambda_{0t}, \qquad (4.22a)$$

$$U_{C_1t} = \lambda_{1t}, \qquad (4.22b)$$

$$\omega_{0t}H_{t-1}U_{C_0t} = -U_{L_0t}, \qquad (4.23a)$$

$$\omega_{1t}G(E_t)H_{t-1}U_{C_1t} = -U_{L_1t}, \qquad (4.23b)$$

$$\omega_{1t}L_{1t}G_t^{'}U_{C_1t} = \left(\varphi_t + \omega_{0t}\right)U_{C_0t}, \qquad (4.24a)$$

$$R_{t+1} = \lambda_{0t}/\lambda_{1t}. \qquad (4.24b)$$

They are constraints in the planner's optimal taxation problem we are about to set up. In the primal approach to optimal taxation these conditions are used to substitute for the four relative prices ω_{0t}, ω_{1t}, φ_t, R_{t+1}, and the two Lagrange multipliers λ_{0t}, λ_{1t}. After substituting, the lifetime budget constraint derived from (4.21a,b) can be written as

$$\sum_{i=0}^{1}\left[C_{it}U_{C_it} + L_{it}U_{L_it}\right] = \eta_t L_{1t}U_{L_1t}. \qquad (\tilde{\lambda}_t\beta^t) \quad (4.25)$$

The condition (4.25) assumes the role of an *implementability constraint* in the planner's second-best problem. Because

$$-\eta_t \frac{L_{1t} U_{L_1 t}}{U_{C_0 t}} \underset{(4.23),(4.24)}{=} (\varphi_t + \omega_{0t}) E_t H_{t-1}, \tag{4.26}$$

the right-hand side of (4.25) can be interpreted as the private cost of education. As it turns out, the marginal increase in H_t is of particular significance when characterizing second-best policies. Let us call the marginal increase the *private marginal cost of human capital*. The formal definition is

$$\begin{aligned} PMC_t^{HC} &\equiv -\frac{d}{dH_t}\left[\eta_t \frac{L_{1t} U_{L_1 t}}{U_{C_0 t}}\right] \\ &= -\frac{L_{1t} U_{L_1 t}}{U_{C_0 t}} \frac{dE_t}{dH_t} \frac{d}{dE_t} \eta(E_t) \\ &= -\frac{L_{1t} U_{L_1 t}}{U_{C_0 t}} \frac{\eta_t'}{G_t' H_{t-1}} \\ &= -\frac{L_{1t} U_{L_1 t}}{U_{C_0 t}} \frac{1}{H_t} \frac{E_t}{\eta_t} \frac{d\eta_t}{dE_t}. \end{aligned} \tag{4.27}$$

The private marginal cost is obviously increasing in the elasticity of the elasticity of $G(E_t)$. If the elasticity $\eta_t = \eta(E_t)$ is constant, $PMC_t^{HC} = 0$ results. If the elasticity is however strictly increasing, PMC_t^{HC} is positive.

The planner maximizes the sum of discounted lifetime utilities (4.3) in C_{0t}, C_{1t}, L_{0t}, L_{1t}, E_t, H_t, and K_t $(t = 0, 1, \ldots)$ subject to the implementability constraint (4.25), the human capital accumulation equation (4.1), and the resource constraint (4.2). In a fully-fledged description of the planner's maximization one would have to include the first-order conditions of profit maximization. However, these conditions can be used to substitute for the endogenous factor prices before taxes and subsidies. Hence, they are not constraining the planner. The solutions are second best in the sense that they have to fulfil the implementability constraint in addition to the first-best constraints (4.1) and (4.2). If lump-sum taxes were available, the planner could ignore (4.25). Inclusion of (4.25) in the set of constraints implies that the planner is restricted in the choice of policy instruments. The restriction is however not an arbitrary one. Quite to the contrary, implicit in the derivation of (4.25) is the assumption that the planner is not constrained in setting consumer prices ω_{0t}, ω_{1t}, φ_t, and R_{t+1}. This means in particular that labour income can be taxed at different rates over an individual's life cycle. If such differentiation is ruled out by assumption, the planner has to respect an additional constraint, which may have strong implications for the design of optimal taxation. See Erosa and Gervais (2002) for a discussion of this point in an OLG model without endogenous education.

To solve the planner's problem set

$$W_t \equiv U_t + \tilde{\lambda}_t \left\{ \sum_{i=0}^{1} \left[C_{it} U_{C_i t} + L_{it} U_{L_i t} \right] - \eta_t L_{1t} U_{L_1 t} \right\}. \tag{4.28}$$

The first-order conditions are as follows:

$$\frac{\partial}{\partial C_{0t}}, \frac{\partial}{\partial L_{0t}} : W_{C_0 t} = \alpha_t = -\frac{W_{L_0 t}}{F_{L_0 t} H_{t-1}}, \tag{4.29}$$

$$\frac{\partial}{\partial C_{1t}}, \frac{\partial}{\partial L_{1t}} : W_{C_1 t} = \alpha_{t+1}\beta = -\frac{W_{L_1 t}}{F_{L_1 t+1} H_t}, \tag{4.30}$$

$$\frac{\partial}{\partial K_t} : \alpha_{t+1}\beta[F_{Kt+1} + 1 - \delta_K] = \alpha_t, \tag{4.31}$$

$$\frac{\partial}{\partial E_t} : \mu_t G_t' H_{t-1} = \tilde{\lambda}_t \eta_t' L_{1t} U_{L_1 t} + \alpha_t (f + F_{L_0 t}) H_{t-1} \Rightarrow$$

$$\frac{\mu_t}{\alpha_t} \underset{(4.27)}{=} \frac{f + F_{L_0 t}}{G_t'} - \frac{\tilde{\lambda}_t}{\alpha_t} U_{C_0 t} PMC_t^{HC}, \tag{4.32}$$

$$\frac{\partial}{\partial H_t} : \alpha_{t+1} \beta \left[F_{L_1 t+1} L_{1t} + F_{L_0 t+1} \cdot (L_{0t+1} - E_{t+1}) - f E_{t+1} \right]$$
$$+ \mu_{t+1} \beta G_{t+1} = \mu_t. \tag{4.33}$$

We wish to derive characterizations of second-best policy with regard to saving, education, and labour. We start with saving. As has been shown by Atkinson and Stiglitz (1972), Sandmo (1974), Atkeson, Chari, and Kehoe (1999), and others, it is efficient not to distort saving if utility is weakly separable between consumption and non-leisure and is homothetic in consumption, $U = U(V(C_0, C_1), L_0, L_1)$ with a linear homogeneous function V. The utility functions defined in (4.14) and (4.15) are examples of weakly separable and homothetic functions. Weak separability and homotheticity implies

$$\begin{aligned} \frac{W_{C_i}}{U_{C_i}} &= 1 + \tilde{\lambda} \Bigg\{ 1 + \sum_{j=0}^{1} \left[C_j \frac{U_{C_j C_i}}{U_{C_i}} + L_j \frac{U_{L_j C_i}}{U_{C_i}} \right] \\ &\quad - \eta L_1 \frac{U_{L_1 C_i}}{U_{C_i}} \Bigg\} \\ &= 1 + \tilde{\lambda} \left\{ 1 + V \frac{U_{VV}}{U_V} + \sum_{j=0}^{1} L_j \frac{U_{VL_j}}{U_V} - \eta L_1 \frac{U_{VL_1}}{U_V} \right\} \\ &= \text{constant in } i = 0, 1. \end{aligned} \tag{4.34}$$

Relying on (4.29) – (4.31) and (4.34) this implies

$$F_{Kt+1} + 1 - \delta_K = \frac{\alpha_t}{\alpha_{t+1} \beta} = \frac{W_{C_0 t}}{W_{C_1 t}} = \frac{U_{C_0 t}}{U_{C_1 t}}. \tag{4.35}$$

This has to be interpreted as saying that it is optimal from the planner's perspective to equate the marginal rate of return to capital with the private marginal rate of substitution in consumption.

Proposition 4.1.
If behaviour is selfish and if utility is weakly separable between consumption and non-leisure and homothetic in consumption, it is second best not to distort saving.

We turn next to education. We first prove that it is efficient not to distort human capital accumulation if the investment function G is isoelastic. We do so by relying on (4.31) – (4.33), which are the first-order conditions with respect to K_t, E_t, and H_t. By making use of (4.31) and (4.32), (4.33) can be written as

$$\begin{aligned}
&\left[F_{L_1 t+1} L_{1t} + F_{L_0 t+1} \cdot (L_{0t+1} - E_{t+1}) - fE_{t+1}\right] \\
&+ \left[\frac{f + F_{L_0 t+1}}{G'_{t+1}} - \frac{\tilde{\lambda}_{t+1}}{\alpha_{t+1}} U_{C_0 t+1} PMC^{HC}_{t+1}\right] G_{t+1} \\
&= \left[\frac{f + F_{L_0 t}}{G'_t} - \frac{\tilde{\lambda}_t}{\alpha_t} U_{C_0 t} PMC^{HC}_t\right] [F_{Kt+1} + 1 - \delta_K]\,. \qquad (4.36)
\end{aligned}$$

Obviously, (4.36) equals (4.9) whenever

$$PMC^{HC}_{t+1} = PMC^{HC}_t = 0, \qquad (4.37)$$

which is the case if $\eta(E_t)$ is constant.

Proposition 4.2.
Assume selfish behaviour. It is second best not to distort education if the human capital investment function $G(E)$ *is isoelastic.*

Proposition 4.2 is a dynamic version of the *education efficiency proposition*, well known from static tax analysis (Jacobs and Bovenberg (2010b); Bovenberg and Jacobs (2005)). An intuitive explanation is the following. The planner cares about two objectives. One objective is to minimize the efficiency loss resulting from distorted choices of consumption and leisure. The other objective is to minimize losses in the rent income generated by education. In general, these two minimizations are not separable, so that the planner has to trade off. Separability is only ensured if the human capital investment function is isoelastic. If this is the case and if the set of policy instruments is sufficiently rich, it is efficient not to distort education and to minimize the efficiency loss resulting from distorted choices of consumption and leisure. According to Proposition 4.2 this result extends to the dynamic framework and it does not explicitly rely on the utility specifications (4.14) and (4.15). Things are different if the private marginal cost of human capital is positive.

To study this case set

$$\Delta_t \equiv \frac{\tilde{\lambda}_t}{\alpha_t} U_{C_0 t} PMC_t^{HC} \cdot (F_{Kt+1} + 1 - \delta_K) - \frac{\tilde{\lambda}_{t+1}}{\alpha_{t+1}} U_{C_0 t+1} PMC_{t+1}^{HC} \cdot G_{t+1}. \quad (4.38)$$

With this definition (4.36) can be written as

$$\Delta_t = \frac{f + F_{L_0 t}}{G_t'}(F_{Kt+1} + 1 - \delta_K) - \frac{f + F_{L_0 t+1}}{G_{t+1}'} G_{t+1} - F_{L_1 t+1} L_{1t} - \left[F_{L_0 t+1} \cdot (L_{0t+1} - E_{t+1}) - f E_{t+1}\right]. \quad (4.39)$$

Comparison of (4.39) and (4.9) reveals that Δ_t is the efficient wedge between the social cost and the social benefit of investing in human capital in period *t* instead of postponing the investment by one period. A positive wedge stands for subsidizing relative to the first best. *A priori* the sign of Δ_t

is indeterminate. This is different if (4.38) is evaluated at a balanced growth path. By definition, balanced growth means that the non-leisure choices $L_{0t} = L_0, L_{1t} = L_1$, and $E_t = E$ are constant in t while consumption, output, and both types of capital all grow at the common gross rate $G = G(E)$, so that we have $H_{t-1} = G^t H_{-1}$, $K_{t-1} = G^t K_{-1}$, $C_{it} = G^t C_{i0} \equiv G^t C_i$. At balanced growth $F_{Kt+1} = F_K$, $G_{t+1} = G$ in t. Because the utility functions are as specified in (4.16) and (4.17), the other variables entering (4.38) take on the following values:

(i)

$$U_{C_0 t} = G^{(d-1)t} U_{C_0 0} \equiv G^{(d-1)t} U_{C_0}.$$

(ii)

$$\begin{aligned} PMC_t^{HC} &\underset{(4.27)}{=} -\frac{L_{1t} U_{L_1 t}}{U_{C_0 t}} \frac{1}{G'(E_t) H_{t-1}} \eta'(E_t) \\ &= -\frac{L_1 U_{L_1 0} G^{dt}}{U_{C_0 0} G^{(d-1)t}} \frac{1}{G'(E) H_{-1} G^t} \eta'(E) \\ &= -\frac{L_1 U_{L_1}}{U_{C_0}} \frac{1}{G' H_{-1}} \eta' = PMC_0^{HC} \equiv PMC^{HC}. \end{aligned}$$

Because U_{C_0}, U_{L_0} are homogeneous of degree $d-1$ in consumption, W_{C_0}, W_{L_0} are likewise homogeneous of degree $d-1$ in consumption. As a result, the growth factor G^t cancels out in equation (4.29): $W_{C_0 t} = -\frac{W_{L_0 t}}{F_{L_0 t} H_{t-1}}$. After cancelling out, the only variable carrying an index t in this equation is the Lagrange multiplier $\tilde{\lambda}_t$. Hence

(iii) $\tilde{\lambda}_t = \tilde{\lambda}$, and *a fortiori*

(iv) $\alpha_t = W_{C_0 t} = G^{(d-1)t} W_{C_0 0} \equiv G^{(d-1)t} W_{C_0}$ and
$$\frac{U_{C_0 t}}{\alpha_t} = \frac{G^{(d-1)t} U_{C_0}}{G^{(d-1)t} W_{C_0}} = \frac{U_{C_0}}{W_{C_0}}.$$

Eventually, setting $R \equiv F_K + 1 - \delta_K$, (4.38) can be written as

$$\Delta = \tilde{\lambda} \frac{U_{C_0}}{W_{C_0}} \cdot PMC^{HC} \cdot (R - G). \tag{4.40}$$

Interpret $\tilde{\lambda} U_{C_0}/W_{C_0}$ as the social cost associated with the implementability constraint. This factor is positive if the implementability constraint is binding, $\tilde{\lambda} > 0$, which is the case if the non-availability of lump-sum taxes is a binding constraint.[2] In this sense the factor measures the cost resulting from the non-availability of lump-sum taxes. PMC^{HC} is the private marginal cost of human capital, which is positive by assumption and increasing in η'. Finally, $R - G$ is the growth gap, which by (4.19) must be positive as well. Hence Δ is the product of three positive factors.

Proposition 4.3.
Assume selfish behaviour, and U *to satisfy* (4.16) *or* (4.17). *At balanced growth it is second best to subsidize education relative to the first best if the private marginal cost of human capital,* PMC^{HC}*, is positive. The strength of positive distortion increases in (i) the private marginal cost of human capital, (ii) the growth gap, and (iii) the cost resulting from the non-availability of lump-sum taxes.*

2 We abstain from proving in detail that the Lagrange multiplier is positive. Jones, Manuelli, and Rossi (1997, p. 109) do this for a maximization which comes close to the present one. The intuition is the following. Paying generation t some positive lump-sum income would show up on the right-hand side of (4.25). The Lagrange multiplier must be positive if increasing such a lump-sum income can be shown to have a negative effect on the planner's objective function. The effect is indeed negative, because such a lump-sum transfer must be paid at the expense of government funds, which are generated by distortive taxes. Although the government budget constraint is not modelled explicitly, it has to be respected. This follows from Walras's law. In summary, the non-availability of lump-sum taxes is the reason why $\tilde{\lambda}$ is positive.

