# Human Capital, Labor Markets, and Public Policy

DISSERTATION
of the University of St. Gallen,
Graduate School of Business Administration,
Economics, Law and Social Sciences (HSG)
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by

**Kai-Joseph Fleischhauer**

from

Germany

Approved on the application of

**Prof. Dr. Christian Keuschnigg**

and

**Prof. Dr. Laszlo Goerke**

Dissertation no. 3370

Difo-Druck GmbH, Bamberg 2007

The University of St. Gallen, Graduate School of Business Administration, Economics, Law and Social Sciences (HSG) hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, May 16, 2007

The President:

Prof. Ernst Mohr, PhD

To my parents
and my wife

# Contents

# List of Figures

# List of Tables

# Acknowledgement

When I decided to do a doctorate and then started to look for an appealing topic some years ago, I often dreamt of this moment: the work is done, my PhD thesis has been accepted, and I still feel reasonably mentally sane... And now, in this illustrious moment, I can announce with certainty: it is even better than I had ever expected! Above all, this success has been made possible by my parents' encouragement who always believed in me although I decided long ago to disappoint my father by not becoming an engineer...

My dissertation owes intellectual debt to many people. The largest share of it has to be attributed to my advisors Christian Keuschnigg and Laszlo Goerke who greatly supported me in my work and suggested various improvements. I am also indebted to Philipp Guyer who inspired me with the fundamentals of economic reasoning. Seminal discussions are due to Michael Lechner and various seminar participants at the University of St. Gallen.

I am looking back already on the time of my dissertation both relieved and filled with melancholy. On the one hand, I am very happy and proud to have achieved this ambitious and challenging goal! During the last years, there were several moments when my desperation was nearly as pronounced as my confidence: doing the same mathematical modeling over and over again - without finding the source of inconsistency concerning economic intuition - is a nightmare probably familiar to all economists...

On the other hand, there have been even more moments of pleasure, satisfaction, and gratification. Primarily, these moments originate from my friends and colleagues at the *Institute of Public Finance and Fiscal Law (IFF-*

*HSG).* Among others, I would like to thank Alex Gulde, Christian Jaag, Mirela Keuschnigg, Benita von Lindeiner, Adrian Oberlin, Stephan Raas, Evelyn Ribi, Christian Stucki, and Niklaus Wallimann, who contributed to my dissertation with professional and moral support. With respect to my PhD year at the Study Center Gerzensee, special thanks go to my fellow sufferers Ingo Borchert, Tariq Hasan, and Conny Wunsch.

Furthermore, I am very much indebted to my wife Carla, my brother Jan, and all my friends in St. Gallen, Zurich, and Hattingen for their patience and loyalty as well as for hundreds of outstanding and unforgettable evenings! Particularly, these are Bernd Brommundt, Jan Falkenhain, Nepomuk Feser, Michael Fischer, Tom Hardtmuth, Sven Kasper, Felix Moldenhauer, Andi Morgenstern, Jörn Morgenstern, Stefan Schneidereit, Lucian Schönefelder, Mark Schumann, Matthias Spitzmüller, and Michael Verhofen.

Finally, my students' fraternity *A.V. Emporia-Alemannia San Gallensis* is responsible for almost making a real Swiss out of this vociferous boy from the *Ruhrpott...* Among many others, I would like to thank my Emporia generation including Kis Baumgartner, Curdin Derungs, Benedict-Christoph Gerth, Lucien Müller, Moritz Rothenbühler, Harald Siegel, Pascal Welschinger, Marc Wermuth, and Pascal Wirth as well as my Emporia family with Stefan Baumgartner, Johannes Braun, Stefan Caliz, Christian Cobbers, and Emmanuel Niklaus.

In a nutshell, it was a great time and an impressive experience that I will never forget!

St. Gallen, June 29, 2007 Kai-Joseph Fleischhauer

# Chapter 1

# Human Capital: Microeconomics

*With the beginning of the new millennium it has become more and more apparent that education and human capital constitute a key element of modern economies. Despite the important role of human capital in modern societies, there are still many unknowns about the process of educational production as well as individual and collective decisions concerning how much and what kind of education to obtain. Chapter 1 of my PhD thesis aims at providing a better understanding of the process of human capital formation and educational attainment. Although human capital plays an important role in both microeconomics and macroeconomics, we focus on the former branch of literature in order to analyze the individual incentives to acquire skills.*

## 1.1 Introduction

With the beginning of the new millennium it has become more and more apparent that education and human capital constitute a key element of modern economies. While some developing countries succeed in attracting industries of the so-called old economy, developed economies have to concentrate on skill-intensive industries in order to defend their leading position. In this context, the change to an information society is occurring rapidly, with information and knowledge the crucial inputs and outputs of nearly all economic processes. Despite the important role of human capital in modern societies, there are still many unknowns about the process of educational production as well as individual and collective decisions concerning how much and what kind of education to obtain.

The contribution of Chapter 1 of my PhD thesis is to provide a better understanding of the process of human capital formation and educational attainment. Although human capital plays an important role in both microeconomics and macroeconomics, we focus on the former branch of literature in order to analyze the individual incentives to acquire skills. Furthermore, we analyze policy instruments and institutional features that may help to increase the aggregate welfare by improving the efficiency of the educational system. In order to structure the vast literature on human capital formation, this review is divided into six sections each of them representing an important stream of human capital literature. In each section, we concentrate on theoretical results or empirical findings depending on the main approaches of the corresponding literature. It is important to note that this literature review provides a general overview of human capital theory and related empirical findings without explicitly referring to those topics of human capital formation dealt with in Chapters 2 to 4 of my PhD thesis.

As an introduction to human capital theory, Section 1.2 introduces the basic concept of human capital that models individuals as investing in skills in response to the expected returns to education. In Section 1.3 of this literature review, we build on this basic approach by distinguishing between general and specific human capital and analyzing the different implications for human capital investments by workers and firms. In perfect labor markets, all costs and benefits of general human capital are borne by the workers, while firms and workers share both the costs and the returns of investments in specific human capital (Becker (1964)). In imperfect labor markets, general training may also be firm-sponsored because the wage structure is compressed, which implies that firms manage to skim labor market rents depending on the amount of training (Acemoglu and Pischke (1998a)).

In order to show the empirical relevance of these theoretical contributions, Section 1.4 provides an overview of various empirical studies measuring the return to education from an individual's point of view according to Mincer (1974). Although the rate of return to education varies significantly in response to several influencing factors, the average estimate for developed economies generally ranges from 5% to 10% (Wilson (2001)). Because these influencing factors may also directly influence the education decision, we review the literature on educational production functions in Section 1.5 and discuss the significance of potential inputs into the process of educational production. While the empirical evidence concerning the impact of school resources is mixed, there is an unambiguous effect of family background and peer groups as well as institutional incentives within the educational system (Hanushek (1997)). Beyond this rather static framework, Section 1.6 takes a dynamic perspective and describes the life-cycle of earnings with endogenous formation of human capital. The two most important approaches by Ben-Porath (1967) and Heckman (1976) manage to replicate the empirical life-cycle patterns with respect to the age-earnings profile of individuals.

Finally, in Section 1.7 of this literature review, we analyze the effects of taxation and education subsidies on human capital formation. The marginal effects of proportional and progressive income taxation on human capital accumulation are generally negative (Heckman, Lochner, and Taber (1999a)). Depending on whether the costs of human capital are direct expenditures or foregone earnings, a "comprehensive income tax" may discriminate either against investments in human or physical capital. In a nutshell, the incentives for human capital formation depend on the net effective tax rate, which implies that education subsidies can improve the efficiency in human capital investment by offsetting tax-induced distortions (Bovenberg and Jacobs (2005)).

Section 1.8 provides an outlook on my research projects, which are presented in detail in Chapters 2 to 4 of my PhD thesis.

## 1.2 The Concept of Human Capital

As an introduction to human capital theory, this section introduces the main characteristics of human capital and the original approach to human capital formation. The basic concept of human capital models individuals as investing in skills in response to the expected returns to education.

"Human capital" can be defined as knowledge, skills, attitudes, aptitudes, and other acquired traits contributing to production (Goode (1959)). Skills represent individual capacities contributing to production as an argument in the production function (Bowles, Gintis, and Osborne (2001)). There are two main components of human capital with strong complementarity: early ability (whether acquired or innate) and skills acquired through formal education (Blundell, Dearden, Meghir, and Sianesi (1999)). Human capital differs from physical capital because it cannot be supplied by oneself and yields market returns only in proportion to the individual's supply of labor (Hall and Johnson (1980)).[1] Ishikawa and Ryan (2002) suggest that it is the stock of human capital that predominantly determines the earnings of individuals. An extensive review of human capital theory is given by Cahuc

[1] The value of human capital positively depends on the number of hours worked in the future. This implies that there are no benefits from investing in human capital if an individual does not intend to work in the future (Heckman, Lochner, and Cossa (2002)). However, if the marginal utility of leisure positively depends on the stock of human capital as assumed by Heckman (1976), there are benefits from human capital investments even in the case of no work.

and Zylberberg (2004).

The first use of the term "human capital" in modern economic literature was by Schultz (1961). He classifies expenditures on human capital as investment rather than consumption.[2] In the same year, Weisbrod (1961) developed a first conceptual framework for estimating the value of assets in the form of human capital. Capital values of people as productive assets are incorporated into an analytical function of sex, age, stock of human capital, etc. The present value of an individual at any given age $a$ is defined as the sum of his discounted expected future earnings $Y_t$ (equal to the value of productivity):

$$V(a) = \sum_{t=a}^{\infty} \frac{P_{at}}{(1+r)^{t-a}} Y_t$$

$P_{at}$ represents the probability of an individual of age $a$ to be alive at age $t$ and $r$ is the discount rate. In general, there are two methods of determining the value of human capital, namely by summing up the costs of production (input-based) and by considering capitalized earnings (output-based) (Kiker (1966)).

The first applications of human capital theory in economics are by Becker and Mincer of the Chicago school. In his original approach, Becker (1964) develops a model of individual investment in human capital. In this view, human capital is similar to "physical means of production". According to Becker (1962), investing in human capital means "all activities that influence future real income through the embedding of resources in people". Human capital investments are expenditures on education, training, health, information, and labor mobility (Weisbrod (1966)).

Human capital accumulation by formal education takes place in three ways: formal schooling (i.e. the individual devotes his whole time to learning), on-the-job training (i.e. post-school training provided by the current

[2] Shaffer (1961) citicizes the application of capital concepts to individuals for three reasons: educational expenditures may be undertaken for other reasons than the expectation of monetary returns, impossibility to relate a certain return to a certain investment, and undesirable basis for the evaluation of policy actions with respect to social welfare.

employer), and off-the-job training (i.e. post-school training provided by "for-profit" proprietary institutions) (Lynch (1991)).[3] These investments involve initial costs (direct education costs such as tuition expenditures, foregone earnings during schooling, and reduced wages during training)[4] in order to gain a return on this investment in the future (Becker (1992)). The return to education is based on two interrelated channels: increased earnings and higher employment probabilities (Bloch and Smith (1977)).[5] In a nutshell, there are two key determinants of the return to education: the costs of education and the employment opportunities after education (Rephann (2002)).

The key element in the model by Becker (1964) is that education is an investment of time and foregone earnings for higher rates of return in later periods. As with investments in physical capital, a human capital investment is only undertaken by wealth-maximizing individuals or firms if the expected return from the investment (which is equal to the net internal rate of return) is greater than the market rate of interest. Regarding the costs of human capital investments, Perri (2003) remarks that - if the best alternative of an investment in specialized human capital is investing in another specialization of human capital - the measure of foregone earnings has to cover the complete opportunity costs of specialized education. These opportunity costs describe what could have been earned with the best alternative specialized education.[6]

According to Haley (1973), there are two streams of human capital lit-

[3] Mincer (1962) estimates the contribution of on-the-job training to the aggregate stock of human capital to be about 50%, while Heckman, Lochner, and Taber (1998a) estimate a fraction of only 23%.

[4] Parsons (1974) distinguishes these three major components of education costs.

[5] Bloch and Smith (1977) find a positive correlation of human capital and labor market employment. Also Mincer (1989) suggests that the probability of being unemployed decreases with the amount of education.

[6] Rosen (1983) suggests increasing rates of return in the utilization of human capital due to fixed investment costs independent of the degree of utilization. This induces private incentives for specialization, i.e. to use one type of human capital as intensively as possible. Hence, each individual has a comparative advantage for a certain occupation that uses the accumulated skill most intensively.

erature based on this original approach to human capital formation. The first analyzes individual investments in human capital in order to estimate the internal rate of return (cf. Section 1.4). The second stream of literature deals with the life-cycle of earnings. The individual faces a trade-off between producing additional human capital and renting his existing stock of human capital in the labor market (cf. Section 1.6).

## 1.3 General and Specific Human Capital

In this section, we distinguish between general and specific human capital and analyze the different implications for human capital investments by workers and firms in perfect and imperfect labor markets. We first discuss the main theoretical contributions and then refer to some empirical findings regarding the theoretical predictions. In a nutshell, there are two key questions considered in this section: First, who pays for investments in human capital depending on the kind of human capital and the competitiveness of labor markets? And second, is the total amount of human capital investment efficient?

General human capital is defined to be not only useful with the current employer but also with other potential employers. In contrast, specific human capital increases the productivity of the worker only in his current job (Becker (1964)).[7] Empirically, it is difficult to distinguish between general and specific training. Loewenstein and Spletzer (1999) try to overcome this problem by directly asking employers whether they assess the provided training to be general or specific.

[7] Parsons (1974) notes that firm-specific human capital is analytically equivalent to transfer costs for adjusting a worker to other firms.

### 1.3.1 General Human Capital

#### Investment in General Human Capital: Perfect Labor Markets

In perfect labor markets, where workers receive wages equal to their marginal product, firms cannot recoup investments in general skills, which implies that they refuse to pay for general training (cf. Table 1.1). This "hold-up" problem arises due to incomplete contracts, which means that one party (i.e. the employer) pays the costs of the investment in human capital, while another party (i.e. the worker) shares in the return (Acemoglu and Shimer (1999)).

However, concerning the aggregate welfare of workers and firms, workers efficiently invest in general human capital because they are the sole beneficiaries of their increased productivity (either with their current or with future employers). Furthermore, workers can finance such investments quite easily by accepting a wage below their productivity during the period of training (the wage may even be negative) (Becker (1962)). For example, this argument can be applied to apprenticeship systems in earlier centuries, where apprentices often paid fees or worked for very low wages until they mastered a certain grade (Hamilton (1996)). Furthermore, the ongoing relevance of this argument for the German apprenticeship system is illustrated in Chapter 2. Hence, if workers are not credit constrained, they efficiently invest in the accumulation of general human capital (cf. Table 1.1).[8]

The empirical evidence of the model by Becker (1964) is mixed. On the one hand, it is supported by the empirical analysis of Veum (1999). By using data from the National Longitudinal Survey of Youth (NLSY), he finds that firm-sponsored training is indeed negatively related to starting wages, but positively related to wage growth. On the other hand, many analyses question the validity of this explanation by showing that there are investments in general human capital which are financed by the employer.

[8] Already Eckaus (1963) criticizes that this result strictly depends on the assumption of perfect labor markets.

| **Skill type** | **Labor markets** | **Firms** | **Workers** | **Total investment** |
|---|---|---|---|---|
| general | perfect | no | yes | efficient |
| general | imperfect | yes | yes | inefficient |
| specific | perfect | yes | yes | inefficient |
| specific | imperfect | yes | yes | inefficient |

Table 1.1: Investment in Human Capital

For example, by further analyzing data from the NLSY, Loewenstein and Spletzer (1999) find that the larger part of firm-sponsored training is general. Other empirical studies also show that firms bear substantial net costs in providing general training to their apprentices. For example, Ryan (1980) examines welder apprentices in the US and Jones (1986) analyzes apprentices in British manufacturing. The costs of apprenticeship training in Germany are discussed in Section 2.2.3.

A number of studies also investigate whether workers taking part in general training programs pay for the costs by accepting lower wages. The majority of these studies do not find evidence of lower wages, at least not in an appropriate amount to fully compensate firms for the costs. An overview of these results is provided by Bishop (1997). Hence, in contradiction to the theoretical predictions of Becker (1964), there is at least some empirical evidence of firm-sponsored investments in the general human capital of their employees.

#### Investment in General Human Capital: Imperfect Labor Markets

In order to give a theoretical explanation for the empirical evidence of firm-sponsored general training, Acemoglu and Pischke (1998a) develop a model with two periods, a training period where workers have identical productivity zero and may receive an amount of general training $t$ at costs $c(t)$, and a second period where workers have an individual productivity $f(t)$ and earn a wage $w(t)$. If labor markets are competitive and workers are not credit constrained, the results of Becker (1964) hold: firms do not invest in general training and workers invest efficiently by equating marginal returns and marginal costs of their investment:

$$f'(t^*) = c'(t^*)$$

However, if labor markets are not perfect or there are other labor market frictions which generate wage compression, the worker's wage is below his marginal product. If the wage structure is compressed, general skills are turned into de facto specific skills and firms manage to skim labor market rents depending on the amount of training. Formally, Acemoglu and Pischke (1999b) express this by assuming $f(t) = w(t) + \triangle(t)$. Hence, the wage function increases with the level of training less steeply than productivity (i.e. the wage structure is compressed), which implies that the firm's profit, equal to the positive gap $\triangle(t)$ between productivity and wage, has a first derivative greater than zero. As a consequence, firms prefer more skilled workers to less skilled ones and invest in general training until the desired level of training satisfies $\triangle'\left(t^f\right) = c'\left(t^f\right)$ (cf. Figure 1.1).

Unfortunately, this advancement in the hold-up problem of the firms (compared to the situation with perfect labor markets) is achieved at the expense of an additional hold-up problem in the training decision of the workers, which implies that the total amount of human capital investment is generally inefficient with respect to the aggregate welfare of workers and firms (Roed and Strom (2002)) (cf. Table 1.1).

Concerning the empirical evidence, Loewenstein and Spletzer (1998) find that general training raises future wages more for workers who change their job than for workers who remain with the training firm. This result is consistent with workers and employers sharing the returns to general training. Furthermore, Brunello (2002) suggests that wage compression and the amount of general training show a positive and significant correlation.

The reasons for inefficiently low investments in general training are summarized in Malcomson, Maw, and McCormick (2003): (1) imperfect capital markets (i.e. the workers are credit constrained), (2) incomplete contracts

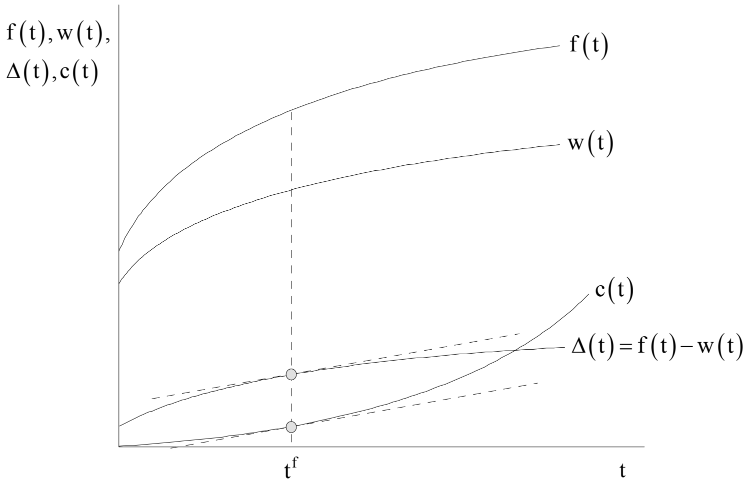


Figure 1.1: Training with a Compressed Wage Structure (Source: Acemoglu and Pischke (1998a), Figure 2)

(i.e. the desired level of training cannot be specified by a contract), (3) positive external effects of human capital investments for potential future employers if there is a positive probability of exogenous separation after the training period, and (4) the presence of labor market imperfections. For this reason, the sources of labor market imperfections are discussed in the following subsection.

**Sources of Labor Market Imperfections**

There are several possible sources of labor market imperfections which generate a compressed wage structure. The first one refers to the presence of transaction costs, for example, due to matching and search frictions. Search frictions derive from imperfect information about potential contractual partners, heterogeneities, the absence of perfect insurance markets, limited mobility, congestion due to large numbers, and other similar factors (Petrongolo and Pissarides (2001)). In practice, it is difficult for workers to quit their existing jobs and find new suitable employers. Similarly, it is costly for firms to replace their employees. The costs of finding new contractual partners create a bilateral monopoly situation in wage determination so that the match-specific surplus has to be shared by bargaining. The Nash bargaining solution implies that the output is distributed between workers and firms according to their bargaining power. If the parameter $0 \leq \beta \leq 1$ indicates the bargaining power of the workers, profits are equal to $\triangle(t) = (1-\beta) f(t)$ (Acemoglu (1997)).[9]

Furthermore, wage compression may arise due to the interaction of general and specific skills. If general and specific skills are complements in the production of output, the presence of specific skills increases the productivity of general human capital. On the other hand, the value of firm-specific skills increases when general skills are acquired (Acemoglu and Pischke (1999b)). Kessler and Luelfesmann (2002) as well as Balmaceda (2001) extend this idea by designing a model with general and specific human capital that constitute strategic complements although returns and costs are technologically disconnected. They find that there is firm-sponsored general training because the hold-up problem of investments in general skills is reduced. According to Bougheas and Georgellis (2004), this interaction of general and specific human capital is the main reason for German firms to offer apprenticeship training positions although training is largely general (cf. Chapter 2).

A third source of wage compression is the presence of asymmetric information between the current firm and other potential employers. There are two possible types of asymmetric information. The first concerns the amount of training the worker has received and is analyzed by Chang and Wang (1996). If potential employers cannot observe the correct productivity and thus pay a wage below the marginal product, the wage structure is compressed. With respect to the German apprenticeship system, this explanation

[9] Nash-bargaining implies that firms and workers maximize the Nash-product $[f(t) - w(t)]^{(1-\beta)} w(t)^{\beta}$. While $w(t)$ is the bargained wage of the worker, the firm is the residual claimant of output so that its profits are equal to the residuum $f(t) - w(t)$. If the fall-back payoffs of firms and workers are zero, the bargained wage is $w(t) = \beta f(t)$ and profits are equal to $\triangle(t) = (1-\beta) f(t)$ (Ortigueira (2006)).

for firm-sponsored general training is less important because each apprenticeship follows a prescribed curriculum (cf. Section 2.2.2). The second possible asymmetry between the current and potential employers is about the innate ability of the worker (hidden knowledge), i.e. the employer learns about the ability of the worker by providing general training (Acemoglu and Pischke (1998b)).

A fourth reason for wage compression is the presence of asymmetric information between the worker and the current employer concerning the worker's effort (hidden action). Hence, wages must satisfy the incentive compatibility constraints which leads to a compressed wage structure (Acemoglu and Pischke (1999b)). In a similar model, Loewenstein and Spletzer (1998) demonstrate that efficiency wages (that are paid to reduce fluctuations) can also induce firms to pay for general training.

Many authors have investigated similar sources of firm-sponsored general training. For example, Bishop (1997) and Lazear (2003) point out that the firm-specific mixture of general skills makes the labor market noncompetitive. Furthermore, wage compression can also be generated by labor market institutions, for example, minimum wages (Acemoglu and Pischke (1999a)) and worker unionization (Freeman and Medoff (1984)).

### 1.3.2 Specific Human Capital

#### Investment in Specific Human Capital: Theory

According to Becker (1964), specific human capital is different from general human capital because workers do not benefit from higher productivity after changing their jobs. Both in perfect and imperfect labor markets, firms can recoup investments in specific skills and thus are willing to share some of the costs of these investments (cf. Table 1.1).

The accumulation of specific human capital leads to lower fluctuations because both firms and workers benefit from keeping their contractual partner (Becker (1962)). In a search model with economic growth and endogenous accumulation of specific human capital, Higashi (2002) confirms that investments in specific human capital reduce the number of quits. This result can be split into two different effects depending on who pays for the investment. Firm-sponsored specific training reduces layoff rates, while worker-financed specific training leads to lower quit rates (Parsons (1972)). Donaldson and Eaton (1976) stress that firms may manipulate the workers' wage profile by investing in specific skills in order to reduce turnover. Hence, the negative relationship between wages and labor turnover creates incentives for firm-sponsored investments in specific human capital (Rosholm and Svarer (2004)).

According to Becker (1964), many sharing rules of costs and returns are possible and the optimal sharing rule depends on the correlation between wage and turnover rate. As a corner solution, specific human capital may be no shared investment if the firms manage to keep the whole return (Donaldson and Eaton (1977)). In this context, Prendergast (1993) describes a dual moral hazard problem: first, workers have an incentive not to accumulate specific skills if it is costly for them; and second, firms have an incentive not to reward the acquired specific skills. As a consequence, the total amount of human capital investment is generally inefficient concerning the aggregate welfare of workers and firms (cf. Table 1.1). By interpreting the distribution of the investment costs in the context of the Coase theorem, Hashimoto (1981) tries to determine the sharing rule of the return to specific training. The key feature is the existence of transaction costs while the wage is set so as to maximize the expected total surplus.[10]

By using a dynamic model of wage determination in the presence of specific human capital, Felli and Harris (1996) show that the worker receives the full value of the match with an alternative employer. More precisely, there are three components of the wage: the worker's expected productivity in the

[10] The Coase theorem predicts the following: If there is an externality and the transaction costs of trading this externality are not too high, (decentral) bargaining will lead to an efficient allocation no matter how the property rights are distributed (Mas-Colell, Whinston, and Green (1995)).

alternative match, a premium reflecting the accumulation of human capital specific to the alternative match that the worker forgoes by staying with the current employer, and a reduction reflecting the human capital specific to the alternative match that the worker also obtains by staying with the current employer. In a dynamic matching model, Arozamena and Centeno (2006) analyze the interaction of job tenure and external labor market conditions in the wage setting process. As the employment relationship evolves (and more match-specific human capital is accumulated), external labor market conditions (particularly unemployment and real growth) exert less influence on the wage.

**Investment in Specific Human Capital: Empirical Evidence**

Empirically, Lynch (1991) finds that individuals with on-the-job training are less likely to leave their current employer while individuals with off-the-job training are more likely to quit. Loewenstein and Spletzer (1997) show that individuals with company training are less likely to quit their job, whereas individuals with school training have mobility characteristics similar to those with no training. If on-the-job training is more specific than off-the-job training (as presumed by Lynch (1991)) and company training is more specific than school training (as suggested by Loewenstein and Spletzer (1997)), these job mobility results are in line with the theoretical predictions of Section 1.3.2. In almost the same manner, Loewenstein and Spletzer (1999) find that specific training and job mobility are negatively correlated.

An empirical study for Switzerland shows that specific training reduces both job search activity and job mobility while general training significantly increases job search (Zweimüller and Winter-Ebmer (2000)). Based on data from the German Socio-Economic Panel (GSOEP), Bougheas and Georgellis (2004) find that labor turnover is negatively correlated with tenure because firm-specific human capital is accumulated during employment.

As summarized in Table 1.1, Section 1.3 of this literature review analyzes the different implications for investments in general and specific human capital by workers and firms. In perfect labor markets, all costs and benefits of general human capital are borne by the workers, while firms and workers share both the costs and the returns of investments in specific human capital (Becker (1964)). In imperfect labor markets, general training may also be firm-sponsored because the wage structure is compressed, which implies that firms manage to skim labor market rents depending on the amount of training (Acemoglu and Pischke (1998a)).

## 1.4 The Rate of Return to Education

In order to show the empirical relevance of the theoretical contributions in Sections 1.2 and 1.3, this section provides an overview of various empirical studies measuring the rate of return to education from an individual's point of view. We first refer to the empirical approach by Mincer (1974) and then report the empirical evidence. Furthermore, we verify whether the returns to education differ systematically, for example, depending on gender or the type of skill and finally discuss the main problems of the empirical analysis.

### 1.4.1 Empirical Approach

The rate of return to education is analyzed by two different branches of human capital literature. According to the micro labor literature, the rate of return to education measures the extra earnings of a worker for an additional year of schooling or training, while the macro growth literature investigates whether the level of education in a cross-section of countries is related to the GDP growth rates (Krueger and Lindahl (2001)). In the following, we focus on the micro literature because it is the rate of return to education that determines the amount of human capital investment at the individual level (cf. Section 1.2).

According to Mincer (1974), if the only costs of an additional year at school are foregone earnings and if the effected proportional income increase is constant over one's lifetime, the logarithm of earnings is linearly dependent on the years of schooling. This approach yields the following Mincerian wage equation:

$$\ln W_i = \beta_0 + \beta_1 S_i + \beta_2 X_i + \beta_3 (X_i)^2 + \varepsilon_i \tag{1.1}$$

$W_i$ denotes the wage of individual $i$, $S_i$ represents the years of schooling, $X_i$ is a measure of work experience, and $\varepsilon_i$ is an individual disturbance term independent of $\beta_0$ and $S_i$. Work experience is also included as a quadratic term in order to capture the concavity of the earnings profile.

The Mincerian wage equation is a log-linear transformation of an exponential function and can be estimated by OLS. The coefficients have a semi-elasticity interpretation and measure the percentage change in $W_i$ for absolute variations in the independent variables. Hence, the parameter $\beta_1$ can be interpreted as the rate of return to investments in education. Harmon, Hogan, and Walker (2003) extend this original approach by including dispersion in the return to schooling and thus treating $\beta_1$ as a random coefficient. However, empirically they do not find any time trend in mean or variance so that the deterministic Mincerian wage equation can be used quite appropriately.

### 1.4.2 Empirical Estimates

By estimating equation (1.1) on cross-sectional data from the 1960 census for the US, Mincer (1974) finds that an additional year of schooling yields a net increase of 11.5% in annual earnings.[11] Subsequently, the Mincerian wage equation has been estimated for many countries by using OLS. The results generally yield estimates of $\beta_1$ between 5% and 15%, with slightly larger estimates for women than men (Psacharopoulos (1994)).

By equating discounted costs and benefits, Becker (1964) estimates an internal rate of return to college and high school education of 13% to 28%. However, Solow (1965) argues that these large estimates are not corrected for

[11] Mincer (1974) converts the 16.2% gross increase in annual earnings to a net increase of 11.5% by factoring out increased labor force participation associated with an increase in education. In a previous paper, Mincer (1958) uses data from the 1950 census.

correlations between education and ability. In order to solve this problem, Ashenfelter and Krueger (1994) estimate the return to schooling by contrasting wage rates of twins with different levels of educational attainment. They find that an additional year of schooling generates a wage increase of about 12% to 16%. In a similar manner, by analyzing a cross-section of twins, Rouse (1999) concludes that the rate of return to education is about 10% per year of schooling. Furthermore, Arias and McMahon (2001) estimate dynamic and expected dynamic rates of return to college and high school education in the US. They find average returns of 13.3% in real terms or 11.7% after correcting for ability, family factors, and measurement errors.

Empirical evidence for developed western economies suggests that the average estimate of the return to an additional year of education ranges from 5% to 10% (Wilson (2001)). For example, for the UK Dearden (1998) finds that the average annual return to an additional year of full-time education is 5.5% for men and 9.3% for women. Comparisons with less-developed countries show that the return to education tends to be higher in latter countries (Acemoglu (2002)). However, at least some of these countries show estimated returns to human capital investments of nearly the same magnitude, for example, Belarus with 10.1% (Pastore and Verashchagina (2006)). In his empirical analysis for Thailand, Hawley (2004) finds increasing returns to education over time which are fluctuating depending on gender and the type of skill.[12]

[12] The general trend of rising returns to human capital in the 20th century is analyzed in a theoretical model by Acemoglu (2002). He suggests that the increasing supply of skills leads to skill-biased technological change. As a consequence, the demand for skills rises and thus the returns to skill and wage inequality increase. According to Lillard and Tan (1992), the individual returns to education change over time due to the increased interaction between demand for and supply of workers at each qualification level. Individuals working in an industry with rapid technological change have above-average returns to education, which can be attributed to the positive correlation of education and adaptability to new technologies in high-tech firms. As suggested by Juhn, Murphy, and Pierce (1993), rising wage inequality stems from two dimensions of inequality which have been growing over time: the return to education and within-group inequality. In this context, Mincer (1997)

In a meta-analysis of the literature on the return to education, Ashenfelter, Harmon, and Oosterbeek (1999) review 96 estimates from 27 studies and nine different countries. They find that the average OLS estimate of the return to schooling is 6.6%, whereas the average IV estimate is 9.3%.[13] Even after adjusting for a possible publication bias (because the probability of being published is higher for statistically significant results), the average IV estimate is 8.1% and still exceeds the average OLS estimate.

The dynamic character of human capital formation has implications for how human capital investments should be optimally distributed over the life-cycle. Figure 1.2 summarizes the major findings of the literature on the return to education and shows that - for a given innate ability - the rate of return to education decreases over the life-cycle. The horizontal axis represents age ($a$), which indicates the individual's position in the life-cycle. The vertical axis represents the rate of return to education assuming the same amount of human capital investment is made at each age. Optimal investment profiles equate the marginal return to an investment with the opportunity costs of funds over the life-cycle. Hence, for constant opportunity costs equal to the real rate of return ($r$), the optimal investment strategy is to invest more at younger ages and less at older ages (Shaw (1989)). For individuals with higher innate ability (i.e. $H_0' > H_0$), the curve in Figure 1.2 lies farther out to the right.

In a nutshell, the most important empirical findings with respect to the average rate of return to education are summarized in Table 1.2. In the following subsections, we refer to some important influencing factors that may have an impact on the magnitude of the return to education.

---

suggests that (log) wages are an increasingly convex function of years of schooling, which implies that growing wage inequality is essentially concentrated at the top of the wage distribution. This result is empirically confirmed by Piketty and Saez (2003), Autor, Katz, and Kearney (2005), and Lemieux (2006).

[13] IV estimates are identified by variability in schooling associated with quarter-of-birth. Individuals who are born early in the year tend to earn less. Krueger and Lindahl (2001) discuss the validity of these instruments.

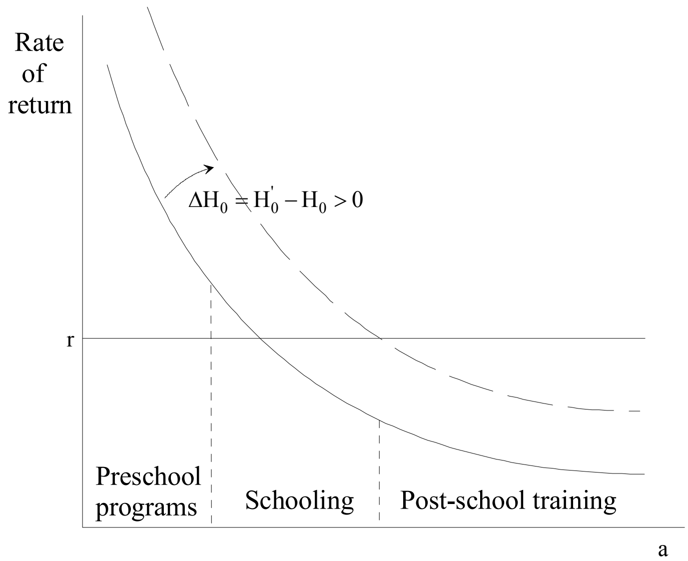


Figure 1.2: Return to Education Depending on Age and Ability (Source: Cunha, Heckman, Lochner, and Masterov (2005), Figure 1A)

| **Study** | **Estimate** |
|---|---|
| Becker (1964) | 13% - 28% |
| Mincer (1974) | 11.5% |
| Ashenfelter and Krueger (1994) | 12% - 16% |
| Psacharopoulos (1994) | 5% - 15% |
| Dearden (1998) | 5.5% - 9.3% |
| Ashenfelter, Harmon, and Oosterbeek (1999) | 6.6% - 9.3% |
| Arias and McMahon (2001) | 11.7% - 13.3% |
| Wilson (2001) | 5% - 10% |

Table 1.2: The Average Rate of Return to Education

### The Rate of Return Depending on Type of Skill

The figures in Table 1.2 represent averages for the population as a whole but the returns to education vary significantly, for example, by the type of acquired skill. According to Wasmer (2006), specific human capital yields a higher return than general human capital investments if the job-finding rate is low. By using data from the Displaced Worker Surveys, Neal (1995) empirically investigates skills that are neither completely general nor fully specific but rather common to firms operating in relatively homogenous economic activities. Because industry-switchers suffer significant wage losses, he suggests that wages strongly reflect industry-specific human capital. In contrast to this result, Cingano (2003) estimates returns to skills by using a special identification strategy and concludes that there are no returns to skills for industrial districts but high returns to firm-specific skills.

Important differences in rate of return to education have been found for different subjects taken in higher education. In the UK, men with chemistry or biology degrees have returns below average while women with education, economics, accountancy or law have significantly higher returns compared to other subjects (Blundell, Dearden, Goodman, and Reed (1997)). In a dynamic model of college and major choice, Arcidiacono (2004) estimates the returns to different majors in order to find reasons for ability sorting across majors.

Furthermore, individuals who completed schooling with some formal qualification have significantly larger returns than individuals with the same amount of schooling but without any formal qualification (Dearden (1998)). Concerning the level of qualification, Blundell, Dearden, Meghir, and Sianesi (1999) report that the average annual return to an O level qualification exceeds the average annual return to an A level or higher education qualification. Evidence from the US also suggests that there are decreasing returns to successive investments in human capital, i.e. the rate of return to education declines with the level of education (Hanoch (1967)).[14]

[14] Duncan and Hoffman (1981) distinguish between received and required level of edu-

Early US studies as well as studies from developing countries find that the return to education is largest for investments in primary education (Psacharopoulos (1994)). With respect to basic skills, Tyler (2004) estimates a model relating cognitive skills (measured by a post-schooling math test), schooling, and earnings. He concludes that basic cognitive skills matter for earnings of young dropouts. Increasing their basic skills by one standard deviation leads to 6.5% higher earnings within the next three years.

### The Rate of Return Depending on Gender

In the UK, the average annual return to a first degree in terms of hourly wages (compared to just A levels) is in the range of 5% to 8% for men and 10% to 13% for women (Blundell, Dearden, Meghir, and Sianesi (1999)). Studies from other countries also find that investments in women's education tend to yield higher rates of return than investments in men's education. For example, Butcher and Case (1994) find higher returns for women in the US. In this context, Mincer and Opek (1982) suggest that the restoration of human capital - after labor market interruptions associated with human capital depreciation - is more efficient than the accumulation of new human capital by men who stay inside the labor market the whole time.

This gender difference in the returns to education arises because the earnings of women are considerably lower than those of men (Blundell, Dearden, Meghir, and Sianesi (1999)). The gender wage gap can be decomposed into three different parts: gender differences in human capital accumulation, occupational sex segregation, and discrimination[15] (as residuum) (Kanazawa

cation. Building on the econometric specification by Verdugo and Verdugo (1989), Rubb (2003) finds that the premium for overeducation is about the same magnitude as the penalty for undereducation but lower than the reward for required education. By using the same econometric specification, Bauer (2002) confirms these results for Germany. In a matching model with ex-post wage bargaining, Charlot, Decreuse, and Granier (2005) demonstrate that both over- and undereducation may take place in equilibrium, depending on the relationship between private and social returns to education.

[15] Bloch and Smith (1977) and Bloch and Smith (1979) find sex discrimination much

(2005)). According to Winter-Ebmer and Zweimüller (1992), occupational sex segregation can have three different reasons: different preferences for various occupations[16], crowding (i.e. disadvantages in "male jobs" leading to oversupply in the more "female jobs"), and human capital theory. With respect to the latter argument, Blackburn (2004) empirically finds that men perform better in math-oriented tests and women better on speed-oriented tests. However, he argues that test score differences explain only a small part of the gender wage gap.

Concerning the gender wage gap due to occupational sex segregation, Polachek (1981) suggests that women tend to choose jobs with low penalties for intermittent employment because they anticipate labor market interruptions due to their fertility decision. Although this argument is rejected by England (1982), Schumann, Ahlburg, and Mahoney (1994) find that the male-female wage differential can be partially attributed to job characteristics. Furthermore, by developing a model of fertility and labor market decisions, Erosa, Fuster, and Restuccia (2002) conclude that fertility decisions generate important gender differences in turnover rates and thus in employment and wages. In an empirical study for apprentices in West Germany, Kunze (2005) verifies a gender wage differential of about 25% that is attributed to occupational segregation. However, Blau (1998) suggests that the convergence in male and female college majors may be responsible for a reduction in the gender wage gap during the 1980s.

### The Rate of Return Depending on Family and School Quality

Empirical research has also highlighted the importance of other factors in affecting the rate of return to education. These factors are, for example, innate ability (Dearden (1998)), family income, parental education, and the

more important than race discrimination.

[16] If different occupations require different skills and people are differently equipped with these skills, individuals have comparative advantages for one occupation leading to self-selection into different occupations (Paglin and Rufolo (1990)). Another explanation for the development of comparative advantages is given by Rosen (1983).

number of siblings (Butcher and Case (1994)) as well as school quality and the proximity to a college (Card and Krueger (1992)). For instance, Lam and Schoeni (1993) show that in Brazil the estimated returns to schooling decrease by about 30% if parental schooling is added as a family background variable to the wage equation.

Card and Krueger (1992) use state-level data to estimate the effect of school quality on the return to education for men born between 1920 and 1949. They find that men educated in states with lower student/teacher ratios, longer average term length, and higher-paid teachers have higher rates of return. By using individual-level data, Altonji and Dunn (1996) estimate the impact of parental education and school quality on educational returns. In most of their specifications, having a more educated parent is associated with a higher return to education while school expenditures per student do not have a positive effect.

In these studies, family and school effects are analyzed with respect to their impact on the return to education but they are not allowed to directly affect the process of human capital formation. The importance of this shortcoming will become more evident when contrasted with the concept of education production functions. This stream of literature takes the opposite position and assumes that the schooling environment directly influences the decision how much education to obtain (cf. Section 1.5). For example, Wilson (2001) estimates a model that allows family, neighborhood, and school characteristics to affect educational attainment both indirectly through the expected return to education and directly as inputs into the process of educational production. She finds that background characteristics predominantly work directly and not via their impact on the rate of return to education.

### 1.4.3 Problems of the Empirical Analysis

There are several problems in empirically estimating the true causal effect of education on individual earnings. The most discussed of them is whether higher observed earnings for better-educated workers are caused by their higher education or whether individuals with greater innate ability and thus higher earning capacity choose to acquire more education. If econometric specifications omit the direct influence of individual ability, the estimated return to education is biased upward, which is referred to as ability bias due to the positive correlation of ability and education.

However, even with ability as an independent variable empirical estimates will be biased because they are unable to separate the contribution of ability from that of education (Heckman and Vytlacil (2000)). IV studies that use exogenous variations in the schooling decision indicate that the estimated OLS values are too small, i.e. the OLS estimates are biased downward. The magnitude of this endogeneity bias is controversial in the literature. While Angrist and Krueger (1991) suggest a limited impact of endogeneity, Harmon and Walker (1995) find a rather large effect. In this context, Murnane, Willet, Duhaldeborde, and Tyler (2000) argue that at least 50% of the full return to higher educational achievement can be attributed to individual ability.

Finally, Lindsay (1971) notes that a correct measure of the return to education would consist in the wealth effect of increased wages due to human capital investments, but not in the pure income difference. The substitution effect of increased wages implies that leisure is substituted for hours worked (if the substitution effect is larger than the income effect), which generates an upward bias in the estimated return to education that is positively related to the size of the investment.

As summarized in Table 1.2, Section 1.4 of this literature review provides an overview of various empirical studies measuring the return to education from an individual's point of view according to Mincer (1974). Although the rate of return to education varies significantly in response to the type of skill, gender, and other influencing factors, the average estimate for developed economies generally ranges from 5% to 10% (Wilson (2001)).

## 1.5 The Education Production Function

In this section, we analyze the literature on educational production functions and discuss the significance of potential inputs into the process of educational production. In the following subsections, we present the findings of various empirical studies that investigate potential arguments of the education production function, for example, the level of resources or the influence of family background and peer groups. Finally, we refer to two models of educational production that take into account some of these inputs discussed before.

According to Hanushek (1971), the severest problem in educational research is the complexity of the educational process. Hence, there is considerable confusion about how empirical studies should be conducted and interpreted (Hanushek (1979)). As suggested by the concept of X-inefficiency in the theory of the firm, inputs in the process of educational production may not be converted efficiently to outputs (Hanushek (1996a)). There are two kinds of potential inefficiencies: the misallocation of resources between different units (although the resources may be used efficiently in each unit) and the inefficient use of given resources (although they may be allocated efficiently) (Bishop and Woessmann (2002)).

Many studies focus on the effects of school characteristics on educational attainment within the framework of education production functions. The relationship between inputs and educational output is assumed to be deterministic and depends on the technology of the education production function. The measurement of school inputs and outputs varies from study to study, but for about two thirds of the studies, the output of educational production is measured by test scores.[17] The other one third focuses on quantity

[17] We do not address the question whether test scores are adequate for measuring educational output. By using data from the National Education Longitudinal Study of 1988, Rose (2006) suggests that, at least for women, test score gains affect both the status of employment and earnings. Also Murnane, Willet, Duhaldeborde, and Tyler (2000) find that one standard deviation increase in mathematics performance at the end of high school translates into 12% higher annual earnings. Furthermore, by using international data on

of schooling achieved, such as high-school graduation or college attendance (Bishop and Woessmann (2002)).

### 1.5.1 The Level of Resources

The relationship between school expenditures and educational outcome reached a lot of public awareness with the so-called Coleman Report that analyzes the equality of educational opportunities. In brief, it concludes that the level of resources has only little impact on educational attainment (Coleman (1966)). Subsequently, various studies have presented mixed conclusions concerning the dependency of student achievement on public expenditures.

In his influential work, Hanushek (1986) reviews 147 regressions (taken from 33 separate published studies) concerning the effects of school characteristics on educational attainment. According to Hanushek, Rivkin, and Taylor (1996), the separate studies are different along two dimensions: the level of data aggregation (in many studies data are aggregated at state level) and the degree of control for other variables that potentially influence the educational performance (for example family background). In a nutshell, Hanushek (1986) compares the sign and the significance of the estimated effects of school inputs. Due to a lack of consistent findings, he concludes that "there appears to be no strong or systematic relationship between school expenditures and student performance", at least after variations in family inputs are taken into account.

Hedges, Laine, and Greenwald (1994) reexamine the same studies like Hanushek (1986) but draw a much different conclusion by summarizing that expenditures are positively related to school outcomes. While Hanushek (1986) uses a "vote counting" selection rule of weighting the separate studies (i.e. the results of each regression receive one vote), Hedges, Laine, and Greenwald (1994) use a more sophisticated method accounting for the size of the estimate, the expected correlation in the error terms from regressions

test scores and wages for eleven countries and two birth cohorts, Bedard and Ferrall (2003) show that test score dispersion and wage dispersion are positively related.

estimated over the same sample, and the potential influence of outliers.

By applying more explicitly defined search criteria to the selection of relevant studies, Greenwald, Hedges, and Laine (1996a) again conclude that educational resources are positively related to student achievement.[18] In a further approach, Hanushek (1997) reviews about 400 studies in order to investigate the relationship between student performance and school resources. He finds that most estimates are not simultaneously positive and significant.

Unfortunately, also more recent studies are not able to resolve whether the level of resources is a significant determinant of educational attainment (cf. Table 1.3). On the one hand, Wilson (2002) uses US data to examine the impact of school expenditures on earnings. By controlling for extensive measures of family background and neighborhood, she finds that school expenditures positively affect both earnings and returns to education. Furthermore, Winter-Ebmer and Wirz (2002) estimate that raising educational expenditures by 1% leads to an increase in the college enrollment rate by 1%. On the other hand, by using data from the Third International Mathematics and Science Study (TIMSS), Hanushek and Luque (2003) compare the performance in different schooling systems and different countries. They find evidence of inefficient input-based schooling policies, independent of the income level of the country and the resource level of the school.

### 1.5.2 Class Size

In line with Greenwald, Hedges, and Laine (1996b), Krueger (2003) criticizes the conclusions of Hanushek (1986) and proposes a different selection rule, namely "equal weight to every study". With this method class size becomes a significant determinant of student achievement. Also Tennessee's Project STAR, a random-assignment experiment, seems to provide a rationale for

[18] Hanushek (1996b) criticizes this "specialized meta-analytic approach" by pointing out systematic distortions towards the desired results. Greenwald, Hedges, and Laine (1996b) reject this reproach by explaining their different understanding of "statistical independence".

| **Input** | **Empirical evidence** |
|---|---|
| level of resources | ambiguous |
| class size | weak |
| private schools | ambiguous |
| school competition | weak |
| teacher quality | strong |
| teacher incentives | strong |
| early education | strong |
| individual ability | strong |
| parental education | strong |
| family income | weak |
| neighborhood | weak |
| peer groups | strong |
| segregation | ambiguous |

Table 1.3: Inputs in Educational Production

the performance-enhancing effect of class size reductions.[19]

Dustmann, Rajah, and Soest (2003) draw a similar conclusion in their analysis of class size, education, and wages. They suggest that lowering the pupil-teacher ratio (as measure for class size) causes an increase in the wage rate. This effect is generated by two channels: a direct effect of reduced class size on wages and an indirect effect by increased probability of staying in school at age 16. In an empirical study for Italy, Brunello and Checchi (2005) analyze the dependency of educational attainment on school quality (measured by the pupil-teacher ratio) and family background (mea-

[19]However, Hanushek (1999) argues that there are several experimental features biasing upward the estimates of Project STAR. His objections are in detail: (1) design and implementation issues, (2) the estimated effect is only valid if students are very young, (3) the estimated effect is only valid for very large reductions in class size, (4) teacher quality is much more important than class size reductions, and (5) the costs of such class size reductions have to be considered.

sured by parental education). They conclude that school quality and family background both positively affect educational achievement and constitute complements in the production of human capital.

Although Heckman, Layne-Farrar, and Todd (1996) find only little empirical evidence for the connection between class size and earnings, Hanushek, Leung, and Yilmaz (2003) agree that class size constitutes the most validated input factor of all school resources. Altogether, the empirical evidence of class size reductions is weak but predominantly positive (cf. Table 1.3).

### 1.5.3 Private Schools and School Competition

Instead of additional school expenditures, Hanushek (1997) proposes to reorganize the process of educational production by providing institutional incentive structures (i.e. output-based schooling policies). Hanushek, Leung, and Yilmaz (2003) suggest two types of efficiency-enhancing institutional features: acceptance of private schools and increased competition between schools (whether public or private).[20]

However, in the empirical literature, it is controversial whether there is a different impact of private and public education on student achievement. For example, Evans and Schwab (1995) find that private schools outperform public schools, but Goldhaber (1996) does not confirm this empirical result. In their empirical analysis of undergraduate students at Ball State University, Horowitz and Spector (2005) suggest that religious private schools perform slightly better than other schools, but this influence is small and seems to disappear in later years. By evaluating the effectiveness of private education across countries with data from the PISA 2000 study, Vandenberghe and

[20] More extensively, Bishop and Woessmann (2002) refer to six institutional features in the process of educational production: the degree of centralization of the examination system, the degree of centralization of school organization, the distribution of responsibilities between different administration levels, the degree of teacher unionization (as measure of teacher influence), the degree of parental influence (for example in choosing which school their children should attend), and the degree of competition between schools.

Robin (2004) conclude that private education does not generate systematic benefits.

Furthermore, the empirical evidence whether private school competition improves the outcome of public schools is not without controversy. While Hoxby (2002b) suggests that private school competition leads to a better performance of public schools, Arum (1996) argues that this result may be only due to increased funding for public schools rather than due to the competition itself. In this context, school vouchers may be an important instrument to allocate monetary stipends so that parents can send their children to the schools (either public or private) they assess to be the most suitable. With respect to the Mikwaukee voucher experiment, Rouse (1998) finds substantial benefits of school vouchers.

In a nutshell, while competition among schools seems to have a positive effect on educational production, the empirical evidence concerning the promotion of private schools is mixed (cf. Table 1.3).

### 1.5.4 Teacher Quality and Teacher Incentives

According to Hanushek (1989), differences in school quality are generated by "teacher skills" that are not strongly related to teacher education, teaching experience, and class size.[21] By using student-level data, Hanushek, Kain, and Rivkin (1998) suggest that at least 7.5% of the total variation in student achievement can be explained by teacher fixed effects. Hence, educational reforms should focus on improving the quality of the teacher force (cf. Table 1.3). This requires a new set of incentives by introducing selective hiring, retention, and pay (Hanushek (2006)).

Furthermore, teachers in public schools face weak performance incentives because tenure is granted and the wage is generally independent of effort and outcome. Ballou and Podgursky (1998) and Ballou (1999) investigate the causes and consequences of weak incentives among public school teachers.

[21] These teacher skills "cannot be described correctly, but possibly can be observed directly" (Hanushek (1986)).

Relative to public schools, private schools adopt hiring and pay practices which favor teachers with better overall academic ability. Hence, school vouchers and competition between public and private schools may generate changes in hiring and pay practices that attract more talented individuals into the teaching profession (Hoxby (2002a)).

In a nutshell, it not only the teacher quality but also the teacher incentives that significantly affect the educational outcome. For this reason, Hanushek (1997) and Hanushek, Leung, and Yilmaz (2003) propose merit pay for teachers and schools in order to provide better institutional incentive structures.[22] Lazear (2000) analyzes existing compensation schemes that are either input-based, output-based, or a combination of them and can be summarized by the following pay structure:

$$w_i = \alpha q_i + \beta \{q_i > q_{\min}\} + \gamma e_i + \delta \{e_i > e_{\min}\}$$

$w_i$ represents the wage of teacher $i$, $q_i$ is a measure of output, and $e_i$ describes the teacher's effort. $\beta$ and $\delta$ are dummy variables equal to one if a certain minimum output $q_{\min}$ or minimum effort $e_{\min}$ are exceeded. In brief, $\alpha$ and $\beta$ are parameters related to an output-based compensation scheme, while $\gamma$ and $\delta$ are related to an input-oriented pay structure. Furthermore, the dummies represent a compensation in discrete steps, while $\alpha$ and $\gamma$ stand for continuously paid wages.

If $\alpha$ is the only parameter different from zero, the pay structure is said to produce "high-powered incentives". Concerning the merit pay for teachers, the optimal compensation scheme depends on the observability of output and the heterogeneity of teachers (Lazear (1986)). Compared to input-based wages, output-based compensation schemes lead to strong sorting of teachers according to their productivity (Lazear (2000)).

[22] The role of economic incentives at universities is empirically analyzed by Lach and Schankerman (2003). They find that increasing the share of license royalties received by academic inventors (i.e. the share not passed on the general university budget) enhances research activity by two channels, namely by increasing the research effort and by sorting (i.e. attraction of better researchers). This effect is much stronger for private universities.

### 1.5.5 Early Education

Ritzen and Winkler (1977) and Psacharopoulos (1994) propose to promote very young and high-ability children because their returns to education exceed those of older and less able students, respectively. According to Heckman and Masterov (2004), the main mechanisms through which early education affects productivity is through its impact on cognitive and non-cognitive skills. Furthermore, the impact of peer abilities appears to be large for the skill development of very young children in preschool (Henry and Rickman (2007)).

Early education may be more effective than offering costly training for those experiencing difficulties graduating from high school (Heckman (2000)). This result is empirically confirmed by Temple and Reynolds (2007) who show that there are consistently positive returns of preschool programs that exceed most other educational interventions, especially later programs such as class size reductions and youth job training (cf. Table 1.3).[23]

### 1.5.6 Family Background and Neighborhood Effects

Haveman and Wolfe (1995) provide an overview of empirical estimates concerning the influence of family and neighborhood variables. The strong effect of parental education on the children's educational success stands out in this research. For example, by estimating the influence of maternal education on their children's achievement, Rosenzweig and Wolpin (1994) find that each additional year of schooling increases the children's test scores significantly by 2.4%. Loury (2006) suggests that also older extended family members (i.e. aunts, uncles, and grandparents) independently affect the schooling of same-gender children. Furthermore, the economic status of the family (measured by total family income) also tends to be positively correlated with educational attainment.

[23] However, the estimated effects on early human capital formation may be only short-lived as suggested by DeCicca (2007) in his empirical analysis of full-day kindergarten.

Concerning the students' choice of alternative programs of study during high school, Zietz and Joshi (2005) suggest that family background and peer group effects constitute the most important determinants of the program choice. Students from families with higher level of education and higher income are more likely to pursue the college program. According to Astone and McLanahan (1991), growing up in an intact family is associated with substantially higher educational achievement. Hanushek (1992) finds that family size negatively affects educational outcome. He attributes this effect to the parental trade-off between the number of children and their schooling performance.

As shown in Table 1.3, neighborhood effects tend to be small relative to the influence of parental education. Nevertheless, the estimated effects are often statistically significant, even when controlling for an extensive number of family characteristics. For example, Brooks-Gunn, Duncan, Klebanov, and Sealand (1993) estimate that the fraction of families with high and low incomes affect the educational achievement of children. By using ZIP code-level data, Corcoran, Gordon, Laren, and Solon (1992) find that living in a neighborhood with more mother-only families and more people on public assistance reduces educational attainment, but that the median income and male unemployment do not have significant effects. In a similar setup, Datcher (1982) concludes that the racial composition of the neighborhood is not significantly correlated with educational outcome.

### 1.5.7 Peer Effects

The Coleman report concludes that peer effects, i.e. the external effects by some students to others, in public schools contribute to differences in the educational achievement of black and white students (Coleman (1966)). According to Epple and Romano (1998), peer effects are defined as the influence of students' mean ability on school quality. Hoxby (2000) concretizes this expression by specifying knowledge spillover, influence on classroom standard, and individual behavior (i.e. self-discipline and disruption). Most empirical studies focus on socioeconomic status indicators of peer quality, such as average income or the percentage of people with college graduation. Some studies also concentrate on the composition of peer groups, for example, the proportion of different ethnic groups.

According to Manski (1993) and Rivkin (2001), there are three problems in estimating peer effects: (1) endogeneity (i.e. self-selection due to family income or educational preferences), (2) the simultaneous interaction of students' mutual influence, and (3) the difficulty to distinguish between peer effects by individual background and peer effects by individual behavior. A number of studies attempt to solve the endogeneity problem, but so far no clear consensus exists concerning its severity.

Evans, Oates, and Schwab (1992) suggest that there are no peer effects once endogeneity is controlled for by estimating simultaneous equations. Furthermore, Arcidiacono and Nicholson (2005) investigate the effect of peer groups with respect to a non-random category of students, namely medical students who graduated from US medical schools between 1996 and 1998. They find that peer effects disappear if school fixed effects are included in order to control for endogeneity. By using fixed-effects models that rely on peer variation between siblings and controlling for parental characteristics, Aaronson (1998) find significant peer effects while Plotnick and Hoffman (1999) do not.

In order to solve the problem of self-selection, Sacerdote (2001) uses a sample of the Dartmouth College with random assignment of roommates. He suggests that - at room level as well as at dormitory level - peers indeed exert a significant influence on the Grade Point Average (GPA). Furthermore, Hoxby (2000) uses some idiosyncratic variation of students and concludes that augmenting the reading scores of the peer group by one point leads to an increase in the individual test score by 0.15 to 0.4 points.

Another approach is pursued by Lefgren (2004) who compares the academic achievement of high-ability students in segregated and integrated schools. He finds that peer effects are small but generally positive and significant. By taking data from the PISA 2000 study, Fertig (2003) analyzes the influence of achievement heterogeneity in a peer group (measured by the variation coefficient) on the individual educational attainment. In order to circumvent the endogeneity problem, he uses an IV approach to estimate the following equation:[24]

$$y_{ijs} = \alpha + \beta' z_i + \gamma C_j^{-i} + \delta' \bar{z}_j^{-i} + \lambda' z_s + \varepsilon_i$$

$y_{ijs}$ describes the educational achievement of individual $i$ in peer group $j$ and school $s$, $z_i$ is a vector of individual characteristics, $\bar{z}_j^{-i}$ is a vector of the corresponding peer group characteristics, $C_j^{-i}$ represents the variation coefficient within the peer group $j$, $z_s$ is a vector of school characteristics, and $\varepsilon_i$ is an unobserved error term. In this context, $\gamma$ measures the endogenous part of the peer effect (i.e. the direct impact of the peer group achievement), while $\delta$ describes the exogenous or contextual effects (i.e. the indirect impact of other peer characteristics).

In a nutshell, Fertig (2003) obtains the following empirical result: the higher the heterogeneity of the peer group, the lower is the individual educational outcome. As a consequence, the aggregate educational output is maximized if schools exhibit perfect homogeneity of students. However, Fertig (2003) admits that the explanatory power of $\gamma$ declines substantially when exogenous effects are controlled for.

### 1.5.8 Models of Educational Production

After the extensive discussion of potential inputs in the previous subsections, we briefly present three models of educational production that take into

[24] The variation coefficient is defined as the standard deviation devided by the mean.

PISA is the abbreviation for "Program for International Student Assessment". The PISA 2000 data contain test scores in reading, mathematics, and science literacy of representative samples of 15-year old students across OECD and non-OECD countries. Furthermore, they provide a rich set of background information about students and schools. For the empirical analysis on hand, Fertig (2003) focuses on the reading scores for US students.

account some of these inputs. By focusing on the level of resources and peer effects, Nechyba (1996) develops the following education production function that describes the educational achievement ($y_{ij}$) of individual $i$ in peer group $j$:

$$y_{ij} = (\bar{y}_j)^{\rho} (e_j)^{1-\rho} \tag{1.2}$$

In this production function (1.2), $e_j$ represents the expenditures per student in peer group $j$, $\bar{y}_j$ is the average educational outcome of peer group $j$, and $\rho$ measures the strength of the peer effect.

If the educational system allows for private schools in combination with private school vouchers, there is endogenous ability-based segregation so that each private school is completely homogenous in ability. This is shown by Nechyba (1996) who undertakes general-equilibrium simulations within the framework of his education production function (1.2). In a theoretical model with public schools (financed by taxes and without tuition) and private schools (completely financed by tuition), Epple and Romano (1998) draw a similar conclusion. They demonstrate that there is a strict hierarchy of school qualities and a two-dimensional student sorting according to ability and income. Quite intuitively, the implementation of tuition vouchers benefits high-ability students relative to students with low ability.

As explained in Section 1.5.7, the presence of peer effects gives rise to the idea of an efficient sorting of students according to their innate ability (Epple and Romano (1998)). In general, there are two possibilities of putting together students with different individual abilities: integration (i.e. randomized mixing) and segregation (i.e. sorting according to the students' abilities). Segregation implies that classes are homogenous and, theoretically, the individual's deviation from mean ability is zero.

Building on the empirical findings in Section 1.5.2, Lazear (2001) makes an important attempt to analyze the effects of segregation in a model with class size as the main determinant of educational production. In his "disruption model", $m_j$ represents class size (i.e. the number of students in class $j$) and $p$ is the probability that a student behaves so that lessons can be given without disruption. The total amount of learning ($L_j$) in class $j$ is defined according to the following education production function:

$$L_j\left(m_j, p\right) = m_j p^{m_j}$$

If there are different ability types of students with different disruption probabilities, the aggregate educational output is generally larger with segregation than with integration. Hence, in line with the empirical findings by Fertig (2003), the aggregate educational output is maximized if schools exhibit perfect homogeneity of students. However, Lazear (2001) also refers to one important exception from this conclusion: if students with low-ability can be transformed into high-ability students by undergoing social contact with these better behaving students, mixed classes may yield higher educational outcomes than segregation. Altogether, the effect of segregation on the educational achievement of students is ambiguous (cf. Table 1.3).

Finally, Caucutt (2002) expands on Nechyba (1996) and Lazear (2001) by considering different ability levels $\theta_k$ (with $n_{kj}$ as the fraction of ability type $k$ in peer group $j$):

$$y_{ij} = \alpha\left(\theta_i\right)^{\beta_1}\left(\sum_k n_{kj}\theta_k\right)^{\beta_2}\left(e_j\right)^{\beta_3} \tag{1.3}$$

Based on this education production function (1.3), Caucutt (2001) and Caucutt (2002) develop a general-equilibrium model (with schools interpreted as clubs) where the change from a public school system (with endogenous number of private schools) to a completely private system of schools with vouchers yields several mixed schools with sorting according to ability and tuition fees.

As summarized in Table 1.3, Section 1.5 of this literature review provides an overview of various empirical studies that discuss the significance of potential inputs into the process of educational production. While the empirical evidence concerning the impact of school resources is mixed, there is an unambiguous effect of family background and peer groups as well as institutional incentives within the educational system (Hanushek (1997)). As shown in Section 1.5.8, the analytical models of educational production focus on a few inputs such as the level of resources, peer groups, and the innate ability of the individual.

## 1.6 Human Capital and the Life-Cycle of Earnings

This section analyzes the life-cycle of earnings with endogenous formation of human capital. We first refer to different models of human capital accumulation over the life-cycle and then discuss their empirical evidence. Finally, we investigate the implications of pension systems for human capital formation.

### 1.6.1 Human Capital over the Life-Cycle

An important stream of human capital literature deals with the life-cycle of earnings. Mincer (1958) points out that the difference between normally distributed abilities and the positively skewed distribution of incomes must be due to investments in human capital over the life-cycle.[25] In a nutshell, Mincer (1970) and Mincer (1997) summarize the empirical evidence concerning the age-earnings profile of individuals. Earnings ($E_t$) positively depend on the stock of human capital ($H_t$) at date $t$; the age-earnings profile is concave and at least for a long time upward-sloping. If human capital investment increases, the age-earnings profile becomes steeper and has its maximum later. These empirical findings are illustrated in Figure 1.3.

Becker (1964) suggests that earnings increase with age but at a declining rate because human capital investment decreases over the life-cycle. Intuitively, this is because younger workers receive the returns to education over a longer period and the investment risk increases with age (which implies

[25] The normal distribution of abilities is also assumed by Becker (1964).

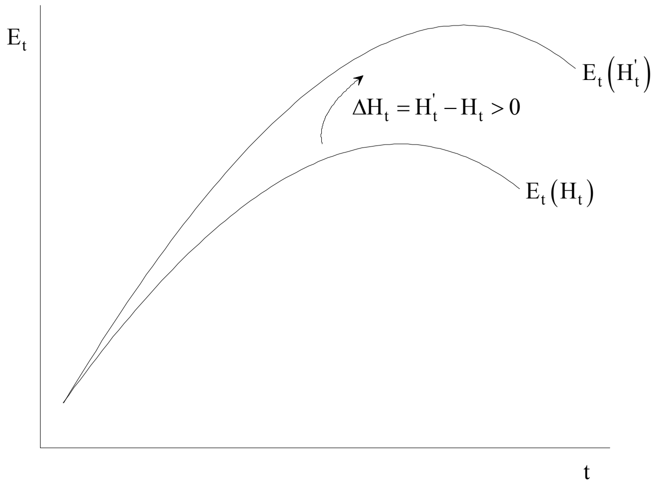


Figure 1.3: Human Capital and the Life-Cycle of Earnings

that older workers discount future earnings more heavily) (Zucker (1967)). Altogether, decreasing marginal returns and increasing marginal costs lead to an optimal amount of human capital investment that negatively depends on age (Mincer (1970)). However, human capital investment may not monotonically decline with age if the accumulated human capital is rather specific than general. While the profitability of general skills depends on the length of working life, the profitability of specific skills only depends on the expected duration of the current job (Bartel and Borjas (1977)).

Models of (general) human capital accumulation over the life-cycle can be attributed to two different branches: earnings maximizing models and utility maximizing models. Earnings maximizing models abstract from the labor-leisure choice problem and only analyze the trade-off between investment and income. At the intensive margin, the individual faces a trade-off between producing additional human capital and renting the existing stock of human capital in the labor market. Utility maximizing models also incorporate the labor-leisure choice so that labor supply becomes endogenous to the model. The difference between these two types of models is illustrated by Snow and

Warren (1990) who explain that the income effect of higher wages (due to investments in human capital) on future labor supply may reduce realized future earnings. However, there are efforts to integrate these two branches, for example, by Blinder and Weiss (1976). Weiss (1986) provides a review of the theoretical literature.

Ben-Porath (1967) develops an earnings maximizing model of human capital accumulation and gives a productivity-based explanation of earnings growing with age. He emphasizes the process by which additions to the individual stock of human capital are produced. The model is similar to Becker (1964) in the sense that each individual is presumed to combine market goods ($D_t$) with some portion ($s_t$) of his own stock of human capital ($H_t$) to produce new human capital ($Q_t$) in period $t$ according to the following production function:

$$Q_t = f(s_t, H_t, D_t) = \beta_0 (s_t H_t)^{\beta_1} (D_t)^{\beta_2}$$

The rate of change of the stock of human capital is equal to

$$\dot{H}_t = Q_t - \sigma H_t$$

where $\sigma$ represents the depreciation rate of human capital. The individual maximizes the sum of all future disposable earnings discounted over the life-cycle. The decision problem consists in choosing that portion of the existing stock of human capital to be used as input in the production of further human capital. This choice is made by comparing the costs of producing an additional unit of human capital to the marginal benefits. The whole decision problem can be divided into two separate parts: first, the trade-off between income and human capital formation, and second, the timing of consumption.

Heckman (1976) designs a utility maximizing model with endogenous labor supply, income, human capital accumulation, consumption, and non-monetary utility of education, which contains the original model of Ben-Porath (1967) as a special case. Human capital accumulation takes place according to the following equation:

$$\dot{H}_t = f(I_t H_t, D_t) - \sigma H_t$$

$H_t$ represents the stock of human capital at date $t$, $I_t$ is the time devoted to human capital production, $D_t$ stands for the direct education costs, and $\sigma$ is the depreciation rate of human capital.

Even the basic model by Ben-Porath (1967) manages to replicate the most important qualitative characteristics of the empirical life-cycle patterns. According to Ben-Porath (1967), there are three different phases of human capital accumulation: an initial phase with no earnings (i.e. full-time human capital production, interpreted as "formal schooling"), a long phase with part-time human capital production, and a last phase with no training.[26] The second phase is characterized by earnings increasing at a declining rate (and eventually decreasing). At any point in time, individuals with more schooling or greater ability invest more in on-the-job training.

A serious problem of the empirical application of human capital theory to life-cycle differences in earnings is that post-school investments are not directly observable. As a consequence, measurement problems arise because a wide range of activities might be viewed as on-the-job training (Hanushek and Quigley (1985)). By using data from the NLLS, Haley (1976) empirically estimates the model by Ben-Porath (1967) and concludes that it fits the data quite well. In contrast, Brown (1976) finds only a poor performance of the model. Finally, Heckman (1976) suggests that his life-cycle model fits the data well and even better than Ben-Porath (1967).

The hypothesis that earnings profiles are driven by human capital investment is further analyzed by Mincer (1997). For the US, he confirms that post-school investment in human capital is indeed the primary factor underlying the wage profiles. He also finds empirical support for the implications

[26] Wallace and Ihnen (1975) develop a model to describe endogenously the "formal schooling" period and the end of this phase depending on the parameters of the model. They suggest a longer "formal schooling" period than Ben-Porath (1967).

by Ben-Porath (1967) concerning the positive relationship between ability, schooling, and on-the-job training.

Besides this vast literature on human capital accumulation over the life-cycle, there is a branch of literature concentrating on the extensive margin. The extensive education decision divides the labor force into different skill groups, such as low-skilled and high-skilled workers. For example, Willis and Rosen (1979) as well as Heckman, Lochner, and Taber (1998a) derive a theoretical model of the demand for higher education and empirically show that the workers' college attendance choice is determined by their expected earnings over the life-cycle.

### 1.6.2 Pension Systems and Human Capital Formation

According to Echevarría (2003), the return to human capital investments is affected by the pension system if finite horizon economies are considered. With a tax-benefit link, the return to education is not restricted to increased labor incomes, but also extends to pensions during retirement. Hence, if workers decide on the optimal amount of human capital investments, they take into account not only the effect on future labor incomes but also on future retirement benefits (Echevarría and Iza (2005)).