This is a remarkable result, for reasons explained before. It is rather evident, and has been noted before, that the *laissez-faire* level of education is inefficient from the first-best perspective. Without government intervention, selfish individuals externalize the positive effect of own education on descendent generations' welfare. Not so evident is the result that human capital accumulation should be distorted along balanced growth while capital accumulation should not be distorted, subject to appropriately chosen utility functions. The sign of the efficient distortion is even less obvious. Note that any revenue needed to subsidize the cost of tuition has to be raised by distortionary labour taxes. With the intuition of Lipsey and Lancaster (1956 - 1957) in mind, one could have hypothesized that it is second best to give negative incentives for human capital accumulation relative to the first best if labour has to be taxed. The contrary, however, is true. The key assumption is the strictly increasing elasticity of the human capital investment function with respect to education. If the elasticity is strictly increasing, the private marginal cost of human capital is positive. With a positive private marginal cost of human capital it is second best to subsidize education. This has been shown before by Richter (2009) to hold in static analysis, and it is shown here to extend to the dynamic framework. The need to subsidize increases in the factors listed in Proposition 4.3. In particular, it increases in the elasticity of the human capital investment function's elasticity.

We finally turn to the study of labour taxation. Of particular interest is the efficient taxation of nonqualified labour relative to qualified labour. As the definition of W_t in (4.28) is structurally asymmetric in L_{0t} and L_{1t}, one may easily conjecture that qualified and nonqualified labour should be taxed differently. To make a clear case for differentiated taxation and to obtain clear-cut results, we focus on balanced growth and specific utility functions. Thus we assume

$$U \equiv \sum_{i=0}^{1} \left[a_i \ln C_i - D_i(L_i)\right]. \tag{4.41}$$

In this particular case the first-order condition (4.29) implies:

$$\begin{aligned} & W_{L_0} + F_{L_0} H_{-1} W_{C_0} = 0 \\ \Leftrightarrow\ & U_{L_0} + F_{L_0} H_{-1} U_{C_0} \\ & \quad = \tilde{\lambda}\left[L_0 D_0''(L_0) + D_0'(L_0)\right] \end{aligned} \tag{4.42}$$

Similarly, (4.30) implies

$$\begin{aligned} & W_{L_1} + F_{L_1} G H_{-1} W_{C_1} = 0 \\ \Leftrightarrow\ & U_{L_1} + F_{L_1} G H_{-1} U_{C_1} \\ & \qquad = \tilde{\lambda}(1-\eta)\left[L_1 D_1''(L_1) + D_1'(L_1)\right]. \end{aligned} \tag{4.43}$$

Denote by $\nu_i \equiv L_i U_{L_i L_i}/U_{L_i} > 0$ the elasticity of marginal utility of leisure in life-period i, and define tax rates τ_i by setting $(1-\tau_0)F_{L_0}H_{-1} = -U_{L_0}/U_{C_0}$, $(1-\tau_1)F_{L_1}GH_{-1} \equiv -U_{L_1}/U_{C_1} \Leftrightarrow (1-\tau_i)F_{L_i} = \omega_i$. Dividing (4.43) through by (4.42) gives us

$$\frac{\tau_1/(1-\tau_1)}{\tau_0/(1-\tau_0)} = (1-\eta)\frac{\nu_1+1}{\nu_0+1}. \tag{4.44}$$

For $\eta = 0$, (4.44) is the familiar (*inverse*) *elasticity rule*. According to this rule, wage taxes τ_i should increase in ν_i. If utility were quasi-linear, the ν_i would be the inverse of the wage elasticity of labour supply in life-period i. Hence taxes would have to vary inversely with the wage elasticities rendering the rule its name. The rule is extended by (4.44) to allow for endogenous education. The effect of education is to reduce the tax on qualified labour relative to the tax on nonqualified labour. The deviation from the elasticity rule increases in the elasticity of the human capital investment function, η. See Richter (2009), who derives a similar rule for the static framework.

Proposition 4.4.
Assume selfish behaviour, and U to satisfy (4.41)*. On a balanced growth path it is then second best to tax labour according to the elasticity rule* (4.44)*. The effect of endogenous education is to reduce the tax on qualified labour relative to the tax on nonqualified labour.*

4.5 EFFICIENT AND EFFECTIVE SUBSIDIZATION OF EDUCATION

As mentioned in the introduction, OECD data suggest that various countries effectively subsidize tertiary education while others effectively tax tertiary education. Before substantiating such a statement one has to clarify the underlying notion of effective subsidization and its relation to efficient subsidization.

In the recent publication of 2009 the OECD reports estimates of the private and public net present values for individuals obtaining tertiary education as part of initial education in 2005. In present notation the private net present value is

$$
\begin{aligned}
\mathrm{NPV}_{priv} &\equiv \omega_1 L_1 G H_{-1} U_{C_1}/U_{C_0} - (\varphi + \omega_0) E H_{-1} \\
&\underset{(20)}{=} (\varphi + \omega_0)\left[\frac{G}{EG'} - 1\right] E H_{-1} = \frac{1-\eta}{\eta}(\varphi + \omega_0) E H_{-1}.
\end{aligned}
$$

For the sake of brevity, the time index t is dropped. The public net present value is the difference between the social and the private net present values where the social value

$$
\mathrm{NPV}_{soc} \equiv F_{L_1} L_1 G H_{-1}/[F_K + 1 - \delta_K] - (f + F_{L_0}) E H_{-1}
$$

captures only the return to education accruing to the investing generation. Denote by

$$
\mathrm{PRR} \equiv \frac{\mathrm{NPV}_{priv}}{(\varphi + \omega_0) E H_{-1}} = \frac{1-\eta}{\eta},
$$

$$
\mathrm{SRR} \equiv \frac{\mathrm{NPV}_{soc}}{(f + F_{L_0}) E H_{-1}} = \frac{F_{L_1} L_1 G}{[F_K + 1 - \delta_K](f + F_{L_0}) E} - 1
$$

the *private rate of return* and the *social rate of return*, respectively. Our suggestion is to speak of effective subsidization only to the extent that the private rate exceeds the social rate. Hence denote by

$$
s \equiv \frac{\mathrm{PRR} - \mathrm{SRR}}{\mathrm{PRR}} \tag{4.45}
$$

the *effective rate of subsidization*. The efficient value s_{eff} of this rate is determined by

$$
\begin{aligned}
(1-\eta) s_{eff} &\underset{def}{=} 1 - \frac{\eta F_{L_1} L_1 G}{[F_K + 1 - \delta_K](f + F_{L_0}) E} \\
&= \frac{[F_K + 1 - \delta_K] f + F_{L_0}) - G' F_{L_1} L_1}{[F_K + 1 - \delta_K](f + F_{L_0})} \\
&\underset{(10),(34)}{=} \frac{\Delta + \mathrm{MEB}}{(F_K + 1 - \delta_K)(f + F_{L_0})/G'}
\end{aligned} \tag{4.46}
$$

where $(f + F_{L_0})/G'$ is the social marginal cost of human capital and $MEB = F_{L_0} L_0 + (F_{L_0} + f)E(1-\eta)/\eta$ the marginal external benefit as specified by (4.10) and (4.11). With Δ and MEB, s_{eff} is positive as well. Equation (4.46) confirms the view that there are two reasons for effective subsidization of education. One is the need to internalize the intergenerational externality and the other is the need to compensate for distortionary labour taxation. Just for the sake of illustration we report the empirical values of s for men as they can be computed by means of the data published by OECD (2009, tables A8.2 and A8.4). Positive values for s are obtained in case of TUR (.47), POL (.34), ESP (.22), POR (.20), AUT (.19), CAN (.18), NOR (.10), ITA (.09), and HUN (.04). Negative values are obtained for SWE (-.03), KOR (-.05), DEN (-.05), FIN (-.06), CZE (-.14), USA (-.16), NZL (-.20), GER (-.20), IRL (-.20), FRA (-.32), BEL (-.32), and AUS (-.40). Such extreme differences in effective rates and even more the opposing signs clearly raise the question of which policy is more efficient. A convincing answer however requires a thorough empirical analysis which has to remain the object of future research. The numbers are only reported to illustrate the empirical relevance of the theoretical investigation undertaken in this chapter.

4.6 Optimal taxation in the OLG model with altruistic individuals

The perfectly altruistic individual is assumed to maximize $\tilde{U}_t \equiv U(C_{0t}, C_{1t}, L_{0t}, L_{1t}) + \beta \tilde{U}_{t+1}$, which by recursive substitution amounts to maximizing the sum of discounted lifetime utilities (4.3) in C_{0t}, C_{1t}, L_{0t}, L_{1t}, E_t, H_t, and K_t

$(t = 0, 1, \ldots)$. This objective is maximized subject to the human capital accumulation constraint (4.1) and the dynasty's budget constraint,

$$\begin{aligned}
&\sum_{t=0}^{\infty}\Big[\pi_{t+1}\omega_{1t}L_{1t}H_t + \pi_t\omega_{0t}\big(L_{0t} - E_t\big)H_{t-1}\Big] \\
&= \sum_{t=0}^{\infty}\Big[\pi_t C_{0t} + \pi_{t+1}C_{1t} + \pi_t\varphi_t E_t H_{t-1} \\
&+ \big(\pi_t - R_{t+1}\pi_{t+1}\big)K_t\Big] \qquad (\lambda). \qquad (4.47)
\end{aligned}$$

The price and cost variables have the same meaning as before. The first-order conditions are $(t = 0, 1, \ldots)$

$$\beta^t U_{C_0 t} = \lambda\pi_t, \beta^t U_{C_1 t} = \lambda\pi_{t+1}, \qquad (4.48)$$

$$\omega_{0t}H_{t-1}U_{C_0 t} = -U_{L_0 t}, \omega_{1t}H_t U_{C_1 t} = -U_{L_1 t}, \qquad (4.49)$$

$$\mu_t G_t' = \big(\varphi_t + \omega_{0t}\big)U_{C_0 t}, R_{t+1} = \pi_t/\pi_{t+1}, \qquad (4.50)$$

$$\begin{aligned}
&\lambda\pi_{t+1}\Big[\omega_{1t}L_{1t} + \omega_{0t+1}\big(L_{0t+1} - E_{t+1}\big) - \varphi_{t+1}E_{t+1}\Big] \\
&\qquad = \beta^t\mu_t - \beta^{t+1}G_{t+1}\mu_{t+1}. \qquad (4.51)
\end{aligned}$$

The last condition implies

$$\begin{aligned}
&\lambda\sum_{t=0}^{\infty}\pi_{t+1}\Big[\omega_{1t}L_{1t} + \omega_{0t+1}\big(L_{0t+1} - E_{t+1}\big) - \varphi_{t+1}E_{t+1}\Big]H_t \\
&\qquad \underset{(4.50)}{=} \sum_{t=0}^{\infty}\Big[\beta^t\mu_t H_t - \beta^{t+1}\mu_{t+1}H_{t+1}\Big] \\
&\qquad = \mu_0 H_0 \underset{(4.50)}{=} \frac{\varphi_0 + \omega_{00}}{G_0'}U_{C_0 0}H_0. \qquad (4.52)
\end{aligned}$$

Multiplying the budget constraint (4.47) through by λ and using (4.48), (4.49), (4.50), and (4.52) to substitute for

$\lambda\pi_t$, $\lambda\pi_{t+1}$, ω_{0t}, ω_{1t}, and R_{t+1} in (4.47) yields the implementability constraint

$$\sum_{t=0}^{\infty} \beta^t \sum_{i=0}^{1} C_{it} U_{C_i t} = B \qquad (\tilde{\lambda}) \qquad (4.53)$$

with

$$B \equiv \left\{ \left[\omega_{00} \left(L_{00} - E_0 \right) - \varphi_0 E_0 \right] H_{-1} + \frac{\varphi_0 + \omega_{00}}{G_0'} H_0 \right\} U_{C_0 0}.$$

Similarly, (4.48), (4.49) and (4.50) can be used to substitute for $\lambda\pi_{t+1}$, ω_{0t+1}, ω_{1t}, and μ_t in (4.51), which leaves us with $(t = 0, 1, ...)$

$$\begin{aligned} & - L_{1t} U_{L_1 t} - \beta \Big[\left(L_{0t+1} - E_{t+1} \right) U_{L_0 t+1} \\ & + \varphi_{t+1} E_{t+1} U_{C_0 t+1} H_t \Big] \\ & \underset{(4.51)}{=} \left\{ \mu_t - \beta G_{t+1} \mu_{t+1} \right\} H_t = \mu_t H_t - \beta \mu_{t+1} H_{t+1} \\ & \underset{(4.50)}{=} \left[\varphi_t U_{C_0 t} - U_{L_0 t} \frac{1}{H_{t-1}} \right] \frac{H_t}{G_t'} \\ & \quad - \beta \left[\varphi_{t+1} U_{C_0 t+1} - U_{L_0 t+1} \frac{1}{H_t} \right] \frac{H_{t+1}}{G_{t+1}'}. \quad (\gamma_t \beta^t) \end{aligned} \qquad (4.54)$$

Interpret (4.54) as the Euler equation for human capital accumulation. The planner maximizes the sum of discounted lifetime utilities (4.3) in C_{0t}, C_{1t}, L_{0t}, L_{1t}, E_t, H_t, K_t, and φ_t $(t = 0, 1, ...)$ subject to the resource constraint (4.2), the accumulation constraint (4.1), the implementability constraint (4.53), and the Euler equation (4.54). It is important to note that the cost of tuition φ_{t+1} $(t = 0, 1, \ldots)$ only appears explicitly in equation (4.54). By contrast, the planner's objective function and the constraints (4.1), (4.2), and (4.53) are independent of φ_{t+1}. The equation (4.54) can therefore be treated as a relationship by which the "free" policy variable φ_{t+1} is determined. This solution procedure is feasible because the coefficient of φ_{t+1} in (4.54) does not vanish. The coefficient equals

$$\beta U_{C_0 t+1} \left[E_{t+1} H_t - \frac{H_{t+1}}{G'_{t+1}} \right] = \beta U_{C_0 t+1} E_{t+1} H_t \left(1 - \frac{1}{\eta_{t+1}} \right) < 0.$$

Hence the planner's problem is equivalent to the simplified version in which (4.3) is maximized in C_{0t}, C_{1t}, L_{0t}, L_{1t}, E_t, H_t, K_t $(t = 0, 1, \ldots)$, and φ_0 subject to (4.1), (4.2), and (4.53).

We first study those first-order conditions of the simplified planner's problem which are associated with variables which do not enter the implementability constraint (4.53) or which drop out when making particular assumptions. The optimization with respect to those variables is not affected by (4.53) and should therefore remain undistorted.

Proposition 4.5.
Assume altruistic behaviour. Then it is second best not to distort education for all generations except the first.

Proposition 4.6.
Assume altruistic behaviour and the utility function to be weakly separable between consumption and non-leisure and homothetic in consumption. Then it is second best not to distort the accumulation of capital for all generations except the first.

Proposition 4.7.
Assume altruistic behaviour and U *to satisfy* (4.16) *or* (4.17). *Then it is second best to tax qualified and nonqualified labour uniformly. This holds for all generations except the first.*

The proof of Proposition 4.5 is rather straightforward. Just note that the variables E_t, H_t, K_t $(t > 0)$ do not enter the implementability constraint. Taking partial derivatives of the Lagrange function with respect to these variables and substituting for the Lagrange multipliers μ_t, α_t yields the efficiency condition (4.9) for $t > 0$. The proof of Proposition 4.6 parallels the one of Proposition 4.1 and is therefore skipped. The proof of Proposition 4.7 is as follows. Set

$$W_t \equiv U_t + \tilde{\lambda} \sum_{i=0}^{1} C_{it} U_{C_i t}.$$

If V is homogeneous of degree $d \neq 0$, then $W_{L_i t} = (1 + \tilde{\lambda} d) U_{L_i t}$ $(i = 0, 1)$. Hence the social and the private marginal rates of intertemporal substitution in non-leisure are equal,

$$\frac{W_{L_1 t}}{W_{L_0 t}} = \frac{U_{L_1 t}}{U_{L_0 t}} \underset{(43)}{=} \frac{\omega_{1t} H_t U_{C_1 t}}{\omega_{0t} H_{t-1} U_{C_0 t}}. \tag{4.55}$$

The equation (4.55) is equally obtained if V is homogeneous of degree zero in the sense of (4.15) with $D \equiv 0$. Taking partial derivatives of the Lagrange function with respect to K_t, L_{0t}, L_{1t}, yields (4.31) and $W_{L_0 t} = -\alpha_t F_{L_0 t} H_{t-1}$, $W_{L_1 t} = -\alpha_{t+1} \beta F_{L_1 t+1} H_t$ $(t > 0)$. Therefore, (4.55) $\Leftrightarrow$

$$\begin{aligned} \frac{\alpha_{t+1} \beta F_{L_1 t+1} H_t}{\alpha_t F_{L_0 t} H_{t-1}} &= \frac{\omega_{1t} H_t U_{C_1 t}}{\omega_{0t} H_{t-1} U_{C_0 t}} \\ \Leftrightarrow \frac{F_{L_1 t+1}}{F_{L_0 t}} &\underset{(4.31)}{=} \left[F_{Kt+1} + 1 - \delta_K\right] \frac{\omega_{1t}}{\omega_{0t}} \frac{U_{C_1 t}}{U_{C_0 t}}. \end{aligned}$$

Define tax rates τ_{it} by setting $1 - \tau_{1t} \equiv \omega_{1t} / F_{L_1 t+1}$, $1 - \tau_{0t} \equiv \omega_{0t} / F_{L_0 t}$. Hence, (4.55) $\Leftrightarrow$

$$\frac{1 - \tau_{0t}}{1 - \tau_{1t}} = \left[F_{Kt+1} + 1 - \delta_K\right] \frac{U_{C_1 t}}{U_{C_0 t}}. \tag{4.56}$$

The utility functions assumed to hold for Proposition 4.7 are weakly separable between consumption and non-leisure and homothetic in consumption. Hence Proposition 4.6 applies and it is second best not to distort saving. As a result, the right-hand of (4.56) equals one and labour tax rates are independent of age.

Proposition 4.6 is just what one would expect in view of the literature. Proposition 4.7 is less obvious, and it even allows us to qualify the main result of Erosa and Gervais (2002) stating that it is generally optimal to differentiate labour taxes across the individual life cycle. The intuitive explanation for this result is that labour supplied in the second life period differs from labour supplied in the first period. While Proposition 4.4 confirms the result of Erosa and Gervais (2002) on assuming selfish individuals, Proposition 4.7 does not. Obviously, in the present framework altruism removes the need to employ age-dependent labour taxes for descendent generations. Age-dependent labour taxes would then be used only as a correcting device if it were second best to distort saving. This becomes clearer when considering utility functions which are additive separable between consumption and non-leisure, $U = V(C_0, C_1) - D(L_0, L_1)$. In this case (4.56) would equally hold but the right-hand side of (4.56) would only equal one in the optimum if V were homothetic. This is a noteworthy qualification of Erosa and Gervais (2002). Above, it is derived from the equality of the social and private marginal rates of intertemporal substitution in non-leisure, (4.55). For this equality to hold we have to assume not only altruism, but also a sufficiently rich set of policy instruments. In particular, the planner must be able to choose ω_{it} independently of φ_t. In other words, the planner must be able to optimize the taxation of labour separately from the subsidization of education.

The generality of Proposition 4.5 is striking. The proposition holds for arbitrary utility and human capital investment functions, and it does not assume balanced growth. This generality is not only remarkable as such. It allows one to qualify related results by Jones, Manuelli, and Rossi (1997) and Atkeson, Chari, and Kehoe (1999). Those results suggest that human capital does not differ that strongly from physical capital to justify different tax policies. In fact, both kinds of accumulation processes should remain undistorted along balanced growth. By contrast, Propositions 4.5 and 4.6 highlight strong differences. The case for leaving education undistorted is much stronger than the case of undistorted saving.

The explanation for such deviating results is as follows. Jones, Manuelli, and Rossi (1997) and Atkeson, Chari, and Kehoe (1999) derive their results in the standard Ramsey model with an infinite planning horizon. At first sight, the proofs show strong parallels to the one given above. In each case, the planner is assumed to maximize the sum of discounted utilities subject to a resource constraint, an accumulation constraint of human capital, an implementability constraint, and an Euler equation for human capital accumulation. Equally, the proof runs in each case by first solving the problem obtained when discarding the Euler equation and by then showing that the Euler equation is fulfilled. Differences come in when solving the relaxed problem and when arguing why the Euler equation is fulfilled. In Jones, Manuelli, and Rossi (1997) and in Atkeson, Chari, and Kehoe (1999, Prop. 5) the Euler equation is one which is not distorted by prices. Hence only undistorted allocations solving the relaxed problem are able to solve the non-relaxed problem. The only solutions fulfilling such requirements are allocations that converge to a balanced growth path and that are obtained when all taxes are zero along such path. The present analysis is much less constraining. The solutions of the relaxed problem need not be undistorted allocations. The Euler equation is not really constraining the planner because it contains the free policy variable. Once more, this demonstrates the importance of modelling two different costs of education. Modelling the cost of foregone earnings only but not the cost of tuition - as Jones, Manuelli, and Rossi (1997) and Atkeson, Chari, and Kehoe (1999) do - has the effect that an instrument is missing allowing the planner to control education independently of labour.

The generality of Proposition 4.5 strongly reminds one of the Production Efficiency Theorem of Diamond and Mirrlees (1971). According to this theorem the allocation of intermediate goods should not be distorted in second best if no lump-sum income accrues to the private sector. This is just what holds in the present model. Investment in human capital is modelled as an intermediate good in the sense that it does not affect the implementability constraint (4.53) for $t > 0$. Furthermore, the only lump-sum income modelled is income earned by the parent generation living in period 0. On setting $\pi_0 = 1$, this income equals $F_{K0}K_{-1} + F_{L_1 0}L_{1-1}H_{-1} + (1 - \delta_K)K_{-1}$. It does not show up in the dynasty's budget constraint (4.47). It must therefore be income accruing to the government budget. The Production Efficiency Theorem is applicable, and Proposition 4.5 can be considered to be a corollary.

The recommendation not to distort education is not easily translated into explicit tax and subsidy rates. The reason is that private incentives are affected by a whole set of tax and subsidy rates, which all must be optimally set. Just inspect the altruist's first-order condition (4.51) determining the optimal amount of human capital. After substituting for the Lagrange multipliers one obtains

$$\omega_{1t}L_{1t} + \omega_{0t+1}L_{0t+1} - \left(\omega_{0t+1} + \varphi_{t+1}\right)E_{t+1} \\ = R_{t+1}\frac{\varphi_t + \omega_{0t}}{G'_t} - G_{t+1}\frac{\varphi_{t+1} + \omega_{0t+1}}{G'_{t+1}}. \quad (4.57)$$

This condition reveals that the altruist's incentive to invest in human capital is not only affected by taxes on own labour income and the subsidy paid to the own cost of tuition. It is additionally affected by the tax on savings, by the next generation's tax on nonqualified labour, and finally by the subsidy paid to the next generation's cost of tuition. More can be said only after making specific assumptions. Just for the sake of illustration, assume U to satisfy (4.16) or (4.17). Hence Propositions 4.6 and 4.7 apply, and it is optimal not to tax saving, $R_{t+1} = F_{Kt+1} + 1 - \delta_K$, and to tax labour independently of age, $1 - \tau_t \equiv \omega_{1t}/F_{L_1t+1} = \omega_{0t}/F_{L_0t}$ $(t > 0)$. Only if optimal wage taxes do neither differentiate across generations, $\tau_t = \tau$, can one infer that it is compatible with efficiency for the cost of tuition to be subsidized at the same rate as labour income is taxed, $\varphi = (1 - \tau)f$. This follows immediately from comparing (4.57) with (4.9). If the mentioned assumptions do not hold, it is difficult to make definite statements about the efficient structural relationship between labour tax rates and education subsidy rates.

The government has to finance the exogenous cash flow of government expenditures A_t $(t > 0)$. If the amount of pure profit earned by the government is insufficient, distortionary taxes have to be employed to balance the budget. In this case, the implementability constraint (4.53) is binding, and it cannot be ruled out that it is efficient to distort the choice of education of generation 0. This raises the question of how to design optimal human capital policy for generation 0. As we are going to learn, the answer comes close to what has been shown to be efficient in the world of selfish individuals. More precisely, generation 0's education should not be distorted if the human capital investment function is isoelastic. If however the private marginal cost of human capital is positive, education should be positively distorted relative to the first best. To show this we maximize (4.3) subject to (4.1), (4.2), (4.53), and (4.54). Taking partial derivatives of the Lagrange function yields the following results after some tedious but straightforward manipulations have been made:

$$\frac{\partial}{\partial \varphi_0} : \gamma_0 = -\tilde{\lambda}(1-\eta_0), \tag{4.58}$$

$$\frac{\partial}{\partial \varphi_1} : \gamma_1 = \gamma_0(1-\eta_1), \tag{4.59}$$

$$\begin{aligned} \frac{\partial}{\partial E_0} : \frac{\mu_0}{\alpha_0} \underset{(4.58)}{=} & \frac{f + F_{L_0 0}}{G'_0} \\ & - \frac{\tilde{\lambda}}{\alpha_0} u_{C_0 0} \frac{\varphi_0 + \omega_{00}}{G'_0} \left[1 + \frac{E_0 G''_0}{G'_0} \right], \end{aligned} \tag{4.60}$$

$$\begin{aligned} \frac{\partial}{\partial E_1} : \frac{\mu_1}{\alpha_1} \underset{(4.58),(4.59)}{=} & \frac{f + F_{L_0 1}}{G'_1} \\ & - \frac{\tilde{\lambda}}{\alpha_1} u_{C_0 1} \frac{\varphi_1 + \omega_{01}}{G'_1} \left[1 - \frac{E_0 G'_0}{G_0} \right] \left[1 + \frac{E_1 G''_1}{G'_1} \right]. \end{aligned} \tag{4.61}$$

The first-order condition with respect to K_0 is the same as (4.31) for $t = 0$. By making use of (4.58)–(4.61) and (4.31) for $t = 0$ we end up with

$$\frac{\partial}{\partial H_0} : \Delta_0 = \frac{f + F_{L_0 0}}{G_0'}\left(F_{K1} + 1 - \delta_K\right) - \frac{f + F_{L_0 1}}{G_1'} G_1$$
$$- F_{L_1 1} L_{10} - \left[F_{L_0 1} \cdot \left(L_{01} - E_1\right) - f E_1\right], \quad (4.62)$$

where

$$\Delta_0 \equiv \frac{\tilde{\lambda}}{\alpha_0} U_{C_0 0} \cdot PMC_0^{HC} \cdot \left(F_{K1} + 1 - \delta_K\right)$$
$$- \frac{\tilde{\lambda}(1 - \eta_o)}{\alpha_1} U_{C_0 1} \cdot PMC_1^{HC} \cdot G_1 \quad (4.63)$$

and $(t = 0, 1)$

$$PMC_t^{HC} \equiv - \frac{L_{1t} U_{L_1 t}}{U_{C_0 t}} \frac{\eta_t^{'}}{G_t' H_{t-1}} \underset{(4.26)}{=} \frac{\varphi_t + \omega_{0t}}{G_t'} \frac{E_t \eta_t'}{\eta_t}$$
$$= \frac{\varphi_t + \omega_{0t}}{G_t'} \left[1 - \frac{E_t G_t'}{G_t} + \frac{E_t G_t''}{G_t'}\right]. \quad (4.64)$$

The variables Δ_0 and PMC_t^{HC} are defined so that the parallels with (4.38) and (4.27) show up. As PMC_t^{HC} vanishes for isoelastic $G(E_t)$, we obtain

Proposition 4.8.
Assume altruistic behaviour and the human capital investment function G to be isoelastic. Then it is second best not to distort the first generation's educational choice.

Proposition 4.8 is just the altruistic analogue to Proposition 4.2. It is a result that one could easily conjecture. Altruism goes beyond selfishness in internalizing efficiency effects. If it is second best not to distort education given that G is isoelastic and behaviour selfish, then it should all the more be second best not to distort education given that G is isoelastic and behaviour altruistic.

Things are less straightforward if the private marginal cost of human capital is positive. Without making further assumptions, it is difficult to sign Δ_0. However, we are able to derive a direct analogue to Proposition 4.3. More precisely, Δ_0 can be shown to be positive if the growth path is balanced and if utility satisfies (4.16) or (4.17). The assumption of balanced growth has the effect of neutralizing the impact of initialization.

The proof is only sketched. First note that $\omega_{0t} = \omega_0$ follows from (4.48). In a second step G^{dt} is shown to be a factor that cancels out of the constraint (4.54), so that φ_t and φ_{t+1} are the only remaining variables in (4.54) carrying an index t. The equation can then be used to solve for$\varphi_t = \varphi_{t+1} \equiv \varphi$. This is a feasible procedure, as the coefficient of φ does not vanish. Just note that after dividing through by G^{dt} the coefficient equals

$$\begin{aligned} &\beta G^d U_{C_0 0}\left[EH - \frac{GH}{G'}\right] + U_{C_0 0}\frac{GH}{G'} \\ &\qquad = U_{C_0 0} EH \left[\beta G^d + \frac{1}{\eta}\left(1 - \beta G^d\right)\right]. \end{aligned}$$

The condition of transversality, $\beta G^d < 1$, implies that the coefficient is positive. Plugging φ into (4.64) yields $PMC_t^{HC} = PMC^{HC}$. Assume $PMC^{HC} > 0$ and prove $\Delta_0 = \Delta > 0$ by inspecting (4.63) and by noting

$$\begin{aligned} &\frac{\tilde{\lambda}}{\alpha_0} U_{C_0 0} \cdot \left(F_{K1} + 1 - \delta_K\right) > \frac{\tilde{\lambda}(1-\eta)}{\alpha_1} U_{C_0 1} \cdot G \\ \underset{(4.31)}{\Leftrightarrow}\ & U_{C_0 0} > \beta(1-\eta) G^{d-1} U_{C_0 0} \cdot G \\ \Leftrightarrow\ & 1 > (1-\eta) \cdot \beta G^d. \end{aligned}$$

The last inequality follows from $\eta < 1$ and, once more, from the condition of transversality.