Lau and Poutvaara (2001a) and Lau and Poutvaara (2001b) study the impact of social security incentives on human capital formation, arguing that actuarial fairness and a tight tax-benefit link increase human capital along with an increase in the retirement age. This is a common result in most theoretical analyses because postponed retirement lengthens the time period at the extensive margin over which individuals can appropriate the benefits from human capital investments, which translates into higher returns to education. Hence, the return to education positively depends on the remaining active years (Trostel (1993)). In a nutshell, postponed retirement raises aggregate human capital because higher returns to education are associated with increased human capital accumulation (Echevarría (2003)).

In their analysis of demographic transition and economic growth, Kalemli-

Ozcan, Ryder, and Weil (2000) show in an overlapping generations framework that augmented life expectancy gives rise to increased human capital formation. In a similar setup, de la Croix and Licandro (1999a) investigate an economy where the workers accumulate human capital as a function of their optimal schooling period. The effect of lower mortality rates on human capital formation is positive because the higher expected flow of future incomes increases human capital per capita. The same result is obtained by Boucekkine, de la Croix, and Licandro (2002) under a setting with uncertain lifetime horizon and endogenous retirement age. Echevarría (2003) argues that an increase in life expectancy translates into higher growth rates by increased human capital formation only if demographic change is accompanied by simultaneous increments in the length of the working life. Hence, if there is a positive correlation between life expectancy and retirement age, an increase in life expectancy will foster the formation of human capital. For the UK, this result is empirically confirmed by Kalemli-Ozcan, Ryder, and Weil (2000).

Furthermore, the positive relationship between retirement age and human capital accumulation also holds in the opposite direction. According to Hernoes, Sollie, and Strom (2000), education is an important determinant of the retirement age. A higher stock of human capital increases the retirement age because the worker's higher productivity implies an increased labor income and makes labor supply more worthwhile compared to retirement. Hence, while early retirement is low among high-skilled workers, low-skilled workers may take the opportunity to retire at the lowest possible date.

As graphically illustrated in Figure 1.3, Section 1.6 of this literature review summarizes the main results of human capital accumulation over the life-cycle. The two most important approaches by Ben-Porath (1967) and Heckman (1976) manage to replicate the empirical life-cycle patterns with respect to the age-earnings profile of individuals. If there is a positive link between contributions and retirement benefits, human capital formation is affected by the pension system because the return to education extends to pensions during retirement.

## 1.7 Fiscal Policy and Human Capital

Weisbrod (1966) assigns to the government the role of using its limited resources for the economic well-being of its people. This section analyzes the effects of taxation and education subsidies on the accumulation of human capital. We first discuss the implications of taxation and different tax reforms on the incentives to acquire skills by the impact on the costs and benefits of investing in human capital. After this, we investigate the effects of education subsidies because tax and expenditure issues constitute "Siamese twins" concerning the maximization of aggregate welfare (Bovenberg and Jacobs (2005)). Finally, we refer to fiscal policy in the context of globalization and factor mobility.

### 1.7.1 The Effects of Taxation on Human Capital

An ideal income tax should define income uniformly as the sum of earnings from all possible sources. Hence, human capital should be taxed like physical or financial assets (Kaplow (1994)). However, the design of tax and education policies must also consider the special characteristics of human capital (Anderberg and Andersson (2003)). Theoretically, this implies that human capital should be taxed in three ways: at birth (i.e. the present value of expected future earnings net of costs should be taxed immediately), over time (reflecting the difference between earnings and depreciation of human capital), and at all moments when uncertainty is resolved (Kaplow (1996)).

Because the ideal income tax is not feasible in reality, we analyze the effects of taxing labor and capital income within a prevalent fiscal system. Due to the special characteristics of human capital there are large differences in the taxation of human and physical capital, which leads to radically different tax burdens (Ill (1984)). In the context of his life-cycle model, Heckman (1976) shows that differences in the tax treatment of human and physical capital may induce the substitution of human capital for physical capital investment.

The effects of taxation on human capital formation can be divided into a substitution effect, an income effect, and a direct effect. While the substitution effect and the income effect exert an indirect influence on human capital investment via their impact on labor supply and the return to education (cf. the argumentation by Lindsay (1971) in Section 1.4.3), the direct effect accrues due to changes in the marginal costs and returns of investments in human capital. The aggregate indirect effect on human capital accumulation (i.e. the sum of substitution and income effect) is ambiguous because the substitution effect (negative) and the income effect (positive) have opposite signs with respect to the taxation of labor income (Heckman, Lochner, and Cossa (2002)). In the context of the life-cycle of earnings as discussed in Section 1.6, the direct effect is the only one operating in earnings maximizing models with exogenous labor supply.

In the absence of direct education costs, there is no direct effect of a proportional income tax on the accumulation of human capital because both marginal returns and marginal costs are scaled down in the same proportion. With direct costs of education, an increase in the tax rate decreases human capital investment if the net financial benefit before taxes is positive (Heckman, Lochner, and Taber (1999a)). In a nutshell, the marginal effects of proportional income taxation on human capital formation may be significantly negative for three reasons. First, some direct inputs into the process of human capital production are not tax-deductible (in contrast to foregone earnings). Second, the negative effect of taxation on labor supply reduces the return to education. And finally, the negative effect of taxation on savings reduces the amount of physical capital and thus - by general-equilibrium effects on interest rates and wages - human capital investments (Trostel (1993)). However, Eaton and Rosen (1980) suggest that the welfare-maximizing tax rate may be greater than zero in order to provide insurance against the riskiness of human capital investments.

A progressive income tax strengthens the negative effects of taxation on human capital formation by workers because marginal returns on future earnings are reduced more than marginal costs of schooling by foregone earnings (Heckman, Lochner, and Taber (1999a)). However, higher tax progression may increase the firms' investments in human capital if the degree of tax progression changes the distribution of the surplus between workers and firms and thus increases the compression of the wage structure (cf. Section 1.3.1). Hence, if labor markets are imperfect and both workers and firms can invest in human capital, higher tax progression increases human capital formation if the positive effect on the firms outweighs the negative effect on the workers (Hungerbühler (2007)). Furthermore, the progressivity of the tax system affects the job mobility of workers. On-the-job search (i.e. the probability of turnover) decreases both with the level of tax rates and with the convexity of the tax system (as measure of progressivity) (Gentry and Hubbard (2002)).

If there is only unskilled labor and physical capital, a "comprehensive income tax" (i.e. the identical taxation of labor and capital income) generates two distortionary effects on resource allocation, namely on the intratemporal leisure-consumption trade-off (due to the taxation of labor) and on the intertemporal saving-consumption trade-off (due to the taxation of capital). However, in the presence of skilled labor and human capital formation, the "comprehensive income tax" generates two additional distortions. On the one hand, it discriminates against physical capital because the capital income tax does not apply to investments in human capital. On the other hand, it discriminates against human capital because the labor income tax reduces the return to education. While the net effect of these two distortions is ambiguous in general, the "comprehensive income tax" discriminates against human capital if there are only direct costs of education and the government does not use the tax revenues to foster human capital formation by education subsidies (Nerlove, Razin, Sadka, and Weizsäcker (1993)).

In contrast, if the only costs of education are foregone earnings, Nielsen and Sorensen (1997) conclude that the "comprehensive income tax" leads to a discrimination against investments in physical capital. Hence, they propose a "dual income tax"[27] (i.e. the combination of a proportional tax on capital income with a progressive tax on labor income) on pure efficiency grounds (and not on the basis of redistribution or insurance arguments) because the progressivity of the labor income tax reduces the return to education.

### 1.7.2 The Effects of Tax Reforms on Human Capital

The first attempt to quantify the importance of human capital formation for the comparative efficiency costs of alternative tax bases is due to Driffill and Rosen (1983). In a partial-equilibrium life-cycle model, they conclude that income taxation according to the "comprehensive income tax" can be dramatically more distortionary than the taxation of consumption. As suggested by Perroni (1995), switching from income to revenue-neutral consumption taxation provides large welfare gains. In these models, endogenous human capital has only little effects because the substitutability between labor supply and human capital is limited by downward rigidities in the stock of human capital (i.e. the stock of human capital can decrease at most with its rate of depreciation). However, Judd (1998) shows that consumption taxation is generally biased against human capital investment.

There are several approaches to analyze the welfare consequences of alternative tax policies in a general-equilibrium framework with endogenous human capital formation. Starting point of various numerical simulations is the current US tax system that is approximated by a proportional tax on capital income and a progressive tax on labor income (Heckman, Lochner, and Taber (1999b)). In a general-equilibrium growth model with overlapping generations and perfect foresight, Perroni (1995) shows that the welfare gains from switching to a revenue-neutral consumption tax are much lower than in a partial-equilibrium setting because prices adjust in response to quantities.

Furthermore, in an overlapping generations dynamic general-equilibrium

[27] Nielsen and Sorensen (1997) refer to it as the "Nordic system of dual income taxation" because it was implemented in all four Nordic countries.

model with endogenous human capital formation and heterogenous agents, Heckman, Lochner, and Taber (1998b) investigate the consequences of two revenue-neutral tax reforms, namely a flat income tax (i.e. a proportional tax on both labor and capital income) and a flat consumption tax. Analytically, each individual solves a two-step decision problem: first, he determines the optimal paths of consumption and post-school human capital investment conditional on the level of schooling, and second, he chooses that level of schooling which maximizes his aggregate welfare over the life-cycle (Heckman, Lochner, and Taber (1998b)). In a nutshell, switching to a flat income tax fosters the accumulation of human capital at the cost of reduced investment in physical capital, while a flat consumption tax is more pro-capital and less favorable to human capital. Similar to Perroni (1995), general-equilibrium effects of these tax reforms are much weaker than partial-equilibrium effects (Heckman, Lochner, and Taber (1999b)).

### 1.7.3 The Effects of Education Subsidies on Human Capital

As in the case of labor income taxation (cf. Section 1.7.1), the effects of education subsidies on human capital accumulation can be divided into a substitution effect (positive), an income effect (negative), and a direct effect. While the aggregate indirect effect through the impact on labor supply is ambiguous, the direct effect is unambiguously positive because the marginal costs of human capital investments are reduced. Hence, education subsidies can improve the efficiency in human capital investment by offsetting tax-induced distortions, i.e. by lowering the negative effect of taxation on human capital formation (Bovenberg and Jacobs (2005)). Furthermore, education subsidies can correct for the hold-up problem in imperfect labor markets as discussed in Section 1.3.1 (Boone and de Mooij (2003)).

Collins and Davies (2003) conclude that the incentives for human capital accumulation depend on the net effective tax rate ($NETR$), i.e. the difference between the effective tax rate ($ETR$) and the effective subsidy rate ($ESR$). This concept is based on the internal rate of return that is determined by comparing the sum of discounted earnings with and without taxation and subsidies, respectively:

$$\begin{aligned} NETR &= ETR - ESR \\ ETR &= \frac{r_g - r_n}{r_g} \\ ESR &= \frac{r_g - r_p}{r_g} \end{aligned}$$

$ETR$, $ESR$, and $NETR$ are calculated by considering the relationship between the gross rate of return ($r_g$), the net rate of return ($r_n$), and the public rate of return ($r_p$). In their empirical analysis for Canada, Collins and Davies (2003) find that the net effective tax rate is smaller than zero.

In a nutshell, Dur and Teulings (2003) favor the implementation of education subsidies for three reasons: (1) redistribution, (2) positive externalities of education, and (3) credit constraints due to capital market imperfections.[28] However, positive external effects of higher education are difficult to establish empirically (for example Krueger and Lindahl (2001)) and also the empirical findings with respect to capital market imperfections are controversial in the literature (for example Shea (2000)).

As a further argument for education subsidies, Bénabou (2002) points out the insurance effect because subsidies to higher education make college attendance more attractive by reducing both the direct costs and the risk, particularly for students with low wealth endowment. If investments in human capital are risky and uninsurable, even small initial changes in college investment may generate large changes in college attendance (Ljungqvist (1995)). In a theoretical model, Akyol and Athreya (2005) suggest large welfare gains of education subsidies relative to the fully decentralized outcome because subsidies lead to nearly mean preserving reductions in college

[28] Capital market imperfections imply that workers may be prevented from investing efficiently in their stock of (general) human capital (cf. Section 1.3.1).

failure risk. Furthermore, if the variance of wages decreases with the level of education, education per se produces an insurance effect (Anderberg and Andersson (2003)). In this context, Bénabou (2002) develops a risk-adjusted, distribution-free measure of general efficiency. By simulating this model with empirical parameter estimates, he shows that the optimal education subsidy should be combined with a consumption tax.

Concerning the optimal level of education subsidies, the aggregate welfare is maximized subject to the public budget constraint (Steuerle (1996)). However, education subsidies also lead to adverse selection (i.e. more low-ability students attend college) and may generate a deadweight loss if they are financed by distortionary taxation. Hence, the optimal level of education subsidies has to maximize the aggregate welfare subject to this trade-off (Dur and Teulings (2003)).

### 1.7.4 Fiscal Policy in the Context of Factor Mobility

According to Boadway, Marceau, and Marchand (1996), mandatory education and the public provision of education can be explained as second-best policies in order to circumvent the hold-up problem of time-consistent optimal taxation by governments with redistributional objectives, i.e. the problem of underinvestment in human capital because workers anticipate the excessive taxation of labor income in the future. However, globalization and increased mobility of high-skilled workers reduce this time-consistency problem because the government's ex post incentives to tax human capital are decreased.

With respect to different policy regimes, the effects of globalization on taxation and education policy are analyzed by Andersson and Konrad (2001) and Andersson and Konrad (2003). As suggested by Boadway, Marceau, and Marchand (1996), closed economies suffer from a hold-up problem of excessive redistribution and governments use education policy as a second-best instrument. If labor mobility is increased by opening the economy, globalization reduces the governments' incentives to provide subsidized education and increases the workers' incentives for human capital formation. Compared to the closed-economy benchmark case, globalization may increase the aggregate welfare if education subsidies are completely directed to immobile low-skilled workers (Andersson and Konrad (2001)).

However, in the presence of "Leviathan governments", the full mobility of high-skilled workers generates a welfare loss because the tax competing governments have incentives to prevent individuals from mobility-increasing investments in human capital. The governments raise taxes on education in order to broaden their tax base, which only comprises the immobile low-skilled workers (Andersson and Konrad (2003)). With respect to time-consistent optimal taxation, Konrad (2001) suggests another solution to the problem of underinvestment in human capital. If the government has only limited information about the effort of workers, then this may work as a commitment mechanism to time-consistent taxation and thus increases the aggregate welfare by reducing the hold-up problem.

In a similar model with education subsidies financed by labor income taxation and possible migration of high-skilled workers, Poutvaara (2001) analyzes the effects of tax competition between countries in a federation. He concludes that this tax competition may lead to inefficiently low taxes and subsidies. In a nutshell, the problem of underinvestment in human capital can be solved by a fiscal system such that taxes are paid to the country of education and not to the country of residence.

Finally, in Section 1.7 of this literature review, we analyze the effects of taxation and education subsidies on human capital formation. The marginal effects of proportional and progressive income taxation on human capital formation are generally negative (Heckman, Lochner, and Taber (1999a)). Depending on whether the costs of human capital are direct expenditures or foregone earnings, a "comprehensive income tax" may discriminate either against investments in human or physical capital. In a nutshell, the incentives for human capital formation depend on the net effective tax rate, which implies that education subsidies can improve the efficiency in human capital investment by offsetting tax-induced distortions (Bovenberg and Jacobs (2005)).

## 1.8 Outlook

The contribution of Chapter 1 of my PhD thesis is to provide a better understanding of the process of human capital formation and educational attainment. Furthermore, this literature review points out some important issues of human capital formation that may shape future research and policy discussions, but it does not focus explicitly on those topics of human capital formation presented in the following Chapters 2 to 4 of my PhD thesis.

Starting from this review in Chapter 1, we will analyze three different policy issues and their implications for human capital formation. In Chapters 2 and 3, we refer to the German system of apprenticeship training and its interactions with the labor market prospects of low-skilled workers. In these two chapters, we use a rather similar model framework but analyze different policy instruments by focusing on the demand side of the labor market in Chapter 2 and on the supply side of the labor market in Chapter 3. In Chapter 4, we discuss the system of old-age provision and its effects on labor force participation.

In Chapter 2, we find that the number of apprenticeship training positions may be inefficiently low compared to the first-best optimum. Many firms decide not to offer training places because trained workers can freely choose to change their employer upon completion of their apprenticeship. Hence, the firms perceive the danger of bearing the costs of training without getting any return. As a consequence, the German apprenticeship system is challenged by an increasing number of unskilled workers. For this reason, we analyze penalty charges to firms that do not offer apprenticeship training positions to school graduates. In our model, the implementation of penalty charges faces a trade-off with respect to overall welfare. While penalty charges increase the number of apprenticeship training positions and thus the fraction of trained workers in the workforce, some firms will leave the market to avoid the financial burden, which implies unemployment among workers with low ability. If the productivity-enhancement of apprenticeship training is low, it is optimal to reject the implementation of penalty charges. However, penalty charges increase the overall welfare if the productivity-enhancement exceeds some lower bound.

In Chapter 3 of my PhD thesis, we refer to the German system of apprenticeship training and social security and show that employment of low-skilled workers is crowded out by the welfare state. Low-skilled workers decide to remain unemployed if their potential labor income falls below the level of social assistance (Sinn, Holzner, Meister, Ochel, and Werding (2006)). Tax credits are a possible policy instrument to reintegrate these low-skilled workers into the labor market by granting an individual subsidy only if the worker decides to enter the labor market. However, it is important to incorporate human capital formation into the analysis of tax credits because subsidies to low-skilled workers increase the opportunity costs of apprenticeship training and thus reduce the workers' incentives to acquire skills (Heckman (2002)). While tax credits reduce the number of unemployed workers, they lower at the same time the number of apprentices, which implies the waiving of increased productivities in the future. Hence, the reintegration of those workers at the bottom of the ability-distribution may not be optimal because the costs in terms of decreased human capital formation would be too high.

Finally, in Chapter 4 of my PhD thesis, we analyze the traditional pay-as-you-go pension system and its implications for retirement and human capital formation. Human Capital is an important factor in the context of pension reforms because the workers' incentives to acquire skills and the desired retirement age are positively related (Trostel (1993)). Unfortunately, there is a trend of declining labor force participation because pension systems provide economic incentives to retire early (Samwick (1998)) and the employment prospects of older workers are significantly reduced (Bingley and Lanot (2004)). For this reason, we analyze two different proposals for pension reform (implementation of individual retirement accounts, increase in the minimum retirement age) and show that the benefits of these reforms strongly depend on the labor market prospects of older workers near retirement age. The firms' employment decision can significantly affect the implications of pension reforms because workers and firms separate as soon as one party decides to leave the market. While high-ability workers may indeed postpone retirement and increase human capital formation as suggested by the incentives of the pension system, low-skilled workers will not be affected by pension reforms if firms refuse to employ them any longer.

# Chapter 2

# Penalty Charges to Non-Training Firms

*There is an ongoing discussion in Germany about the implementation of penalty charges if firms refuse to offer apprenticeship training positions to school graduates. Chapter 2 of my PhD thesis aims at analyzing penalty charges to non-training firms in a theoretical model that systematically compares the costs and benefits of this policy instrument. Building on recent training literature, a two-period partial-equilibrium model is developed that allows for worker heterogeneity in ability and covers special features of the German apprenticeship system.*

*In our model, the implementation of penalty charges faces a trade-off with respect to overall welfare. While penalty charges increase the number of apprenticeship training positions and thus the fraction of trained workers in the workforce, some firms will leave the market to avoid the financial burden, which implies unemployment among workers with low ability. Furthermore, the model suggests that optimal penalty charges increase the overall welfare compared to the laissez-faire equilibrium only if the productivity-enhancement of apprenticeship training exceeds some lower bound.*

## 2.1 Introduction

There is an ongoing discussion in Germany about the implementation of penalty charges if firms refuse to offer apprenticeship training positions to school graduates. The German apprenticeship system provides basic skills to a large share of the workforce and is thus considered an exemplary model for vocational education (Harhoff and Kane (1997)).[29] But in the last several years, there are more and more youths unable to find apprenticeship training positions, who therefore remain unskilled after graduation.[30] In 2002, only 31.3% of all firms were providing apprenticeship training positions (BMBF (2004)). Many firms decide not to offer training places because trained workers can freely choose to change their employer upon completion of their apprenticeship. Hence, the firms perceive the danger of bearing the costs of training without getting any return. As a consequence, the famous German apprenticeship system is challenged by unemployment among young workers and an increasing number of unskilled workers.

In order to approach the problem of missing apprenticeship training positions, a compulsory training quota for each firm is proposed depending on the number of full-time workers. If a firm does not satisfy this quota, it has to pay a previously defined penalty for each training place missing to meet the quota. Unfortunately, a theoretical analysis of penalty charges is still lacking. This chapter aims at closing this gap by presenting a two-period partial-equilibrium model that systematically compares the costs and benefits of this policy instrument. In the literature, there are two theoretical explanations for firms providing general training. First, the basic approach of Becker (1964) concentrates on the firms' current incentives during the training period. If the training wage is low enough - i.e. the gap between

---

[29] In Germany, two thirds of the age-group from 15 to 24 are provided with vocational training through the apprenticeship system (BMBF (2004)).

[30] From 1991 to 2003, the number of new apprenticeship contracts has decreased from 571,206 to 564,493. The number of registered apprenticeship training positions has decreased from 1,629,312 to 1,581,629 (BMBF (2004)).

apprentice output and training wage is large enough - to compensate for the costs of training activities, firms decide to provide general training to their workers. Second, Acemoglu and Pischke (1999b) point out the firms' incentives with respect to future returns because training firms manage to skim a rent from the increased output if labor markets are imperfect and the trained workers stay with the training firm after the apprenticeship has been completed.[31]

According to Acemoglu and Pischke (1998a), the training decision of firms bears two different kinds of inefficiencies. First, firms take into account only their own gains from higher productivity and neglect the gains for the workers through higher wages in the future. Second, firms further underinvest in the human capital of their workers if there is a positive probability of separation after the training period because firms do not take into account higher profits of potential employers in the future. One policy measure to reduce these inefficiencies in the provision of general training are firing costs. Because the probability of separation is reduced, employment protection increases the incentives to invest in human capital for both workers and firms (Fella (2005)).[32] The total effect of firing costs on unemployment is ambiguous because fewer separations lead to lower unemployment but some firms decide to leave the market and thus it becomes harder for unemployed workers to find a job (Belot, Boone, and Ours (2002)).

The contribution of Chapter 2 of my PhD thesis is twofold because the formal analysis of penalty charges, which is based on recent training literature with oligopolistic labor markets, is extended in two important ways. First, our model allows for worker heterogeneity in ability and manages to explain

[31] Cf. also Acemoglu (1997), Acemoglu and Pischke (1998a), Acemoglu and Pischke (1998b), and Acemoglu and Pischke (1999a).

[32] The empirical evidence shows a positive relationship between training and job tenure (Lynch (1991) and Loewenstein and Spletzer (1999)). According to Stähler (2006), employment protection increases the fraction of skilled workers if workers decide ex-ante on their human capital investment and if they receive a sufficiently large fraction of the return to this investment.

the extensive training decision of firms depending on the individual ability of workers. Only those workers above some critical level of individual ability are offered an apprenticeship training position while workers with low ability remain unskilled. Second, we bring together the two theoretical explanations of firm-sponsored general training in the context of the institutional setting of the German apprenticeship system. We show that firms may provide general training both because they are looking forward to future returns (as in Acemoglu and Pischke (1999b)) and because they currently benefit from doing so (as in Becker (1964)). Beyond Acemoglu and Pischke (1998a), our analysis indicates a third kind of inefficiency, which is generated by the fixed training wage during the apprenticeship.

In a nutshell, there are three key questions considered in this chapter: First, what is the impact of penalty charges on the number of apprenticeship training positions? Second, what are the effects of penalty charges on the number of unemployed workers? And third, what is the optimal level of penalty charges subject to the employment decision of the firms? In our model, the implementation of penalty charges faces a trade-off with respect to overall welfare. While penalty charges increase the number of apprenticeship training positions and thus the fraction of trained workers in the workforce, some firms will leave the market to avoid the financial burden, which implies unemployment among workers with low ability. By also considering administration costs of implementation, our formal analysis suggests that the optimal policy depends on the productivity-enhancement of training. If the productivity-enhancement of apprenticeship training is low, it is optimal to reject the implementation of penalty charges and the economy will attain the laissez-faire equilibrium. However, penalty charges increase the overall welfare if the productivity-enhancement exceeds some lower bound.

Chapter 2 of my PhD thesis proceeds as follows: the next section discusses human capital theory and its two approaches to firm-sponsored human capital formation. After this, the institutional setting of the German apprenticeship system is illustrated. In Section 2.4, our partial-equilibrium model is developed and the equilibrium without penalty charges is presented. In Section 2.5, the implementation of penalty charges is analyzed and optimal penalty charges are derived. Section 2.6 concludes.

## 2.2 The German Apprenticeship System

### 2.2.1 Historical Relevance

Historically, there have been several characteristics of apprenticeship training. First, the length of the apprenticeship was specified contractually in advance and independent of individual ability. For example, this applied to the *métier* in France, *arte* in Italy, *craft guild* in England, and *Zunft* or *Innung* in Germany. Furthermore, apprenticeships were characterized by a low starting wage and the acquired skills were largely general to a particular trade or industry (Booth and Satchell (1994)).

Finally, apprenticeship training has been intensely regulated, for example by guilds in medieval times. This regulation typically implied a minimum length of the apprenticeship and the monitoring of training. The craft guilds of the middle ages had supervisory functions that included the right of search to insure that good materials and appropriate processes of manufacture were employed. In Germany, a range of institutions funded collectively by firms controlled the working of the apprenticeship system (Pirenne (1936)).

### 2.2.2 The Institutional Setting

The educational system of Germany is one of the most segregated among industrialized countries. There are four types of German secondary schools: lower (*Hauptschule*), middle (*Realschule*), upper (*Gymnasium*), and mixed (*Gesamtschule*). Upon their conclusion, all of these school tracks require the successful completion of exams which indicate whether students are qualified to enter into an apprenticeship, other vocational training, or the university (Cooke (2003)).

Apprenticeship training can be undertaken in a variety of skilled blue or white collar positions. It combines part-time schooling with a work-based component (the so-called "dual system") and is largely general. Training firms have to follow a prescribed curriculum and apprentices take a rigorous exam at the end of the apprenticeship. Industry or craft chambers certify whether firms fulfill the requirements to train apprentices adequately, while worker councils in the firms monitor the training. During the apprenticeship, workers receive a low training wage, which is set for each industrial sector by negotiation of the collective bargaining parties.[33] After having passed the exam, apprentices receive a formal skill certificate that is accepted nationwide (Bougheas and Georgellis (2004)). An overview of the German system of apprenticeship training is given by Soskice (1994) and Harhoff and Kane (1995).

In our formal analysis, we impose two simplifications on the German apprenticeship system. First, we assume that the training wage is identical for all apprentices and thus independent of the worker's individual productivity. This assumption is empirically justified because the bargained wage works as a lower bound for the training wage in each industrial sector and is binding for about 80% of the training firms (Krekel (2005)). Note that the training wage does not constitute a decision variable of the government unless we allow to restrict the level of tariff autonomy in the economy. Second, we assume that the length of the apprenticeship is fixed and identical for each apprentice. Although those school graduates with upper secondary education (*Abitur*) can shorten the apprenticeship period, our assumption is warranted because the extent to which the apprenticeship can be shortened is rather small and does not vary continuously with the worker's individual productivity.[34]

[33] In 2004, the average training wage per month was 617 Euros in West Germany and 526 Euros in East Germany (Krekel (2005)). Note that even those firms that are not covered by collective bargaining are not allowed to pay less than 80% of the bargained wage of their industrial sector.

[34] In their analysis of the regulation of apprenticeship contracts, Malcomson, Maw, and

Following from the second assumption, firms (and workers) decide at the extensive margin (i.e. participation) whether to provide apprenticeship training or to employ workers regularly, i.e. to provide full-time work without formal qualification. At the intensive margin (i.e. hours of training), the accumulation of human capital by apprenticeship training is determined by the fixed length of the apprenticeship and the specified curriculum rather than by investment decisions of the training firms. Hence, we focus on the extensive margin in order to analyze inefficiencies in the number of apprenticeship training positions.

### 2.2.3 The Costs of Apprenticeship Training

In 1991, the *Bundesanstalt für Berufsbildung* investigated training firms with respect to accounting costs and apprentice productivity in order to assess the net costs of training, which are defined as the difference between gross costs and apprentice output. The results of this study are described in Von-Bardeleben, Beicht, and Fehér (1995). The first step is to calculate gross costs as the sum of payroll costs, training personnel, material, and equipment used in the training process as well as direct costs of any external training. However, in many firms trainers are not engaged in training full-time but also work in productive activities. The study takes two different approaches to this problem. The first is to prorate time spent on training by part-time personnel ($A$), the second is to exclude the costs of part-time

McCormick (2003) assume that the training wage and the length of the apprenticeship are endogenously determined. Because workers earn wages equal to their marginal product upon completion of the apprenticeship, the training wage must be sufficiently low in order to induce firms to provide apprenticeship training. The optimal length of the apprenticeship solves a trade-off between current returns for the training firm during the apprenticeship and future returns for all potential employers after the apprenticeship. Malcomson, Maw, and McCormick (2003) show that regulation to increase the length of the apprenticeship, combined with a subsidy for each completed apprenticeship if the efficiency loss from distortionary taxation to finance the subsidy is sufficiently low, can increase the number of apprentices and the amount of training per apprentice.

trainers completely from the calculation of gross costs ($B$). Hence, the latter approach serves as a lower bound for the training costs borne by the firms.

In a second step, the output of apprentices is estimated. A measure of output is designed by multiplying time spent in productive activities by the payroll costs of a skilled worker and the relative efficiency of apprentices ($C$). However, this calculation implicitly assumes that the wages of skilled workers are set competitively and thus reflect the marginal product. If labor markets are imperfect, the output of apprentices is underestimated because the wage falls below the marginal product. For this reason, an alternative measure of apprentice output is designed by assuming imperfect labor markets with a markdown of 50% ($2C$).[35] Table 2.1 illustrates the role of these assumptions by using data from VonBardeleben, Beicht, and Fehér (1995) for Germany in 1991 (in German Marks per apprentice per year).[36]

From Table 2.1, it is evident that at least large firms bear significant costs in providing general training to their apprentices. As a consequence, many firms do not offer apprenticeship training positions in order to avoid this financial burden. However, depending on the method of measuring training costs and apprentice output, there may be net benefits from apprenticeship training for small and middle-sized firms.

## 2.3 The Model

We consider a discrete-time model with two types of agents, namely workers and firms. In line with Acemoglu and Pischke (1998a), there are two periods, a training period (period 1) and a working period (period 2). The length of both periods is normalized to unity. Production takes place in worker-firm pairs and no capital is needed. According to Eissa, Kleven, and Kreiner

[35] Otherwise, if the wage exceeded the marginal product, the output of apprentices is overestimated and the measure of apprentice output should be designed with a markup.

[36] Krekel (2005) provides an overview of how the costs of apprenticeship training in Germany depend on the industrial sector, the region, and the year of training. The training wage slightly increases with the year of training.

| | **All firms** | **Number of employees** | | | |
|---|---|---|---|---|---|
| | | 1 - 9 | 10 - 49 | 50 - 499 | 500 + |
| 1) Gross costs ($A$) | 29,573 | 27,473 | 28,176 | 30,344 | 35,692 |
| 2) Gross costs ($B$) | 18,051 | 13,867 | 15,074 | 20,283 | 28,197 |
| *Perfect markets* | | | | | |
| Appr. output ($C$) | 11,711 | 12,221 | 11,465 | 12,099 | 10,311 |
| 1) Net costs ($A - C$) | 17,862 | 15,252 | 16,711 | 18,245 | 25,381 |
| 2) Net costs ($B - C$) | 6,340 | 1,646 | 3,609 | 8,184 | 17,886 |
| *Imperfect markets* | | | | | |
| Appr. output ($2C$) | 23,422 | 24,442 | 22,930 | 24,198 | 20,622 |
| 1) Net costs ($A - 2C$) | 6,151 | 3,031 | 5,246 | 6,146 | 15,070 |
| 2) Net costs ($B - 2C$) | -5,371 | -10,575 | -7,856 | -3,915 | 7,575 |

Table 2.1: The Costs of Apprenticeship Training in Germany (Source: von Bardeleben, Beicht, and Fehér (1995), Chart 27 and Table 12)

(2006), a model of extensive labor supply requires some type of heterogeneity, either in preferences or in ability. In our approach, workers have identical preferences but are heterogeneous in their initial ability, which is exogenously given.

At the beginning of period 1, each firm meets one worker whose individual ability is drawn randomly from a distribution that is common knowledge. At the extensive margin, firms and workers decide whether to engage in apprenticeship training, to produce with regular work, or not to produce at all. An apprenticeship takes place only if both parties agree on apprenticeship training. In the second period, all workers can be employed regularly, but only those workers who were trained in period 1 have an increased productivity. In line with Malcomson, Maw, and McCormick (2003), the two periods are connected by an exogenous separating probability. Altogether, the economy evolves over time as illustrated in Figure 2.1.

In our model, firms may provide general training both because they are looking forward to future returns in period 2 (as in Acemoglu and Pischke

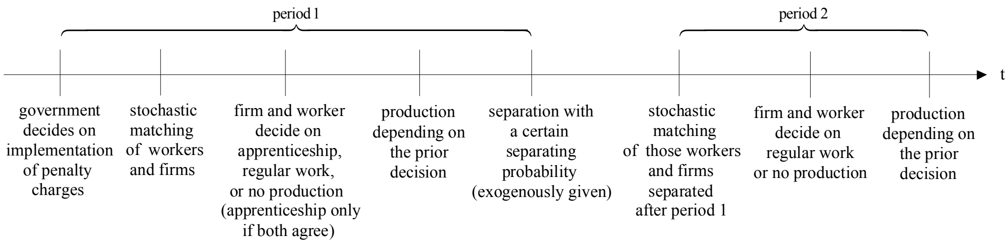


Figure 2.1: The Evolution over Time

(1999b)) and because they currently benefit from doing so in period 1 (as in Becker (1964)). Especially the latter argument has to be considered because the implementation of penalty charges affects the current training decision by altering the opportunity costs of not offering apprenticeship training positions. The model assumptions and the training decisions of firms and workers are described in the following subsections.

### 2.3.1 The Workers

At the beginning of period 1, workers differ in their individual ability, which is assumed to be uniformly distributed on the interval $[\theta_L, \theta_H]$.[37] After the match has taken place, firms can unambiguously observe the worker's ability.[38] In conformity with Malcomson, Maw, and McCormick (2003) and

[37] This heterogeneity of workers is an important extension compared to the model by Malcomson, Maw, and McCormick (2003). The continuous distribution of abilities allows obtaining a smooth participation decision at the individual level (Mirrlees (1971)). In order to keep the following calculations as simple as possible, we assume a uniform distribution of abilities.

[38] This assumption is in line with Boone and Bovenberg (2006). Furthermore, it is implicitly included in the whole literature on human capital and the life-cycle of earnings. Each worker offers his individual stock of human capital to the firms and is rewarded by a rental price per unit of human capital. Hence, we rule out asymmetric information (hidden knowledge) as a source of wage compression as described in Section 1.3.1. If the worker's productivity were not observed by the firm, there would be adverse selection as modeled e.g. by DeMeza and Webb (2001).

Hanushek, Leung, and Yilmaz (2003), the mass of workers is normalized to unity by defining $\theta_L \equiv 0$ and $\theta_H \equiv 1$. By assumption, the mass of firms is also one, which implies that each firm is matched with one worker whose ability is uniformly distributed on $[0, 1]$.