Proposition 4.9.

Assume altruistic behaviour, and U *to satisfy* (4.16) *or* (4.17). *At balanced growth it is second best to subsidize the first generation's educational choice relative to the first best if the private marginal cost of human capital,* PMC^{HC}, *is positive.*

It would be nice if one could similarly characterize second-best policy with regard to the first generation's choice of labour and saving. However, analogues to Propositions 1 and 4 seem not to hold. In particular, it seems that the first generation's saving decision is systematically distorted. The reason is the factor $U_{C_0 0}$ entering the right-hand side of (4.53). This factor implies a lack of symmetry when taking partial derivatives of B with respect to C_{i0} $(i = 0, 1)$. As a result it is second best to distort saving.

The parallelism between Propositions 4.9 and 4.3 allows us to tell a unifying story for selfish and altruistic individuals. Altruism well reduces the need to subsidize education relative to *laissez-faire*. Altruism also implies that the second-best tax policy for descendent generations is more like the first-best policy. The accumulation of human capital should remain undistorted, and – if utility functions are well selected – labour taxes need not be differentiated across the individual life cycle. The short-run policy recommendations for altruism, however, parallel the long-run recommendations for selfishness. Labour has to be taxed, and – given that the elasticity of the human capital investment function is strictly increasing – education should be subsidized relative to the first best. Whether saving should be taxed is not a matter of selfishness or altruism. With regard to descendent generations it primarily depends on assumptions made with regard to the marginal rate of intertemporal substitution in consumption. In any case, the recommendation not to distort education is better founded in dynamic analysis than the recommendation not to distort saving.

4.7 SUMMARY

The accumulation of human capital may suffer from all sorts of potential inefficiencies. Most of them have simply been assumed away in the present study. Such a procedure is, no doubt, debatable. Critical is the ignoring of possible causes of capital market or policy failure. Even more critical is the ignoring of individual heterogeneity and informational asymmetry. Still, the procedure is defended with the objective of studying efficient taxation in Ramsey's tradition. More precisely, this chapter aims at bridging the gap that separates the two strands of Ramsey tax analyses which exist for the finite and the infinite planning horizon. Our knowledge of efficient human capital policy in Ramsey's tradition is largely shaped by incompatible results derived for the different horizons. The results derived for the infinite horizon suggest that education should not be distorted in the long run, just as saving should not be distorted in the long run. Hence it seems as if efficient policy does not differentiate between human and nonhuman capital. By way of contrast, the results in finite horizon strongly suggest differentiated policies. Whether education should be distorted or not appears to depend primarily on how education affects the individual's earning potential. More precisely, only if the earnings function is weakly separable in qualified labour supply and education and if the elasticity with respect to the latter is constant, should the choice of education be not distorted by second-best policy (Jacobs and Bovenberg, 2010b). By way of contrast, the question of whether saving should be distorted or not primarily has to be answered with regard to the taxpayer's preferences. More precisely, saving should not be taxed if the taxpayer's utility is weakly separable between consumption and labour/non-leisure and homothetic in consumption (Atkinson and Stiglitz, 1972).

The model filling the gap between finite and infinite Ramsey tax analyses is one with overlapping generations. The present chapter studies second-best policy for education, saving, and labour in such an overlapping-generations model with endogenous growth. There have been earlier attempts to do the same. In view of the present study, two attempts deserve to be cited more than others. These are by Atkeson, Chari, and Kehoe (1999) on one side and by Wigger (2002, Sec. 3.4) and Docquier, Paddison, and Pestieau (2007) on the other side. The most conspicuous differences to the present study are the following ones. The focus of the present study is on human capital accumulation, while the focus of Atkeson, Chari, and Kehoe (1999) is on nonhuman capital. Their paper contains extensions to both endogenous education and overlapping generations, but it fails to integrate the two. The work of Wigger (2002) and Docquier, Paddison, and Pestieau (2007) does integrate them. However, it does not allow for endogenous labour supply and second-best taxation. The authors assume the availability of non-distorting tax instruments, which the present study does not. In a sense, the present chapter starts where Atkeson, Chari, and Kehoe (1999)and where Wigger (2002) and Docquier, Paddison, and Pestieau (2007) stop. It goes beyond Atkeson, Chari, and Kehoe (1999) by integrating endogenous education and overlapping generations, and it goes beyond Wigger (2002) and Docquier, Paddison, and Pestieau (2007) by endogenizing labour supply and by doing second-best tax analysis.

The present chapter studies two possible reasons for allocational inefficiency. One is the non-availability of non-distorting tax instruments. The other is individual selfishness. Taxpayers are assumed to externalize the positive effect that their human capital investments have on the productivity of descendent generations. As stressed by Wigger (2002) and by Docquier, Paddison, and Pestieau (2007), selfishness is the source of an intergenerational externality. It gives reason to subsidize education relative to *laissez-faire*. Such subsidization, however, requires government revenues. In the framework studied by Wigger (2002) and by Docquier, Paddison, and Pestieau (2007) it is efficient to subsidize education up to the first-best level where marginal social costs equal marginal social returns. The result assumes the availability of non-distortionary tax instruments. The key assumption of the present study, however, is that no tax instruments are available that would allow the government to raise the revenue needed to subsidize education without creating distortions. As it turns out, it is still second best not to distort education if only the human capital investment function is isoelastic. This result can be considered to be the dynamic version of the education efficiency proposition known from static Ramsey analysis.

It is, however, argued that an isoelastic investment function has the unappealing implication that all human capital accumulated by past generations melts down to zero if only one generation stops investing. If, by way of contrast, human capital depreciates just by some fraction and if the investment function's elasticity is strictly increasing, then investment incentives should overshoot the first best at balanced growth. In other words, it is efficient in the long run to combine positive tax wedges in the labour market with an effective subsidy wedge for education. The need to subsidize is shown to increase in (i) the private marginal cost of human capital, (ii) the cost resulting from the non-availability of lump-sum taxes, and (iii) the growth gap. Furthermore, it turns out to be efficient to tax labour such that qualified labour is less distorted than nonqualified labour.

If taxpayers are altruists with respect to descendent generations, one clear reason for government intervention does not apply. The effect that education has on descendent generations' productivity is internalized by altruists. The only remaining inefficiency modelled in this chapter is caused by the need to employ distortionary taxes for financing government expenditures. As it turns out, all generations except the first one should still be given non-distorted incentives for accumulating human. Furthermore, if utility functions are well selected, saving should not be distorted and labour should be taxed uniformly across the individual life cycle. These results contrast with results derived by Jones, Manuelli, and Rossi (1997), Atkeson, Chari, and Kehoe (1999), and Erosa and Gervais (2002). While Jones, Manuelli, and Rossi (1997) and Atkeson, Chari, and Kehoe (1999, Prop. 5) suggest that human capital does not differ that strongly from physical capital to justify different dynamic tax policies, the present analysis highlights strong differences. If altruism is assumed, the case for leaving education undistorted is much stronger than the case of undistorted saving. Our results equally allow us to qualify the main result of Erosa and Gervais (2002), who stress the need to employ age-dependent labour taxes in second best. In the present framework, however, altruism has the effect of implying equality of the social and private marginal rates of intertemporal substitution in non-leisure. The optimality of uniform labour taxation is an immediate though intriguing corollary to this equality.

The results on non-distortionary taxation do not require removing all distortions. On the contrary, the labour supply of descendent generations will be distorted if the government has to finance exogenous government expenditures by relying on distortionary instruments. Nor do the results on non-distortionary taxation extend to the dynasty's first generation, indexed by zero in the present chapter. A more precise characterization of optimal policy for generation 0 is difficult, as the specific features not only depend on the shape of the human capital investment function but also on initial values of key variables. As in the case with selfish individuals, it is efficient not to distort education if the investment function is isoelastic in education. If, however, the elasticity is strictly increasing and if the impact of initialization is suppressed by assuming balanced growth, it is second best to subsidize education relative to the first best. The reason is the same as the one given before in the scenario with selfish individuals. A strictly increasing elasticity of the investment function has the effect that it is second best to subsidize education in static analysis, and this effect extends to the dynamic framework. At balanced growth the need to subsidize increases in the derivative of the investment function's elasticity, and it is the stronger, the more binding the non-availability of lump-sum taxes is and the more deficient growth is.

The unifying bottom line for selfish and altruistic individuals is as follows. Altruism well reduces the need to subsidize education relative to *laissez faire*, and altruism also implies that descendent generations should be given non-distorted incentives for accumulating human capital. The short-run policy recommendations for altruism, however, agree with the long-run recommendations for selfishness. Labour has to be taxed, and – given that the elasticity of the human capital investment function is strictly increasing – education should be subsidized relative to the first best. Whether saving should be taxed is not a matter of selfishness or altruism. It primarily depends on assumptions made with regard to the marginal rate of intertemporal substitution in consumption.

5

EFFICIENT SUBSIDIZATION OF HUMAN CAPITAL ACCUMULATION WITH OVERLAPPING GENERATIONS AND ENDOGENOUS GROWTH: A NUMERICAL EXAMPLE

5.1 INTRODUCTION

Chapter 4 studies second best policies for education, saving, and labor in a two-period OLG model in which endogenous growth results from human capital accumulation. It identifies two effects that serve to justify the subsidization of education if individuals are selfish: (i) Internalize the external intergenerational effect and (ii) account for distortionary labor taxation. Propositions 4.2 and 4.3 are the pivotal results:

Proposition 4.2. *Assume selfish behaviour. It is second best not to distort education if the human capital investment function* $G(E)$ *is isoelastic.*

Proposition 4.3. *Assume selfish behaviour, and* U *to satisfy balanced growth properties. At balanced growth it is second best to subsidize education relative to the first best if the private marginal cost of human capital,* PMC^{HC}*, is positive. The strength of positive distortion increases in (i) the private marginal cost of human capital, (ii) the growth gap, and (iii) the cost resulting from the non-availability of lump-sum taxes.*

This chapter numerically analyzes these propositions and studies the importance of the intergenerational external effect and of the distortionary labor taxation effect as reasons to justify the subsidization of education. As it turns out, the case for subsidizing education to account for distortionary labor taxation is rather weak. The still dominant justification for subsidizing education is to internalize intergenerational externalities. This result is robust and holds for a wide range of parameter values.

5.2 Restrictions on the Utility and Production Functions

To have a model that exhibits balanced growth, certain restrictions on the utility function and on the production functions must be imposed.

5.2.1 *Utility Function*

Section 4.3 derives the balanced growth properties that the utility function must satisfy. The individual's utility function used in the present analysis is as follows:

$$
\begin{aligned}
U_t &= U(C_{0t}, C_{1t}, L_{0t}, E_t, L_{1t}) \\
&= \begin{cases} \frac{(C_{0t}(1-L_{0t}-E_t)^{\gamma_0})^{1-\psi}}{1-\psi} + \rho \frac{(C_{1t}(1-L_{1t})^{\gamma_1})^{1-\psi}}{1-\psi}, \\ \hfill 0 < \psi \neq 1. \\ \ln C_{0t} + \gamma \ln(1 - L_{0t} - E_t) \\ \qquad + \rho(\ln C_{1t} + \gamma \ln(1 - L_{1t})), \quad \psi = 1. \end{cases}
\end{aligned} \tag{5.1}
$$

ρ is the individual's discount factor, which may equal the planner's discount factor β. γ_0 and γ_1 are taste parameters for leisure. The individual's time endowment is normalized to one, that is, the time constraints read $1 = L_0 + E + \text{leisure}$ and $1 = L_1 + \text{leisure}$. $1/\psi$ is the individual's intertemporal elasticity of substitution in consumption.

(5.1) is a special case of conditions (4.16) and (4.17). It is straightforward to verify the balanced growth path properties as explained in section 4.3.

The utility function used here is time-separable. Without loss of generality it explicitly includes the time E spent on education when young. The whole analysis in chapter 4 can be done using the present specification of the utility function without affecting any results. It is used here to facilitate the calibration of how the individual spends his discretionary time on leisure, working and education.

5.2.2 *Production Function*

The firm's technology is given by the following Cobb-Douglas production function:

$$\begin{aligned} F_t = F\big(K_{t-1}, L_{0t}H_{t-1}, L_{1t-1}H_{t-1}\big) = \\ \mathcal{A}K_{t-1}^{\alpha}\big(L_{0t}H_{t-1} + L_{1t-1}H_{t-1}\big)^{1-\alpha}. \end{aligned}$$

$\mathcal{A}$ is the total factor productivity, which is a scaling constant that can be arbitrarily set. Following Auerbach, Kotlikoff, and Skinner (1983, p. 87-88) and Erosa and Gervais (2002), the labor input is the sum of effective units of labor supply of the young and old individual. This specification implies that a young and old individual's marginal product of labor is equal, $F_{L_0 t} = F_{L_1 t} = F_{Lt}$.

As one can see, along the balanced growth path at which the stock of physical capital increases at the same rate as the stock of human capital, the marginal products of capital and labor stay constant.

5.2.3 *Human Capital Investment Function*

The accumulation of human capital is described by the following law of motion:

$$G(E_t)H_{t-1} = H_t.$$

The function expressing the stock H_t of human capital in the next period features constant returns to scale with respect to the current stock H_{t-1} of human capital. This specification is another requirement for the economy to grow along a balanced-growth path (Lucas (1988), Caballe and Santos (1993)).

The human capital investment function $G(E)$ reads

$$G(E) = DE^{\tilde{\eta}} + 1 - \delta_H, \ \delta_H \leqslant 1,$$

where $D > 0$ is a shift parameter and δ_H is the rate at which human capital depreciates. Let $\eta(E)$ denote the elasticity of the function $G(E)$. Then:

$$\eta(E) = \left[1 - \frac{1-\delta_H}{G(E)}\right]\tilde{\eta}.$$

As long as the stock of human capital does not fully depreciate, the elasticity η is increasing in E.