In line with Malcomson, Maw, and McCormick (2003), workers are risk-neutral and maximize the sum of their discounted incomes over both periods:[39]

$$V(\theta) = w + \delta w' \tag{2.1}$$

The discount factor $\delta \equiv \frac{1}{1+r}$ with $r$ as the market interest rate expresses the preference for current and future wealth. The higher $\delta$, the higher is the weighting of period 2 and the lower is the preference for period 1. In the first period, the worker's wage $w$ is equal to

$$w = \begin{cases} w_A & \text{if apprenticeship in } t = 1 \\ w(\theta) = \beta\theta & \text{if regular work in } t = 1 \\ 0 & \text{if unemployment in } t = 1 \end{cases} \tag{2.2}$$

In the case of apprenticeship training, the worker receives the training wage $w_A \geq 0$, which is independent of his ability and identical for all apprentices. This assumption is motivated by the institutional setting of the German apprenticeship system as discussed in Section 2.2.2. With regular work, the wage corresponds to the Nash bargaining solution of oligopolistic labor markets. By defining the output good as numéraire and assuming an identical, linear one-to-one production function for the connection of output and labor (which is the only factor of production), the marginal product of each worker corresponds to his productivity $\theta$.[40] According to Acemoglu (1997), the parameter $0 \leq \beta \leq 1$ indicates the (identical) bargaining power

[39] The wage corresponds to the worker's income because labor supply is implicitly normalized to unity. In line with Ben-Porath (1967), we do not analyze a more general utility function of workers. Note that there is no uncertainty for workers because they receive a predetermined wage in period 2, either from their current or from another employer.

[40] The production function exhibits constant returns to scale. From the firm's point of view, the worker's ability can be interpreted as individual productivity.

of workers concerning the division of output. Because the worker receives zero income in the case of unemployment, his wage is equal to $w(\theta) = \beta\theta$, which implies that there are labor market frictions because the worker's wage falls below his marginal product for $\beta < 1$.[41]

In period 2, the worker's wage $w'$ is given by[42]

$$w' = \begin{cases} w(\theta') = \beta\theta' & \text{if regular work in } t = 2 \\ 0 & \text{if unemployment in } t = 2 \end{cases} \tag{2.3}$$

$$\theta' = \begin{cases} (1+\alpha)\,\theta & \text{if apprenticeship in } t = 1 \\ \theta & \text{if regular work in } t = 1 \\ (1-\sigma)\,\theta & \text{if unemployment in } t = 1 \end{cases} \tag{2.4}$$

Like in the first period, the worker receives zero income in the case of unemployment. With regular work, the wage corresponds to the Nash bargaining solution.[43] Hence, it depends on the worker's productivity $\theta'$, which is determined by the status of employment in period 1 according to equation

[41] Wages in both periods are determined by Nash-bargaining, which implies that the worker's wage is a fraction of his marginal product $\theta$. The reason is that the firm and the worker maximize the Nash-product $(\theta - w)^{(1-\beta)}\, w^{\beta}$. While $w$ is the bargained wage of the worker, the firm is the residual claimant of output so that its profits are equal to the residuum $(\theta - w)$ (cf. Section 2.3.2). Because there is only one chance for a worker-firm match in period 1 (cf. Figure 2.1) and workers become unemployed in the case of no agreement, the failure to agree on a wage yields an income and profit level of zero. Hence, because the fall-back payoffs are zero, the bargained wage is equal to $w(\theta) = \beta\theta$ (Ortigueira (2006)). This result also holds when there is no unemployment in equilibrium because the fall-back payoffs would only be affected if there were more than one chance for a match. Hence, zero unemployment is not assumed but may be the result of the matching process. Note that our analytical results will not change qualitatively if the worker receives the full marginal product, which implies $\beta = 1$ and $w(\theta) = \theta$.

[42] Our model does not consider search frictions which implies identical matching probabilities equal to one for all workers. Alternatively, the probability of a match in period 2 may explicitly depend on the worker's status of employment in period 1. Different matching probabilities can be justified by different frictions in searching for employment (Mincer (1989)).

[43] Similar to period 1, the bargained wage is equal to $w(\theta') = \beta\theta'$ in the second period. The reason is that the fall-back payoffs are zero because the bargaining takes place in

(2.4). In the second period, the productivity of all workers employed regularly in period 1 is unchanged (i.e. $\theta' = \theta$). The productivity of previously unemployed workers declines to $\theta' = (1 - \sigma)\,\theta$ because the fraction $\sigma > 0$ of skills not employed in period 1 is lost and thus no longer available in period 2 (Pissarides (1992)).

For all trained workers, the productivity increases to $\theta' = (1 + \alpha)\,\theta$, whereas the parameter $\alpha \geq 0$ represents the productivity-enhancement of apprenticeship training. While the relative productivity gain from apprenticeship training is equal to $\alpha$ and thus identical for all trained workers, the absolute productivity gain $\theta' - \theta = \alpha\theta$ is proportional to the worker's innate productivity $\theta$.[44] This assumption is motivated by the literature on human capital formation at the intensive margin (cf. Ben-Porath (1967) and Heckman (1976)). Because the productivity-enhancement unambiguously depends on the amount of training, the assumption that $\alpha$ is identical for each worker and independent of $\theta$ implies that the amount of training is the same for all apprentices. This implication is in conformity with the German system of apprenticeship training, where the amount of training per apprentice is defined by the prescribed curriculum rather than by the training firm (cf. Section 2.2.2).[45] Furthermore, the identical productivity-enhancement simplifies the formal analysis because the fraction of trained workers can be aggregated more easily.

---

worker-firm pairs (either with the same contractual partner as in period 1 or with a new partner in the case of separation) and this is the only chance for an agreement in period 2 (cf. Figure 2.1). The failure to agree on a wage yields an income and profit level of zero.

[44] Formally, this means $\frac{\partial(\theta' - \theta)}{\partial\theta} = \alpha \geq 0$. Intuitively, the accumulation of new skills is easier when more skills are already available. This relationship is also suggested by Ben-Porath (1967) and Mincer (1997). Because the parameter $\alpha$ determines the productivity and thus the wage in period 2, it constitutes the key determinant of the return to education as analyzed in the theory of human capital (c.f. Section 1.4).

[45] If the firms could decide on their human capital investment not only at the extensive but also at the intensive margin and if the optimal amount of training depended on the workers' innate ability, the productivity-enhancement would not be the same for all workers.

Altogether, the total income of a worker with initial ability $\theta$ and regular employment in period 2 is obtained by substituting equations (2.2) to (2.4) into equation (2.1):

$$V(\theta) = \begin{cases} w_A + \delta\beta(1+\alpha)\theta & \text{if apprenticeship in } t=1 \\ (1+\delta)\beta\theta & \text{if regular work in } t=1 \\ \delta\beta(1-\sigma)\theta & \text{if unemployment in } t=1 \end{cases} \tag{2.5}$$

### 2.3.2 The Firms

As modeled by Malcomson, Maw, and McCormick (2003), firms are risk-neutral and maximize the discounted sum of their expected profits over both periods:

$$\pi(\theta) = \pi + \delta E[\pi'] \tag{2.6}$$

In period 1, the firm can invest in the worker's human capital by offering an apprenticeship training position. Due to the prescribed curriculum of the apprenticeship, the firm only has this discrete choice at the extensive margin but cannot decide on the amount of training at the intensive margin. However, the firm can reject to provide firm-sponsored general training by employing the worker regularly or leaving the market. Hence, the firm's profits in period 1 are equal to[46]

$$\pi = \begin{cases} \chi\theta - w_A & \text{if apprenticeship in } t=1 \\ (1-\beta)\theta & \text{if regular work in } t=1 \\ 0 & \text{if no production in } t=1 \end{cases}$$

In both periods, the firm's profits are equal to the difference between revenue and costs per worker. In the case of apprenticeship training, the relative efficiency of the apprentice is reduced compared to a regular full-time worker, i.e. $0 < \chi < 1$. As discussed in Section 2.2.3 with respect to

[46] To hold calculations simple, we assume the fixed costs of production to be zero. Analytically, the fixed training wage works like fixed costs of apprenticeship training. There is only one difference: with direct costs of apprenticeship training it would never be optimal to train all workers in the first-best optimum (cf. Section 2.3.3).

the German apprenticeship system, the efficiency parameter $\chi$ refers to the allocation of time between training and work in the first period and represents the fraction of time spent in productive activities.[47] The profit of a firm that regularly produces with a worker of productivity $\theta$ corresponds to the Nash bargaining solution and is equal to $(1-\beta)\,\theta$.[48]

At the end of period 1, three situations can arise: the match continues, firm and worker are separated for an exogenous reason, or firm and worker definitively separate because the firm decides to fire the worker. As shown in Figure 2.1, at the beginning of period 2 there is a stochastic matching of those firms and workers separated after the first period. After the match has taken place, firms can unambiguously observe the worker's ability and thus each firm faces a worker of a well defined productivity. If the worker's pro-

[47] The reduced efficiency of apprentices is due to external schooling and internal seminars. The fraction $(1-\chi)$ of time spent in training activities is determined by the prescribed curriculum of the German apprenticeship system. Note that our anayltical results would be strengthened (without changing qualitatively) if we assumed an inverse relationship between the productivity-enhancement $\alpha$ and $\chi$. However, we neglect this relationship in the following analysis because both parameters do not constitute decision variables of the firm.

For technical reasons, we restrict the training wage to satisfy $w_A \leq \chi - (1-\beta)$. This assumption can also be justified by economic intuition: the training wage should not exceed the difference between the output of the most productive apprentice and the firm's profit by regularly employing the most productive worker. Because of $w_A \geq 0$ this asumption implies $\chi \geq 1-\beta$. Note that the restriction $w_A \leq \chi - (1-\beta)$ constitutes a sufficient condition for the pivotal productivity $\theta^{LF}$ to be non-negative (cf. Section 2.3.4). If this restriction were not imposed, we would have to extend our formal analysis to the additional case $\theta^{LF} < 0$ without apprenticeship training. However, this case is empirically less important.

[48] Because the firm's profits are equal to $(\theta - w)$ and wages are determined by $w(\theta) = \beta\theta$ (cf. Section 2.3.1), the Nash bargaining solution implies profits equal to $\pi = (1-\beta)\,\theta$. Note that the firm's fall-back payoff in period 1 is zero because there is only one chance for a worker-firm match (cf. Figure 2.1). The same holds in period 2, which implies that profits are equal to $\pi' = (1-\beta)\,\theta'$ in the case of regular work. Altogether, the results of the bargaining process concerning the division of output allow comparing the firm's different employment opportunities at the extensive margin.

ductivity is high and the firm wants to keep the worker, there is an exogenous separating probability $\rho$, which implies that the employment relationship of period 1 is maintained in period 2 with probability $(1-\rho)$.[49] However, if the worker's productivity falls below the critical level $\bar{\theta}$, the firm decides to fire the worker (i.e. there is separation with probability one). Because the workers can be employed regularly, the expected profits of all firms correspond to the Nash bargaining solution and thus depend on the worker's expected productivity in period 2:[50]

$$E\left[\pi'\right] = \begin{cases} (1-\beta)\left[(1-\rho)\,\theta' + \rho E[\theta']\right] & \text{if } \theta \geq \bar{\theta} \\ (1-\beta)\,E[\theta'] & \text{if } \theta < \bar{\theta} \end{cases}$$

The expected productivity $E[\theta']$ of a separated worker corresponds to the average productivity in the pool of all separated workers. Hence, it depends on the fraction of firms that fire their worker after period 1 and the fraction $(1-\bar{\theta})$ of firms that decide to keep their worker as well as on the fraction $(1-\theta_A)$ of trained workers, the fraction $(\theta_A-\theta_U)$ of regular workers, and the fraction $\theta_U$ of those workers unemployed in period 1. In the following, we assume that all firms providing apprenticeship training in period 1 want to keep their worker (i.e. $\theta_A > \bar{\theta}$) while all firms leaving the market in period 1 want to meet another worker (i.e. $\theta_U < \bar{\theta}$).[51] Hence, there must be one firm with regular work in period 1 that is just indifferent between keeping its worker and firing him after the first period. This implies that the expected

[49] Note that the magnitude of $\rho$ is not crucial for the following analysis. On the one hand, in the extreme case $\rho = 1$, firms and workers definitively separate after period 1, which maximizes the second kind of inefficiency as discussed in Section 2.3.4. On the other hand, in the case $\rho = 0$, there is no exogenous separation so that this kind of inefficiency cancels out. However, our qualitative conclusions remain unaffected in both cases.

[50] Like in the first period, the firm's profits in period 2 are zero if no production takes place.

[51] The first assumption is empirically confirmed by Booth and Satchell (1994) who show that completed apprenticeships significantly increase the probability of job tenure for both workers and firms.

productivity of a separated worker is generally composed of four parts:[52]

$$E[\theta'] = \rho \int_{\theta_A}^{1} (1+\alpha)\,\theta d\theta + \rho \int_{\bar{\theta}}^{\theta_A} \theta d\theta + \int_{\theta_U}^{\bar{\theta}} \theta d\theta + \int_{0}^{\theta_U} (1-\sigma)\,\theta d\theta$$

Taken together both periods, firms can choose four different actions: (1) train the worker and try to keep him, (2) employ the worker regularly and try to keep him, (3) employ the worker regularly, fire him after the first period, and meet a new worker with the expected productivity $E[\theta']$, and (4) leave the market in period 1 and meet a new worker with $E[\theta']$ in period 2. Altogether, depending on the worker's productivity in the first period, total profits of a firm are equal to

$$\pi(\theta) = \begin{cases} (1) \quad \chi\theta - w_A + \delta(1-\beta)\left[(1-\rho)(1+\alpha)\theta + \rho E[\theta']\right] & \text{if } \theta \geq \theta_A \\ (2) \quad (1-\beta)\theta + \delta(1-\beta)\left[(1-\rho)\theta + \rho E[\theta']\right] & \text{if } \bar{\theta} \leq \theta < \theta_A \\ (3) \quad (1-\beta)\theta + \delta(1-\beta)E[\theta'] & \text{if } \theta_U \leq \theta < \bar{\theta} \\ (4) \quad \delta(1-\beta)E[\theta'] & \text{if } \theta < \theta_U \end{cases} \tag{2.7}$$

In the following analysis, it is not necessary to explicitly determine $E[\theta']$ and its components. The firm decides at the extensive margin whether to provide apprenticeship training, to employ the worker regularly, or to leave the market (cf. Sections 2.3.4 and 2.4.1). Concerning the discrete choice between apprenticeship training and regular work, the firm compares total profits (1) and (2) in equation (2.7), which implies that the expected productivity cancels out. The same holds with respect to the choice between regular work and market exit because the firm compares total profits (3) and (4).

Furthermore, the expected productivity $E[\theta']$ does not show up in the welfare analysis of Sections 2.3.4 and 2.4.2 because uncertainty is resolved

[52] As it is the case with laissez-faire (cf. Section 2.3.4), the fraction of unemployed workers might be zero (i.e. $\theta_U = 0$), which implies that there are only the cases (1) to (3) in equilibrium. Furthermore, if there is only exogenous separation, $E[\theta']$ equals the average productivity of all workers in the economy.

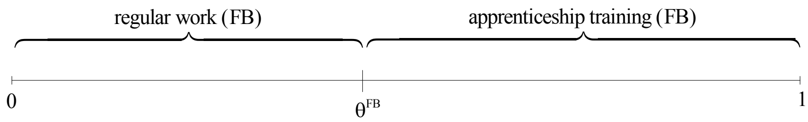


Figure 2.2: The First-Best Optimum

after the match has taken place and each firm faces a worker of a well defined productivity. This worker may be either the same as in the first period if the firm wants to keep the worker (with probability $(1-\rho)$) or this worker may be hired randomly from the pool of all separated workers. However, the average productivity of all workers in this pool is important only for the ex ante expected profits of a specific firm, but it does not enter into the ex post welfare equations. In aggregate, all firms and workers are matched in the second period (either randomly from the pool of all separated workers or because the firm wants to keep the worker) and it is not relevant which firm and which worker are matched.

### 2.3.3 The First-Best Optimum

In the first-best optimum (FB), the overall welfare (i.e. the common surplus of workers and firms) is maximized. Obviously, there is no unemployment in the first-best optimum (i.e. $u^{FB} = 0$) because each unemployed worker would be equivalent to lost productivity.

In the following, we assume that $\theta^{FB}$ describes the welfare maximizing pivotal productivity between apprenticeship training and regular work.[53] Hence, the optimal number of apprentices in the first period is $n^{FB} = 1 - \theta^{FB}$.[54] This situation is illustrated in Figure 2.2. The overall welfare

[53] Note that there are two possibilities for $\theta^{FB}$: first, it can be an interior solution within the interval $[0, 1]$, and second, it can be a corner solution, which implies that it is either equal to zero or one.

[54] More accurately, $n^{FB}$ describes the *mass* of apprentices in the first-best optimum. In the following, we will neglect this inaccuracy.

in the first-best optimum unambiguously depends on $\theta^{FB}$:[55]

$$W^{FB} = \underbrace{\int_{\theta^{FB}}^{1} (\chi + \delta (1 + \alpha))\, \theta d\theta}_{\text{apprenticeship training}} + \underbrace{\int_{0}^{\theta^{FB}} (1 + \delta)\, \theta d\theta}_{\text{regular work}} \tag{2.8}$$

The first integral in (2.8) describes the aggregate output of those workers who receive apprenticeship training in the first period. The productivity of these workers is reduced by the efficiency parameter $\chi$ in period 1, but it is increased by the factor $(1 + \alpha)$ in period 2. The second integral is equal to the aggregate output of regular workers without an apprenticeship training position. These unskilled workers with $\theta < \theta^{FB}$ have an identical productivity $\theta$ in both periods, which adds up to a present value of $(1 + \delta)\, \theta$.

Maximizing (2.8) with respect to $\theta^{FB}$ yields the optimal pivotal productivity between apprenticeship training and regular work. The partial derivative of $W^{FB}$ with respect to $\theta^{FB}$ is given by

$$\frac{\partial W^{FB}}{\partial \theta^{FB}} = -(\chi - 1 + \delta\alpha)\, \theta^{FB} \begin{cases} > 0 & \text{if } \alpha < \frac{1-\chi}{\delta} \\ \leq 0 & \text{if } \alpha \geq \frac{1-\chi}{\delta} \end{cases} \tag{2.9}$$

Because, the first derivative in (2.9) is either larger or smaller than zero for all productivities $\theta \in [0, 1]$, the pivotal productivity $\theta^{FB}$ represents a corner solution, which is either equal to zero or one. For $\alpha < \frac{1-\chi}{\delta}$, the overall welfare strictly increases with $\theta^{FB}$ and thus decreases with the number of apprenticeship training positions. Hence, the overall welfare is maximal for $\theta^{FB} = 1$, i.e. no workers should be trained ($n^{FB} = 0$). For $\alpha \geq \frac{1-\chi}{\delta}$, the overall welfare decreases with $\theta^{FB}$. Therefore, the overall welfare is maximal for $\theta^{FB} = 0$, i.e. all workers should be trained ($n^{FB} = 1$). Both cases are summarized in the following proposition.

[55] The density function of individual abilities is $f(\theta) = 1$. Note that the labeling at the bottom in (2.8) refers to the workers' status of employment in period 1.

**Proposition 2.1** *Depending on the productivity-enhancement $\alpha$, the number of apprenticeship training positions in the first-best optimum is equal to*

$$n^{FB} = \begin{cases} 0 & \text{if } \alpha < \frac{1-\chi}{\delta} \\ 1 & \text{if } \alpha \geq \frac{1-\chi}{\delta} \end{cases} \tag{2.10}$$

*Therefore, the overall welfare in the first-best optimum is equal to*

$$W^{FB} = \begin{cases} \frac{1}{2}(1+\delta) & \text{if } \alpha < \frac{1-\chi}{\delta} \\ \frac{1}{2}(\chi + \delta(1+\alpha)) & \text{if } \alpha \geq \frac{1-\chi}{\delta} \end{cases} \tag{2.11}$$

### 2.3.4 The Benchmark without Penalty Charges

#### The Training Decision of Firms

At the extensive margin, each firm decides whether to offer an apprenticeship training position, to employ the worker regularly, or to leave the market. Without penalty charges, no firm will leave the market because it is always possible to make positive profits by employing the worker regularly.[56] In the following, we characterize the training decision of firms depending on the individual productivities of the workers.

**Definition 2.1** *The pivotal productivity $\theta^{LF}$ is defined to be the lowest productivity of those workers to whom firms decide to offer an apprenticeship training position. A firm prefers apprenticeship training to regular work if its profits over both periods solve*

$$\begin{aligned} \chi\theta - w_A + \delta(1-\beta)\left[(1-\rho)(1+\alpha)\theta + \rho E[\theta']\right] \\ \geq (1-\beta)\theta + \delta(1-\beta)\left[(1-\rho)\theta + \rho E[\theta']\right] \\ \theta \geq \theta^{LF} \equiv \frac{w_A}{\chi - (1-\beta) + \delta(1-\rho)(1-\beta)\alpha} \end{aligned} \tag{2.12}$$

[56] Note that regular employment does not generate positive profits for $\theta_L = 0$ so that the firm ist just indifferent between employing the worker and leaving the market. Because this borderline case has no consequences for the overall welfare, we simply assume that this worker with the lowest productivity is also employed regularly.

Concerning the decision between apprenticeship training and regular work, the firm compares total profits in the first and the second row of equation (2.7), which implies that the expected productivity $E[\theta']$ cancels out (cf. Section 2.3.2). The pivotal productivity $\theta^{LF}$ increases with the training wage ($w_A$) and the separating probability ($\rho$) and decreases with the relative efficiency of apprentices ($\chi$) and the productivity-enhancement of apprenticeship training ($\alpha$).[57]

### The Training Decision of Workers

Additionally, we have to analyze the training decision of workers. Workers never prefer to remain unemployed because they receive zero income in this case.[58] Workers prefer apprenticeship training in period 1 if the discounted sum of their expected earnings over both periods is larger with training than with regular work. Hence, the following participation constraint must be satisfied:

$$w_A + \delta\beta\,(1+\alpha)\,\theta \geq (1+\delta)\,\beta\theta \tag{2.13}$$

**Proposition 2.2** *For $\alpha \geq \frac{1}{\delta}$, the participation constraint (2.13) is satisfied for all productivities $\theta \in [0,1]$. However, if the productivity-enhancement of apprenticeship training is low, workers with high ability prefer to remain unskilled. For $\alpha < \frac{1}{\delta}$, a worker prefers apprenticeship training to regular work if*

$$\theta \leq \theta^W \equiv \frac{w_A}{\beta\,(1-\delta\alpha)} \tag{2.14}$$

*The pivotal productivity $\theta^W$ is defined to be the highest productivity of those workers who decide to accept an apprenticeship training position if the productivity-enhancement of training is low.*

---

[57] Booth and Satchell (1994) empirically confirm that firms offer apprenticeship training positions to those workers with higher individual ability. Note that $0 \leq \theta^{LF} \leq 1$ is implied by $0 \leq w_A \leq \chi - (1-\beta)$.

[58] We assume that there are no unemployment benefits. If there were unemployment benefits greater than zero, workers with very low ability may prefer to remain unemployed.

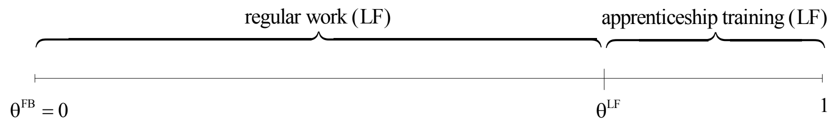


Figure 2.3: The Laissez-Faire Equilibrium 1

If the productivity-enhancement of apprenticeship training exceeds the lower bound $\frac{1}{\delta}$, all workers prefer apprenticeship training to regular work because their wages in period 2 increase by more than they have to forgo in the first period. However, for $\alpha < \frac{1}{\delta}$, only low-ability workers will adopt an apprenticeship training position. Workers with individual ability above $\theta^W$ prefer to work regularly because the productivity-enhancement in period 2 is too low to compensate for the low training wage in period 1.

### The Laissez-Faire Equilibrium

In a first step, we concentrate on the case $\alpha \geq \frac{1}{\delta}$, i.e. we assume that all workers prefer apprenticeship training to regular work. According to the training decision of the firms, $n^{LF} = 1 - \theta^{LF}$ workers receive apprenticeship training while all other workers are employed regularly. Like in the first-best optimum, there is no unemployment because firms and workers always prefer regular work to market exit. This laissez-faire equilibrium (LF) is illustrated in Figure 2.3.

The aggregate welfare of workers over both periods is equal to the discounted sum of aggregate incomes. Hence, in the laissez-faire equilibrium, the aggregate welfare of workers (W) is equal to

$$W_W^{LF} = \underbrace{\int_{\theta^{LF}}^{1} \left[w_A + \delta\beta\left(1+\alpha\right)\theta\right] d\theta}_{\text{apprenticeship training}} + \underbrace{\int_{0}^{\theta^{LF}} \left(1+\delta\right)\beta\theta d\theta}_{\text{regular work}} \tag{2.15}$$

The first integral in (2.15) is equal to the discounted aggregate wage sum of trained workers who earn $w_A$ in the first period and $w\left(\theta\right) = \beta\left(1+\alpha\right)\theta$ in period 2. The second integral describes the discounted aggregate wage sum of those workers who are employed regularly in both periods and earn wages equal to the Nash bargaining solution, i.e. $w(\theta) = \beta\theta$.

In almost the same manner, the aggregate welfare of firms (F) over both periods is equal to the discounted sum of aggregate profits:

$$W_F^{LF} = \underbrace{\int_{\theta^{LF}}^{1} [(\chi + \delta(1-\beta)(1+\alpha))\theta - w_A]\, d\theta}_{\text{apprenticeship training}} + \underbrace{\int_{0}^{\theta^{LF}} (1+\delta)(1-\beta)\theta d\theta}_{\text{regular work}} \tag{2.16}$$

The first integral is equal to the discounted aggregate profits of training firms (period 1) and those firms producing with a trained worker (period 2). In the first period, the profits of the training firms are determined by the efficiency parameter $\chi$ and the training wage $w_A$. In period 2, the productivity of trained workers is augmented by the factor $(1+\alpha)$, which increases the firms' profits by the output share $(1-\beta)$ according to the Nash bargaining solution. The second integral is equal to the discounted aggregate profits of firms employing unskilled workers regularly in both periods. As suggested in Section 2.3.2, the pool of all separated workers and the expected productivity of a separated worker do not show up in the welfare analysis. The reason is that all firms and workers are matched in the second period (either randomly or because the firm wants to keep the worker) and, concerning the overall welfare, it is not relevant which firm and which worker are matched.

Altogether, the overall welfare for $\alpha \geq \frac{1}{\delta}$ is determined by the aggregate welfare of workers and firms as described by equations (2.15) and (2.16):

$$W^{LF} = \underbrace{\int_{\theta^{LF}}^{1} (\chi + \delta(1+\alpha))\theta d\theta}_{\text{apprenticeship training}} + \underbrace{\int_{0}^{\theta^{LF}} (1+\delta)\theta d\theta}_{\text{regular work}} \tag{2.17}$$

Note that the parameter $\beta$ cancels out because the bargaining power only determines how the output is divided between workers and firms. In a


Figure 2.4: The Laissez-Faire Equilibrium 2

second step, we have to analyze the laissez-faire equilibrium for $\alpha < \frac{1}{\delta}$. In this case, all workers with individual ability $\theta > \theta^W$ prefer regular work to apprenticeship training. In brief, there are two different subcases that have to be considered. On the one hand, if the productivity-enhancement $\alpha$ is very small (i.e. $\alpha < \alpha_0 \equiv \frac{1-\chi}{\delta(1-\rho(1-\beta))}$), the pivotal productivity $\theta^W$ falls below $\theta^{LF}$. Hence, there are no workers that both prefer apprenticeship training and are offered a training place (i.e. $n^{LF} = 0$). All workers are employed regularly, which implies that the overall welfare for $\alpha < \alpha_0$ is equal to

$$W^{LF} = \underbrace{\int_0^1 (1+\delta)\,\theta d\theta}_{\text{regular work}} \tag{2.18}$$

On the other hand, for $\alpha_0 \leq \alpha < \frac{1}{\delta}$, the pivotal productivity $\theta^W$ is larger than $\theta^{LF}$ (cf. Figure 2.4).[59] Hence, $n^{LF} = \theta^W - \theta^{LF}$ workers are trained and the overall welfare is equal to

$$W^{LF} = \underbrace{\int_{\theta^W}^1 (1+\delta)\,\theta d\theta}_{\text{regular work}} + \underbrace{\int_{\theta^{LF}}^{\theta^W} (\chi + \delta\,(1+\alpha))\,\theta d\theta}_{\text{apprenticeship training}} + \underbrace{\int_0^{\theta^{LF}} (1+\delta)\,\theta d\theta}_{\text{regular work}} \tag{2.19}$$

By summarizing (2.17), (2.18), and (2.19), the different cases of laissez-faire are presented in the following proposition.

[59] Note that $\frac{1-\chi}{\delta(1-\rho(1-\beta))} < \frac{1}{\delta}$ is implied by $\chi > \rho\,(1-\beta)$.

**Proposition 2.3** *Depending on the productivity-enhancement $\alpha$, the number of apprenticeship training positions in the laissez-faire equilibrium is equal to*

$$n^{LF} = \begin{cases} 0 & \text{if } \alpha < \alpha_0 \equiv \frac{1-\chi}{\delta(1-\rho(1-\beta))} \\ \theta^W - \theta^{LF} & \text{if } \alpha_0 \leq \alpha < \frac{1}{\delta} \\ 1 - \theta^{LF} & \text{if } \alpha \geq \frac{1}{\delta} \end{cases} \tag{2.20}$$

*Therefore, the overall welfare in the laissez-faire equilibrium is equal to*

$$W^{LF} = \begin{cases} \frac{1}{2}(1+\delta) & \text{if } \alpha < \alpha_0 \\ \frac{1}{2}(1+\delta) + \frac{1}{2}(\chi - 1 + \delta\alpha)\left[(\theta^W)^2 - (\theta^{LF})^2\right] & \text{if } \alpha_0 \leq \alpha < \frac{1}{\delta} \\ \frac{1}{2}(\chi + \delta(1+\alpha)) - \frac{1}{2}(\chi - 1 + \delta\alpha)(\theta^{LF})^2 & \text{if } \alpha \geq \frac{1}{\delta} \end{cases} \tag{2.21}$$

In the following subsection, the laissez-faire equilibrium is compared to the first-best optimum in order to evaluate the possibilities of welfare-enhancing government interventions.

### Inefficiencies in the Number of Apprenticeship Training Positions

By comparing the number of apprenticeship training positions in the first best optimum (2.10) and in the laissez-faire equilibrium (2.20), we conclude that it is efficient only for small values of $\alpha$. With laissez-faire, it is always less or equal to the first-best number of apprenticeship training positions.

**Proposition 2.4** *Depending on $\alpha$, the deviation from the number of apprenticeship training positions in the first-best optimum is equal to*

$$\triangle n = n^{LF} - n^{FB} = \begin{cases} 0 & \text{if } \alpha < \frac{1-\chi}{\delta} \\ -1 & \text{if } \frac{1-\chi}{\delta} \leq \alpha < \alpha_0 \\ \theta^W - \theta^{LF} - 1 & \text{if } \alpha_0 \leq \alpha < \frac{1}{\delta} \\ -\theta^{LF} & \text{if } \alpha \geq \frac{1}{\delta} \end{cases} \tag{2.22}$$

For $\alpha \geq \frac{1-\chi}{\delta}$, the number of apprenticeship training positions is inefficiently low because the first-best optimum requires $\theta^{FB} = 0$. Unfortunately, the training decision of firms bears three different kinds of inefficiencies, which are in line with the results of Acemoglu and Pischke (1999b). In the following, we focus on the case $\alpha \geq \alpha_0$ where the deviation from the first-best optimum is driven by the pivotal productivity $\theta^{LF}$. The higher $\theta^{LF}$, the lower is the number of apprenticeship training positions and the larger are the inefficiencies in the training decision of the firms.[60] In order to distinguish between the three kinds of inefficiencies, we label the workers' bargaining power by an index for period 1 and period 2, respectively:

$$\theta^{LF} = \frac{w_A}{\chi - (1 - \beta_1) + \delta (1 - \rho) (1 - \beta_2) \alpha}$$

The first kind of inefficiency stems from the fact that firms take into account only their own gains from higher productivity and neglect the gains for the workers by higher wages in the second period. Hence, firms underinvest in the workers' human capital by offering an insufficient number of apprenticeship training positions. This inefficiency could be eliminated only if the firms were the sole beneficiaries of apprenticeship training (i.e. $\beta_2 = 0$). The higher the workers' bargaining power in period 2, the higher is the pivotal productivity $\theta^{LF}$ and the larger is the difference compared to the number of apprenticeship training positions in the first-best optimum.

Second, there is a probability of exogenous separation after the training period. The pivotal productivity $\theta^{LF}$ increases with $\rho$ because the training firms bear the risk of not participating in the workers' higher productivity after the apprenticeship has been completed. Hence, firms do not take into account higher profits of potential future employers in period 2. The higher the exogenous separating probability, the larger is the inefficiency in the provision of apprenticeship training. This is in line with the empirical results of Lynch (1991) and Loewenstein and Spletzer (1999) who find a positive relationship between training and job tenure. This kind of inefficiency could be eliminated only if there were no exogenous separation after the first period

[60] Note that the inefficiencies are reduced if the productivity-enhancement of apprenticeship training becomes larger: $\frac{\partial \theta^{LF}}{\partial \alpha} < 0$.

($\rho = 0$) or if future employers could be identified in advance and included into the apprenticeship contract at the beginning of period 1.

Beyond the approach of Acemoglu and Pischke (1999b), our model covers a third source of inefficiency that is generated by the fixed training wage during the apprenticeship.[61] This kind of inefficiency can be completely different depending on the productivity-enhancement $\alpha$. For $\alpha < \alpha_0$, the distortion $\triangle n$ is not affected by the magnitude of the training wage. For $\alpha_0 \leq \alpha < \frac{1}{\delta}$, the deviation from the first-best number of apprenticeship training positions is reduced if $w_A$ increases. The reason is that an increase in the training wage raises the number of workers who prefer apprenticeship training by more than it decreases the number of training firms (i.e. $\triangle\theta^W > \triangle\theta^{LF}$). This effect is directly opposed for $\alpha \geq \frac{1}{\delta}$ because all workers prefer apprenticeship training to regular work. However, the firms consider $w_A$ as training costs but do not take into account that it also determines the income of apprentices in period 1. In this case, $\theta^{LF}$ increases with $w_A$, which implies that the training wage is inefficiently high:

$$\frac{\partial \theta^{LF}}{\partial w_A} > 0 \quad \Leftrightarrow \quad \frac{\partial n^{LF}}{\partial w_A} < 0 \quad \Leftrightarrow \quad \frac{\partial \triangle n}{\partial w_A} < 0$$

Hence, for $\alpha \geq \frac{1}{\delta}$ and thus $\theta^{FB} = 0$, the government could lower $\theta^{LF}$ and thus reduce the third kind of inefficiency by regulating the training wage towards zero. Without generating unemployment, even the first-best level

[61] For $\alpha \geq \frac{1-\chi}{\delta}$, the first-best number of apprenticeship training positions could be achieved if the training wage were proportional to the worker's innate ability. For $\tilde{w}_A = s\theta$ with $\beta_1 \leq s \leq \chi - (1-\beta_1) + \delta\alpha$, all workers and firms prefer apprenticeship training to regular work so that the number of trained workers is equal to $n^{FB} = 1$. This condition is obtained by excluding the other two kinds of inefficiencies, i.e. by assuming $\beta_2 = 0$ and $\rho = 0$. Therefore, the fixed training wage $w_A$ is inefficient for those workers whose innate ability does not lie within the following range:

$$\beta_1 \theta \leq w_A \leq [\chi - (1-\beta_1) + \delta\alpha]\, \theta$$

If the training wage is too high relative to the output of an apprentice, the firm is deterred from training the worker.

of apprenticeship training could be achieved by $w_A = 0$. Because all workers should be trained and the productivity of the apprentice with the lowest ability is equal to $\chi\theta_L = 0$, the training wage has to satisfy $w_A = 0$ in order to guarantee the first-best number of apprenticeship training positions. However, as discussed in Section 2.2.2, the training wage is set by the collective bargaining parties and does not constitute a decision variable of the government. Because regulating $w_A$ would require to restrict the level of tariff autonomy in the economy, we assume that this policy instrument is not available and the government takes the training wage as given when deciding on the optimal level of penalty charges.