To numerically analyze the model and to solve for the balanced-growth path allocation, it is necessary to transform the model with growth into one without growth. To do so, divide all growing variables by the current stock H_{t-1} of human capital:

$$\hat{K}_{t-1} = \frac{K_{t-1}}{H_{t-1}},\ \hat{C}_{0t} = \frac{C_{0t}}{H_{t-1}},\ \hat{C}_{1t} = \frac{C_{1t}}{H_{t-1}},\ \hat{A}_t = \frac{A_t}{H_{t-1}}. \tag{5.2}$$

This procedure ensures that all hat variables are constant along the balanced growth path. The remaining variables L_{0t}, L_{1t} and E_t remain constant. The normalized conditions of the selfish individual's problem, evaluated along the balanced growth path, are as follows:

$$\begin{aligned}
&(1-\tau_0)(1-\alpha)\mathcal{A}\hat{K}^{\alpha}(L_0+L_1)^{-\alpha}L_0 \\
&\quad = \hat{C}_0 + (1-\tau^H)fE + (\hat{K}+\hat{B})G(E), &(5.3)\\
&(1-\tau_1)(1-\alpha)\mathcal{A}\hat{K}^{\alpha}(L_0+L_1)^{-\alpha}L_1 \\
&\quad + \left[(1-\tau^K)\left(\alpha\mathcal{A}\hat{K}^{\alpha-1}(L_0+L_1)^{1-\alpha}\right)+1-\delta_K\right]\times \\
&\quad (\hat{K}+\hat{B}) = \hat{C}_1/G(E), &(5.4)\\
&(1-\tau^K)\alpha\mathcal{A}\hat{K}^{\alpha-1}(L_0+L_1)^{1-\alpha}+1-\delta_K \\
&\quad = \rho^{-1}\left(\frac{\hat{C}_1}{\hat{C}_0}\right)^{\psi}\left(\frac{(1-L_0-E)^{\gamma_0}}{(1-L_1)^{\gamma_1}}\right)^{1-\psi}, &(5.5)\\
&(1-\tau_0)(1-\alpha)\mathcal{A}\hat{K}^{\alpha}(L_0+L_1)^{-\alpha} = \gamma_0\frac{\hat{C}_0}{1-L_0-E}, &(5.6)\\
&(1-\tau_1)(1-\alpha)\mathcal{A}\hat{K}^{\alpha}(L_0+L_1)^{-\alpha} = \gamma_1\frac{\hat{C}_1/G(E)}{1-L_1}, &(5.7)\\
&\frac{(1-\tau_1)(1-\alpha)\mathcal{A}\hat{K}^{\alpha}(L_0+L_1)^{-\alpha}L_1 D\tilde{\eta}E^{\tilde{\eta}-1}}{(1-\tau^K)\alpha\mathcal{A}\hat{K}^{\alpha-1}(L_0+L_1)^{1-\alpha}+1-\delta_K}
\end{aligned}$$

$$= (1-\tau^H)f + (1-\tau_0)(1-\alpha)\mathcal{A}\hat{K}^{\alpha}(L_0+L_1)^{-\alpha}, \quad (5.8)$$

$$\mathcal{A}\hat{K}^{\alpha}\big((L_0+L_1)\big)^{1-\alpha} + (1-\delta_K)\hat{K} =$$
$$\hat{C}_0 + \hat{C}_1/G(E) + fE + G(E)\hat{K} + \hat{A}. \quad (5.9)$$

The equations are specializations of the equations of the general model in section 4.4. (5.3) and (5.4) are the individual's budget constraints when young and old. (5.5) - (5.8) are the individual's marginal conditions. (5.9) is the resource constraint. The government's budget constraint is satisfied because of Walras's law.

The parameters of the model are chosen as follows. The utility function is logarithmic, i.e., the intertemporal elasticity of substitution is equal to one, $\psi = 1$, which is a value frequently used in the literature, for instance in a related analysis by Chari, Christiano, and Kehoe (1994). The annual time preference rate is ρ is 0.96, which corresponds to a quarterly value of 0.99 ($= 0.96^{1/4}$).

Total factor productivity is equal to 8. Varying the parameter $\mathcal{A}$ only has level effects and does not affect the qualitative results. The capital's share α is set equal to 0.29, which is a value derived by Gomme and Rupert (2007). It is assumed that the stock of physical capital is completely depreciated after 30 years (De la Croix and Michel, 2002, p. 338).

Tax rates on labor and capital income are uniformly set equal to 30%. Several methodologies have been put forward to estimate average tax rates on capital and labor income. See Gomme and Rupert (2007) for a discussion. Tax rates ranging from a low 17% to a high 30% on labor income and tax rates on capital income between 27% and 50% are reported by Mendoza, Razin, and Tesar (1994). Gomme and

Rupert (2007) use 22% and 29% as tax rates on capital and labor income. And finally, Bouzahzah, De la Croix, and Docquier (2002) set the tax rate on labor income equal to 29%. The present choice of tax rates therefore is not too far away from what the literature suggests.

The level of government consumption $\hat{A}$ is set to equal 40% of output:

$$\hat{A} = 0.4 \times \mathcal{A}\hat{K}^{\alpha}\big((L_0 + L_1)\big)^{1-\alpha}. \tag{5.10}$$

This is a high number compared to a low 20%, which is often used in the literature (Lucas (1990), Chari, Christiano, and Kehoe (1994)). The perhaps unusual choice may be justified as follows. As the theory suggests, it is optimal to subsidize education to account for distortionary labor taxation. If the need for tax revenue is high, the negative effects of distortionary taxation can be expected to be large. The results characterizing the need to subsidize education, that will be presented and discussed below, therefore do not underestimate the negative effect of distortionary taxation. If at a high level of government consumption the case for subsidizing education is weak, then this is also true for lower levels.

(5.10) implies that the level of government consumption grows at the common growth rate. This means that the *level* of government consumption is not fixed but only the *share*. If one fixes the level, government consumption drops to a negligible fraction of output as the economy grows along the balanced growth path. To avoid this unrealistic property, it is necessary to extend the setup of the model such that in equilibrium the share of government expenditures remains constant along the balanced growth path. Section 5.4 provides the details.

The choice of the two leisure parameters γ_0 and γ_1 ensures that total non-leisure time is not larger than 0.5. More specifically, when young the individual spends a fraction of about 46% on working and the remainder of about 4% on education. In a related analysis, Bouzahzah, De la Croix, and Docquier (2002) numerically solve a six-periods-OLG model with endogenous labor, saving and education. They use data from time use surveys to calibrate their model. If I use their data and compute it to a two-period model, I end up with the numbers 13/30 and 1/15 for the fractions of time spent on working when young and old, and on education.

One could also use data from the American Time Use Survey . First, one builds the age groups 15-44 years and 45-74 years. Second, one computes time allocated to working and education. The last step involves relating these numbers to discretionary time, which is 24h minus time spent on sleeping and personal use. I used data from the waves 2003 to 2009 and ended up with the following numbers: $L_0 = 0.2620$, $E = 0.0530$ and $L_1 = 0.2323$.

The present choice comes somewhat close to these numbers. Unfortunately, it was not possible to precisely match these numbers. As it turned out, other parameters, in particular the parameters of the human capital investment function, responded very sensitive to the choice of the time inputs, which then gave rise to meaningless results. Calibrating the two leisure parameters in a way such that total non-leisure time is not larger than 50% of discretionary time allows for the greatest flexibility. The proper calibration of time is one of the most difficult issues. Even the information provided Lucas (1990) how he has done the job is a bit blurry.

The human capital investment function is specified as follows. First, the elasticity parameter $\tilde{\eta}$ is fixed to 0.5. To my knowledge, nobody so far has tried to seriously estimate a human capital investment function that represents a period of 30 years. The choice of the parameter therefore cannot claim to match an observable characteristic of the economy. Bouzahzah, De la Croix, and Docquier (2002) use values in the range of 0.1 and 0.3 also without further justification. Lucas (1990) uses 0.8 by referring to Rosen (1976), who estimates a value of 0.76. Theory only requires the parameter to lie between zero and one. Other than that, I have no priors and choose a value that lies somewhat in-between. Given this choice and the amount of education the shift parameter D is chosen to match an annual growth rate of 1.8%, which is a value used by Bouzahzah, De la Croix, and Docquier (2002):

$$DE^{\tilde{\eta}} + 1 - \delta_H = G(E) = (1 + 0.018)^{30} \tag{5.11}$$

Finally, it is assumed that education expenses are fully subsidized. Initially, the government bears the burden of education related expenses. The fee parameter is one, $f = 1$. This choice implies that in equilibrium education expenses amount to about 1% of output. Because education expenses are multiplied by the current stock of human capital, the share of education expenses stays constant in the steady state.

Table 5.1: Initial equilibrium

		$\delta_H = 1$	$\delta_H = 0.5$	$\delta_H = 0$
Annual growth rate			0.018	
Debt/output			-0.004	
Gov. spending/output			0.400	
Edu. expenses/output		0.013	0.009	0.005
Consumption/output		0.545	0.548	0.552
Investment/output		0.042	0.042	0.043
Young labor tax rate	τ_0^L		0.300	
Old labor tax rate	τ_1^L		0.300	
Capital tax rate	τ^K		0.300	
Subsidy rate	τ^H		1	
Annual interest rate		0.089	0.088	0.088
Young labor Supply	L_0	0.448	0.463	0.478
Old labor Supply	L_1		0.500	
Education	E	0.052	0.037	0.022

Because the theory suggests that for human capital depreciation rates lower than one education should be subsidized, the initial equilibrium is calibrated to three different choices of the human capital depreciation rate. This gives a set of parameters for three initial economies that share the following characteristics: Consumption, investment and education as a fraction of output amount to about 55%, 4% and 1%, respectively. The remaining 40% is spent on government consumption. The economies grow at the annual rate of 1.8%. Initial government's assets are worth about 0.4% of output. The interest rate is about 9%. Tables 5.1 and 5.2 summarize the results of the calibration exercise.

Table 5.2: Initial choice of parameters

		$\delta_H = 1$	$\delta_H = 0.5$	$\delta_H = 0$
Time preference rate	ρ		0.960	
Inter. elast. of subst.	ψ		1.000	
TFP	$\mathcal{A}$		8.000	
Capital's share	α		0.290	
Capital dep. rate	δ_K		1.000	
Education elast.	$\tilde{\eta}$		0.500	
Fee parameter	f		1.000	
Shift parameter	D	7.513	6.283	4.784
Young leisure par.	γ_0	1.178	1.145	1.113
Old leisure par.	γ_1	0.813	0.799	0.786

5.4 ENDOGENOUS GOVERNMENT CONSUMPTION

As Jones, Manuelli, and Rossi (1993) and Chari and Kehoe (1999, p. 1714) note, government consumption as a fraction of output converges to zero as the economy grows if the *level* of government consumption is held constant. For this reason the present analysis assumes that the *share* of government consumption is constant by setting $A_t = aF_t$, where a denotes the share. This assumption ensures that the level of government consumption grows at the common growth rate, which is determined endogenously. As a result, the level of government consumption is endogenous. This section serves to show how this assumption can be rationalized.

There are several routes available to endogenize government consumption. One way, which is taken by Jones, Manuelli, and Rossi (1993), is to include government consumption as a factor in the production function. Then the limiting tax on capital income may be positive. This approach will not be pursued here because it makes the comparability of the implied results to the ones derived in chapter 4 difficult. Another way is to include government consumption in the individual's utility function as proposed by Chari and Kehoe (1999, p. 1714). In an infinite horizon model with a single individual it is then easy to see that the level of government consumption is optimally chosen in a way such that it grows at the common endogenously determined growth rate. This extension does not affect the results concerning the optimal choice of tax rates.

This idea may be directly applied to the present overlapping generations model. Suppose the function

$$V(A_t, A_{t+1}) = b\frac{A_t^{1-\psi}}{1-\psi} + \rho b\frac{A_{t+1}^{1-\psi}}{1-\psi}$$

is added to the utility function (5.1). $b > 0$ is a scaling constant. Then the Ramsey problem consists of two parts. First, as before, the planner chooses the optimal allocation given a stream of government consumption. Then he chooses the optimal sequence of government consumption, i.e., maximizes over the choice of A_t. The following first-order conditions emerge:

$$\begin{aligned}
bA_t^{-\psi} &= C_{0t}^{-\psi}(1-L_{0t}-E_t)^{\gamma_0(1-\psi)}\times \\
&\qquad \left(1+\phi(1-\psi)\left(1-\gamma_0\frac{L_{0t}+E_t}{1-L_{0t}-E_t}\right)\right) \\
\rho bA_t^{-\psi} &= \rho C_{1t-1}^{-\psi}(1-L_{1t-1})^{\gamma_1(1-\psi)}\times \\
&\qquad \left(1+\phi(1-\psi)\left(1-(1-\eta)\gamma_1\frac{L_{1t-1}}{1-L_{1t-1}}\right)\right)
\end{aligned}$$

When one divides these equations by $F_t^{-\psi}$, and evaluates them along the balanced growth path at which labor supply is constant, one ends up with the rules that describe the optimal choice of government consumption:

$$b\left(\frac{A_t}{F_t}\right)^{-\psi} = \left(\frac{C_{0t}}{F_t}\right)^{-\psi} \times \xi_1$$
$$b\left(\frac{A_t}{F_t}\right)^{-\psi} = \left(\frac{C_{1t-1}}{F_t}\right)^{-\psi} \times \xi_2$$

ξ_1 and ξ_2 capture all remaining terms. Consumption and output grow at the same rate, and the RHS are constant. Hence, the LHS must be constant, which means that government consumption as a share of output must be constant along the balanced growth path.

5.5 RAMSEY PROBLEM

The solution to the Ramsey problem is described by equations (4.29) - (4.33) in section 4.4. The normalized specializations of the equations, evaluated along the balanced growth path, are as follows:

$$(1-\alpha)\mathcal{A}\hat{K}^{\alpha}(L_0+L_1)^{-\alpha} = \gamma_0 \frac{\hat{C}_0}{1-L_0-E}\frac{\Omega_{L_0}}{\Omega_{C_0}}, \tag{5.12}$$

$$(1-\alpha)\mathcal{A}\hat{K}^{\alpha}(L_0+L_1)^{-\alpha} = \gamma_1 \frac{\hat{C}_1/G}{1-L_1}\frac{\Omega_{L_1}}{\Omega_{C_1}}, \tag{5.13}$$

$$R = \rho^{-1}\left(\frac{\hat{C}_1}{\hat{C}_0}\right)^{\psi}\left(\frac{(1-L_0-E)^{\gamma_0}}{(1-L_1)^{\gamma_1}}\right)^{1-\psi}\frac{\Omega_{C_0}}{\Omega_{C_1}} \tag{5.14}$$

$$= \beta^{-1}G(E)^{\psi}, \tag{5.15}$$

$$\begin{aligned}(1-\alpha)\mathcal{A}\hat{K}^{\alpha}(L_0+L_1)^{1-\alpha} - fE \\ = (R-G)\frac{f+(1-\alpha)\mathcal{A}\hat{K}^{\alpha}(L_0+L_1)^{-\alpha}}{D\tilde{\eta}E^{\tilde{\eta}-1}} \\ -(R-G)\phi\Omega_{C_0}^{-1}PMC^{HC},\end{aligned} \tag{5.16}$$

$$
\begin{aligned}
\mathrm{PMC}^{\mathrm{HC}} &= \rho\gamma_1 \frac{L_1}{1-L_1}\left(\frac{\hat{C}_1}{\hat{C}_0}\right)^{-\psi}\hat{C}_1/G\times \\
&\left(\frac{(1-L_1)^{\gamma_1}}{(1-L_0-E)^{\gamma_0}}\right)^{1-\psi}\frac{E}{\eta}\eta', && (5.17)\\
&\hat{C}_0^{1-\psi}(1-L_0)^{\gamma_0(1-\psi)}\left(1-\gamma_0\frac{L_0}{1-L_0-E}\right) \\
&\quad+\rho\hat{C}_1^{1-\psi}(1-L_1)^{\gamma_1(1-\psi)}\times \\
&\quad\quad\left(1-(1-\eta)\gamma_1\frac{L_1}{1-L_1}\right)G^{1-\psi}=0, && (5.18)\\
&\mathcal{A}\hat{K}^{\alpha}\big((L_0+L_1)\big)^{1-\alpha}+(1-\delta_K)\hat{K} \\
&\quad\quad=\hat{C}_0+\hat{C}_1/G+fE+G\hat{K}+\hat{A}, && (5.19)\\
&R=\alpha\mathcal{A}\hat{K}^{\alpha-1}(L_0+L_1)^{1-\alpha}+1-\delta_K, && (5.20)\\
&G=DE^{\tilde{\eta}}+1-\delta_H. && (5.21)
\end{aligned}
$$

The Ω-terms result from the differentiation of the pseudo-welfare function W, see appendix 5.A. Equations (5.12) and (5.13) are the marginal conditions for the individual's labor-leisure choices when young and old. (5.14) and (5.15) are the intra- and inter-generational Euler equations. (5.16) is the dynamic Euler equation for human capital. This equation and the term PMC defined by (5.17), which is the private marginal cost, will be further discussed below. (5.18) and (5.19) are the implementability and resource constraint, respectively. R and G are the rate of return to capital net of deprecation and the growth rate, respectively.