In a nutshell, all three kinds of inefficiencies are generated by the "hold-up" problem of incomplete contracts, which means that one party (i.e. the current employer) bears the costs of apprenticeship training while another party (i.e. the worker and the future employer) shares in the return. These inefficiencies would be solved if firms and workers could write a complete contract in which the external effects (i.e. both the costs and benefits) of apprenticeship training are internalized. Unfortunately, the positive probability of exogenous separation after period 1 makes it necessary to take into account the profits of the new employer in period 2. Obviously, this is not possible because the future employer is unknown in advance. Furthermore, the third kind of inefficiency would require to bargain on $w_A$, which is made unfeasible by the downward rigidities in the training wage as discussed in the context of the German apprenticeship system. Hence, Coasian bargaining on the training wage and writing a complete contract covering all beneficiaries of apprenticeship training are not feasible. Taken together, these inefficiencies constitute the necessary condition for welfare-enhancing government interventions. Penalty charges are a possible policy instrument to move the economy towards its first-best optimum by increasing the number of apprenticeship training positions.

Compared to the implementation of penalty charges, there are other policy instruments which may be even superior concerning the maximization of overall welfare. For example, the government could reverse the statutory incidence by imposing penalty charges on those workers who do not adopt an apprenticeship training position. However, because workers indeed prefer apprenticeship training to regular work (at least for $\alpha \geq \frac{1}{\delta}$, cf. Section 2.3.4) and firms are not allowed to pay a training wage below $w_A$, this measure has no effect on the number of trained workers but may induce some workers to leave the market in the worst case. Furthermore, the government could decide to pay firm-subsidies for each completed apprenticeship as suggested by Malcomson, Maw, and McCormick (2003). With these subsidies, the welfare gain of more apprenticeship training positions has to be weighted against the efficiency loss from distortionary taxation to finance the subsidy. However, we restrict the range of available policy instruments and do not refer to this policy instrument in the following analysis because we exclusively focus on the effects of penalty charges and their implications for overall welfare.

## 2.4 The Welfare Analysis of Penalty Charges

### 2.4.1 The Pivotal Productivities with Penalty Charges

If identical penalty charges $T \geq 0$ are imposed on the firms in the case of regular work, the opportunity costs of not offering apprenticeship training positions are increased.[62] Hence, the pivotal productivity between apprenticeship training and regular work is decreased.

**Definition 2.2** *The pivotal productivity* $\theta_A^{PC}$ *is defined to be the lowest productivity of those workers to whom firms decide to offer an apprenticeship training position if they are exposed to penalty charges. With penalty charges (PC), a firm prefers apprenticeship training to regular work if its profits over*

[62] The training quota for the identical firms is one, i.e. every firm is assigned to train the worker it meets. Consequently, the penalty charges are identical for each untrained worker. Note that penalty charges are analytically the same as fixed costs of production that only accrue in the case of regular work.

*both periods solve*

$$\begin{aligned}
&\chi\theta - w_A + \delta(1-\beta)\left[(1-\rho)(1+\alpha)\theta + \rho E[\theta']\right] \\
&\geq (1-\beta)\theta - T + \delta(1-\beta)\left[(1-\rho)\theta + \rho E[\theta']\right] \\
&\theta \geq \theta_A^{PC} \equiv \frac{w_A - T}{\chi - (1-\beta) + \delta(1-\rho)(1-\beta)\alpha} \qquad (2.23)
\end{aligned}$$

As in Section 2.3.4, $E[\theta']$ cancels out because the firm wants to keep the worker both in the case of apprenticeship training and in the case of regular work. Concerning the pivotal productivity between apprenticeship training and regular work, penalty charges work like a reduction in the training wage. Like $\theta^{LF}$, the pivotal productivity $\theta_A^{PC}$ increases with $\rho$ because the expected return to apprenticeship training is reduced for the training firm if separation becomes more likely.

However, some firms decide to leave the market in period 1 (with zero profits) and to reenter in period 2 in order to avoid the financial burden of penalty charges. Hence, workers with very low ability remain unemployed in period 1 because firms do no longer make profits by employing them regularly.

**Definition 2.3** *The pivotal productivity $\theta_U^{PC}$ is defined to be the lowest productivity of those workers to whom firms decide to offer regular work compared to zero profits in the case of market exit. With penalty charges, a firm prefers regular work to market exit if its profits over both periods solve*

$$\begin{aligned}
&(1-\beta)\theta - T + \delta(1-\beta)E[\theta'] \geq \delta(1-\beta)E[\theta'] \\
&\theta \geq \theta_U^{PC} \equiv \frac{T}{1-\beta} \qquad (2.24)
\end{aligned}$$

Again, $E[\theta']$ cancels out because the firm decides to meet a new worker both in the case of regular work and in the case of no production. Because separation is sure after the first period, the pivotal productivity $\theta_U^{PC}$ is independent of $\rho$.[63]

[63] Note that the right hand side in Definition 2.2 and the left hand side in Definition 2.3 are not identical because the first one refers to those firms with regular work that want

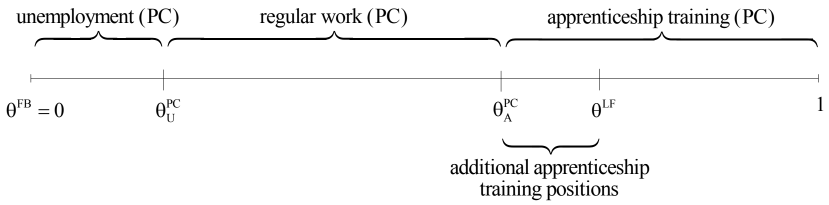


Figure 2.5: The Equilibrium with Penalty Charges

Before analyzing the equilibrium with penalty charges, we have to determine the relationship of the three pivotal productivities $\theta^{LF}$, $\theta_A^{PC}$ and $\theta_U^{PC}$.

**Proposition 2.5** *If penalty charges do not exceed some upper bound*

$$T \leq \bar{T} \equiv \frac{(1-\beta)\, w_A}{\chi + \delta\,(1-\rho)\,(1-\beta)\,\alpha}$$

*the relationship of the pivotal productivities is the following (cf. Appendix A):*

$$\theta^{LF} \geq \theta_A^{PC} \geq \theta_U^{PC}$$

### 2.4.2 The Equilibrium with Penalty Charges

With penalty charges, $n^{PC} = 1 - \theta_A^{PC}$ workers are trained in period 1. Hence, there are $n^{PC} - n^{LF} = \theta^{LF} - \theta_A^{PC}$ additional apprenticeship training positions compared to the laissez-faire equilibrium. All workers with individual ability between $\theta_A^{PC}$ and $\theta_U^{PC}$ are employed regularly while the low-ability workers with $\theta < \theta_U^{PC}$ remain unemployed.[64] This situation is illustrated in Figure 2.5.

---

to keep their worker while the second one refers to those firms that decide to meet a new worker in period 2.

[64]Note that high penalty charges $T = \bar{T}$ imply $\theta_A^{PC} = \theta_U^{PC}$ so that the number of apprenticeship training positions is maximized at the cost of suppressed regular work. In this case, also the number of unemployed workers is maximal.

Altogether, the aggregate welfare of workers with penalty charges is equal to

$$W_W^{PC} = \underbrace{\int_{\theta_A^{PC}}^{1} \left[w_A + \delta\beta\left(1+\alpha\right)\theta\right] d\theta}_{\text{apprenticeship training}} + \underbrace{\int_{\theta_U^{PC}}^{\theta_A^{PC}} \left(1+\delta\right)\beta\theta d\theta}_{\text{regular work}} + \underbrace{\int_{0}^{\theta_U^{PC}} \delta\beta\left(1-\sigma\right)\theta d\theta}_{\text{unemployment}} \tag{2.25}$$

The first integral in (2.25) is equal to the discounted aggregate wage sum of trained workers while the second one describes the discounted aggregate wage sum of workers who are employed regularly in both periods. The third integral is equal to the discounted aggregate wage sum of low-ability workers who are employed regularly in period 2 but remain unemployed in the first period. These workers with $\theta < \theta_U^{PC}$ earn nothing in period 1. In the second period, there is no unemployment. All workers are employed regularly and compensated according to the Nash bargaining solution.

Concerning the aggregate welfare of firms, we must also consider the penalty charges imposed on all $(\theta_A^{PC} - \theta_U^{PC})$ firms that employ workers regularly in period 1. Compared to the laissez-faire equilibrium, the profits of these firms are reduced by $T$ in the first period. Hence, the aggregate welfare of firms with penalty charges is equal to

$$\begin{aligned} W_F^{PC} &= \underbrace{\int_{\theta_A^{PC}}^{1} \left[\left(\chi + \delta\left(1-\beta\right)\left(1+\alpha\right)\right)\theta - w_A\right] d\theta}_{\text{apprenticeship training}} \\ &\quad + \underbrace{\int_{\theta_U^{PC}}^{\theta_A^{PC}} \left[\left(1+\delta\right)\left(1-\beta\right)\theta - T\right] d\theta}_{\text{regular work}} + \underbrace{\int_{0}^{\theta_U^{PC}} \delta\left(1-\beta\right)\left(1-\sigma\right)\theta d\theta}_{\text{unemployment}} \end{aligned} \tag{2.26}$$

The first integral is equal to the discounted aggregate profits of training firms (period 1) and firms producing with a trained worker (period 2). While the second integral describes the discounted aggregate profits of firms that decide to employ workers regularly in both periods, the third one is equal to the discounted aggregate profits of those firms producing with low-ability workers in period 2. As in Section 2.3.4, the expected productivity $E[\theta']$ does not show up in equation (2.26).

In order to determine the overall welfare with penalty charges, we have to consider the total amount of penalty charges which is not lost but may be spent by the government. Penalty charges exert an indirect influence on the overall welfare in (2.27) by changing the pivotal productivities compared to the laissez-faire equilibrium, but they cancel out as a direct determinant because they represent pure lump-sum transfer from the firms to the government. Additionally, the implementation of penalty charges generates administration costs, which are assumed to increase with the level of penalty charges:[65]

$$C(T) = \frac{c}{2}T^2$$

Altogether, the overall welfare with penalty charges is equal to

$$W^{PC}(T) = \underbrace{\int_{\theta_A^{PC}}^{1} (\chi + \delta(1+\alpha))\,\theta d\theta}_{\text{apprenticeship training}} + \underbrace{\int_{\theta_U^{PC}}^{\theta_A^{PC}} (1+\delta)\,\theta d\theta}_{\text{regular work}} + \underbrace{\int_{0}^{\theta_U^{PC}} \delta(1-\sigma)\,\theta d\theta}_{\text{unemployment}} - C(T) \tag{2.27}$$

[65] Our analytical results will be strengthened if $C(T)$ has to be financed by distortionary taxation and thus generates an efficiency loss of $\lambda$ per unit of money. However, we neglect these additional efficiency costs, which implies that our results are not directly comparable to the policy instrument of firm-subsidies for each completed apprenticeship as mentioned at the end of Section 2.3.4. Note that our analytical results concerning the optimal penalty charges were qualitatively the same if we excluded any costs of administration and completely focused on the efficiency costs of unemployment among low-ability workers. In this case, the advantages of penalty charges would be even larger.

In order to determine the optimal penalty charges, we have to maximize (2.27) with respect to $T$:

$$\underbrace{(\chi - 1 + \delta\alpha)\,\theta_A^{PC}\left(-\frac{\partial\theta_A^{PC}}{\partial T}\right)}_{\text{more apprenticeship training positions}} = \underbrace{(1 + \delta\sigma)\,\theta_U^{PC}\frac{\partial\theta_U^{PC}}{\partial T}}_{\text{more unemployment}} + \underbrace{cT}_{\text{administration costs}} \tag{2.28}$$

The first-order condition (2.28) compares the marginal benefits (on the left hand side) and the marginal costs (on the right hand side) of an increase in $T$. On the one hand, the welfare gains arise because the number of apprenticeship training positions is increased by $(-\frac{\partial\theta_A^{PC}}{\partial T})$, which generates additional productivity of $(\chi - 1 + \delta\alpha)$ per unit of initial ability. On the other hand, there are two negative welfare effects which are shown on the right hand side of (2.28). First, the number of unemployed workers is increased by $\frac{\partial\theta_U^{PC}}{\partial T}$, which implies a reduction in productivity of $(1 + \delta\sigma)$ per unit of initial ability. Second, there is an additional welfare loss of $cT$ due to the administration costs.

Altogether, the implementation of penalty charges faces a trade-off with respect to overall welfare. On the one hand, penalty charges are welfare-enhancing because they increase the number of apprenticeship training positions and thus the fraction of skilled workers in the second period. Hence, there are higher wages and higher profits after the apprenticeship because the output is shared between workers and firms. On the other hand, penalty charges are costly because some firms will leave the market (which implies unemployment among workers with low ability) and administration costs are raised.

Given the inefficiencies of the laissez-faire equilibrium as discussed in Section 2.3.4, small penalty charges generate a first-order welfare gain by stimulating apprenticeship training, while a small deviation of unemployment from its first-best level only creates a second-order welfare loss. Hence, by taking into account both welfare effects, small penalty charges give rise to an overall welfare gain. However, as the penalty charges are raised further, the welfare loss from increasing unemployment becomes larger while the efficiency gain from additional apprenticeship training positions becomes smaller as the economy is moved towards the first-best level of training. Including the administration costs of penalty charges, there must be a well determined optimal level of penalty charges that maximizes the overall welfare by equating the welfare gain on the left hand side and the welfare loss on the right hand side in (2.28).

The optimal penalty charges $T^*$ are obtained by substituting the pivotal productivities $\theta_A^{PC}$ and $\theta_U^{PC}$ into equation (2.28) and solving for $T$:[66]

$$T^* = \frac{(1-\beta)^2 (\chi - 1 + \delta\alpha)}{(1-\beta)^2 (\chi - 1 + \delta\alpha) + (v_1)^2 v_2} w_A \tag{2.29}$$

with $v_1 \equiv \chi - (1-\beta) + \delta (1-\rho)(1-\beta)\alpha$ and $v_2 \equiv 1 + (1-\beta)^2 c + \delta\sigma$. In order to guarantee $T^* \geq 0$, the productivity-enhancement of apprenticeship training has to exceed some lower bound: $\alpha \geq \frac{1-\chi}{\delta}$. If this condition is satisfied, apprenticeship training raises the overall welfare as it is already shown in equation (2.9). Hence, the implementation of penalty charges increases the overall welfare only if all workers should be trained in the first-best optimum (cf. equation (2.10)).[67]

Finally, in order to describe the optimal penalty charges depending on the level of the productivity-enhancement, we have to consider the participation constraint of workers as illustrated in Section 2.3.4. Although the number of apprenticeship training positions in the laissez-faire equilibrium deviates from the first-best optimum for $\alpha \geq \frac{1-\chi}{\delta}$ (cf. equation (2.22) in Section 2.3.4), the optimal penalty charges are greater than zero not until the productivity-enhancement exceeds the lower bound $\alpha_0 \equiv \frac{1-\chi}{\delta(1-\rho(1-\beta))}$, which lies above $\frac{1-\chi}{\delta}$. The reason is that the implementation of penalty charges increases the overall welfare only if the participation constraint is satisfied for a sufficient number

[66] The calculation of $T^*$ is shown in Appendix B.1. Note that $T^*$ describes a maximum because $\alpha > 0$ guarantees that the second derivative is negative, i.e. $\frac{\partial^2 W^{PC}(T)}{\partial T^2} < 0$.

[67] Note that $T^*$ never reaches the upper limit $\bar{T}$ if the administration costs of penalty charges do not fall below some critical level. These considerations for $0 \leq T^* \leq \bar{T}$ are explained in Appendix B.2.

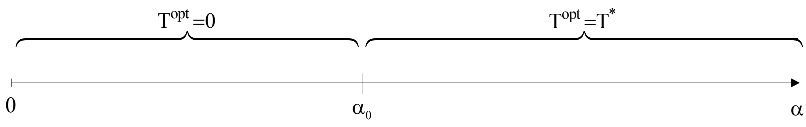


Figure 2.6: The Optimal Penalty Charges

of workers (cf. Appendix B.3). These conclusions are summarized in the following proposition and illustrated in Figure 2.6.

**Proposition 2.6** *Depending on the level of the productivity-enhancement, the optimal penalty charges are equal to*

$$T^{opt} = \begin{cases} 0 & \text{if} \quad \alpha < \alpha_0 \\ T^* & \text{if} \quad \alpha \geq \alpha_0 \end{cases} \tag{2.30}$$

The optimal level of penalty charges explicitly depends on the productivity-enhancement of apprenticeship training. If $\alpha$ is low, it is optimal to reject the implementation of penalty charges, i.e. it is optimal to choose $T^{opt} = 0$ because an additional number of apprenticeship training positions cannot compensate for the costs of administration and the welfare loss of unemployment among low-skilled workers. However, if $\alpha$ exceeds the critical level $\alpha_0$, it is optimal to implement penalty charges according to $T^{opt} = T^*$. The critical level $\alpha_0$ decreases with the relative efficiency of apprentices, the discount factor, and the bargaining power of workers. However, it increases with the probability of exogenous separation after the first period.

### 2.4.3 Interpretation

It is important to analyze in which way the optimal penalty charges are affected by changes in the key parameters of the model.

**Proposition 2.7** *For $\alpha \geq \alpha_0$, the comparative statics of the optimal penalty charges are as follows:*[68]

$$\frac{\partial T^*}{\partial \alpha} > 0 \qquad \textit{if} \;\; \alpha < \bar{\alpha}$$
$$\frac{\partial T^*}{\partial \sigma} < 0$$
$$\frac{\partial T^*}{\partial \rho} > 0$$
$$\frac{\partial T^*}{\partial c} < 0$$
$$\frac{\partial T^*}{\partial w_A} > 0$$

In general, the optimal penalty charges are high if the benefits of apprenticeship training are large and if the costs of unemployment are low. Hence, $T^*$ primarily increases with $\alpha$ because a greater productivity-enhancement raises the benefits of additional apprenticeship training positions. However, this conclusion is only true up to some upper limit of $\alpha$ because the pivotal productivity $\theta^{LF}$ decreases with $\alpha$ (cf. Section 2.3.4). For $\alpha \geq \bar{\alpha}$ (the critical value $\bar{\alpha}$ is defined in Appendix B.4), the number of apprenticeship training positions in the laissez-faire equilibrium is already close to the first-best optimum and the inefficiencies are rather small. Therefore, the welfare gain from stimulating apprenticeship training is rather low compared to the welfare loss from increased unemployment among low-ability workers, which implies that the optimal penalty charges become lower for high values of $\alpha$.

Technically, this relationship between $T^*$ and $\alpha$ is illustrated by the left hand side of the first-order condition (2.28) because the marginal benefits of penalty charges depend on the productivity-enhancement in three ways. While the discounted relative productivity gain $\delta\alpha$ proportionally increases with $\alpha$, both the number of additional apprenticeship training positions

[68]The calculations of the comparative statics are presented in Appendix B.4. Note that the comparative statics with respect to $\alpha$ will be even more positive if we assume an inverse relationship between $\chi$ and $\alpha$.

$\left(-\frac{\partial\theta_A^{PC}}{\partial T}\right)$ and the innate ability $\theta_A^{PC}$ of the additional apprentices are reduced. Taken together, the effect of an increase in $\alpha$ on the marginal benefits of penalty charges is ambiguous. For $\alpha < \bar{\alpha}$, the marginal benefits increase with $\alpha$ because the first effect outweighs the latter, which is initially small. However, for high values of $\alpha$, the negative effect becomes larger while the relative productivity gain remains constant. Hence, there must be a critical level $\bar{\alpha}$ where the two opposite effects are just equal.

Furthermore, the optimal penalty charges decrease with $\sigma$ and $c$ because the marginal costs on the right hand side of (2.28) become larger. While an increase in $\sigma$ and thus a greater depreciation of skills during unemployment increases the welfare loss of unemployed workers, the parameter $c$ directly raises the costs of administration and makes the implementation of penalty charges more costly.

Altogether, the optimal penalty charges are higher, the larger the inefficiencies in the training decision of the firms and thus the larger the welfare gain by moving the number of apprenticeship training positions towards its first-best optimum. As discussed in Section 2.3.4, the degree of distortion increases with the probability of exogenous separation because the firms consider less the workers' higher productivity in period 2. Hence, the marginal benefits of penalty charges increase with $\rho$, which can be seen on the left hand side of equation (2.28).

For the same reason, the comparative statics of $T^*$ with respect to the training wage are positive because penalty charges reduce the inefficiencies in the number of apprenticeship training positions by affecting $\theta_A^{PC}$ and $\left(-\frac{\partial\theta_A^{PC}}{\partial T}\right)$ like a reduction in $w_A$.[69] Note that the optimal penalty charges would be smaller if the government could reduce $\rho$ and $w_A$, for example, by institutionally weakening the bargaining power of unions with respect to the training wage.

[69] This effect is also true for $\alpha_0 \leq \alpha < \frac{1}{\delta}$ because the implementation of penalty charges increases the opportunity costs of regular work for the firms without reducing the workers' incentives to receive apprenticeship training.

| Worker type | Welfare difference |
|---|---|
| Workers with high ability ($\theta^{LF} \leq \theta \leq 1$) | $\pm 0$ |
| Workers with middle ability ($\theta_A^{PC} \leq \theta < \theta^{LF}$) | $w_A + \beta(\delta\alpha - 1)\theta > 0$ |
| Workers with middle ability ($\theta_U^{PC} \leq \theta < \theta_A^{PC}$) | $\pm 0$ |
| Workers with low ability ($0 \leq \theta < \theta_U^{PC}$) | $-\beta(1+\delta\sigma)\theta < 0$ |

Table 2.2: Welfare Effects for Workers Depending on Individual Ability

By substituting the optimal penalty charges (2.30) into equation (2.27), we obtain the overall welfare with optimal penalty charges.

**Proposition 2.8** *With optimal penalty charges, the overall welfare is equal to*[70]

$$(W^{PC})^* = \begin{cases} W^{LF} & \text{if } \alpha < \alpha_0 \\ \frac{1}{2}(\chi + \delta(1+\alpha)) - \frac{1}{2}\frac{(\chi-1+\delta\alpha)v_2}{(1-\beta)^2(\chi-1+\delta\alpha)+(v_1)^2 v_2}(w_A)^2 & \text{if } \alpha \geq \alpha_0 \end{cases}$$

Therefore, the overall welfare is increased by the implementation of penalty charges if the productivity-enhancement exceeds the lower bound $\alpha_0$, which makes the optimal penalty charges (2.30) greater than zero.

Finally, while the previous analysis focuses on the welfare consequences of penalty charges in aggregate, it is also important to investigate the implications for different groups of workers. High-ability workers are not affected by the implementation of penalty charges because they are trained anyway. While workers with middle ability above average benefit from penalty charges because firms now offer them apprenticeship training positions, low-ability workers suffer from unemployment in period 1 and reduced wages in period 2. These different implications are summarized in Table 2.2. In a nutshell, the increased number of apprenticeship training positions is achieved at the cost of unemployment among workers with low ability.

[70] The calculation of the overall welfare with optimal penalty charges is presented in Appendix B.5. The comparative statics of $(W^{PC})^*$ are shown in Appendix B.6.

## 2.5 Conclusion

Chapter 2 of my PhD thesis presents a two-period partial-equilibrium model that systematically compares the costs and benefits of penalty charges. As discussed in Section 2.2, there are two theoretical explanations in the literature for firms providing general training. Our formal analysis is based on recent training literature with oligopolistic labor markets but the model is adapted to the German system of apprenticeship training with a fixed length of the apprenticeship and an identical training wage. Furthermore, the model incorporates worker heterogeneity in ability which allows to analyze the welfare implications of penalty charges for different groups of workers.

In the laissez-faire equilibrium, the training decision of firms bears three different kinds of inefficiencies. As a consequence, the number of apprenticeship training positions may be too low compared to the first-best optimum. In line with the results of Acemoglu and Pischke (1999b), firms do not take into account the benefits from increased productivity accruing both for workers and other employers in the future. These two inefficiencies increase with the workers' bargaining power and the exogenous separating probability after the training period. In our model, there is a third kind of inefficiency, which is generated by the fixed training wage during the apprenticeship. However, penalty charges can reduce this inefficiency and increase the number of apprenticeship training positions towards the first-best optimum.

In our model, the implementation of penalty charges faces a trade-off with respect to overall welfare. While penalty charges increase the number of apprenticeship training positions and thus the fraction of trained workers in the workforce, some firms will leave the market to avoid the financial burden, which implies unemployment among workers with low ability. Additionally, the implementation of penalty charges generates administration costs. The formal analysis suggests that the optimal policy depends on the productivity-enhancement of apprenticeship training. If the increase in productivity due to an apprenticeship is low, it is optimal to reject the implementation of penalty charges. In this case, the laissez-faire equilibrium corresponds to the first-best optimum. However, the implementation of penalty charges increases the overall welfare if the productivity-enhancement exceeds some lower bound.

As mentioned in Section 2.3.4, there may be better policy instruments than penalty charges to induce apprenticeship training without generating detrimental effects on the number of unemployed workers. For example, the government could regulate the training wage towards zero (at zero overall welfare costs but with political restrictions by the tariff autonomy) or subsidize firms for each completed apprenticeship (at welfare costs from distortionary taxation to finance the subsidy). However, in our analysis we restrict the range of policy instruments in order to focus on the welfare effects of penalty charges. Hence, our results only suggest that penalty charges may increase the overall welfare compared to the laissez-faire equilibrium. Future research should take into account other policy instruments and compare their welfare implications with those of penalty charges.

In this model, the number of apprenticeship training positions and the number of unemployed workers are endogenously determined and depend on the individual ability of workers. This extension advances the analysis of Malcomson, Maw, and McCormick (2003), where the number of trained workers is determined by some fixed costs of training. Nevertheless, our model has been kept simple for expositional and calculational reasons. For example, we assume a uniform distribution of abilities and do not explicitly refer to the process of collective wage setting for apprentices. The theoretical results of our stylized model only allow for qualitative conclusions concerning the implementation of penalty charges in the context of the German system of apprenticeship training. In order to assess the quantitative magnitude of these effects, we would have to estimate the elasticities of the firms' labor demand and training responses at the extensive margin. However, the underlying insights into the model presented here are robust to various types of generalization. Hence, they constitute a promising basis for policy recommendations and future research.

# Chapter 3

# Tax Credits for Low-Skilled Workers

*There is an ongoing discussion in Germany about the implementation of tax credits in order to reintegrate low-skilled workers into the labor market. Chapter 3 of my PhD thesis aims at analyzing the policy instrument of tax credits in a theoretical model that systematically compares the costs and benefits in the context of the German system of apprenticeship training and social security. Building on recent training literature, a two-period partial-equilibrium model is developed that allows for worker heterogeneity in ability.*

*In our model, the implementation of tax credits in terms of a negative income tax faces a trade-off with respect to overall welfare. While tax credits reduce the number of unemployed workers at the extensive margin, they increase at the same time the opportunity costs of apprenticeship training, which implies that human capital formation is decreased. Furthermore, the model suggests that it may not be optimal to reintegrate those workers at the bottom of the ability distribution into the labor market.*

## 3.1 Introduction

There is an ongoing discussion in Germany about the implementation of tax credits in order to reintegrate those low-skilled workers into the labor market who would remain unemployed otherwise. Since the seventies, Germany has experienced a dramatic increase in structural unemployment.[71] The German labor markets are particularly challenged by increasing unemployment among workers with low qualification.[72] Compared to other OECD countries, the unemployment of low-skilled workers is very high and is largely made up of individuals suffering long spells (OECD (2004)).

According to recent analyses of the IMF that concentrate on the supply side of the labor market, one major reason lies in the compressed wage structure, i.e. in the downward rigidity in wages across skill categories (IMF (2005)).[73] This wage compression accrues because labor markets are heavily regulated and the level of social assistance is high. The German labor markets are especially characterized by high levels of employment protection and strong unions in the process of collective wage setting as well as downward rigidities in wages of unskilled workers. In brief, low-skilled workers decide to remain unemployed because their potential labor income falls below the level of social assistance as defined by the German system of social security (Sinn, Holzner, Meister, Ochel, and Werding (2006)).

In order to approach the problem of rising unemployment among low-skilled workers, tax credits are proposed depending on the level of individual labor income. With tax credits, a worker receives an individual subsidy if he decides to enter the labor market. This policy instrument is part of the so-called welfare-to-work strategy, which should reduce poverty by raising

[71] Cf. e.g. Bertola (2001) and Nickell (1997).

[72] In 2000, the rate of unemployment among workers without formal education was 19.4% in West Germany and 50.3% in East Germany (Reinberg and Hummel (2002)).

[73] Note that wage compression in imperfect labor markets is considered to be the major source of firm-sponsored general training (cf. Section 1.3.1). However, this aspect is neglected in our analysis because we focus on the labor supply and training decision of workers.

employment of low-skilled workers rather than by increasing welfare benefits for the unemployed. One subsidy scheme, which has been extensively discussed in the literature, is the *Earned Income Tax Credit* (EITC) in the US. Unfortunately, the theoretical analysis of tax credits in the context of both labor supply and human capital formation has been fragmentary so far. While previous research on tax credits has mainly focused on the effects on labor supply and employment (for example Meyer and Rosenbaum (2001)), the impact on skill formation has been widely neglected. There are only few contributions that also consider the accumulation of human capital (for example Heckman, Lochner, and Cossa (2002)).

Chapter 3 of my PhD thesis aims at closing this gap by developing a two-period partial-equilibrium model that systematically compares the costs and benefits of tax credits. It is important to incorporate the training decision into the analysis of tax credits because subsidies to low-skilled workers increase the opportunity costs of training and thus reduce the workers' incentives to acquire skills (Heckman (2002)). In a nutshell, there are three key questions considered in this chapter: First, what is the impact of tax credits on labor supply at the extensive margin? Second, what are the effects of tax credits on human capital formation? And third, what is the optimal level of tax credits subject to the training decision of workers?

The contribution of this chapter is twofold because the formal analysis of tax credits, which is based on recent training literature with oligopolistic labor markets, is extended in two important ways. First, we bring together the theoretical explanations of labor supply and human capital formation in the context of the German system of apprenticeship training and social security. Second, our model allows for worker heterogeneity in ability and manages to explain endogenously the labor supply and training decision of workers at the extensive margin.

In our model, the implementation of tax credits in terms of a negative income tax faces a trade-off with respect to overall welfare. While tax credits reduce the number of unemployed workers at the extensive margin, they increase at the same time the opportunity costs of apprenticeship training, which implies that human capital formation is decreased. Furthermore, the model suggests that it may not be optimal to reintegrate those workers at the bottom of the ability distribution into the labor market. Because the costs in terms of decreased human capital formation would be too high, it may be more efficient to leave aside those workers with the lowest productivities.

Chapter 3 of my PhD thesis proceeds as follows: the next section discusses the concept of tax credits and its implications for labor supply and human capital formation. In Section 3.3, the institutional setting of the German system of apprenticeship training and social security is illustrated. In Section 3.4, our partial-equilibrium model is developed and the laissez-faire equilibrium without tax credits is presented. We show that the system of social security generates unemployment among low-skilled workers because private employment is crowded out by the welfare state. In Section 3.5, the implementation of tax credits is analyzed and the welfare effects are derived. Section 3.6 concludes.

## 3.2 Tax Credits, Labor Supply, and Human Capital

### 3.2.1 The Concept of Tax Credits

As a consequence of social assistance and downward rigidities in wages across skill categories, workers with low qualification have become less employable (Phelps (1997)).[74] Hence, economic policy should promote wage flexibility in order to reduce unemployment among low-skilled workers. For example, incentives for unemployment could be reduced by lowering the level of social assistance. However, the implementation of these policy instruments is difficult due to political restrictions or because taxes and social assistance are set by distinct political institutions based on rather different interests (Boone

[74] This result is formally derived in Section 3.4.5.

and Bovenberg (2004)).[75]

In order to approach the problem of rising unemployment among low-skilled workers, tax credits are proposed depending on the level of individual labor income. This implies that the government pays a subsidy to those workers who are in employment and whose income does not exceed some critical level. By granting additional income only in the case of employment, some formerly unemployed workers will thus be motivated to enter the labor market because the combination of their own labor income plus the subsidy makes them better off than in the case of unemployment (Phelps (1997)). To some extent tax credits can be classified within the general set of wage subsidies. According to Phelps (1997), wage subsidies are individually based, not means-tested, and with limited duration. Eligibility usually depends on a certain duration of receipt for the unemployment insurance. In contrast, tax credits are typically means-tested and do not show a time limitation (Blundell (2005)).[76]

Compared to reductions in the level of social assistance, tax credits are a policy instrument to increase employment without lowering the standard of living because regular wages of unskilled workers are very low (the so-called "working poor") (Snower (1994)). However, compared to traditional social assistance, tax credits are less targeted at those individuals who suffer from involuntary unemployment (Boone and Bovenberg (2006)). In a nutshell, Immervoll, Kleven, Kreiner, and Saez (2005) conclude that unemployment due to labor market imperfections increases the attractiveness of tax credits and reduces the desirability of traditional social assistance.

In recent years, several countries have introduced tax credits in various forms. In Europe, the most important examples are the *Working Families Tax Credit* (WFTC) in the UK, the *In-Work Tax Credit* in Belgium, the *Fam-*

---

[75] For example, in some federal countries local governments determine social benefits while the central government decides on the tax system (Boone and Bovenberg (2004)).

[76] Orszag and Snower (2003) distinguish between wage subsidies, which are paid to all workers with low income and not limited in time, and hiring subsidies, which are targeted exclusively at unemployed workers and limited in time.

*ily Income Support Programme* in Ireland, and the *Employment Tax Credit* in the Netherlands (OECD (2005)). However, theoretical and empirical research has mainly focused on the *Earned Income Tax Credits* (EITC) in the US. This transfer scheme works as a tax credit for workers with low labor incomes up to some critical level beyond which the subsidy is phased out. The EITC aims at reducing working poverty as well as generating greater labor supply incentives for unemployed workers (Steuerle (1990)). In the context of the institutional setting in Germany, we refer to the concept of "combined wages" (cf. Section 3.3.3), which represents the German approach to the implementation of tax credits.

In a general equilibrium model with continuous distribution of abilities, Hanushek, Leung, and Yilmaz (2003) investigate different redistribution policies and their implications for labor supply and human capital formation concerning the trade-off between equity and efficiency. They conclude that wage subsidies, which are directed to uneducated workers, dominate the implementation of education subsidies and the traditional negative income tax (NIT).[77] By incorporating both margins of labor supply into the original approach of Mirrlees (1971), Saez (2002) analyzes the welfare consequences of the traditional NIT and the EITC. He shows that subsidizing low-income

[77] Based on Mirrlees (1971), optimal income tax theory shows that redistribution should take the form of a NIT in order to reduce the high marginal tax rates on traditional social assistance (Immervoll, Kleven, Kreiner, and Saez (2005)). In general, the traditional NIT implies the following tax scheme, which depends on the level of individual labor income:

$$T(I) = tI - y \begin{cases} \leq 0 & \text{if } I \leq \bar{I} \equiv \frac{y}{t} \\ > 0 & \text{if } I > \bar{I} \end{cases}$$

The NIT implies that all individuals (independent of their status of employment) receive a basic lump-sum transfer $y$, which corresponds to the guaranteed income of each individual. Hence, individuals with low income $I$ below the critical level $\bar{I}$ face a negative amount of tax liability, i.e. they receive an income subsidy from the government. With increasing individual labor income, the subsidy is phased out according to the tax rate $t$. Note that the NIT is not restricted to workers in employment but also applies to unemployed workers. This is the main difference compared to the concept of tax credits.

workers by tax credits is welfare-enhancing if the labor supply response is concentrated along the extensive margin.

### 3.2.2 Implications of Tax Credits for Labor Supply

A central finding of recent empirical literature in public finance and labor economics is that labor supply responses are concentrated more at the extensive margin (participation) than at the intensive margin (hours of work) (Blundell and MaCurdy (1999)). Eissa, Kleven, and Kreiner (2004) point out that the extensive margin is particularly important for labor supply incentives at the bottom of the income distribution.

The reason for the dominance of the extensive margin could be that hours worked are fixed, which implies that workers face quantity restrictions when they decide to enter the labor market (Hausman (1985)). Simplifying, workers have the discrete choice between work and unemployment but can hardly decide on hours worked at the intensive margin (Zabalza, Pissarides, and Barton (1980)). Indeed, empirical evidence shows that workers decide either to remain unemployed or to work at least some minimum number of hours (Eissa, Kleven, and Kreiner (2006)).[78]

Theoretically, tax credits give rise to a trade-off between low-skilled employment at the extensive margin and work effort of high-skilled workers at the intensive margin (Boone and Bovenberg (2004)). However, evidence from the EITC in the US (Eissa and Liebman (1996)) and the WFTC in the UK (Blundell and Hoynes (2001)) shows substantial positive effects at the extensive margin, but only small negative effects on hours of work for those workers who are in employment. With respect to the EITC, the empirical analysis by Meyer and Rosenbaum (2001) shows that the annual employment of single mothers has increased by 9% between 1984 and 1996.[79] Because hours worked

[78] This discrete labor supply behavior is theoretically explained by non-convexities due to fixed costs of working (Cogan (1981)). Eissa and Hoynes (2005) summarize further possible reasons why the extensive margin is more responsive than the intensive margin.

[79] Single mothers represent over three-quarters of all EITC recipients (Eissa and Hoynes

have only slightly fallen at the intensive margin, Meyer (2002) suggests that labor supply adjustments mainly take place at the extensive margin. In almost the same manner, Heckman (1993) concludes that the extensive margin is empirically much more important than the intensive margin. With respect to European countries, Immervoll, Kleven, Kreiner, and Saez (2005) use the EUROMOD micro-simulation model to demonstrate that tax credits generate positive labor supply responses at the extensive margin.

Concerning the evaluation of tax reforms, Eissa and Hoynes (2005) suggest that ignoring the participation margin may lead to even the wrong sign of the welfare effect. Unfortunately, the theoretical public finance literature has largely ignored the participation decision and instead has focused on labor supply at the intensive margin. For this reason, our model explicitly allows for labor supply responses at the extensive margin if the income of low-skilled workers is raised by the implementation of tax credits.

### 3.2.3 Implications of Tax Credits for Human Capital

The implications of tax credits for human capital formation can be very different and crucially depend on the design of the transfer scheme. With respect to the EITC, the consequences for human capital accumulation are neglected by many contributions. For example, Orszag and Snower (2003) admit that implications for human capital formation lie beyond the scope of their paper.

According to the analysis of Heckman, Lochner, and Cossa (2002), tax credits place important disincentives on human capital accumulation. They suggest that the average skill-level declines because current workers reduce

(2005)). Meyer and Rosenbaum (2001) estimate that 60% of this increase in extensive labor supply is due to the EITC. For single mothers, this labor supply response at the extensive margin is confirmed by Eissa and Hoynes (2004). However, this result may be changed if an integrated model of family labor supply is considered (Hausman (1985)). Indeed, Eissa and Hoynes (2004) argue that their results are different for married couples. While labor force participation of the head of families is increased, it declines for secondary earners.

their investment in human capital at the intensive margin.[80] As a consequence, Blundell (2005) stresses the importance of these interactions between labor supply and human capital formation for the evaluation of tax credits. He suggests that a dynamic analysis of optimal income transfer programs has to take into account the incentive effects on human capital investment.

## 3.3 The Institutional Setting in Germany

### 3.3.1 The Apprenticeship System

Because the institutional setting of the German apprenticeship system has already been discussed in Section 2.2, we only refer to the main assumption we impose in the context of our formal analysis in this chapter. We assume that the length of the apprenticeship is fixed and identical for each apprentice. Although those school graduates with upper secondary education (*Abitur*) can shorten the apprenticeship period, our assumption is warranted because the extent to which the apprenticeship can be shortened is rather small and does not vary continuously with the worker's individual productivity.

Following from this assumption, workers (and firms) decide at the extensive margin (i.e. participation) whether to receive apprenticeship training or to be employed regularly, i.e. to work full-time without formal qualification. At the intensive margin (i.e. hours of training), the accumulation of human capital by apprenticeship training is determined by the fixed length of the apprenticeship and the specified curriculum rather than by investment decisions of the apprentices.

---

[80] The effects on human capital investment are divided into a substitution effect, an income effect, and a direct effect (which accrues due to changes in marginal costs and returns) (Heckman, Lochner, and Cossa (2002)).

### 3.3.2 Labor Market Regulation and Social Security

In Germany, the degree of labor market regulation is high compared to other OECD countries (OECD (2005)). Employment protection and firing costs make it difficult for firms to respond flexibly to changing market conditions. With firing costs, job growth in response to GDP growth is diminished because firms account for the possibility of worsening business prospects in the future. Hence, they hire fewer workers or even decide to leave the market in order to avoid the costs of possibly having to fire them. As a consequence, it becomes harder for unemployed workers to find a job (Heckman (2002)). However, the total effect of firing costs on unemployment is ambiguous because fewer separations lead to lower unemployment (Belot, Boone, and Ours (2002)).[81] Furthermore, employment protection increases the incentives to invest in human capital for both workers and firms because the probability of separation is decreased (Fella (2005)).

Wage bargaining is conducted by the collective bargaining parties. Although Germany does not have a legally mandated minimum wage, union wage floors effectively operate as wage minimums for certain groups of workers. Furthermore, replacement rates by social insurance are substantial. In the sixties, two kinds of social assistance, the *Arbeitslosenhilfe* and the *Sozialhilfe,* were implemented in order to insure workers against the risk of unemployment. According to Immervoll, Kleven, Kreiner, and Saez (2005), total social benefits constituted 72.6% of the disposable income for the lowest decile group in Germany in 1998.

This system of traditional social assistance is criticized for keeping persons on welfare and out of the labor market.[82] The poverty trap accrues because

[81] The empirical evidence concerning the relationship between firing costs and unemployment is mixed. While Scarpetta (1996) and Elmeskov, Marint, and Scarpetta (1998) find a negative correlation, Nickell (1998) does not. Concerning workers with different skill levels, Stähler (2006) finds that high-skilled workers benefit more from the same level of employment protection than low-skilled workers.

[82] In 1999, 2,790,000 people draw benefits from the *Sozialhilfe* and 1,300,000 from the *Arbeitslosenhilfe*. Subject to health condition, family obligation and training measures,

the level of social assistance works as de facto minimum wage. This downward rigidity in wages across skill-categories leads to a compressed wage structure (Schöb and Weimann (2003)). As a consequence, unemployment is generated especially among low-skilled workers because private employment is crowded out by the welfare state as long as the replacement rate exceeds the market wage rate for unskilled workers (Sinn (2003)). Because social benefits are generously granted only in the case of no work, the labor supply incentives of low-skilled workers are reduced by the welfare state (Sinn (2002)). According to Reinberg and Hummel (2002), the rate of unemployment in 2000 among workers without formal education was 19.4% in West Germany and 50.3% in East Germany. Empirical evidence is also shown by Layard and Nickell (1999). These negative effects become even stronger in a dynamic context, for the longer workers are unemployed and the more their skills depreciate (Snower (1994)).

In the context of the German labor market reforms in 2005, both kinds of social benefits were merged to the uniform social assistance *Arbeitslosengeld II*, which falls below the previous replacement rates and is independent of earnings in the past.[83] Furthermore, the requirements of eligibility and sanctions in the case of misuse were aggravated. However, the negative impact on employment of low-skilled workers is reduced but still existent (Ochel (2005)).

### 3.3.3 The German Concept of "Combined Wages"

In Germany, two different concepts of tax credits are discussed which both aim at reintegrating low-skilled workers into the labor market. The policy instrument of tax credits is referred to as "combined wages" because the income of low-skilled workers in employment is augmented by combining

---

the labor force potential was 2,200,000 or about 58% of all beneficiaries (Raffelhüschen (2001)).

[83] According to Peter Hartz, the chairman of the committee working on these labor market reforms, this part of the reform proposals is also referred to as *Hartz IV*.

their wage with an individual subsidy.

The first concept of "combined wages" concentrates on the demand side of the labor market and proposes to pay employment subsidies to those firms that hire a formerly unemployed worker. In this context, some part of the unemployment benefits could be used as vouchers for the firms in order to reduce the net labor costs (Snower (1994)). With respect to the institutional setting in Germany, Schöb and Weimann (2003) suggest to exempt firms from paying social security contributions for low-skilled workers in order to decrease the labor costs of firms by 34% (this proposition is referred to as the *Magdeburger Alternative).*

However, in our formal analysis we will focus on the second concept of "combined wages" that refers to the supply side of the labor market and corresponds to the initial explanations in Section 3.2.1. This approach aims at increasing the labor supply incentives of low-skilled workers by paying them a subsidy if they decide to enter the labor market.

In recent years, there have been various approaches in Germany to implement tax credits according to the second concept of "combined wages". The focus of most approaches has been on long-time unemployed, welfare recipients, or generally workers with low qualification. Besides some regionally defined projects[84], two approaches have been applied nationwide. First, the *Arbeitnehmerhilfe* determines wage subsidies for unemployed workers of at most 13 Euro per day if the working time exceeds 15 hours per week. Second, the *Mainzer Modell* was regionally designed in 2000, extended nationwide in 2002, and terminated by the end of 2003.[85] It was composed of subsidies to

[84] An overview is given by Kaltenborn (2001). For example, nine local authorities in the federal state Baden-Württemberg applied the *Einstiegsgeld* between 1999 and 2002 (Dann, Kirchmann, Spermann, and Volkert (2002)). In Hessen, "combined wages" were introduced by seven local authorities in 2000. It was hardly engaged and replaced by the *Kasseler Modell Kombilohn* (KAMOKO) in 2001. Like the *Mainzer Modell*, it was terminated in 2003.

[85] At the beginning, the *Mainzer Modell* was only designed for the federal states Rheinland-Pfalz and Brandenburg.

social security contributions and child benefits for a minimum working time of 15 hours per week and a gross income of at least 325 Euro per month. For example, for a family with two children, the two components add up to a maximum subsidy of 283 Euro per month.[86] However, the main problem of both approaches has been their limited time horizon. The maximum duration of advancement has been three years in the *Mainzer Modell* and only three months in the *Arbeitnehmerhilfe* (Kaltenborn (2003)). As a consequence, the demand of workers for these wage subsidies has been very low considering the enormous amount of more than two millions of unemployed workers with low qualification.[87]

The limited time horizon of these tax credit programs is the main criticism brought forward by the ifo institute, an economic research institute in Germany. Low-skilled workers should be subsidized permanently because their income permanently falls below the reservation wage as defined by the level of social assistance (Sinn, Holzner, Meister, Ochel, and Werding (2006)). Furthermore, the previous concepts still face high marginal tax rates on social welfare benefits so that working incentives for low-skilled workers are reduced (Sinn (2002)). Hence, tax credits for workers in employment are proposed according to the concept of a NIT in order to generate stronger labor supply incentives. Note that this subsidy scheme is fundamentally different from the traditional concept of a NIT (cf. Section 3.2.1) because no assistance is granted to unemployed workers. Only those workers who are in employment are eligible for tax credits (Sinn, Holzner, Meister, Ochel, and Werding (2006)).[88]

---

[86] Cf. Jülicher (2002) and Bittner, Hollederer, Kaltenborn, Rudolph, Vanselow, and Weinkopf (2001).

[87] The *Arbeitnehmerhilfe* covers about 8.000 workers per year and there were 6.137 participants in the *Mainzer Modell* (Dann, Kirchmann, Spermann, and Volkert (2002)).

[88] The ifo institute defines this concept as "Activating Social Assistance". This means that social benefits are paid in order to activate formerly unemployed workers for the labor market. The whole concept is explained in more detail in Sinn, Holzner, Meister, Ochel, and Werding (2006).

Starting from these considerations, our model analyzes labor supply and training responses of workers at the extensive margin. We refer to the concept of the ifo institute by designing tax credits in terms of a NIT for workers in employment. Concerning the labor force participation of low-skilled workers, it is crucial to consider all distortions that are generated by traditional social assistance and move the equilibrium away from the first-best optimum. For this reason, we first discuss the laissez-faire equilibrium (cf. Section 3.4.4) and the equilibrium with social assistance (cf. Section 3.4.5) in order to point out the analytical basis of comparison. Subsequently, the welfare analysis of tax credits is presented in Section 3.5.

## 3.4 The Model

We consider a discrete-time model with two types of agents, namely workers and firms. In line with Acemoglu and Pischke (1998a), there are two periods, a training period (period 1) and a working period (period 2). The length of both periods is normalized to unity. Production takes place in worker-firm pairs and no capital is needed. According to Eissa, Kleven, and Kreiner (2006), a model of extensive labor supply requires some type of heterogeneity, either in preferences or in ability. In our approach, workers have identical preferences but are heterogeneous in their initial ability which is exogenously given.

At the beginning of period 1, each firm meets one worker whose individual ability is drawn randomly from a distribution that is common knowledge. At the extensive margin, workers and firms decide whether to engage in apprenticeship training, to produce with regular work, or not to produce at all. An apprenticeship takes place only if both parties agree on apprenticeship training.[89] In the second period, all workers can be employed regularly, but only

[89] In line with Acemoglu and Pischke (1999b), there is no exogenous separation after the first period. The implementation of an exogenous separating probability as in Malcomson, Maw, and McCormick (2003) does not change our analytical results because we concentrate

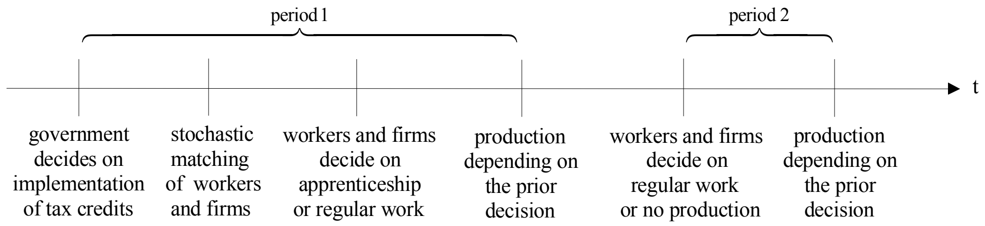


Figure 3.1: The Evolution over Time

those workers who were trained in period 1 have an increased productivity. Altogether, the economy evolves over time as illustrated in Figure 3.1. The model assumptions and the labor market decisions of firms and workers are described in the following subsections.

### 3.4.1 The Workers

At the beginning of period 1, workers differ in their individual ability, which is assumed to be uniformly distributed on the interval $[\theta_L, \theta_H]$.[90] After the match has taken place, firms can unambiguously observe the worker's ability.[91] In conformity with Hanushek, Leung, and Yilmaz (2003) and Malcomson, Maw, and McCormick (2003), the mass of workers is normalized to unity by defining $\theta_L \equiv 0$ and $\theta_H \equiv 1$. By assumption, the mass of firms is also

on the supply side of the labor market. For the workers who face the training decision in period 1, it is irrelevant whether their higher wages in period 2 are paid by their current or by another employer.

[90] The continuous distribution of abilities allows obtaining a smooth participation decision at the individual level (Mirrlees (1971)). In order to keep the following calculations as simple as possible, we assume a uniform distribution of abilities.

[91] This assumption is in line with Boone and Bovenberg (2006). Furthermore, it is implicitly included in the whole literature on human capital and the life-cycle of earnings. Each worker offers his individual stock of human capital to the firms and is rewarded by a rental price per unit of human capital. Hence, we rule out asymmetric information (hidden knowledge). If the worker's productivity were not observed by the firm, there would be adverse selection as modeled e.g. by DeMeza and Webb (2001).

one, which implies that each firm is matched with one worker whose ability is uniformly distributed on $[0, 1]$.

In line with Malcomson, Maw, and McCormick (2003), workers are risk-neutral and maximize the sum of their discounted utilities over both periods:[92]

$$V(\theta) = v + \delta v' \tag{3.1}$$

The discount factor $\delta \equiv \frac{1}{1+r}$ with $r$ as the market interest rate expresses the preference for current and future wealth. The higher $\delta$, the higher is the weighting of period 2 and the lower is the preference for period 1. In the first period, the worker's utility $v$ is equal to the difference between his wage $w$ and potential training costs $e$:

$$v = w - \kappa e \tag{3.2}$$

$$w = w(\theta) = \theta \tag{3.3}$$

By defining the output good as numéraire and assuming an identical, linear one-to-one production function for the connection of output and labor (which is the only factor of production), the marginal product of each worker corresponds to his productivity $\theta$.[93] To simplify matters, we assume that the worker's productivity is the same both in the case of apprenticeship training and in the case of regular work and that there is no unemployment in period 1. Furthermore, in line with Becker (1964), we assume that the worker is the full claimant of his marginal product, which implies that the worker's wage is equal to $w(\theta) = \theta$.[94]

[92] In line with Ben-Porath (1967), we do not analyze a more general utility function of workers. Note that the wage corresponds to the worker's labor income because labor supply is implicitly normalized to unity.

[93] The production function exhibits constant returns to scale. From the firm's point of view, the worker's ability can be interpreted as individual productivity.

[94] Note that most of our analytical results will not change qualitatively if the worker does not receive the full marginal product, but a fraction $\beta$ of his marginal product $\theta$. In this case, wages will be determined by Nash-bargaining. If the parameter $0 \leq \beta \leq 1$ indicates the (identical) bargaining power of workers concerning the division of output and the fall-

$\kappa$ is a dummy variable which is one in the case of apprenticeship training and zero otherwise. In line with the literature on human capital accumulation over the life-cycle, there are training costs $e$ that have to be borne by the worker if apprenticeship training takes place (cf. equation (3.2)).[95] The training costs are identical for each worker because the fixed length of the apprenticeship is defined by the German system of apprenticeship training (cf. Section 2.2). We do not consider Coasian bargaining on the distribution of $e$ between workers and firms because we want to focus on the supply side of the labor market. This simplification is justified because an apprenticeship is assumed to have no negative effect on the worker's wage in the first period. Hence, the worker receives the full wage and does not have to bear additional education costs in terms of a reduced training wage during the apprenticeship.

In period 2, the worker's utility $v'$ is equal to his wage $w'$:

$$v' = w' \tag{3.4}$$

$$w' = \begin{cases} w(\theta') = \theta' & \text{if regular work in } t = 2 \\ 0 & \text{if unemployment in } t = 2 \end{cases} \tag{3.5}$$

$$\theta' = (1 + \kappa\alpha)\,\theta \tag{3.6}$$

In the case of unemployment, the worker receives zero income because there is no social assistance (this assumption is modified in Section 3.4.5).

back payoffs are zero, the bargained wage will be equal to $w(\theta) = \beta\theta$. For $\beta < 1$, there will be labor market frictions because the worker's wage falls below his marginal product. As a consequence, the pivotal ability between apprenticeship training and regular work will exceed its first-best level, which implies that the number of trained workers will be too small in the laissez-faire equilibrium (cf. Section 3.4.4).

[95] Most models that analyze the accumulation of human capital over the life-cycle completely concentrate on the investment decision of workers (for example Ben-Porath (1967) and Heckman (1976)). Although there are two major components of education costs, namely foregone earnings and direct expenditures (Parsons (1974)), we omit the former and focus on the latter by assuming direct costs of education equal to $e$ (cf. equation (4.2)). This assumption is in line with Sheshinski (1971), Atkinson (1973), and Nerlove, Razin, Sadka, and Weizsäcker (1993) who consider only direct costs of education in their analyses of income taxation.

With regular work, the worker's wage corresponds to his marginal product $\theta'$, which depends on the status of employment in period 1 according to equation (3.6). In the second period, the productivity of all workers employed regularly in period 1 is unchanged (i.e. $\theta' = \theta$).

For all trained workers, the productivity increases to $\theta' = (1 + \alpha)\,\theta$, whereas the parameter $\alpha \geq 0$ represents the productivity-enhancement of apprenticeship training. While the relative productivity gain from apprenticeship training is equal to $\alpha$ and thus identical for all trained workers, the absolute productivity gain $\theta' - \theta = \alpha\theta$ is proportional to the worker's innate ability $\theta$.[96] This assumption is motivated by the literature on human capital formation at the intensive margin (cf. Ben-Porath (1967) and Heckman (1976)). Because the productivity-enhancement unambiguously depends on the amount of training, the assumption that $\alpha$ is identical for each worker and independent of $\theta$ implies that the amount of training is the same for all apprentices. This implication is in conformity with the German system of apprenticeship training, where the amount of training per apprentice is defined by the prescribed curriculum rather than by the worker (cf. Section 3.3.1).[97] Furthermore, the identical productivity-enhancement simplifies the formal analysis because the fraction of trained workers can be aggregated more easily.

Altogether, the total utility of a worker with ability $\theta$ (who is in employment in both periods) is obtained by substituting equations (3.2) to (3.6)

[96] Formally, this means $\frac{\partial(\theta' - \theta)}{\partial\theta} = \alpha \geq 0$. Intuitively, the accumulation of new skills is easier when more skills are already available. This relationship is also suggested by Ben-Porath (1967) and Mincer (1997). Because the parameter $\alpha$ determines the productivity and thus the wage in period 2, it constitutes the key determinant of the return to education as analyzed in the theory of human capital (c.f. Section 1.4).

[97] If the firms could decide on their human capital investment not only at the extensive but also at the intensive margin and if the optimal amount of training depended on the workers' innate ability, the productivity-enhancement would not be the same for all workers.

into equation (3.1):

$$V(\theta) = (1 + \delta(1 + \kappa\alpha))\theta - \kappa e \tag{3.7}$$

### 3.4.2 The Firms

Firms are risk-neutral and maximize the sum of their discounted profits over both periods.[98] In both periods, the firm's profits are equal to the difference between revenue and costs per worker. Because the worker always receives a wage equal to his full marginal product, the firm's profits are zero in each possible case, i.e. with apprenticeship training, with regular work as well as without any production.[99] This implies $\pi = 0$ in period 1 and $\pi' = 0$ in period 2 so that total profits are always equal to

$$\pi(\theta) = \pi + \delta\pi' = 0$$

Because the firms never make positive profits by training the worker or employing him regularly, they are just indifferent between production and leaving the market. However, in order to focus on the supply side of the labor market, we simply assume that the firms always agree with the workers' decision. For example, if the worker decides to receive apprenticeship training, the firm accepts this decision and offers an apprenticeship training position. As a consequence, we can neglect the demand side of the labor market in the following analysis because only the workers determine their status of employment in both periods.

[98] The production side is modeled similar to Malcomson, Maw, and McCormick (2003). Note that there is no uncertainty because the probability of exogenous separation is zero.

[99] If the worker's wage is determined by Nash bargaining, the firm will be the residual claimant of output so that its profits will be equal to the residuum $(\theta - w)$. If the fallback payoffs are zero, the Nash bargaining solution implies profits equal to $\pi = (1 - \beta)\theta$ in period 1. The same holds in period 2, which implies that profits are equal to $\pi' = (1 - \beta)\theta'$ in the case of regular work.

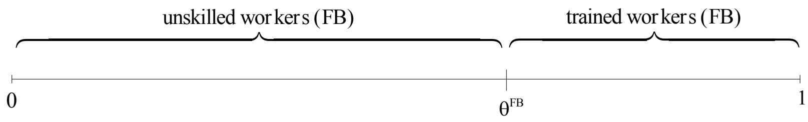


Figure 3.2: The First-Best Optimum

### 3.4.3 The First-Best Optimum

In the first-best optimum (FB), the overall welfare (i.e. the common surplus of workers and firms) is maximized. Obviously, there is no unemployment in the first-best optimum (i.e. $u^{FB} = 0$) because each unemployed worker would be equivalent to lost productivity. However, only those workers should receive apprenticeship training whose increase in discounted productivity in period 2 exceeds the training costs in period 1.

In the following, we assume that $\theta^{FB}$ describes the welfare maximizing pivotal ability between apprenticeship training and regular work.[100] Hence, the optimal number of apprentices in the first period is $n^{FB} = 1 - \theta^{FB}$.[101] This situation is illustrated in Figure 3.2. The overall welfare in the first-best optimum unambiguously depends on $\theta^{FB}$:[102]

$$W^{FB} = \underbrace{\int_0^{\theta^{FB}} (1+\delta)\,\theta d\theta}_{\text{unskilled workers}} + \underbrace{\int_{\theta^{FB}}^{1} \left[(1+\delta(1+\alpha))\,\theta - e\right] d\theta}_{\text{trained workers}} \tag{3.8}$$

The first integral in (3.8) describes the aggregate output of regular workers

[100] Hanushek, Leung, and Yilmaz (2003) refer to it as "ability cutoff". Note that there are two possibilities for $\theta^{FB}$: first, it can be an interior solution within the interval $[0, 1]$, and second, it can be a corner solution, which implies that it is either equal to zero or one. In order to guarantee an interior solution for both $\theta^{FB}$ and $\theta^{LF}$ (cf. Section 3.4.4), we assume that the training costs do not exceed some upper bound: $e \leq \delta\alpha$.

[101] More accurately, $n^{FB}$ describes the *mass* of apprentices in the first-best optimum. In the following, we will neglect this inaccuracy.

[102] The density function of individual abilities is $f(\theta) = 1$. Note that the labeling at the bottom in (3.8) refers to the workers' qualification in period 2.

without an apprenticeship training position. These unskilled workers have an identical productivity $\theta$ in both periods, which adds up to a present value of $(1+\delta)\,\theta$. The second integral is equal to the aggregate output of those workers who receive apprenticeship training in the first period. These workers with $\theta \geq \theta^{FB}$ generate training costs $e$ in period 1. In the second period, their productivity is increased by the factor $(1+\alpha)$. Maximizing (3.8) with respect to $\theta^{FB}$ yields the optimal pivotal ability between apprenticeship training and regular work.

**Proposition 3.1** *In the first-best optimum, the lowest ability of those workers who should receive apprenticeship training is equal to*

$$\theta^{FB} = \frac{e}{\delta\alpha} \tag{3.9}$$

In the first-best optimum, the pivotal ability between apprenticeship training and regular work increases with $e$ because higher training costs make apprenticeship training less profitable. However, $\theta^{FB}$ decreases with the productivity-enhancement $\alpha$ because the productivity of trained workers becomes larger and thus the return to education is increased. Just as well, $\theta^{FB}$ decreases with the discount factor because an increase in $\delta$ is equivalent to a decrease in $r$. Hence, the welfare in period 2 is discounted less and thus weighted to a greater extent. By substituting (3.9) into equation (3.8), we obtain the first-best level of overall welfare.

**Proposition 3.2** *In the first-best optimum, the overall welfare is equal to*

$$W^{FB} = \frac{1}{2}\left(1+\delta\left(1+\alpha\right)\right) - e + \frac{1}{2}\frac{e^2}{\delta\alpha} \tag{3.10}$$

### 3.4.4 The Laissez-Faire Equilibrium

In the laissez-faire equilibrium (LF) without government intervention, workers never prefer to remain unemployed since this would be equivalent to receiving zero income.[103] Workers prefer apprenticeship training in period 1 if their total utility in equation (3.7) is larger with apprenticeship training (i.e. $\kappa = 1$) than with regular work (i.e. $\kappa = 0$).

**Definition 3.1** *The pivotal ability $\theta^{LF}$ is defined to be the lowest ability of those workers who decide to receive apprenticeship training. A worker prefers apprenticeship training to regular work if his total utility over both periods solves*

$$(1 + \delta(1 + \alpha))\theta - e \geq (1 + \delta)\theta$$
$$\theta \geq \theta^{LF} \equiv \frac{e}{\delta\alpha} = \theta^{FB} \tag{3.11}$$

By comparing this pivotal productivity with the first-best optimum in Section 3.4.3, we conclude that the laissez-faire equilibrium is efficient (cf. Figure 3.3). The reason is that the worker is the full claimant of his marginal product, which implies that he takes into account all costs and benefits of apprenticeship training. Hence, there is no "hold-up" problem such that one party (i.e. the worker) pays the training costs, while another party (i.e. the firm) shares in the return (Acemoglu and Shimer (1999)).[104]

[103] Note that regular work does not generate a positive income for $\theta_L = 0$. Because this borderline case has no consequences for the overall welfare, we simply assume that this worker with the lowest ability also decides to work regularly.

[104] A "hold-up" problem would arise due to incomplete contracts. For example, if wages are determined by Nash-bargaining and $\beta$ indicates the worker's bargaining power, the pivotal ability between apprenticeship training and regular work would be inefficiently high:

$$\theta^{LF} = \frac{e}{\delta\beta\alpha} \geq \theta^{FB}$$

In this case, $\theta^{LF}$ decreases with the workers' bargaining power $\beta$. The case $\beta = 1$ implies zero rents for the firms and the absence of labor market frictions so that the number of apprenticeship training positions achieves its first-best level, i.e. $n^{LF} = n^{FB}$. However, for $\beta < 1$ the training decision of workers implies $\theta^{LF} > \theta^{FB}$ and thus distortions compared to the first-best optimum. In this case, the deviation from the first-best number of

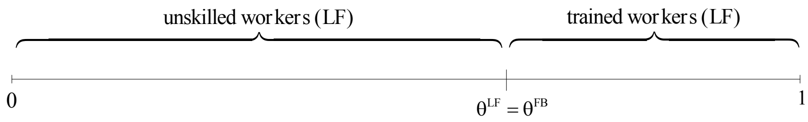


Figure 3.3: The Laissez-Faire Equilibrium

**Proposition 3.3** *In the laissez-faire equilibrium, the number of unemployed workers and the number of trained workers correspond to the first-best optimum:*

$$\begin{aligned} u^{LF} &= u^{FB} = 0 \\ n^{LF} &= n^{FB} = 1 - \frac{e}{\delta\alpha} \end{aligned}$$

*Hence, the laissez-faire equilibrium is efficient and the overall welfare is maximal:*

$$W^{LF} = W^{FB} \tag{3.12}$$

### 3.4.5 The Equilibrium with Social Assistance

Unfortunately, the German labor markets are hardly characterized by the laissez-faire equilibrium as described in the previous subsection. As discussed in Section 3.3.2, labor markets are heavily regulated and the labor supply of workers is distorted by the system of social security. For this reason, we

---

apprenticeship training positions is equal to

$$\triangle n = n^{LF} - n^{FB} = -\frac{(1-\beta)\, e}{\delta\beta\alpha} < 0 \qquad \text{if } \beta < 1$$

This inefficiency stems from the fact that workers only consider their higher wages in period 2 but do not take into account increased profits of firms by the output share $(1-\beta)$. Hence, workers underinvest in the accumulation of human capital. The difference $\triangle n$ and thus the degree of distortion become larger, the lower the worker's bargaining power $\beta$. The inefficiency would be solved if firms and workers could bargain on the distribution of $e$ and write a complete contract in which the external effects (i.e. both the costs and benefits) of apprenticeship training are internalized.

incorporate into our model social benefits $z$ that are paid to all unemployed workers in period 2.[105]

In accordance with the *Arbeitslosengeld II* of the German system of social security, $z$ is identical for all workers and thus independent of labor income in period 1. We assume that the government cannot observe why workers remain unemployed and thus also supports those workers who voluntarily refuse to work in order to become eligible for social assistance. The public expenditures for social benefits in the second period are financed by a lump-sum tax $T$, which is imposed on all workers independent of their status of employment. Hence, social benefits $z$ describe the gross transfer while unemployed workers receive net payments of $z - T$.

### The Pivotal Abilities with Social Assistance

Obviously, the labor demand of firms is not affected by the system of social security. Also, the pivotal ability of workers between apprenticeship training and regular work remains the same because $\theta^{LF}$ is independent of $z$. Hence, there are no distortions in the number of apprenticeship training positions.

However, there are some low-ability workers who now decide to remain unemployed in the second period because their utility is higher by receiving social assistance than by working regularly. These distortions arise because $z$ defines the reservation wage and thus affects the labor force participation of workers. In other words, social assistance (SA) generates an individual participation tax rate $\tau^{SA}$, which is higher the lower the individual productivity $\theta'$ in the second period:[106]

$$\tau^{SA} = \frac{z}{\theta'}$$

[105] In order to keep the following calculations as simple as possible, we concentrate on period 2 and do not consider social assistance in period 1.

[106] The participation tax rate is defined as the average tax rate on labor market participation (Keuschnigg (2005)). In line with Eissa, Kleven, and Kreiner (2006), labor supply at the extensive margin is linked to the participation tax rate. Note that the participation tax rate is zero in the laissez-faire equilibrium, i.e. $\tau^{LF} = 0$.

In line with Boone and Bovenberg (2004), the participation constraint is binding at the bottom of the ability distribution.[107] In this context, the pivotal ability between regular work and unemployment is described by the following definition.

**Definition 3.2** *With social assistance, the pivotal ability $\theta^{SA}$ is defined to be the lowest ability of those workers who decide to work regularly. A worker prefers regular work to unemployment if his total utility over both periods solves*

$$(1+\delta)\,\theta - \delta T \geq \theta + \delta\,(z - T)$$
$$\theta \geq \theta^{SA} \equiv z \tag{3.13}$$

The pivotal ability between regular work and unemployment increases with $z$ because higher social benefits make it less attractive to enter the labor market. The relationship of the pivotal abilities is clarified by the following proposition.

**Proposition 3.4** *If the level of social assistance does not exceed some upper bound*

$$z \leq \theta^{FB}$$

*the relationship of the pivotal abilities is the following (cf. Appendix C):*

$$\theta^{LF} \geq \theta^{SA}$$

### Overall Welfare with Social Assistance

The number of trained workers is the same as in the laissez-faire equilibrium, i.e. $n^{SA} = n^{LF} = 1 - \theta^{LF}$. However, the system of social security generates (voluntary) unemployment in the second period among those workers with

[107] In Immervoll, Kleven, Kreiner, and Saez (2005), labor supply at the extensive margin is driven by fixed costs of work effort. In our approach, each worker decides on his labor market participation depending on his labor income and the level of social assistance.

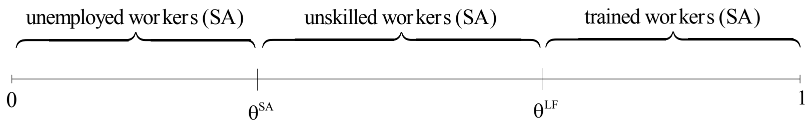


Figure 3.4: The Equilibrium with Social Assistance

individual ability below $\theta^{SA}$, i.e. $u^{SA} = \theta^{SA}$. Hence, labor market participation is too low and unemployment is too high compared to the first-best optimum. This situation is illustrated in Figure 3.4.

Because, the aggregate profits of firms are zero (cf. Section 3.4.2), the overall welfare with social assistance corresponds to the sum of the total utilities of all workers over both periods:[108]

$$W^{SA} = \underbrace{\int_0^{\theta^{SA}} [\theta + \delta (z - T)] \, d\theta}_{\text{unemployed workers}} \tag{3.14}$$

$$+ \underbrace{\int_{\theta^{SA}}^{\theta^{LF}} [(1 + \delta) \theta - \delta T] \, d\theta}_{\text{unskilled workers}} + \underbrace{\int_{\theta^{LF}}^{1} [(1 + \delta (1 + \alpha))\theta - e - \delta T] \, d\theta}_{\text{trained workers}}$$

The first integral in (3.14) is equal to the aggregate utility of those workers who remain unemployed in period 2, the second one describes the aggregate utility of workers employed regularly in both periods, and the third integral is equal to the aggregate utility of trained workers. The public expenditures for social assistance are financed by lump-sum taxation of all workers independent of their status of employment, which implies that each worker has to pay the lump-sum tax $T$. Hence, we omit further labor market distortions that would be generated by taxing only those workers who are in employment.

[108] As before, the labeling at the bottom refers to the workers' qualification or status of employment in period 2.

In order to determine the overall welfare with social assistance, we have to consider the total amount of social benefits and taxes. Although $z$ and $T$ affect the aggregate welfare of workers in (3.14), they cancel out in the overall welfare equation (3.15) because they represent pure lump-sum transfers between the workers and the government. Hence, $W^{SA}$ simplifies to

$$W^{SA} = \underbrace{\int_0^{\theta^{SA}} \theta d\theta}_{\text{unemployed workers}} + \underbrace{\int_{\theta^{SA}}^{\theta^{LF}} (1+\delta)\,\theta d\theta}_{\text{unskilled workers}} + \underbrace{\int_{\theta^{LF}}^{1} [(1+\delta(1+\alpha))\theta - e]\, d\theta}_{\text{trained workers}} \quad (3.15)$$

The overall welfare with social assistance is obtained by substituting the pivotal abilities $\theta^{LF}$ and $\theta^{SA}$ into equation (3.15).