This system of equations will be solved for the Ramsey allocation given the parameters as derived in section 5.3. To understand the results presented in Table 5.3, it is helpful to recapitulate the theoretical analysis in chapter 4. Recall (4.36) in section 4.4, which is the dynamic Euler equation for human capital. This equation describes the optimal choice of education in the Ramsey equilibrium:

$$\begin{aligned}&\left[F_{L_1t+1}L_{1t} + F_{L_0t+1}L_{0t+1}) - fE_{t+1}\right] \\ &+ \left[\frac{f + F_{L_0t+1}}{G'_{t+1}} - \frac{\phi_{t+1}}{\alpha_{t+1}}U_{C_0t+1}\mathrm{PMC}^{HC}_{t+1}\right] G_{t+1} \\ &\quad = \left[\frac{f + F_{L_0t}}{G'_t} - \frac{\phi_t}{\alpha_t}U_{C_0t}\mathrm{PMC}^{HC}_t\right][F_{Kt+1} + 1 - \delta_K]. \quad (5.22)\end{aligned}$$

The dynamic Euler equation is the corner stone of the analysis to explore how and why education is optimally subsidized in the Ramsey equilibrium. Because of (4.26) in section 4.4,

$$-\eta_t \frac{L_{1t}U_{L_1t}}{U_{C_0t}} = \left(\varphi_t + \omega_{0t}\right)E_t H_{t-1}, \quad (5.23)$$

which is derived from the selfish individual's optimality condition (5.8) for education, PMC denotes the private marginal cost of human capital. It measures how the private cost of education changes as the stock H_t increases:

$$\begin{aligned}\mathrm{PMC}_t = -\frac{d}{dH_t}\left(\eta_t \frac{L_{1t}U_{L_1t}}{U_{C_0t}}\right) = \\ -\frac{L_{1t}U_{L_1t}}{U_{C_0t}}\frac{1}{H_t}\frac{E_t}{\eta_t}\frac{d\eta_t}{dE_t} \quad (5.24)\end{aligned}$$

Essentially, PMC is a quantity that captures how the accumulation of human capital is modeled. If the elasticity of the elasticity η is zero, which is the case if human capital fully depreciates, PMC $= 0$. Otherwise, if $d\eta_t/dE_t > 0$, PMC > 0. The variable Δ_t collects the terms related to PMC:

$$\Delta_t \equiv \frac{\phi_t}{\alpha_t}U_{C_0t}\mathrm{PMC}^{HC}_t\left(F_{Kt+1} + 1 - \delta_K\right) -$$

$$\frac{\phi_{t+1}}{\alpha_{t+1}} U_{C_0 t+1} PMC^{HC}_{t+1} G_{t+1}. \quad (5.25)$$

Then the dynamic Euler equation (5.22) can be further simplified by help of this notation:

$$\Delta_t = R_{t+1} \frac{f + F_{L_0 t}}{G'_t} - F_{L_1 t+1} L_{1t}$$
$$- F_{L_0 t+1}(L_{0t+1} + E_{t+1}) - (f + F_{L_0 t+1}) E_{t+1} \left(\frac{1}{\eta_{t+1}} - 1 \right). \quad (5.26)$$

The interpretation of (5.26) is the same as on page 111. Δ_t is a wedge between the social benefit and social cost of education. A positive wedge stands for subsidizing education relative to the first best.

Note that the terms on the RHS of (5.26) relate to different generations. Whereas the first line is the difference between the social marginal cost and the social marginal benefit of the individual born in period t, the second line is related to the individual born in period $t + 1$. Therefore, the second line of (5.26) is the marginal external effect, which is denoted MEB:

$$MEB_{t+1} = F_{L_0 t+1}(L_{0t+1} + E_{t+1}) +$$
$$(f + F_{L_0 t+1}) E_{t+1} \left(\frac{1}{\eta_{t+1}} - 1 \right). \quad (5.27)$$

Then (5.26) can be written as:

$$\Delta_t + MEB_{t+1} = R_{t+1} \frac{f + F_{L_0 t}}{G'_t} - F_{L_1 t+1} L_{1t}. \quad (5.28)$$

One can see that two quantities drive a wedge between the social marginal cost and social marginal benefit of education of the individual born in period t. Δ is the tax wedge and MEB is the intergenerational externality. Both quantities serve as a justification for the effective subsidization of education. The question now is what is the size of both quantities.

Section 4.5 explains the notion of effective subsidization of education. Effective subsidization refers to the extent to which the private rate of return exceeds the social rate of return to education from the perspective of the individual. The private rate of return is the private net present value to education that accrues to the individual over the private cost that has to be borne by him. Similarly, the social rate of return is the social net present value over the social cost. Hence, the effective rate of subsidization is

$$s = \frac{\mathrm{PRR} - \mathrm{SRR}}{\mathrm{PRR}}. \tag{5.29}$$

One can show that in the Ramsey equilibrium, evaluated along the balanced growth path, the efficient rate of subsidization includes both the tax wedge and the marginal external effect:

$$s_{\mathrm{eff}} = \frac{\Delta + \mathrm{MEB}}{R(f + w)(1 - \eta)/G'}, \tag{5.30}$$

where Δ is the tax wedge, evaluated along the balanced growth path:

$$\Delta = \phi \frac{U_{C_0}}{W_{C_0}} \times \mathrm{PMC}^{HC} \times (R - G). \tag{5.31}$$

ϕ is the Lagrange multiplier associated with the selfish individual's implementability constraint. It as a measure of the non-availability of lump-sum taxes. The fraction U_{C_0}/W_{C_0} equals one if the individual's intertemporal elasticity of substitution equals one. PMC is the private marginal cost of human capital as discussed before. The last term $R - G$ is the growth gap, which necessarily must be positive. It again becomes clear that the question of how to optimally subsidize education depends on the tax wedge Δ and the marginal external effect MEB. This closes the recapitulation of the theoretical analysis that was necessary to appreciate the following results.

Table 5.3 presents the results for the three calibrated economies. The first part of the table is structured in the same way as Table 5.1 and summarizes the characteristics of the economy in the Ramsey equilibrium. The second part presents the results related to the distortion of the education decision.

Clearly, the growth rate is significantly higher in the Ramsey equilibrium than in the initial equilibrium. It increases to about 3.4%, which can be attributed to the fact that the individual spends more time on education. Government's assets are worth 4.4% of output. Labor tax rates decrease slightly and are different for a young and old individual, see Proposition 4.4. As expected, capital income is not taxed, see Proposition 4.1. The subsidy rate increases from 100% to about 188%. To understand this optimal government policy, consider the government's budget constraint:

$$\hat{A} + \tau^H fE + RB = \tau_0^L wL_0 + \tau_1^L wL_1 + GB. \tag{5.32}$$

Note that B is negative. Then the government's income consists of labor income taxes and the net yield to the asset B which is $(R - G)B$. Thus, the gap between the rate of return to physical capital net of depreciation and the growth

Table 5.3: Results

		$\delta_H = 1$	$\delta_H = 0.5$	$\delta_H = 0$
Annual growth rate		0.036	0.034	0.031
Debt/output		−0.044	−0.044	−0.045
Gov. spend./output		0.400	0.400	0.400
Edu. exp./output		0.035	0.028	0.020
Consumption/output		0.479	0.485	0.493
Investment/output		0.087	0.087	0.087
Young labor tax rate	τ_0^L	0.258	0.258	0.253
Old labor tax rate	τ_1^L	0.211	0.233	0.252
Capital tax rate	τ^K	0.000	0.000	0.000
Subsidy rate	τ^H	1.722	1.871	2.060
Annual interest rate		0.079	0.077	0.073
Young labor supply	L_0	0.323	0.348	0.379
Old labor supply	L_1	0.659	0.646	0.632
Education	E	0.152	0.128	0.098
Lagrange multiplier	ϕ	0.184	0.182	0.177
Private marginal cost of human capital	PMC	0	0.016	0.038
Return to capital	R	9.770	9.186	8.347
Growth rate	G	2.925	2.750	2.499
Growth gap	$R - G$	6.845	6.436	5.848
Human capital dist.	Δ	0	0.019	0.039
Human capital dist.	MEB	2.129	2.327	2.608
Efficient subsidy rate	s_{eff}	1.012	0.895	0.792
Efficient subsidy rate without Δ		1.012	0.887	0.780

rate, called the growth gap, is the interest rate that government bonds earn in equilibrium. The annualized growth gap amounts to about 10%. The total yield to government bonds is about 28% of output $(= (R - G) \times B)$ and therefore does not suffice to finance government consumption, which is 40% of output. The government taxes labor income at rates of about 26% and 23% to finance this level of government consumption. As a result, total taxes amount to more than 40% of output. Then the government uses the education subsidy to transfer government revenue back to the individual.

The education policy can be described by the efficient rate of subsidization. See section 5 for the details. The efficient rate of subsidization s_{eff} accounts for two effects. One is the distortionary taxation effect Δ as defined by (5.31). The other is the marginal external effect MEB as defined by (5.27). If human capital fully depreciates, $\delta_H = 1$, $\Delta = 0$. Thus, only the marginal external effect matters and the efficient subsidy rate is positive. The distortionary taxation effect Δ is positive when the private marginal cost of education is positive which is the case when human capital does not fully depreciate. With $\delta_H = 0$ the distortionary taxation effect Δ slightly increases. The result is that the efficient rate of subsidization s_{eff} decreases from 1.012 to 0.792. But as one can see the share of s_{eff} that can be attributed to Δ is small. If one did not take it into consideration, the efficient rate of subsidization is only slightly smaller. As a result, most importantly education is effectively subsidized to internalize intergenerational effects. To additionally account for the distortionary taxation effect only slightly increases the efficient rate of subsidization. Thus, the theoretical analysis identifies a quantity that also serves to justify the subsidization of education. But as the numerical analysis shows this quantity does not play a big role.

A sensitivity analysis shows that this result holds for lower shares of government consumption, for varying elasticity parameters $\tilde{\eta}$ between zero and one, for higher values of the intertemporal elasticity of substitution and for higher fee parameters. Still the marginal external is more dominant than the distortionary taxation effect.

The distortionary taxation effect to rationalize the subsidization of education has been put forth by Trostel (1996) for the first time. He finds that subsidizing education by making the related cost tax-deductible substantially reduces the adverse effects of distortionary taxation on investing in education. More than 78% of the the cost should be tax-deductible. Beyond this rate, distortionary taxation effects cease to exist. But the analysis differs with regard to various aspects. He sets up a representative-agent model, in which an intergenerational externality does not exist. The human capital production function does not feature constant returns to scale with respect to the current stock of human capital and market goods necessary for production. This refers to the perhaps most important modelling aspect that subsidization refers to market goods, not time devoted to education as in the present context.

5.6 DISCUSSION AND CONCLUSION

Chapter 4 studies two effects that serve to justify the subsidization of education if individuals are selfish. As is well-known, a subsidy is called for to internalize intergenerational external effects of education. The analysis highlighted the negative effects distortionary labor taxation has on education. For this reason, education should be subsidized relative to the first best if the elasticity of the human capital investment function is increasing. The present chapter disentangles these two effects and evaluates their magnitudes. The result is that the case for subsidizing education to account for distortionary labor taxation is rather weak. The still dominant justification for subsidizing education is to internalize intergenerational externalities. This result is robust and holds for a wide range of parameter values.

How serious can this result considered to be? Caveats are in order because the timing of the model may be inappropriate or the human capital investment function may be ill-specified. The model assumes that life may be divided into only two periods, each of which lasts for 30 years. Thus one might object that questions of lifelong learning are not modeled. To describe this process of lifelong learning one would have to extend the model to a, say, 60-periods overlapping generation model in the fashion of Auerbach, Kotlikoff, and Skinner (1983) or Erosa and Gervais (2002). Undoubtedly, this further complicates the analysis and more questions arise.

As mentioned in the calibration section, it is difficult to imagine something like a human capital investment function that describes the accumulation process of the stock of human capital over a period of 30 years. A first step would be to empirically derive the human capital investment function. Most likely this will be done using annual data. Then the question arises how this information could be aggregated to derive a 30-years human capital investment function.

Obviously, there is still a lot to do for econometricians that could try to identify the distortionary taxation effect and assess its significance.

The partial derivatives of the pseudo-welfare function

$$W = U + \phi\Big(C_0 U_{C_0} + C_1 U_{C_1} + (L_0 + E) U_{L_0} + (1-\eta) L_1 U_{L_1}\Big)$$

are as follows:

$$\begin{aligned} W_{C_0} &= U_{C0} + \phi\big(U_{C0} + C_0 U_{C0C0} + L_0 U_{L0C0} + E_0 U_{L0C0}\big) \\ &= U_{C0}\Omega_{C_0} \end{aligned}$$

with

$$\Omega_{C_0} = \left(1 + \phi(1-\psi)\left(1 - \gamma_0 \frac{L_0 + E}{1 - L_0 - E}\right)\right);$$

$$\begin{aligned} W_{C_1} &= U_{C1} + \phi\big(C_1 U_{C1C1} + U_{C1} + (1-\eta) U_{C1L1} L_1\big) \\ &= U_{C1}\Omega_{C_1} \end{aligned}$$

with

$$\Omega_{C_1} = \left(1 + \phi(1-\psi)\left(1 - (1-\eta)\gamma_1 \frac{L_1}{1 - L_1}\right)\right);$$

$$\begin{aligned} W_{L_0} &= U_{L0} + \phi\big(C_0 U_{C0L0} + L_0 U_{L0L0} + U_{L0} + E U_{L0L0}\big) \\ &= U_{L0}\Omega_{L_0} \end{aligned}$$

with

$$\Omega_{L_0} = \left(1 + \phi\left[1 - \big(\gamma_0(1-\psi) - 1\big)\frac{L_0 + E}{1 - L_0 - E} + 1 - \psi\right]\right);$$

$$
\begin{aligned}
W_E = U_{L0} + \phi\big(C_0 U_{C0L0} + L_0 U_{L0L0} + U_{L0} + E U_{L0L0} - \eta' L_1 U_{L1}\big) \\
= W_{L_0} - \phi\eta' L_1 U_{L1};
\end{aligned}
$$

$$
\begin{aligned}
W_{L_1} = U_{L1} + \phi\big(C_1 U_{C1L1} + (1-\eta)(L_1 U_{L1L1} + U_{L1})\big) \\
= U_{L1} L_1 \Omega_{L_1}
\end{aligned}
$$

with

$$
\Omega_{L_1} = \left(1 + \phi\left[(1-\eta)\left(1 - \big(\gamma_1(1-\psi) - 1\big)\frac{L_1}{1-L_1}\right) + 1 - \psi\right]\right).
$$

6

CONCLUSION

This dissertation provided an analysis of three models of optimal taxation of human capital in Ramsey's tradition. Chapters 2 and 3 presented models with only a single individual that lived for two periods or until infinity, respectively. Chapter 4 presented two overlapping generations model with a selfish or altruistic individual that lived for two periods. The models also differed with respect to how the accumulation of human capital was modelled. The first two models had a human capital production function in which the stock of human capital did not enter as a production factor. By contrast, the human capital production function in the overlapping generations models featured constant returns to scale with respect to the stock of human capital.