**Proposition 3.5** *With social assistance, the overall welfare is equal to*

$$W^{SA} = W^{LF} - \frac{1}{2}\delta z^2 \quad (3.16)$$

Comparing equations (3.12) and (3.16) shows that the system of social security decreases the overall welfare compared to the laissez-faire equilibrium (i.e. $W^{SA} < W^{LF}$). This gap increases with $z$ because the labor supply incentives of low-ability workers are reduced. For all workers with individual ability below $\theta^{SA}$, the wage that could be earned by working regularly falls below the level of social assistance. Hence, these workers can increase their utility by remaining unemployed. This inefficiency in the labor supply of workers constitutes the necessary condition for welfare-enhancing government interventions.

Obviously, labor market distortions could be reduced by lowering the level of social assistance. However, if social benefits cannot be reduced by the government as suggested by Boone and Bovenberg (2004), tax credits are a possible policy instrument to move the economy towards the first-best optimum by decreasing the number of unemployed workers. Based on the equilibrium with social assistance, the welfare implications of tax credits are analyzed in the following section.

## 3.5 The Welfare Analysis of Tax Credits

As discussed in Section 3.2.1, tax credits imply that the government pays a subsidy to each worker who is employed regularly in the second period and whose labor income $I$ does not exceed the critical income level $\bar{I}$. By paying low-skilled workers a subsidy $S$ in addition to their individual labor income, formerly unemployed workers are motivated to enter the labor market because this "combined wage" makes them better off than the level of social assistance.

Following the propositions of Sinn, Holzner, Meister, Ochel, and Werding (2006), tax credits are designed in terms of a NIT for those workers who are employed regularly in period 2. This design implies that all workers in employment receive a basic lump-sum transfer $y$, which corresponds to the guaranteed income of each worker. The individual subsidy $S$ decreases continuously with labor income, i.e. it phases out at rate $s$ up to the critical income level $\bar{I}$. This smooth reduction of the subsidy avoids undesirable jumps in the distribution of net incomes.[109] In contrast to the traditional NIT, tax credits are only directed to net beneficiaries, while workers with labor income above $\bar{I}$ are not affected.

Formally, the subsidy $S$ in period 2 depends on the individual labor income $I$ in the following manner:

$$S(\theta') = \begin{cases} y - s\theta' & \text{if } I(\theta') = \theta' \leq \bar{I} \equiv \frac{y}{s} \\ 0 & \text{if } I(\theta') > \bar{I} \end{cases} \tag{3.17}$$

with $y$ as the basic transfer and $s$ as the phase-out rate of the tax credits. As can be seen in (3.17), individuals with low income $I$ below the critical

[109] If abilities are continuously distributed and not observable by the government, this transfer scheme is analytically equivalent to a lump-sum subsidy which is constant up to some critical level of labor income. Each worker with labor income above this critical level can imitate a lower ability in order to become eligible for the subsidy. Hence, the additional income due to the subsidy becomes smaller the higher the true ability of the worker and thus the higher the income loss necessary to fall below the critical level of income.

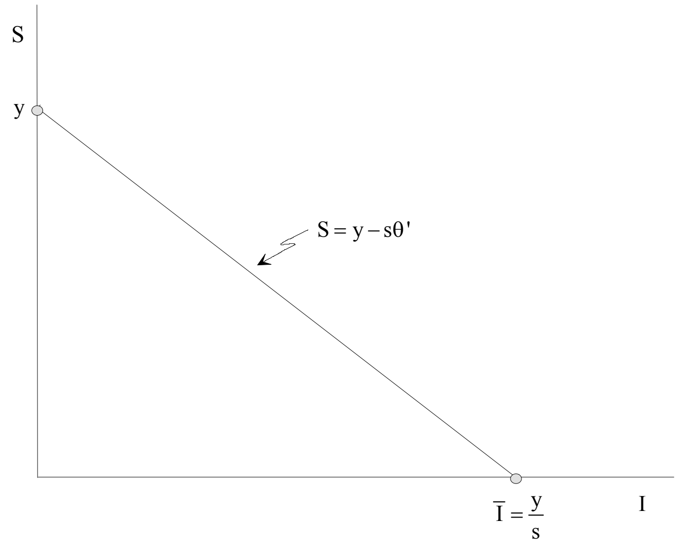


Figure 3.5: Tax Credits Depending on Labor Income

level $\bar{I}$ receive a wage subsidy from the government. According to Hausman (1985), the phase-out rate $s$ mainly determines labor supply at the intensive margin while $y$ affects the participation decision of workers. Hence, the basic transfer is much more important for the labor supply response of low-income workers because labor supply adjustments take place at the extensive margin (cf. Section 3.2.2). Depending on $I$, this subsidy scheme is graphically illustrated in Figure 3.5.

There is a critical ability $\bar{\theta}$ that indicates the worker just receiving zero subsidy, i.e. all workers with higher ability than $\bar{\theta}$ receive no subsidy. The critical income level $\bar{I}$ is equal to $\frac{y}{s}$ so that the critical ability is defined by $\bar{\theta} \equiv \frac{y}{s}$. In line with Hanushek, Leung, and Yilmaz (2003), we assume that each unskilled worker is subsidized. This assumption implies the relationship $\bar{\theta} \geq \theta^{LF}$, which allows focusing on those workers who receive a positive amount of subsidy without formal qualification but no subsidy in the case of apprenticeship training (cf. the pivotal ability between apprenticeship training and regular work in Definition 3.3). This case is empirically

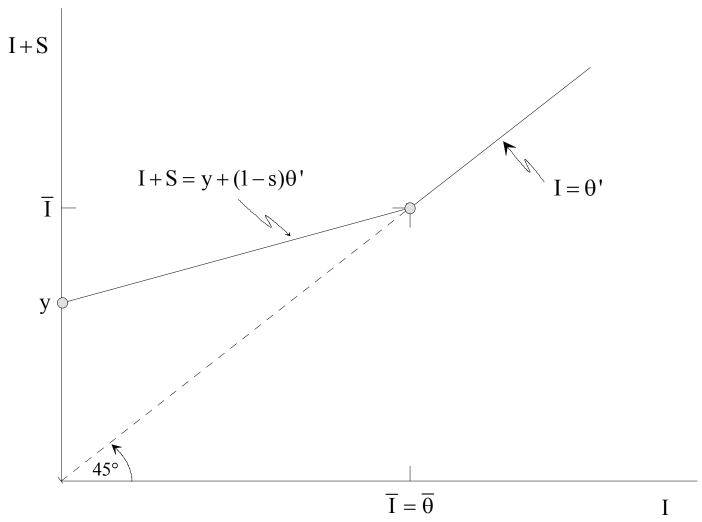


Figure 3.6: Total Income with Tax Credits

relevant because many trained workers with medium ability would receive a much lower wage and would thus be eligible for tax credits if they had remained untrained. Furthermore, our assumption is justified because we want to concentrate on the most interesting case with a significant reduction in unemployment.[110]

The relationship between individual labor income ($I$) and total income including the tax credits ($I + S$) is graphically illustrated in Figure 3.6. Depending on $\bar{\theta}$, the worker's total income in period 2 is equal to

[110] Obviously, this is a crucial assumption for our analytical results in Section 3.5.2. There are two other possible cases that are neglected in the following analysis. First, we could concentrate on those workers who always receive a positive amount of subsidy irrespective of whether they train or not. In this case, the basic transfer cancels out in their training decision according to Definition 3.3. And second, we could focus on those workers who are never eligible for tax credits because their income exceeds the critical level even without apprenticeship training. In this case, the pivotal ability between apprenticeship training and regular work is independent of both $y$ and $s$.

$$I(\theta') + S(\theta') = \begin{cases} y + (1-s)\,\theta' & \text{if } \theta' \leq \bar{\theta} \\ \theta' & \text{if } \theta' > \bar{\theta} \end{cases}$$

### 3.5.1 The Pivotal Abilities with Tax Credits

As in Section 3.4.5, public expenditures for social assistance and tax credits are financed by lump-sum taxation of all workers independent of their status of employment. In line with Boone and Bovenberg (2004) and Boone and Bovenberg (2006), the level of social assistance is taken as given by the government. Hence, workers still receive social benefits $z$ in the case of unemployment. In consequence of the tax credits, the workers' decisions between apprenticeship training, regular work, and unemployment are modified.

**Definition 3.3** *With tax credits (TC), a worker prefers apprenticeship training to regular work if*

$$\theta - e + \delta\left[(1+\alpha)\,\theta - T\right] \geq \theta + \delta\left[(1-s)\,\theta + y - T\right]$$

$$\theta \geq \theta_A^{TC} \equiv \frac{e + \delta y}{\delta\,(\alpha + s)} \tag{3.18}$$

*With tax credits, a worker prefers regular work to unemployment if*

$$\theta + \delta\left[(1-s)\,\theta + y - T\right] \geq \theta + \delta\left[z - T\right]$$

$$\theta \geq \theta_U^{TC} \equiv \frac{z - y}{1 - s} \tag{3.19}$$

Before analyzing the equilibrium with tax credits, we have to determine the relationship of the four pivotal abilities $\theta^{LF}$, $\theta^{SA}$, $\theta_A^{TC}$, and $\theta_U^{TC}$.

**Proposition 3.6** *The relationship of the pivotal abilities is the following (cf. Appendix C):*

$$\theta_A^{TC} \geq \theta^{LF} \geq \theta^{SA} \geq \theta_U^{TC}$$

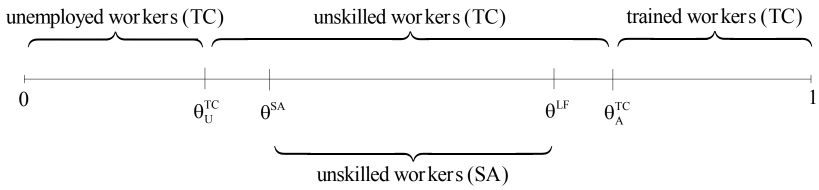


Figure 3.7: The Equilibrium with Tax Credits

### 3.5.2 The Equilibrium with Tax Credits

Tax credits aim at reducing unemployment among low-skilled workers. Indeed, the pivotal ability between regular work and unemployment is decreased (i.e. $\theta_U^{TC} \leq \theta^{SA}$), which implies that labor market distortions are reduced and the equilibrium approaches the first-best optimum with zero unemployment. The reason is that the subsidy $S$ lowers the participation tax rate for all beneficiaries of the tax credits:

$$\tau^{TC} = \frac{z - S(\theta')}{\theta'} < \tau^{SA}$$

This reduction in the participation tax rate is equivalent to an increase in the opportunity costs of unemployment. Hence, some low-ability workers who were unemployed in the equilibrium without tax credits now decide to enter the labor market.

However, there is also an adverse effect of tax credits on the behavior of high-ability workers because the pivotal ability between apprenticeship training and regular work is increased (i.e. $\theta_A^{TC} \geq \theta^{LF}$). Hence, there are less workers that prefer apprenticeship training to regular work because the opportunity costs of training are increased. Compared to the equilibria with laissez-faire and with social assistance, the number of trained workers is distorted, i.e. $n^{TC} = 1 - \theta_A^{TC} < n^{LF} = n^{SA}$. This situation is is illustrated in Figure 3.7.

As in Section 3.4.5, the overall welfare with tax credits corresponds to the sum of the total utilities of all workers over both periods:

$$W^{TC} = \underbrace{\int_{0}^{\theta_U^{TC}} [\theta + \delta (z - T)] \, d\theta}_{\text{unemployed workers}} + \underbrace{\int_{\theta_U^{TC}}^{\theta_A^{TC}} [(1 + \delta (1 - s)) \, \theta + \delta (y - T)] \, d\theta}_{\text{unskilled workers with tax credits}} + \underbrace{\int_{\theta_A^{TC}}^{1} [(1 + \delta (1 + \alpha))\theta - e - \delta T] \, d\theta}_{\text{trained workers without tax credits}} \tag{3.20}$$

The first integral in (3.20) is equal to the aggregate utility of low-ability workers who remain unemployed in period 2 and the second one describes the aggregate utility of workers employed regularly in both periods. While these unskilled workers are eligible for tax credits because their labor income falls below the critical income level, the third integral refers to the aggregate utility of trained workers who receive no subsidy. Similar to Section 3.4.5, the public expenditures for tax credits and social assistance are financed by lump-sum taxation of all workers, which implies that each worker has to pay the lump-sum tax $T$. Again, we omit further labor market distortions that would be generated by taxing only those workers in employment.

In order to determine the overall welfare with tax credits, we have to consider the total amount of expenditures and taxes. Tax credits affect the aggregate welfare of workers in (3.20) and exert an indirect influence on the overall welfare in (3.21) by changing the pivotal abilities compared to the equilibrium with social assistance, but they cancel out as a direct determinant because they represent pure lump-sum transfer between the workers and the government. Altogether, the overall welfare with tax credits is equal to

$$W^{TC}(y,s) = \underbrace{\int_0^{\theta_U^{TC}} \theta d\theta}_{\text{unemployed workers}} + \underbrace{\int_{\theta_U^{TC}}^{\theta_A^{TC}} (1+\delta)\,\theta d\theta}_{\text{unskilled workers}} + \underbrace{\int_{\theta_A^{TC}}^{1} \left[(1+\delta(1+\alpha))\theta - e\right] d\theta}_{\text{trained workers}} \quad (3.21)$$

In order to determine the optimal tax credits, we have to maximize (3.21) with respect to $y$ and $s$:

$$\underbrace{\delta\theta_U^{TC}\left(-\frac{\partial\theta_U^{TC}}{\partial y}\right)}_{\text{less unemployed workers}} = \underbrace{\delta\alpha\theta_A^{TC}\frac{\partial\theta_A^{TC}}{\partial y}}_{\text{less trained workers}} - \underbrace{\frac{\partial\theta_A^{TC}}{\partial y}e}_{\text{reduced training costs}} \quad (3.22)$$

$$\underbrace{\delta\alpha\theta_A^{TC}\left(-\frac{\partial\theta_A^{TC}}{\partial s}\right)}_{\text{more trained workers}} - \underbrace{\left(-\frac{\partial\theta_A^{TC}}{\partial s}\right)e}_{\text{increased training costs}} = \underbrace{\delta\theta_U^{TC}\frac{\partial\theta_U^{TC}}{\partial s}}_{\text{more unemployed workers}} \quad (3.23)$$

The first-order conditions (3.22) and (3.23) compare the marginal benefits and the marginal costs of an increase in $y$ and $s$. Raising the basic transfer $y$ increases the number of regular workers at the lower end by $(-\frac{\partial\theta_U^{TC}}{\partial y})$, which generates additional productivity $\delta$ per unit of initial ability. The (net) marginal costs of an increase in $y$ are shown on the right hand side of (3.22). These costs are composed of two effects that both accrue because the number of regular workers is increased at the upper end by $\frac{\partial\theta_A^{TC}}{\partial y}$. According to the first summand, there is a productivity loss of $\delta\alpha$ per unit of initial ability due to those workers who now reject apprenticeship training and thus no longer have an increased productivity in the second period. However, this decline in the number of trained workers also implies that the aggregate training costs are reduced by $e$ per former apprentice.

As shown in the first-order condition (3.23), the opposite welfare effects hold for an increase in $s$. Raising the phase-out rate $s$ decreases the number of regular workers and thus raises the number of unemployed workers at the lower end by $\frac{\partial \theta_U^{TC}}{\partial s}$. At the same time, the number of trained workers is increased by $\left(-\frac{\partial \theta_A^{TC}}{\partial s}\right)$.

Altogether, the implementation of tax credits faces a trade-off with respect to overall welfare. On the one hand, tax credits are welfare-enhancing because they manage to decrease the number of unemployed workers in period 2 (either by an increase in $y$ or by a reduction in $s$). On the other hand, tax credits increase the opportunity costs of training, which implies that less workers are willing to receive apprenticeship training.

### 3.5.3 Interpretation

The implementation of tax credits aims at increasing the labor supply of low-skilled workers at the extensive margin by raising the opportunity costs of unemployment and thus fostering the workers' incentives to enter the labor market. However, in line with Heckman, Lochner, and Cossa (2002), a significant increase in labor force participation is achieved only at the cost of reduced human capital formation by workers in employment because the opportunity costs of apprenticeship training are raised at the same time.

If the intended reduction in unemployment is rather small, it will be possible to avoid any distortions in the training decision of workers (i.e. $\theta_A^{TC} = \theta^{FB}$) by choosing the basic transfer equal to $\tilde{y} = \frac{e}{\delta\alpha}s$. In this case, there is a constant relationship of the two variables of the subsidy scheme so that they just balance the negative effects on the pivotal ability between apprenticeship training and regular work. However, by considering the welfare effects of both unemployment and apprenticeship training, it is welfare-enhancing to further increase the basic transfer.

Given the distortions of the equilibrium with social assistance as discussed in Section 3.4.5, small tax credits with a basic transfer slightly above $\tilde{y}$ generate a first-order welfare gain by the reduction in unemployment while a small deviation of apprenticeship training from its first-best level only creates a second-order welfare loss. Hence, by taking into account both welfare effects, these small tax credits give rise to an overall welfare gain. However, as the tax credits are raised further, the welfare loss from reduced training becomes larger while the efficiency gain from additional labor force participation becomes smaller as the economy is moved towards the first-best level of unemployment. Consequently, there must be a well determined optimal level of tax credits ($y^*$, $s^*$) that maximizes the overall welfare by equating the welfare gain on the left hand side and the welfare loss on the right hand side in the first-order conditions (3.22) and (3.23).

**Proposition 3.7** *With optimal tax credits, unemployment is reduced to*

$$(u^{TC})^* = \frac{z - y^*}{1 - s^*} < u^{SA}$$

*At the same time, the number of trained workers is reduced to*

$$(n^{TC})^* = 1 - \frac{e + \delta y^*}{\delta\,(\alpha + s^*)} < n^{SA}$$

By substituting the pivotal abilities $\theta_U^{TC}$ and $\theta_A^{TC}$ into (3.22) and solving for $y$, we obtain the welfare maximizing basic transfer $y^*$ depending on the optimal phase-out rate $s^*$:[111]

$$y^* = \frac{(\alpha + s^*)^2}{(1+\alpha)\,(\alpha + (s^*)^2)} z + \frac{(1 - s^*)^2\, s^*}{\delta\,(1+\alpha)\,(\alpha + (s^*)^2)} e \qquad (3.24)$$

Because the optimal basic transfer (3.24) is positive (cf. Appendix D.2), the implementation of tax credits increases the overall welfare compared to the equilibrium with social assistance. Furthermore, by choosing the basic transfer as high as the level of social assistance (i.e. $y = z$), it is even possible to achieve zero unemployment (i.e. $u^{TC} = 0$). However, the optimal basic transfer falls below $z$ if the welfare gain of tax credits is too low to compensate for the costs of less apprenticeship training. For a given level of $s^*$, the optimal basic transfer falls below the level of social assistance if

[111] The calculation of $y^*$ is presented in Appendix D.1. Note that $y^*$ describes a welfare maximum because the second derivative is negative, i.e. $\frac{\partial^2 W^{TC}(y,s)}{\partial y^2} < 0$.

the productivity-enhancement of apprenticeship training exceeds some lower bound (cf. Appendix D.2):

$$\alpha > \frac{e}{\delta z} s^*$$

In this case, the reintegration of those workers at the bottom of the ability distribution into the labor market is not optimal. Because the costs in terms of decreased human capital formation would be too high, it is more efficient to leave aside those workers with the lowest productivities.

Obviously, by lowering the level of social assistance, it would be possible to reduce unemployment without generating distortions in the training decision of workers. However, in line with Boone and Bovenberg (2004), we assume that this policy instrument is not available and the government takes $z$ as given when deciding on the optimal level of tax credits. Note that, for a given level of $s^*$, the optimal basic transfer $y^*$ increases with $z$. This result is in line with Boone and Bovenberg (2006) who conclude that, at high levels of social assistance, tax credits and traditional social benefits constitute complements because tax credits aim at offsetting the negative impact of social assistance on the labor supply of low-skilled workers. With respect to the UK, Blundell and Hoynes (2001) empirically show that the positive impact of tax credits on labor force participation is reduced by the existence of other social benefits. For this reason, Sinn, Holzner, Meister, Ochel, and Werding (2006) propose to combine the implementation of tax credits with a reduction in the level of social assistance. This reduction in $z$ would decrease the optimal basic transfer and thus lower the negative effects of tax credits on human capital formation.

## 3.6 Conclusion

Chapter 3 of my PhD thesis ranks among a new line of research that transcends the boundaries of labor economics and public finance. With respect to the German labor markets, Sinn (2002) and Schöb and Weimann (2003) conclude that the only way to lower the poverty trap and to reactivate the low-skilled part of the labor force is to subsidize work instead of unemployment. This chapter presents a two-period partial-equilibrium model that systematically compares the costs and benefits of tax credits. As proposed by Sinn, Holzner, Meister, Ochel, and Werding (2006), tax credits are designed in terms of a NIT for those low-skilled workers who decide to enter the labor market.

Our formal analysis is based on recent training literature with oligopolistic labor markets but the model is adapted to the German system of apprenticeship training and social security. In the laissez-faire equilibrium, there is no unemployment and the number of trained workers corresponds to the first-best optimum. As suggested by Sinn (2003), the system of social security generates unemployment among low-ability workers because private employment is crowded out by the welfare state. The implementation of tax credits faces a trade-off with respect to overall welfare. While tax credits reduce the number of unemployed workers at the extensive margin, they increase at the same time the opportunity costs of apprenticeship training, which implies that the training decision of high-ability workers is distorted and the number of trained workers is lowered. In conformity with Heckman, Lochner, and Cossa (2002), a significant increase in labor force participation is achieved only at the cost of reduced human capital formation by workers in employment.

Nevertheless, the implementation of tax credits increases the overall welfare compared to the equilibrium with social assistance because the positive effect on labor force participation outweighs the negative effect on apprenticeship training, which implies that the optimal basic transfer is positive. Hence, in line with the empirical results of Eissa, Kleven, and Kreiner (2004) and Immervoll, Kleven, Kreiner, and Saez (2005), the introduction of tax credits is justified on theoretical grounds. However, if the optimal basic transfer falls below the level of social assistance, it is not optimal to reintegrate those workers at the bottom of the ability distribution into the labor market. Because the costs in terms of decreased human capital formation would be too high, it is more efficient to leave aside those workers with the lowest productivities.

In our model, the number of unemployed workers and the number of apprentices are endogenously determined and depend on the individual ability of workers. Nevertheless, the model has been kept simple for expositional and calculational reasons. The theoretical results of our stylized model only allow for qualitative conclusions concerning the implementation of tax credits in the context of the German system of apprenticeship training and social security. In order to assess the quantitative magnitude of these effects, we would have to estimate the elasticities of the workers' labor supply and training responses at the extensive margin. However, the underlying insights into the model presented here are robust to various types of generalization. Hence, they constitute a promising basis for policy recommendations and future research.

# Chapter 4

# Pension Reform

*The demographic transition in industrialized countries poses challenges to the pension system, which is essentially organized according to the pay-as-you-go principle in most countries. Chapter 4 of my PhD thesis aims at analyzing two proposals for pension reform in a theoretical model that endogenously explains the retirement and training decision of workers who are heterogeneous in ability. Because the economic benefits of motivating late retirement strongly depend on the employment prospects of workers near retirement age, the model includes the firms' employment decision at the extensive margin.*

*The first reform proposal, the implementation of individual retirement accounts, increases the workers' incentives to acquire skills and to postpone retirement. However, if the capital funded pillar of the pension system becomes strong, low-ability workers may not attain their optimal retirement age because firms refuse to employ them any longer. In a similar manner, the second reform proposal to increase the minimum retirement age may not work for low-ability workers if their separation date is determined by the firms before the minimum retirement age is achieved.*

## 4.1 Introduction

Starting from the "golden age of retirement" (Cremer and Pestieau (2000)), the demographic transition in industrialized countries poses challenges to the pension system and the economy as a whole (Bovenberg and Knaap (2005)). Declining fertility and increasing life expectancy give rise to prolonged demographic change that continuously changes the size and the composition of the labor force.[112] This topic has received much attention in recent years, mainly due to the expected sharp increase in the ratio of retirees per active worker. For example, in Switzerland, the ratio of people above 65 relative to the active population of age 20-64 is expected to increase from 30.1% in 2000 to 50.5% in 2060 (KommissionfürKonjunkturfragen (2005)).

The economic impact of this demographic change mainly derives from the reduction in aggregate labor supply and the impact on fiscal and social security budgets, necessitating either a significant reduction in old-age benefits, an increase in social security contribution rates, an increase in the minimum retirement age, or other measures to improve labor market participation (Disney (2000)).[113] Börsch-Supan (2003) and Martín (2003) argue that in order to finance the current pension system, contributions and tax rates would have to rise substantially, which reduces labor force participation of younger cohorts and destabilizes the pension system even more. For example, while the average contribution rate in the European Union was 16% in 2000, it will increase to 27% in 2050 if the present benefit rules are kept unchanged (EuropeanCommission (2001)).

[112] In the European Union, life expectancy at age 65 has increased by more than one year per decade since 1950 (Cremer, Lozachmeur, and Pestieau (2004)).

[113] Acording to Hines and Taylor (2005), the US Social Security trust fund will be empty in 2044. Martín (2003) uses a CGE model to analyze how pension reforms may alleviate the expected financial difficulties of current PAYG systems. In the long-run, the financial sustainability of pension systems may also be improved by immigration or family policy (Kirchgässner (2005)). However, Börsch-Supan (2003) argues that the decrease in the relative size of the economically active population cannot be balanced by higher capital intensity.

The trend of declining labor force participation is aggravated by the fact that an increasing fraction of older workers decides to retire early (Conde-Ruiz and Galasso (2004)). The reason is that existing pension systems impose considerable negative accrual rates of pension wealth and thus provide economic incentives to leave the active population at younger and younger ages (Samwick (1998)). Gruber and Wise (2005) provide an extensive source on retirement behavior in industrialized countries. However, there is a second argument why labor force participation is low among older workers. In most industrial countries, the number of older workers in unemployment is disproportionately high, which implies that the employment prospects of workers near retirement age are significantly reduced (Bingley and Lanot (2004)). Hence, workers may be constrained by poor labor market conditions from working until their desired retirement date (Coile and Levine (2006)).

Besides its adverse effects on the pension system, population aging may induce individuals to invest more in their human capital if aging is accompanied with postponed retirement and longer working periods (Echevarría (2003)). According to consistent findings in the literature, human capital accumulation and economic growth are increased via these channels, possibly even without changes in the system of old-age provision.[114] However, most of the existing analysis of demographic change and pension reforms has been cast in a framework of flexible labor markets with endogenous labor supply.[115] Pension reform in the presence of labor market frictions (for example Keuschnigg and Keuschnigg (2004)) or with endogenous human capital formation (for example Jensen, Lau, and Poutvaara (2004)) is considered quite rarely.

Chapter 4 of my PhD thesis aims at closing this gap by developing a

[114] Cf. de la Croix and Licandro (1999b), Kalemli-Ozcan, Ryder, and Weil (2000), Kalemli-Ozcan (2002), Boucekkine, de la Croix, and Licandro (2002), Soares (2005), and Echevarría (2004)

[115] Cf. Lindbeck and Persson (2003), Feldstein and Liebman (2002), Cutler, Poterba, Sheiner, and Summers (1990), Feldstein and Samwick (1999), Fehr (2000), Kotlikoff (2002), and Fehr, Jokisch, and Kotlikoff (2004).

two-period partial-equilibrium model that analyzes the implications of two proposals for pension reform with a particular focus on retirement age and human capital formation. It is important to incorporate human capital formation into the analysis of pension reforms because extending the working life increases the return to education and thus fosters the workers' incentives to acquire skills (Trostel (1993)). However, the partial-equilibrium framework of this chapter implies that the analysis does not take into account the budget constraint of the pension system but focuses on the individual incentives of workers subject to the system of old-age provision.

The contribution of this chapter is twofold because the formal analysis of pension systems is extended in two important ways. First, we derive endogenously the workers' training intensity and the date of retirement, which both depend on the individual ability of the workers. Second, our model also refers to the demand side of the labor market by incorporating the employment decision of firms that decide how long to continue production with the workers. The firms' employment decision can significantly affect the implications of pension reforms because workers and firms separate as soon as one party decides to leave the market. Hence, the effects of pension reforms strongly depend on the workers' employment prospects near retirement age. In a nutshell, there are two key questions considered in this chapter: First, starting from the current pension system of traditional pay-as-you-go, what is the impact of different pension reforms on the workers' incentives with respect to retirement age and human capital formation? And second, what are the actual effects of these pension reforms subject to the employment decision of the firms?

As a first proposal for pension reform, we analyze the introduction of individual retirement accounts which imply to (partly) move the pension system from a pay-as-you-go towards a capital funded system. In aggregate, there is a double benefit from such a reform: reductions in labor market distortions at the extensive margin (retirement age) as well as increased human capital formation at the intensive margin (training intensity). Hence, individual retirement accounts increase the workers' incentives to acquire skills and to postpone retirement. However, if the capital funded pillar of the pension system becomes strong and workers want to significantly postpone retirement, low-ability workers may not attain their desired retirement age because firms refuse to employ them any longer. In this case, the benefits of pension reform mainly accrue for high-ability workers while the benefits for low-ability workers are reduced once their employment prospects near retirement age are controlled for. This conclusion will be even strengthened and there will be no effect at all on the retirement date of low-skilled workers if human capital formation is exogenous. Hence, the superiority of capital funding as postulated by Feldstein (2005a) is lowered and may even vanish in the presence of labor market imperfections.

As a second reform proposal, we analyze an increase in the minimum retirement age without changing the fundamental nature of the pay-as-you-go system. Again, only high-ability workers may be affected because they are forced to stay inside the labor market until they have reached the minimum retirement age. In contrast, there may be no impact on low-ability workers if their retirement date is determined by the firms before the minimum retirement age is achieved.

Chapter 4 of my PhD thesis proceeds as follows: the next section discusses the theory of old-age provision, labor supply, and human capital formation. In Section 4.3, our partial-equilibrium model is developed and the decision problems of workers and firms are discussed. In Section 4.4, we describe the current pension system with traditional pay-as-you-go and its implications for the workers' behavior with respect to retirement and training. In Sections 4.5 and 4.6, the two proposals for pension reforms are analyzed depending on the individual ability of the workers. Section 4.7 concludes.

## 4.2 Pension Systems and Labor Supply

### 4.2.1 Different Pension Systems

Two important arguments for the existence of pension systems are the myopia of those individuals who do not save adequately for their old-age provision (Feldstein (1985)) and the asymmetric information between government and workers with respect to voluntarily chosen poverty (Kotlikoff, Spivak, and Summers (1982)).[116] According to Feldstein (1974), social security wealth is equal to "the present actuarial value of the social security benefits to which the current adult population will be entitled at age 65 minus the present actuarial value of the social security taxes that they will pay before reaching that age". Obviously, this term does not mean real wealth but corresponds to a claim on current and future taxpayers (Feldstein (1996)).

In general, there are two different concepts of old-age provision, namely capital funding and the pay-as-you-go system. Feldstein (2005a) and Keuschnigg (2005) provide an overview.

#### Capital Funding and Pay-As-You-Go

With a capital funded pension system, the worker's contributions are paid into an individual account where they accumulate with interest until they are paid out during retirement in the form of actuarially fair pensions. The individual rate of return on one's own contributions corresponds to the market rate of interest, reduced by an administrative fee. Hence, every generation finances its old-age provision from its own savings accumulated during the previous working life (Feldstein (1974)). The accumulated assets of pension funds are a major source of aggregate savings and can easily run up to 100 percent of GDP and more, depending on the size of contributions allocated to the system (Feldstein (1996)).

[116] According to the empirical analysis by Reimers and Honig (1996), at least men behave myopically because they respond to current retirement benefits rather than to their social security wealth.

In theory, old-age provision according to the capital funding principle constitutes a perfect substitute for private saving. Because the pension system generates the same rate of return that workers could earn via private investments on the capital markets, there is no distortive tax involved. The contributions to the funded system simply replace private savings that would otherwise have been necessary to provide for old-age income. Hence, a forced increase in social security will reduce private savings by an equal amount so that consumption, bequests, and aggregate savings will be unaffected (Barro (1974)). However, this offset between private and pension wealth may be less than one-for-one due to potential counter-effects such as bequest motives, myopia, liquidity constraints, and political risks (Bottazzi, Jappelli, and Padula (2006)).[117]

In contrast to capital funding, the pay-as-you-go (PAYG) system is based on a so-called inter-generational contract. In every period, the contributions of the current active population finance the pension entitlements of the retired population. No capital stock is accumulated because old-age provision entirely rests on unfunded intergenerational transfers. Hence, private savings are crowded out because people save less if they need not to provide for their own old-age income (Feldstein (1974)). In theory, the PAYG system constitutes a perfect substitute for private bequests (Barro (1974)).

The PAYG principle has two advantages compared to the capital funded pension system: protection against the risk of inflation and protection against fluctuations of financial markets because the aggregate risk is diversified over generations (Diamond and Orszag (2005)).[118] With capital funding,

[117] Sheshinski and Weiss (1981) argue that the result of Barro (1974) strictly depends on the absence of uncertainty. Otherwise, an increase in social security will be only partially compensated by a decrease in private savings, which implies an increase in aggregate savings. There has been substantial research on the effect of social security on private savings. Although there is no agreement on the magnitude of this effect, most studies suggest that social security reduces the amount of private savings. Cf. e.g. Blinder, Gordon, and Wise (1981), Feldstein (1982), and Kotlikoff (1979).

[118] According to Sinn (2004), the PAYG pension system also works as insurance against the risk of not having children and as an enforcement device for ungrateful children who

the higher expected return on one's own contributions has to be balanced against the higher riskiness of these investments (Feldstein (1996)). The risks of PAYG only refer to long-run political factors, future demographic evolutions, and the future of productivity and wages to which contributions and benefits are related (Miles, Timmermann, de Haan, and Pagano (1999)).

The main drawback of the PAYG system is its dependency on the relative size of the active population (Kotlikoff (1996)). For example, if demographic changes raise the dependency ratio (i.e. the ratio of retirees per active worker), sustainability of the system demands either higher individual contributions, lower old-age benefits, or an increase in the minimum retirement age (Disney (2000)). The primary costs of the PAYG principle are lower private savings and thus reduced capital accumulation (Feldstein (1985)). This theoretical conclusion is empirically confirmed by Samwick (2000) who uses a panel of countries over 25 years to analyze the effects of pension systems on aggregate savings. Another important cost is the deadweight loss of implicit taxation on labor supply at both the intensive and the extensive margins (cf. Section 3.2.2).

Altogether, the optimal level of PAYG benefits solves a trade-of between protection against the risk of elderly poverty and distortions concerning private savings and labor supply (Feldstein (1985)). According to Feldstein (2005a), the capital funded pension system provides a better solution to this trade-off.

#### Transition from Pay-As-You-Go to Capital Funding

Nearly all developed countries have adopted a pension system of the PAYG type. In general, pension systems are financed by a payroll tax imposed on the labor income of the active population. The size of these unfunded pension systems has increased over the last decades. In 1995, old-age provision absorbed 4.5% of GDP in the US, 13% in Italy, 16.5% in France, and over 20% in Sweden (Galasso and Profeta (2002)).

---

are unwilling to pay a pension to their parents.

Since the 1970s, various proposals for pension reform have suggested moving the prevalent PAYG system (at least partially) towards a capital funded system of old-age provision.[119] According to these reform proposals, reallocating resources from the PAYG system to financial assets will eliminate many shortcomings of the current system. First, demographic change will no longer affect the financial viability of the pension system. Second, the implementation of individual retirement accounts will reduce labor market distortions. And third, private savings and economic growth will be fostered, which implies an increase in the present value of expected future consumption (Feldstein (2005b)).