The most important results of this dissertation are:

- If the single individual only lives for two periods and is endowed with a given stock of human capital, it is second best to tax capital and to subsidize human capital.

- If the single individual lives until infinity and the human capital production function does not include the current stock of human capital as a production factor, it is second best in the long run to not tax capital and to tax human capital.

- If the single individual is selfish and does not internalize his education decisions, it is second best to effectively subsidize human capital for two reasons: (i) Internalize the external intergenerational effect and (ii) account for distortionary taxation.

- A numerical analysis however shows that the justification of subsidizing human capital to account for distortionary taxation is rather weak.

The perhaps most interesting issue that further research could focus on is to extend the overlapping generations model of chapter 4 in two directions. First, Erosa and Gervais (2002) build a fairly general overlapping generations model of optimal taxation where the individual lives for more than two periods. The authors derive the optimal tax policy and relate their work to the influential paper by Auerbach, Kotlikoff, and Skinner (1983). These authors study the effects of switching from a proportional income tax to either a proportional tax on consumption or a proportional tax on labor income. Hence, the analysis focusses on the effect of the switch to a given tax system, which must not be necessarily the optimal one. This exactly is done by Erosa and Gervais (2002), who analyze the switch to the optimal tax system. But it is important to bear in mind that the productivity profile of the individual over the life cycle is given. Both Auerbach, Kotlikoff, and Skinner (1983) and Erosa and Gervais (2002) take the productivity profile from Welch (1979). The contribution then could be to further generalize the model of chapter 4 to a time horizon with, say, 55 years, or, put differently, to add endogenous education to the model by Erosa and Gervais (2002) and thus to endogenize the productivity profile. Then a calibration and simulation exercise similar to that of chapter 5 can be executed to assess the relevance of subsidizing education to account for distortionary taxation in relation to internalizing the intergenerational externality.

Second, the model studies an individual that is either selfish or altruistic à la Barro (1974) and Becker (1974). These are certainly two polar cases. One can think of other preferences that are somewhat in between where an individual does care about his descendants to a certain degree. Andreoni (1989) was the first to set up an overlapping generations model in which altruism is not "pure" but "impure". The individual gets a "warm glow" from giving to his immediate descendant. The idea of deriving utility from leaving a bequest to descendants may be applied to the stock of human capital. Glomm and Ravikumar (1992) and Cremer and Pestieau (2006) take up this idea and provide a normative analysis of education policies such as public vs. private provision of education and taxation/subsidization of education, respectively.

Which kind of policy advice can be drawn from the present research? The analysis clearly shows that thinking about education efficiency involves calculating the social and private benefits and costs of education over the entire life cycle. One then has to bear in mind that earnings when educated are usually higher than foregone earnings when being educated. Then, if the marginal tax rate is increasing in earnings, the tax treatment of future earnings and foregone earnings differ, which may translate into an inequality between the social and private rate of return to education. Moreover, the individual has to pay for direct cost that affect the rate of return of education. The government may choose to tax or subsidize these costs to adjust the private rate of return such that it equals the social rate of return. To achieve education efficiency, a first step could be to make the direct cost of education fully tax-deductible from future earnings, which then means that direct costs and future earnings are taxed at the same rate.

This issue has been and still is under scrutiny by the Federal Fiscal Court in Germany. As a general rule, under the provisions of section 12 (5) of the German Income Tax Act education related costs are considered privately induced and therefore may not be deductible as earnings-related or special expenses.[1] The Federal Fiscal Court however ruled in 2009 that education related costs may be deductible as earnings-related expenses if the taxpayer has already completed vocational education before tertiary education was started. Section 12 (5) of the German Income Tax Code then does not apply (Federal Fiscal Court, Judgment, Ref. VI R 14/07, June 18, 2009). An appeal is currently pending with the Federal Fiscal Court concerning the question whether the aforementioned judgement also applies to taxpayers who have started studying right after having finished school (Federal Fiscal Court, Ref. VI R 7/10).[2]

1 The crucial difference between earnings-related and special expenses is the following: Earnings-related expenses (section 9 (1) of the German Income Tax Act) are deducted from earnings and a possibly resulting loss may be carried forward to the following fiscal year under the provisions of section 10d (2) of the German Income Tax Act. Special expenses, e.g. education expenses (section 10 (1) number 7 of the German Income Tax Code) are deducted from positive overall earnings, which is the sum of all net earnings, up to 4,000 Euro. This difference, which is taxable income, cannot become negative. A possibly resulting loss may not be carried forward. Moreover, if overall earnings are negative, special expenses are not tax-deductible at all within the fiscal year in question.

2 The Fiscal Court of Hamburg, that has granted this appeal, however denied the applicability of this judgment (Fiscal Court of Hamburg, Judgment, Ref. 5 K 193/08, November 25, 2009).

The Advisory Board to the Federal Ministry of Finance (2010) even goes further than the aforementioned judgment of the Federal Fiscal Court and makes two far-reaching proposals to achieve education efficiency, which holds when the social and private rates of return to education are equal. This equality holds when the return and the direct and indirect costs of education are taxed at the same rate. First, the Advisory Board suggests to consider education expenses as anticipated earnings-related expenses and to allow for an interest-bearing carry forward. The effect is that direct costs and benefits of education are taxed at the same rate. Second, as has been said before, the government taxes the return to education at a higher rate than foregone earnings if the marginal tax rate increases in taxable income. To compensate the taxpayer it is therefore necessary that the government furthermore bears an additional fraction of the foregone earnings amounting to the difference between the higher tax rate when educated and the lower tax rate when being educated. Based on the assumption that foregone earnings amount to 20,000 Euros per year, which implies a marginal tax rate of 29%, and that initial earnings when educated amount to 45,000 Euro, which implies a marginal tax rate of 41%, the government shall grant a subsidy worth 2,400 Euro $(= (0.41 - 0.29) \times 20,000)$ to students.

This dissertation contributes to the discussion insofar as it sheds light on the circumstances under which education efficiency shall prevail in a second-best world. In particular, the analysis has demonstrated that there are not too implausible circumstances under which an inequality between the private and social rate of return to education is optimal.

BIBLIOGRAPHY

Advisory Board to the Federal Ministry of Finance (2010): "Beitragspflichtiger Hochschulzugang," Federal Ministry of Finance.

Ammermüller, A., and A. M. Weber (2005): "Educational Attainment and Returns to Education in Germany: An Analysis by Subject of Degree, Gender and Region," ZEW Discussion Papers No. 05-17, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Anderberg, D. (2009): "Optimal policy and the risk properties of human capital reconsidered," *Journal of Public Economics*, 93(9-10), 1017–1026.

Andersson, F., and K. A. Konrad (2003): "Human capital investment and globalization in extortionary states," *Journal of Public Economics*, 87(7-8), 1539–1555.

Andreoni, J. (1989): "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy*, 97(6), 1447–58.

Atkeson, A., V. V. Chari, and P. J. Kehoe (1999): "Taxing capital income: a bad idea," *Federal Reserve Bank of Minneapolis Quarterly Review*, 23(3), 3–17.

Atkinson, A. B., and J. E. Stiglitz (1972): "The structure of indirect taxation and economic efficiency," *Journal of Public Economics*, 1(1), 97–119.

——— (1980): *Lectures on Public Economics*. McGraw-Hill Inc., London, United Kingdom and New York (NY), United States of America.

ATUS (2003-2009): "Annual results," `http://www.bls.gov/tus/`.

Auerbach, A. J., and L. J. Kotlikoff (1987): *Dynamic Fiscal Policy*. Cambridge University Press, New York (NY), United States of America.

Auerbach, A. J., L. J. Kotlikoff, and J. Skinner (1983): "The Efficiency Gains from Dynamic Tax Reform," *International Economic Review*, 24(1), 81–100.

Barbie, M., and C. Hermeling (2006): "Optimal Taxation in a Simple Model of Human Capital Accumulation," *Economics Bulletin*, 5(5), 1–8.

Barro, R. J. (1974): "Are Government Bonds Net Wealth?," *Journal of Political Economy*, 82(6), 1095–1117.

——— (1991): "Economic Growth in a Cross Section of Countries," *The Quarterly Journal of Economics*, 106(2), 407–43.

——— (1998): *Determinants of Economic Growth: A Cross-Country Empirical Study*, vol. 1 of *MIT Press Books*. The MIT Press, Cambridge (MA), United States of America.

Barro, R. J., and J.-W. Lee (1994): "Sources of economic growth," *Carnegie-Rochester Conference Series on Public Policy*, 40(1), 1–46.

Bassanini, A., and S. Scarpetta (2001): "The Driving Forces of Economic Growth: Panel Data Evidence for the OECD Countries," *OECD Economic Studies*, 2001(2).

Becker, G. S. (1974): "A Theory of Social Interactions," *Journal of Political Economy*, 82(6), 1063–93.

——— (1993): *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. The University of Chicago Press, Chicago (IL), United States of America, 3rd edn.

Ben-Porath, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *The Journal of Political Economy*, 75(4), 352–365.

Benhabib, J., and M. M. Spiegel (1994): "The role of human capital in economic development evidence from aggregate cross-country data," *Journal of Monetary Economics*, 34(2), 143–173.

Blankenau, W. F., and N. B. Simpson (2004): "Public education expenditures and growth," *Journal of Development Economics*, 73(2), 583–605.

Boadway, R., N. Marceau, and M. Marchand (1996): "Investment in Education and the Time Inconsistency of Redistributive Tax Policy," *Economica*, 63(250), 171–89.

Boskin, M. J. (1975): "Notes on the Tax Treatment of Human Capital," NBER Working Papers 0116, National Bureau of Economic Research.

Bouzahzah, M., D. De la Croix, and F. Docquier (2002): "Policy reforms and growth in computable OLG economies," *Journal of Economic Dynamics and Control*, 26(12), 2093–2113.

Bovenberg, A. L., and B. Jacobs (2005): "Redistribution and education subsidies are Siamese twins," *Journal of Public Economics*, 89(11-12), 2005–2035.

Bowman, M. J. (1964): "Schultz, Denison, and the Contribution of "Eds" to National Income Growth," *Journal of Political Economy*, 72(5), 450–464.

Bringhurst, R. (2002): *The Elements of Typographic Style*, Version 2.5. Hartley & Marks, Publishers, Point Roberts (WA), United States of America.

Bull, N. (1993): "When all the optimal dynamic taxes are zero," Working Paper Series / Economic Activity Section 137, Board of Governors of the Federal Reserve System (U.S.).

Caballe, J., and M. S. Santos (1993): "On Endogenous Growth with Physical and Human Capital," *Journal of Political Economy*, 101(6), 1042–67.

Cahuc, P., and A. Zylbergerg (2004): *Labor Economics*. MIT, Cambridge (MA), United States of America.

Cameron, S. V., and J. J. Heckman (2001): "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy*, 109(3), 455–499.

Card, D. (1999): "The causal effect of education on earnings," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3 of *Handbook of Labor Economics*, chap. 30, pp. 1801–1863. Elsevier.

Carneiro, P., and J. J. Heckman (2002): "The Evidence on Credit Constraints in Post–secondary Schooling," *Economic Journal*, 112(482), 705–734.

CDU, C. u. F. (2009): "Wachstum. Bildung. Zusammenhalt. Der Koalitionsvertrag zwischen CDU, CSU und FDP," `http://www.cdu.de/doc/pdfc/091026-koalitionsvertrag-cducsu-fdp.pdf`, accessed December 15, 2010.

Chamley, C. (1986): "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54(3), 607–22.

Chari, V. V., L. J. Christiano, and P. J. Kehoe (1994): "Optimal Fiscal Policy in a Business Cycle Model," *Journal of Political Economy*, 102(4), 617–52.

Chari, V. V., and P. J. Kehoe (1999): "Optimal fiscal and monetary policy," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1 of *Handbook of Macroeconomics*, chap. 26, pp. 1671–1745. Elsevier.

Christiano, L. J. (2010): "Lecture Notes on Optimal Fiscal and Monetary Policy," `http://faculty.wcas.northwestern.edu/~lchrist/papers/optimal.pdf`, accessed December 9, 2010.

Ciriani, S. (2007): "Education policy in a general equilibrium model with heterogeneous agents," *Economics Bulletin*, 9(1), 1–7.

Cremer, H., and P. Pestieau (2006): "Intergenerational Transfer of Human Capital and Optimal Education Policy," *Journal of Public Economic Theory*, 8(4), 529–545.

da Costa, C. E., and L. J. Maestri (2007): "The risk properties of human capital and the design of government policies," *European Economic Review*, 51(3), 695–713.

Davies, J., and J. Whalley (1991): "Taxes and Capital Formation: How Important is Human Capital?," in *National Saving and Economic Performance*, NBER Chapters, pp. 163–200. National Bureau of Economic Research, Inc.

De la Croix, D., and P. Michel (2002): *A Theory of Economic Growth*, Cambridge Books. Cambridge University Press, New York (NY), United States of America.

de la Fuente, A., and R. Doménech (2006): "Human Capital in Growth Regressions: How Much Difference Does Data Quality Make?," *Journal of the European Economic Association*, 4(1), 1–36.

Denison, E. F. (1962): *The Sources of Economic Growth in the United States and the Alternatives before us*. Committee for Economic Development.

——— (1967): *Why Growth Rates Differ - Postwar Experience in Nine Western Countries*. The Brookings Institution, Washington D.C., United States of America.

Diamond, P. A. (2007): "Comment on Mikhail Golosov, Aleh Tsyvinski, Iván Werning, New Dynamic Public Finance: A Users Guide," in *NBER Macroeconomics Annual 2006, Volume 21*, ed. by K. R. Daron Acemoglu, and M. Woodford, NBER Chapters, chap. 5, pp. 365–379. National Bureau of Economic Research.

Diamond, P. A., and J. A. Mirrlees (1971): "Optimal Taxation and Public Production I: Production Efficiency," *American Economic Review*, 61(1), 8–27.

Docquier, F., O. Paddison, and P. Pestieau (2007): "Optimal accumulation in an endogenous growth setting with human capital," *Journal of Economic Theory*, 134(1), 361–378.

Dupor, B., L. Lochner, C. Taber, and M. B. Wittekind (1996): "Some Effects of Taxes on Schooling and Training," *American Economic Review*, 86(2), 340–46.

Ellwood, D., and T. Kane (2000): "Who is Getting a College Education?: Family Background and the Growing Gaps in Enrolment," in *Securing the Future: Investing in Children from Birth to College*, ed. by S. Danzinger, and J. Waldfogel, chap. 10, pp. 283–324. Russel Sage Foundation, New York (NY), United States of America.

Englander, S., and A. Gurney (1994): "Medium-term determinants of OECD productivity," *OECD Economic Studies*, 22, 49–109.

Erosa, A., and M. Gervais (2002): "Optimal Taxation in Life-Cycle Economies," *Journal of Economic Theory*, 105(2), 338–369.

Fischer, S. (1980): "Dynamic inconsistency, cooperation and the benevolent dissembling government," *Journal of Economic Dynamics and Control*, 2(1), 93–107.

Gemmell, N. (1996): "Evaluating the Impacts of Human Capital Stocks and Accumulation on Economic Growth: Some New Evidence," *Oxford Bulletin of Economics and Statistics*, 58(1), 9–28.

Glomm, G., and B. Ravikumar (1992): "Public versus Private Investment in Human Capital Endogenous Growth and Income Inequality," *Journal of Political Economy*, 100(4), 813–34.

Golosov, M., N. Kocherlakota, and A. Tsyvinski (2003): "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, 70(3), 569–587.

Golosov, M., and A. Tsyvinski (2008): "optimal fiscal and monetary policy (without commitment)," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf, and L. E. Blume. Palgrave Macmillan, Basingstoke, United Kingdom.

Golosov, M., A. Tszvinski, and I. Werning (2007): "New Dynamic Public Finance: A User's Guide," in *NBER Macroeconomics Annual 2006, Volume 21*, ed. by K. R. Daron Acemoglu, and M. Woodford, NBER Chapters, chap. 5, pp. 317–388. National Bureau of Economic Research.

Gomme, P., and P. Rupert (2007): "Theory, measurement and calibration of macroeconomic models," *Journal of Monetary Economics*, 54(2), 460–497.

Grochulski, B., and T. Piskorski (2010): "Risky human capital and deferred capital income taxation," *Journal of Economic Theory*, 145(3), 908–943.

Hall, R. E., and C. I. Jones (1999): "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *The Quarterly Journal of Economics*, 114(1), 83–116.

Hanushek, E. A., and D. Kim (1995): "Schooling, Labor Force Quality, and Economic Growth," NBER Working Papers 5399, National Bureau of Economic Research, Inc.

Harmon, C., H. Oosterbeek, and I. Walker (2003): "The Returns to Education: Microeconomics," *Journal of Economic Surveys*, 17(2), 115–156.

Heckman, J. J., L. Lochner, and C. Taber (1998): "Tax Policy and Human-Capital Formation," *American Economic Review*, 88(2), 293–97.

——— (1999): "Human capital formation and general equilibrium treatment effects: a study of tax and tuition policy," *Fiscal Studies*, 20(1), 25–40.

Jacobs, B. (2005): "Optimal Income Taxation with Endogenous Human Capital," *Journal of Public Economic Theory*, 7(2), 295–315.

Jacobs, B., and A. L. Bovenberg (2010a): "Human Capital and Optimal Positive Taxation of Capital Income," *International Tax and Public Finance*, 17(5), 451–478.

——— (2010b): "Optimal Taxation of Human Capital and the Earnings Function," *Journal of Public Economic Theory*, forthcoming.

Jacobs, B., D. Schindler, and H. Yank (2010): "Optimal Taxation of Risky Human Capital," *Scandinavian Journal of Economics*, forthcoming.

Jones, L. E., R. E. Manuelli, and P. E. Rossi (1993): "Optimal Taxation in Models of Endogenous Growth," *Journal of Political Economy*, 101(3), 485–517.

——— (1997): "On the Optimal Taxation of Capital Income," *Journal of Economic Theory*, 73(1), 93–117.

Jorgenson, D. W., and B. M. Fraumeni (1992): "Investment in Education and U.S. Economic Growth," *Scandinavian Journal of Economics*, 94(0), S51–70.

Judd, K. L. (1985): "Redistributive taxation in a simple perfect foresight model," *Journal of Public Economics*, 28(1), 59–83.

——— (1999): "Optimal taxation and spending in general competitive growth models," *Journal of Public Economics*, 71(1), 1–26.

——— (2007): "Comment on Mikhail Golosov, Aleh Tsyvinski, Iván Werning, New Dynamic Public Finance: A Users Guide," in *NBER Macroeconomics Annual 2006, Volume 21*, ed. by K. R. Daron Acemoglu, and M. Woodford, NBER Chapters, chap. 5, pp. 381–384. National Bureau of Economic Research.

Judson, R. (1998): "Economic Growth and Investment in Education: How Allocation Matters," *Journal of Economic Growth*, 3(4), 337–59.

Kane, T. J. (1994): "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education," *Journal of Political Economy*, 102(5), 878–911.

King, R. G., C. I. Plosser, and S. T. Rebelo (1988): "Production, growth and business cycles : I. The basic neoclassical model," *Journal of Monetary Economics*, 21(2-3), 195–232.

——— (2002): "Production, Growth and Business Cycles: Technical Appendix," *Computational Economics*, 20(1-2), 87–116.

KING, R. G., AND S. REBELO (1990): "Public Policy and Economic Growth: Developing Neoclassical Implications," *Journal of Political Economy*, 98(5), S126–50.

KLENOW, P., AND A. RODRÍGUEZ-CLARE (1997): "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?," in *NBER Macroeconomics Annual 1997, Volume 12*, NBER Chapters, pp. 73–114. National Bureau of Economic Research, Inc.

KOCHERLAKOTA, N. R. (2005): "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, 73(5), 1587–1621.

KREBS, T. (2003): "Human Capital Risk And Economic Growth," *The Quarterly Journal of Economics*, 118(2), 709–744.

KRUEGER, A. B., AND M. LINDAHL (2001): "Education for Growth: Why and for Whom?," *Journal of Economic Literature*, 39(4), 1101–1136.

KYDLAND, F. E., AND E. C. PRESCOTT (1977): "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85(3), 473–91.

LEVINE, R., AND D. RENELT (1992): "A Sensitivity Analysis of Cross-Country Growth Regressions," *American Economic Review*, 82(4), 942–63.

LIPSEY, R. G., AND K. LANCASTER (1956 - 1957): "The General Theory of Second Best," *The Review of Economic Studies*, 24(1), 11–32.

LJUNGQVIST, L., AND T. J. SARGENT (2004): *Recursive Marcoeconomic Theory*. MIT Press, Cambridge (MA), United States of America, 2nd edn.

LUCAS, R. E. J. (1988): "On the mechanics of economic development," *Journal of Monetary Economics*, 22(1), 3–42.

——— (1990): "Supply-Side Economics: An Analytical Review," *Oxford Economic Papers*, 42(2), 293–316.

LUCAS, R. E. J., AND N. L. STOKEY (1983): "Optimal fiscal and monetary policy in an economy without capital," *Journal of Monetary Economics*, 12(1), 55–93.

Mankiw, N. G., D. Romer, and D. N. Weil (1992): "A Contribution to the Empirics of Economic Growth," *The Quarterly Journal of Economics*, 107(2), 407–437.

Mendoza, E. G., A. Razin, and L. L. Tesar (1994): "Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption," *Journal of Monetary Economics*, 34(3), 297–323.

Merkel, A. (2009): "Wachstum. Bildung. Zusammenhalt. Deutschland zu neuer Stärke führen - Regierungserklärung von Bundeskanzlerin Angela Merkel vor dem Deutschen Bundestag am 10.11.2009," `http://www.bundesregierung.de/Content/DE/Publikation/Bestellservice/__Anlagen/2009-11-26-regierungserklaerung-bkin.html`, accessed December 15, 2010.

Milesi-Ferretti, G. M., and N. Roubini (1998): "On the taxation of human and physical capital in models of endogenous growth," *Journal of Public Economics*, 70(2), 237–254.

Mincer, J. A. (1974): *Schooling, Experience, and Earnings*. Columbia University Press, New York (NY), United States of America.

Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38(114), 175–208.

——— (1986): "The theory of optimal taxation," in *Handbook of Mathematical Economics*, ed. by K. J. Arrow, and M. Intriligator, vol. 3 of *Handbook of Mathematical Economics*, chap. 24, pp. 1197–1249. Elsevier.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1991): "The Allocation of Talent: Implications for Growth," *The Quarterly Journal of Economics*, 106(2), 503–30.

Myles, G. D. (2007): "Economic Growth and the Role of Taxation," Report prepared for the OECD, `http://www.people.ex.ac.uk/gdmyles/papers/pdfs/OECDfin.pdf`, accessed October 25, 2010.

——— (2009a): "Economic Growth and the Role of Taxation-Theory," OECD Economics Department Working Papers 713, OECD, Economics Department.

——— (2009b): "Economic Growth and the Role of Taxation - Aggregate Data," OECD Economics Department Working Papers 714, OECD, Economics Department.

——— (2009c): "Economic Growth and the Role of Taxation - Disaggregate Data," OECD Economics Department Working Papers 715, OECD, Economics Department.

Nielsen, S. B., and P. B. Sørensen (1997): "On the optimality of the Nordic system of dual income taxation," *Journal of Public Economics*, 63(3), 311–329.

Obama, B. (2009): "The President addresses a Joint Session of Congress, February 24, 2009," `http://www.whitehouse.gov/the_press_office/Remarks-of-President-Barack-Obama-ADDRESS-to-Joint-Session-of-Congress`, accessed December 15, 2010.

OECD (2009): *Education at a Glance*. OECD Publications, Paris, France.

——— (2010): *Education at a Glance*. OECD Publications, Paris, France.

Pecorino, P. (1993): "Tax structure and growth in a model with human capital," *Journal of Public Economics*, 52(2), 251–271.

Psacharopoulos, G. (1973): *Returns to Education: An International Comparison*. Elsevier, Jossey Bass, San Francisco (CA), United States of America.

——— (1994): "Returns to investment in education: A global update," *World Development*, 22(9), 1325–1343.

Ramsey, F. P. (1927): "A Contribution to the Theory of Taxation," *The Economic Journal*, 37(145), 47–61.

Reis, C. (2007): "Essays on Optimal Taxation," Ph.D. thesis, Massachusetts Institute of Technology, Department of Economics.

Rey, E. D., and M. del Mar Racionero (2002): "Optimal educational choice and redistribution when parental education matters," *Oxford Economic Papers*, 54(3), 435–448.

Richter, W. F. (2009): "Taxing Education in Ramsey's Tradition," *Journal of Public Economics*, 93(11-12), 1254–1260.

Ritter, F. E., and L. J. Schooler (2001): "The learning curve," in *International Encyclopedia of the Social and Behavioral Sciences*, ed. by P. B. W. Kintch, N. Smelser, pp. 8602–8605. Pergamon.

Rosen, S. (1976): "A Theory of Life Earnings," *Journal of Political Economy*, 84(4), S45–S67.

Sandmo, A. (1974): "A Note on the Structure of Optimal Taxation," *American Economic Review*, 64(4), 701–06.

Schultz, T. W. (1961a): "Investment in Human Capital," *American Economic Review*, 51(1), 1–17.

——— (1961b): "Education and Economic Growth," in *Social Forces Influencing American Education*, ed. by N. B. Henry, and R. W. Tyler, chap. 3, pp. 46–88. The University of Chicago Press, Chicago (IL), United States of America.

Sianesi, B., and J. V. Reenen (2003): "The Returns to Education: Macroeconomics," *Journal of Economic Surveys*, 17(2), 157–200.

Solow, R. M. (1957): "Technical Change and the Aggregate Production Function," *The Review of Economics and Statistics*, 39(3), 312–320.

Taber, C. (2002): "Tax Reform and Human Capital Accumulation: Evidence from an Empirical General Equilibrium Model of Skill Formation," *Advances in Economic Analysis & Policy*, 2(1).

Topel, R. (1999): "Labor markets and economic growth," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3 of *Handbook of Labor Economics*, chap. 44, pp. 2943–2984. Elsevier.

Trostel, P. A. (1993): "The Effect of Taxation on Human Capital," *Journal of Political Economy*, 101(2), 327–50.

——— (1996): "Should Education Be Subsidized?," *Public Finance Review*, 24(1), 3–24.

Uzawa, H. (1965): "Optimum Technical Change in An Aggregative Model of Economic Growth," *International Economic Review*, 6(1), 18–31.

Welch, F. (1979): "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust," *Journal of Political Economy*, 87(5), S65–S97.

Wigger, B. U. (2002): *Public pensions and economic growth*. Springer-Verlag, Berlin, Germany.

Wigger, B. U., and R. K. von Weizsäcker (2001): "Risk, Resources, and Education: Public Versus Private Financing of Higher Education," *IMF Staff Papers*, 48(3), 6.

Willis, R. J. (1986): "Wage determinants: A survey and reinterpretation of human capital earnings functions," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and R. Layard, vol. 1 of *Handbook of Labor Economics*, chap. 10, pp. 525–602. Elsevier.

Yuen, C.-W. (1991): "Taxation, Human Capital Accumulation, and Economic Growth," Ph.D. thesis, University of Chicago, Department of Economics.

COLOPHON

This thesis was typeset with $\LaTeX\,2_\varepsilon$ using Hermann Zapf's *Palatino* and *Euler* type faces (Type 1 PostScript fonts *URW Palladio L* and *FPL* were used). The listings are typeset in *Bera Mono*, originally developed by Bitstream, Inc. as "Bitstream Vera". (Type 1 PostScript fonts were made available by Malte Rosenau and Ulrich Dirr.) The typographic style was inspired by Bringhurst's genius as presented in *The Elements of Typographic Style* (Bringhurst, 2002). It is available for LaTeX via CTAN as "`classicthesis`".

NOTE: The custom size of the textblock was calculated using the directions given by Mr. Bringhurst (pages 26–29 and 175/176). 10 pt Palatino needs 133.21 pt for the string "abcdefghijklmnopqrstuvwxyz". This yields a good line length between 24–26 pc (288–312 pt). Using a "*double square textblock*" with a 1:2 ratio this results in a textblock of 312:624 pt (which includes the headline in this design). A good alternative would be the "*golden section textblock*" with a ratio of 1:1.62, here 312:505.44 pt. For comparison, `DIV9` of the `typearea` package results in a line length of 389 pt (32.4 pc), which is by far too long. However, this information will only be of interest for hardcore pseudo-typographers like me.

To make your own calculations, use the following commands and look up the corresponding lengths in the book:

```
\settowidth{\abcd}{abcdefghijklmnopqrstuvwxyz}
\the\abcd\ % prints the value of the length
```

Please see the file `classicthesis.sty` for some precalculated values for Palatino and Minion.

159.85193pt

Final Version as of May 8, 2011 at 10:44.

DECLARATION

Ich versichere, dass ich diese Dissertation selbstständig verfasst habe. Bei der Erstellung der Arbeit habe ich mich ausschließlich der angegebenen Hilfsmittel bedient. Die Dissertation ist nicht bereits Gegenstand eines erfolgreich abgeschlossenen Promotions- oder sonstigen Prüfungsverfahrens gewesen.

Dortmund, 28.01.2011

Christoph Braun