However, it is widely accepted that shifting the current PAYG pension systems towards capital funding will not generally be neutral (Cremer and Pestieau (2000)). Feldstein (2005a) provides a list of four issues determining whether a shift from unfunded to funded systems of old-age provision will raise social welfare: (1) the costs of the transition process, (2) the level of administrative costs, (3) the riskiness of financial markets, and (4) the implications for those workers at the bottom of the income distribution.

The magnitude of the transition costs is controversial in the literature. While Feldstein and Samwick (1999) and Lindbeck and Persson (2003) estimate moderate costs of moving from PAYG to capital funding, Miles, Timmermann, de Haan, and Pagano (1999) find rather high expenses. By simulating the effects of different pension reforms for Germany and Austria, Fehr (2000) and Keuschnigg and Keuschnigg (2004) conclude that all reforms redistribute towards future generations at the cost of currently active generations that have to "pay double".

Finally, instead of moving from PAYG towards capital funding, Diamond and Orszag (2005) propose only slight changes within the current framework of traditional PAYG, namely a mixture of increased contributions and reduced benefits in order to restore the financial sustainability of the pension

[119] Cf. Feldstein (1996), Feldstein and Samwick (1999), Feldstein and Samwick (2002), Kotlikoff (1996), and Mitchell and Zeldes (1996).

system. Imrohoroglu, Imrohoroglu, and Joines (1995) derive optimal social security replacement rates and associated benefits by means of an applied general equilibrium model. According to Boldrin, Dolado, Jimeno, Peracchi, Breyer, and Fernandez (1999) and Lindbeck and Persson (2003), an efficient pension system should be a mixture of PAYG and capital funded systems in order to diversify both political risks and the volatility of financial markets. However, if the initial system of old-age provision distorts endogenous labor supply and provides incentives to retire early, Bovenberg and Sorensen (2004) show that the introduction of compulsory retirement accounts can be Pareto-improving even in the presence of intragenerational heterogeneity.

### 4.2.2 Labor Market Effects of Pension Systems

#### Theoretical Implications

The PAYG system generates an implicit tax on labor supply at both the intensive and the extensive margins (Keuschnigg (2005)). While labor supply at the intensive margin means the continuous decision of how much to work during the active life, the extensive margin reflects the discrete labor supply choice by comparing costs and benefits of continuing work and postponing retirement by another period.

From an individual perspective, the rate of return on the contributions to the PAYG pension system is the internal rate of return equal to the sum of population and productivity growth rates, which is below the real interest rate in a dynamically efficient economy (Feldstein (2005b)).[120] The demographic change with declining labor force participation further reduces the rate of return to the PAYG system because it strictly depends on the relative size of the active population. The foregone interest margin is considered as an implicit tax on labor earnings which includes both the payroll marginal tax and foregone benefits (Cremer, Lozachmeur, and Pestieau (2004)). The

[120] This must hold at least in the long run. Otherwise, intertemporal budget constraints would no longer be defined (Keuschnigg (2005)).

size of this implicit tax is proportional to the difference between the market rate of interest on private savings and the rate of return on PAYG contributions. The implicit tax generates distortionary effects on labor supply of active workers, job search, and unemployment.[121]

Whether contributions to a PAYG system are perceived as an implicit tax depends on whether pensions are linked to one's own past contributions or not. A PAYG system of the Beveridge type pays a flat pension that is uncoupled from one's own contributions. In this case, there is no tax-benefit link at all so that individuals must perceive their contribution rates as a 100% tax because they receive the same pension anyway (Feldstein (2005a)). In contrast, a PAYG system of the Bismarck type includes a tax-benefit link that relates the size of the pension to the size of one's own contributions in the past.

Unfortunately, contributions tend to be actuarially unfair so that old-age insurance yields a much lower return than private savings (Feldstein (2005a)). According to calculations for Germany, about 50% of the contribution is a tax on labor while the rest is a price for individually received services (Fenge and Werding (2003)). Without changing the fundamental nature of the PAYG system, the implicit tax can be reduced by strengthening the individual tax-benefit link, i.e. by relating pensions more closely to one's own past contributions (Lindbeck and Persson (2003)). However, the capital funded system computes benefits in an actuarially fair way by definition. Hence, it provides the fullest possible tax-benefit link and thus avoids any type of implicit taxation. Moving to capital funding will not only raise aggregate capital accumulation, but will also eliminate implicit taxes on intensive and extensive labor supply (Feldstein (2005b)).

At the extensive margin, the implicit tax stems from the fact that most PAYG systems do not adjust the size of retirement benefits in an actuarially fair way (Crémer and Pestieau (2003)). According to Lau and Poutvaara

[121] Demmel and Keuschnigg (2000), Corneo and Marquardt (2000), and Keuschnigg (2004) show that unemployment increases with the level of implicit taxes.

(2001b), the unfair adjustment of benefits can be interpreted as a subsidization of early retirement. Indeed, the empirical evidence suggests a negative effect of pension contributions on employment and labor force participation (Scarpetta (1996)).

**Empirical Evidence**

There is a considerable literature on the economics of old-age insurance arguing that there is indeed a small but significant impact of pension systems on the average age of retirement and the likelihood of going on pension.[122] For example, Stock and Wise (1990) suggest that increasing the minimum retirement age from 55 to 60 has reduced the probability of retirement before age 60 by over a third.

Influential work of Börsch-Supan (2000) and Börsch-Supan (2003) shows the importance of the extensive margin for Germany. Börsch-Supan (2000) estimates that a decrease of retirement benefits by 12% would reduce the retirement probability of the 60 year old from 39.3% to 28.1% for a given labor income. According to Gruber and Wise (2005), each year of later retirement should be rewarded by a 6% increase in future benefits for the system to be actuarially fair. Empirical evidence concerning the sensitivity of the extensive margin is summarized in Diamond and Gruber (1997) and Gruber and Wise (2005).

### 4.2.3 Employment Prospects of Older Workers

In industrialized countries, an increasing number of older workers leave the labor force at younger and younger ages. For example, from 1950 to 1989 the labor force participation in the US declined from 46% to 17% for men over

[122] Cf. e.g. Blau (1994), Burtless and Moffitt (1985), Burtless (1986), Fields and Mitchell (1984), Hurd and Boskin (1984), Hall and Johnson (1980), Mitchell and Fields (1984), Rust and Phelan (1997), Samwick (1998), Mitchell and Phillips (2000), Crémer and Pestieau (2003), and Stock and Wise (1990). According to Samwick (1998), weak empirical results may be attributed to cross-sectional variation in retirement benefits.

65 and from 87% to 67% for men between 55 and 64 (Lumsdaine and Wise (1990)). In some European countries (for example France and Italy), male labor force participation between 60 and 64 fell from above 70% in 1960 to below 20% in 2002 (Conde-Ruiz and Galasso (2004)).

For older workers, rates of job loss have significantly increased in recent years. In the US, the 3-year job loss rate for workers over age 55 rose from 11% in 1981 to 16% in 1993 (Farber (1997)). According to Chan and Stevens (2001), losing a job has significant negative effects on future employment probabilities. Only 61% of displaced men and 55% of displaced women over age 50 are reemployed two years after a job loss. By using the National Longitudinal Study of Older Men, Diamond and Hausman (1984) confirm that older men face long periods of unemployment and increased retirement probabilities after a job loss. Furthermore, an empirical study by CongressionalBudgetOffice (1993) shows that about 50% of displaced workers over age 60 leave the labor force by early retirement.

Due to these poor labor market conditions, workers may be constrained from working until their desired retirement date (Coile and Levine (2006)). The employment prospects of older workers strongly depend on the firms' incentives to hire and to employ workers near retirement age (Chan and Stevens (2001)). Unfortunately, firm behavior has received much less attention in the retirement literature than the worker's decision to go on pension. However, profit-maximizing firms play an active role because they decide on the termination of production by comparing marginal benefits and marginal costs of employing workers for another time period (Hutchens (1999)). Hence, Hakola and Uusitalo (2005) conclude that it is important to consider retirement as a joint decision of workers and firms.

According to Bingley and Lanot (2004) and Heywood, Ho, and Wei (1999), there are three possible explanations why the employment prospects of older workers are reduced compared to those of younger workers: (1) a decline in the worker's productivity over time, (2) a steep age-earnings profile due to delayed compensation schemes, and (3) an increase in the level of non-wage labor costs over time.

As suggested in the life-cycle model of human capital accumulation by Heckman (1976), the worker's productivity may decline because his human capital depreciates over time. On the one hand, older workers have more work experience, which generally increases their productivity. But on the other hand, skills may decrease with age after a certain point so that productivity is reduced (Johnson and Neumark (1997)). Furthermore, unexpected positive technology shocks may accelerate the skill depreciation of older workers, which implies that firm-sponsored retraining is more costly for older workers because the returns to this investment are recouped over a shorter period of time (Bartel and Sicherman (1993)). Hence, firms may use early retirement as a possible way to renew their workforce (Hakola and Uusitalo (2005)).

The empirical evidence concerning the relationship between age and productivity is mixed (Haltiwanger, Lane, and Spletzer (1999)). In their empirical study for manufacturing firms in France between 1994 and 1997, Crépon, Deniau, and Pérez-Duarte (2002) find that the productivity of older workers declines by about 10% on average. By using wages as a proxy for individual productivity, Kotlikoff and Wise (1989) conclude that the productivity of salesmen increases until age 52 and then declines by 16% until age 60. In a similar study for workers of a trading concern, Kotlikoff and Gokhale (1992) find that there is a productivity loss of nearly 20% compared to the productivity maximum at age 47. An overview is provided by Börsch-Supan, Duzgun, and Weiss (2005).

Concerning the second explanation for reduced employment prospects of older workers, optimal long-term labor contracts according to Lazear (1981) imply that low-tenure workers earn less than their productivity and high-tenure workers earn more than their productivity in order to alleviate the monitoring problems of firms. However, this upward-sloping wage profile of delayed compensation schemes introduces a new source of inefficiency because workers have an incentive to stay on the job past the efficient age of retirement (Lazear (1979)). Hence, firms need to limit the time period over which older workers receive wages above their productivity (for example by "mandatory retirement") (Leigh (1984)).[123] This theoretical result is empirically confirmed by Daniel and Heywood (2007) who find that indicators of delayed compensation are associated with a lower likelihood of firms hiring older workers. In brief, delayed compensation schemes create incentives to reduce labor costs by terminating the employment of older workers (Heywood, Ho, and Wei (1999)).

With respect to the third explanation, Hutchens (1988) finds that newly hired older workers are clustered in a smaller set of occupations than newly hired younger workers and older workers in general. He concludes that firms may employ older workers but dislike to hire them, which implies that job opportunities decline with age. This result is empirically confirmed by Chan and Stevens (2001) and can be attributed to an increase in the level of non-wage labor costs over time. According to Straka (1992), employer contributions to health insurance negatively affect the employment prospects of older workers. Because the costs of health insurance raise with age, the employment prospects of older workers are reduced compared to those of younger workers (Scott, Berger, and Garen (1995)).[124]

In our formal analysis, we refer to the third explanation. We assume that the costs of production increase with the worker's age over the life-cycle, which implies that it is more costly for the firm to obtain the same output with an older worker than with a younger worker of the same initial productivity.

Besides the greater probability of job loss, the poor labor market conditions of older workers may have another implication. Workers may experience wage cuts, which increases the incentives to retire early. Although wage cuts are a theoretical possibility, in practice it seems unlikely based on previous

[123] In this context, Lazear (1979) suggests that the firm's age of mandatory retirement is positively correlated with the worker's education.

[124] Concerning the empirical analysis of Scott, Berger, and Garen (1995), Heywood, Ho, and Wei (1999) argue that the third explanation (increasing non-wage labor costs) cannot be easily disentangled from the second one (delayed compensation).

research concerning the cyclicality of real wages. According to empirical estimates by Devereaux (2001), the wage elasticity with respect to labor market prospects is low. Hence, in line with Coile and Levine (2006), we do not consider the possibility of wage cuts in our analysis.

Before analyzing two proposals for pension reform in Sections 4.5 and 4.6, we discuss the assumptions of our model (cf. Section 4.3) and the current pension system with traditional PAYG (cf. Section 4.4) in order to point out the analytical basis of comparison.

## 4.3 The Model

We consider a discrete-time model with two types of agents, namely workers and firms. In line with Feldstein (1985), there are two periods with fixed length. As modeled by Lau and Poutvaara (2001b) and Crémer and Pestieau (2003), the first period is fully active while the second period is endogenously split into a working subperiod and a retirement subperiod. In conformity with Lau and Poutvaara (2001a), our model allows for endogenous human capital formation, which is limited to the first period. Production takes place in worker-firm pairs and no capital is needed.

At the beginning of period 1, each firm meets one worker whose individual ability is drawn randomly from a distribution that is common knowledge.[125] In the second period, all workers can be employed regularly until at least one party decides to separate, either the worker or the firm. The remaining time of period 2 defines the length of the retirement subperiod. Altogether, the economy evolves over time as shown in Figure 4.1. The model assumptions

[125] In line with Acemoglu and Pischke (1999b), there is no exogenous separation after the first period. Incorporating an exogenous separating probability as in Malcomson, Maw, and McCormick (2003) does not change our analytical results because we focus on the supply side of the labor market. For the workers who face the training decision in period 1, it is irrelevant whether their higher wages in period 2 are paid by their current or by another employer.

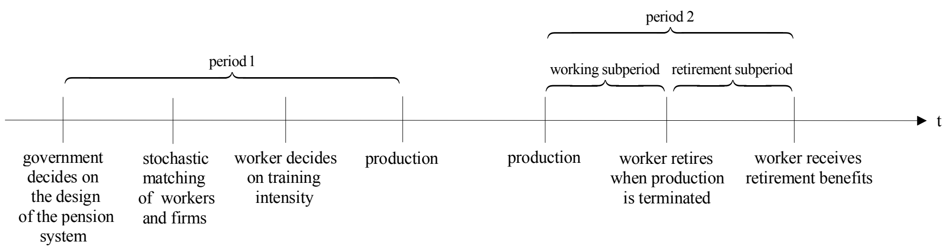


Figure 4.1: The Evolution over Time

and the labor market decisions of firms and workers are described in the following subsections.

### 4.3.1 The Workers

By assumption, each worker is matched with one firm that can unambiguously observe the worker's ability.[126] In line with Crémer and Pestieau (2003) and Jensen, Lau, and Poutvaara (2004), there are two types of workers with individual ability $\theta_i = \{\theta_L, \theta_H\}$ at the beginning of period 1 and identical lifetime $(1 + h)$. The length of period 1 is normalized to unity. The length of period 2 is equal to $h$ and covers both a working and a retirement subperiod. In accordance with Malcomson, Maw, and McCormick (2003), workers are risk-neutral and maximize the sum of their discounted utilities over both periods:

$$U(\theta_i) = u_{1i} + \delta u_{2i} \tag{4.1}$$

The discount factor $\delta \equiv \frac{1}{1+r}$ with $r$ as the market interest rate expresses the preference for current and future welfare. The higher $\delta$, the higher is the

[126] This assumption is in line with Boone and Bovenberg (2006). Furthermore, it is implicitly included into the whole literature on human capital and the life-cycle of earnings. Each worker offers his individual stock of human capital to the firms and is rewarded by a rental price per unit of human capital. Hence, we rule out asymmetric information (hidden knowledge). If the worker's productivity were not observed by the firm, there would be adverse selection as modeled e.g. by DeMeza and Webb (2001).

weighting of the following periods and the lower is the preference for period 1.

In period 1, workers can invest in their stock of human capital.[127] As in the literature on human capital accumulation over the life-cycle, the training costs have to be borne by the worker.[128] Although there are two major components of education costs, namely foregone earnings and direct expenditures (Parsons (1974)), we omit the former and focus on the latter by assuming direct costs of education that increase with the training intensity $e_i$ (cf. equation (4.2)). This assumption is in line with Lau and Poutvaara (2001a), where the individual stock of human capital depends on the amount of resources devoted to the process of educational production.[129] Furthermore, the worker can transfer wealth from period 1 to period 2 by reducing his consumption and saving an amount $s_i$. Altogether, the worker's utility

[127] This is an advancement compared to Martín (2003) who excludes human capital formation by assuming an exogenous distribution of education types.

[128] Most models that analyze the accumulation of human capital over the life-cycle completely concentrate on the investment decision of workers (for example Ben-Porath (1967) and Heckman (1976)).

[129] In a similar manner, also Sheshinski (1971), Atkinson (1973), and Nerlove, Razin, Sadka, and Weizsäcker (1993) consider only direct costs of education in their analyses of income taxation. In contrast to other contributions in the human capital literature (for example Trostel (1993)), this approach implies that the education costs are not deductible from the tax (i.e. the contribution rate $\tau$). Note that the tax-deductibility of training costs would reduce but not reverse the positive effect of an increase in the tax-benefit link on the desired retirement age and the level of training (cf. Appendix F). The effect would be lower because a stronger link between contributions and benefits mitigates the positive impact of the tax-deductibility on human capital formation. However, the qualitative conclusions of our analysis in Sections 4.5 and 4.6 would remain unchanged.

in the first period is equal to[130]

$$u_{1i} = (1 - \tau) \, w_{1i} - s_i - \frac{1}{2}(e_i)^2 \tag{4.2}$$

$$w_{1i} = \beta\theta_i \tag{4.3}$$

In both periods, the worker's wage corresponds to the Nash bargaining solution of oligopolistic labor markets. By defining the output good as numéraire and assuming an identical, linear one-to-one production function for the connection of output and labor (which is the only factor of production), the marginal product of each worker corresponds to his productivity $\theta_i$.[131] According to Acemoglu (1997), the parameter $0 \leq \beta \leq 1$ indicates the (identical) bargaining power of workers concerning the division of output. Because the worker receives zero income in the case of unemployment, the wage of worker $i$ in period 1 is equal to $w_{1i} = \beta\theta_i$, which implies that there are labor market frictions because the worker's wage falls below his marginal product for $\beta < 1$.[132] The contributions to the pension system are equal to payroll taxes $\tau$ that are proportional to the worker's labor income.

In period 2, the worker's productivity increases to $\theta_i'$ depending on the training intensity in the first period. Hence, in conformity with Lau and

[130] In line with Ben-Porath (1967), we do not analyze a more general utility function of workers. Note that the wage corresponds to the worker's labor income because labor supply is implicitly normalized to unity. By assumption, there are no effort costs of labor in period 1.

[131] The production function exhibits constant returns to scale. From the firm's point of view, the worker's ability can be interpreted as individual productivity.

[132] Wages in both periods are determined by Nash-bargaining, which implies that the worker's wage is a fraction of his marginal product $\theta_i$. The reason is that the firm and the worker maximize the Nash-product $(\theta_i - w_i)^{(1-\beta)} (w_i)^{\beta}$. While $w_i$ is the bargained wage of the worker, the firm is the residual claimant of output so that its profits are equal to the residuum $(\theta_i - w_i)$ (cf. Section 4.3.4). Because there is only one chance for a worker-firm match in period 1 (cf. Figure 4.1) and workers become unemployed in the case of no agreement, the failure to agree on a wage yields an income and profit level of zero. Hence, because the fall-back payoffs are zero, the bargained wage is equal to $w_i = \beta\theta_i$ (Ortigueira (2006)).

Poutvaara (2001a), there is no uncertainty about the return to education.[133] Like in the first period, the bargained wage corresponds to the Nash bargaining solution $w_{2i} = \beta\theta_i'$.[134] The labor income in period 2 is equal to the product of the worker's wage and labor supply $t_i$. The worker endogenously determines $t_i$, which defines the length of the working subperiod and thus the date of retirement. In line with Crémer and Pestieau (2003), there are convex effort costs of labor that increase with $t_i$.[135] Hence, the worker's utility in the second period is equal to

$$u_{2i} = (1-\tau)\, t_i w_{2i} + R s_{1i} - \frac{\gamma}{2}(t_i)^2 + (h - t_i)\, b \tag{4.4}$$

$$w_{2i} = \beta\theta_i' = (1+e_i)\,\theta_i \tag{4.5}$$

$$b = m(t_i)\, k + n(t_i)\left[(1+p)\,\tau w_{1i} + \tau t_i w_{2i}\right] \tag{4.6}$$

The amount of private savings $s_i$ from the first period is augmented by the factor $R = (1+r)$. The retirement subperiod has the length $(h - t_i)$ because it starts at date $(1 + t_i)$ and continues until the worker dies at date $(1 + h)$. According to Fisher and Keuschnigg (2007), retirement benefits are determined by the benefit rule as specified in equation (4.6). In general, pensions are composed of two parts. The first part refers to the pension system of the Beveridge type, which pays a flat pension $k$ uncoupled from own contributions. The second part refers to the Bismarck type and includes a tax-benefit link that relates the size of the pension $b$ to the size of one's own contributions in periods 1 and 2. Note that the parameter $p$ may be smaller than the market interest rate $r$. Whether contributions are actuarially fair, depends on the design of $m(t_i)$ and $n(t_i)$ that define to which

[133] Because $e_i$ determines the amount of training in period 1 as well as the increased productivity in period 2, it constitutes the key determinant of the return to education as analyzed in the theory of human capital (c.f. Section 1.4).

[134] The reason is that the fall-back payoffs are zero because the bargaining takes place in worker-firm pairs and this is the only chance for an agreement in period 2 (cf. Figure 4.1).

[135] Mitchell and Fields (1984) point out that, in addition to the income opportunities as defined by the benefit rule of the pension system, the worker's preferences for income relative to leisure are the most important determinant of the retirement age.

extent individual benefits are adjusted in response to the retirement decision $t_i$.

Altogether, the total utility of a worker with ability $\theta_i$ is obtained by substituting (4.2) to (4.5) into equation (4.1):

$$\begin{aligned} U(\theta_i) &= (1-\tau)\beta\theta_i - \frac{(e_i)^2}{2} \\ &\quad +\delta\left[(1-\tau)t_i\beta(1+e_i)\theta_i - \frac{\gamma}{2}(t_i)^2 + (h-t_i)b\right] \\ b &= m(t_i)k + n(t_i)\left[(1+p)\tau w_{1i} + \tau t_i w_{2i}\right] \end{aligned} \tag{4.7}$$

Note that private savings cancel out because they are increased and discounted by the same market interest rate $r$. Hence, they only represent transfers from one period to another without influence on total utility. In line with Lau and Poutvaara (2001a), total utility (4.7) is separable in lifetime consumption and lifetime leisure.

### 4.3.2 Retirement and Training with Laissez-Faire

In the laissez-faire equilibrium (LF) without government intervention, there are no social security contributions and no pension benefits. Hence, the parameters $\tau$ and $b$ are equal to zero in equation (4.7). Each worker maximizes the present value of his total utility with respect to the decision variables $t_i$ and $e_i$:

$$\max_{t_i,e_i} \beta\theta_i - \frac{(e_i)^2}{2} + \delta\left[t_i\beta(1+e_i)\theta_i - \frac{\gamma}{2}(t_i)^2\right] \tag{4.8}$$

Maximizing (4.8) with respect to $t_i$ and $e_i$ yields

$$t_i : \quad \delta\beta(1+e_i)\theta_i = \delta\gamma t_i \tag{4.9}$$

$$e_i : \quad \delta t_i\beta\theta_i = e_i \tag{4.10}$$

The first-order conditions compare the marginal benefits (on the left hand side) and the marginal costs (on the right hand side) of an increase in $t_i$ and $e_i$, respectively. Equation (4.9) represents the worker's optimization problem with respect to $t_i$. While the additional effort costs of labor are shown on the right hand side, the increase in labor income is equal to $\beta(1+e_i)\theta_i$, which corresponds to the wage in the second period. In (4.10), each additional unit of training generates costs equal to $e_i$, but this investment makes the worker more productive and thus increases the labor income in period 2 by $t_i\beta\theta_i$.

Furthermore, the first-order conditions (4.9) and (4.10) represent the optimal values $t_i = R^{LF}(e_i)$ and $e_i = R^{LF}(t_i)$ as a function of the other decision variable:

$$t_i = R^{LF}(e_i) = \frac{\beta\theta_i}{\gamma}(1+e_i) \tag{4.11}$$

$$e_i = R^{LF}(t_i) = \delta\beta\theta_i t_i \tag{4.12}$$

These optimality functions (4.11) and (4.12) are upward sloping in a $(e_i, t_i)$-diagram. Hence, as discussed in Section 1.6.2, retirement age and training intensity constitute complements concerning the worker's optimal choice of labor supply and human capital formation. Additionally, both $t_i$ and $e_i$ positively depend on the individual ability $\theta_i$.[136] This theoretical result is in line with the empirical findings of Fields and Mitchell (1984) who suggest that those workers gaining more by postponing retirement will retire later. In Figure 4.2, the two optimality functions are graphically illustrated for both ability types of workers.

In order to determine the worker's optimal values for retirement age and training, we have to combine (4.11) and (4.12).

[136] The complementarity of ability and human capital formation is in line with Blundell, Dearden, Meghir, and Sianesi (1999).

**Proposition 4.1** *With laissez-faire, the desired retirement age and the optimal training intensity of a worker with ability $\theta_i$ are equal to*[137]

$$(t_i)^{LF} = \frac{\beta\theta_i}{\gamma - \delta\left(\beta\theta_i\right)^2} \tag{4.13}$$

$$(e_i)^{LF} = \frac{\delta\left(\beta\theta_i\right)^2}{\gamma - \delta\left(\beta\theta_i\right)^2} \tag{4.14}$$

The worker's optimal values for retirement age and training increase with the bargaining power $\beta$ and initial ability $\theta_i$ because these parameters determine the wage and thus the benefits of both training in period 1 and work in period 2.[138] This result is in line with Cremer, Lozachmeur, and Pestieau (2004) and Martín (2003) who suggest that workers with high productivity will retire later. Furthermore, both $(t_i)^{LF}$ and $(e_i)^{LF}$ increase with $\delta$ because an increase in $\delta$ is equivalent to a decrease in $r$. Hence, the additional income in period 2 is discounted less and thus weighted to a greater extent. Finally, both values decrease with $\gamma$ because higher effort costs of labor imply that the net marginal benefits of work and education are reduced.

[137] In order to guarantee an interior solution for $t_i$ and $e_i$, we assume that the total length of period 2 is not too small, i.e. $h \geq \frac{\beta\theta_H}{\gamma-\delta(\beta\theta_H)^2}$. Alternatively, we could include the parameter $h$ in the effort costs of labor. However, in line with Cremer and Pestieau (2000), we abstain from this possibility in order to keep the calculations as simple as possible. The calculation of the optimal values for retirement age and training intensity with laissez-faire is presented in Appendix E. Concerning the following sections, the calculation of the optimal values proceeds in the same manner.

Although the problem of reduced labor force participation is raised by demographic change (cf. Section 4.1), the optimal values in (4.13) and (4.14) are independent of the worker's lifetime because utility in the second period considers the sum $(h - t_i)\, b$ of all retirement benefits. $(t_i)^{LF}$ and $(e_i)^{LF}$ would depend on $h$ if the worker's utility were not linearly affected by the income at different dates of the retirement subperiod. In this case, the marginal utilities of income could differ between different dates and an increase in $h$ (by exogenous demographic change) would require to postpone retirement in order to maximize total utility by equalizing the marginal utilities.

[138] Note that there are two opposite effects of the individual ability on the retirement age, namely an income effect with negative correlation and a substitution effect with positive correlation (Jensen, Lau, and Poutvaara (2004)).

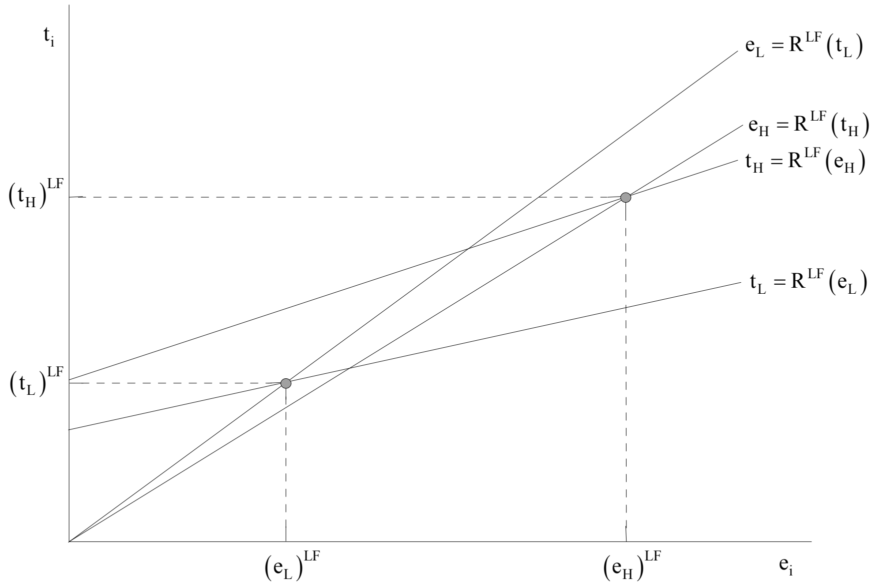


Figure 4.2: The Workers with Laissez-Faire

### 4.3.3 Retirement and Training with Pension Systems

With pension systems, each worker has to pay contributions during his working life and receives retirement benefits $b$ that depend on the type of the pension system. Note that the partial-equilibrium framework of this analysis implies that the model does not take into account the budget constraint of the pension system and the interactions between the contribution rate $\tau$ and the level of benefits.

With a pension system, each worker maximizes his total utility (4.7) subject to the benefit rule (4.6) as defined by the pension system:

$$\max_{e_i,t_i} (1-\tau)\beta\theta_i - \frac{(e_i)^2}{2} + \delta\left[(1-\tau)t_i\beta(1+e_i)\theta_i - \frac{\gamma}{2}(t_i)^2 + (h-t_i)b\right]$$
$$s.t. \quad b = m(t_i)k + n(t_i)\left[(1+p)\tau w_{1i} + \tau t_i w_{2i}\right]$$

The first-order conditions are the following:

$$t_i \quad : \qquad \delta\left[(1-\tau)\,\beta(1+e_i)\theta_i + (h-t_i)\,\frac{\partial b}{\partial t_i}\right] = \delta\left[\gamma t_i + b\right] \qquad (4.15)$$

$$e_i \quad : \qquad \delta\left[t_i\,(1-\tau)\,\beta\theta_i + (h-t_i)\,\frac{\partial b}{\partial e_i}\right] = e_i \qquad (4.16)$$

As in Section 4.3.2, the first-order conditions compare the marginal benefits (on the left hand side) and the marginal costs (on the right hand side) of an increase in $t_i$ and $e_i$, respectively. According to equation (4.15), the marginal benefits of an increase in $t_i$ are changed compared to (4.9) due to the pension system. First, the wage for an additional unit of labor supply is decreased by the contributions $\tau$. Second, if the retirement benefits are adjusted in response to the length of the working subperiod, $b$ increases with $t_i$. Furthermore, the marginal costs on the right hand side are increased because the worker forgoes the retirement benefits $b$ if he decides to postpone retirement by another unit of time. In (4.16), the marginal benefits of an increase in $e_i$ are modified compared to laissez-faire in (4.10). First, the additional labor income in period 2 is reduced by the contributions. And second, if the pension system contains a tax-benefit link, more training increases the level of benefits during the retirement subperiod.

Compared to the situation with laissez-faire, the worker's decision with respect to retirement age and training is distorted by the pension system. As discussed in Section 4.2.2, the degree of these distortions depends on the type of the pension system and can be illustrated by the implicit tax rates on labor supply and training intensity. According to Gruber and Wise (2005), the implicit tax on labor supply at the extensive margin is equal to the ratio of the change in the present value of retirement benefits and labor income.

**Definition 4.1** *Depending on the type of the pension system, the implicit tax rates $\tau_t^*$ on labor supply at the extensive margin and $\tau_e^*$ on training intensity at the intensive margin are defined by*

$$\tau_t^* \equiv \tau + \frac{b - (h - t_i) \frac{\partial b}{\partial t_i}}{\beta(1 + e_i)\theta_i} \tag{4.17}$$

$$\tau_e^* \equiv \tau - \frac{(h - t_i) \frac{\partial b}{\partial e_i}}{t_i \beta \theta_i} \tag{4.18}$$

### 4.3.4 The Firms

Firms are not affected by the system of old-age provision so that they solve the same decision problem with laissez-faire and with a pension system. As modeled by Malcomson, Maw, and McCormick (2003), firms are risk-neutral and maximize the sum of their discounted profits over both periods:

$$\pi(\theta_i) = \pi_{1i} + \delta \pi_{2i} = (1 - \beta)\,\theta_i + \delta \left[ t^F (1 - \beta)(1 + e_i)\,\theta_i - \frac{f}{2\theta_i} (t^F)^2 \right] \tag{4.19}$$

In period 1, the profits correspond to the Nash bargaining solution, i.e. $\pi_{1i} = (1 - \beta)\,\theta_i$. In the second period, each firm decides how long production with the worker is continued. The firm leaves the market at the separation date $t^F$, which implies that production is terminated. As shown in (4.19), firms have to bear some time-dependent costs of production, which increase with $t^F$ because non-wage labor costs rise over time (cf. Section 4.2.3). Hence, the firm's incentives to terminate an employment contract increase with the worker's age because it is more costly for the firm to obtain the same output with an older worker than with a younger worker of the same initial ability.[139]

[139] As mentioned at the end of Section 4.2.3, we do not consider wage cuts of older workers as a consequence of poor labor market conditions. Hence, in line with Coile and Levine (2006), we rule out that older workers respond to their reduced employment prospects by a reduction in their wage. This implies that Coasian bargaining on the time-dependent costs of production and writing a complete contract covering all costs and benefits of

Maximizing (4.19) with respect to $t^F$ yields the first-order condition

$$\delta (1 - \beta) (1 + e_i) \theta_i = \delta \frac{f}{\theta_i} t^F \tag{4.20}$$

Equation (4.20) compares the marginal revenue (on the left hand side) and the marginal costs (on the right hand side) of an increase in $t^F$. While production in period 2 generates a marginal revenue equal to the Nash bargaining solution $(1 - \beta) (1 + e_i) \theta_i$, there are additional time-dependent costs that have to be borne by the firm. By solving (4.20) for $t^F$, we obtain the optimal separation date of the firm as a function of the worker's training intensity $e_i$ (cf. Figure 4.3).

**Proposition 4.2** *The optimal separation date of the firm is equal to*

$$t^F (e_i) = \frac{1 - \beta}{f} (1 + e_i) (\theta_i)^2 \tag{4.21}$$

The optimal separation date of the firm increases with $e_i$ and $\theta_i$ because the worker's productivity and thus the firm's profits are raised. Hence, not only the worker's desired retirement age (cf. Section 4.3.2) but also the firm's optimal separation date positively depend on the worker's productivity. Furthermore, $t^F (e_i)$ decreases with $f$ because higher time-dependent costs of production make it is less profitable to employ the worker. Obviously, $t^F (e_i)$ decreases with $\beta$ because the firm's output share is equal to $(1 - \beta)$.

## 4.4 The Traditional PAYG Pension System

With a PAYG system of old-age provision, contributions to the pension system are not treated like private savings and benefits are not refunded from the stock of own savings (cf. Section 4.2.1). In most countries, the current

postponed retirement are not feasible. However, even if workers could lower their wage to finance some fraction of the time-dependent costs, the actual retirement date would fall below (4.13) because the worker's incentives for early retirement are increased by the lower wage.


Figure 4.3: The Separation Decision of the Firms

PAYG pension system represents a mixture of the Beveridge type and the Bismarck type with actuarially fair adjustment of benefits. In the following, this mixture is analyzed in detail because it constitutes the analytical basis of comparison for evaluating the proposals for pension reform in Sections 4.5 and 4.6.

With the Beveridge type system, benefits are completely uncoupled from own contributions and independent of the length of the working life:[140]

$$b^{Beveridge} = k \tag{4.22}$$

In contrast, retirement benefits according to the Bismarck type depend on the worker's contributions in the past and thus include a tax-benefit link:

$$b^{Bismarck} = \frac{\alpha}{h - t_i} \left[ \tau \beta \theta_i + t_i \tau \beta \left(1 + e_i\right) \theta_i \right] \tag{4.23}$$

[140]These properties implies $m(t_i) = 1$ and $n(t_i) = 0$ in the benefit rule (4.6).

The parameter $\alpha$ provides a measure of the tax-benefit link because it determines to which extent the retirement benefits depend on one's own contributions. The benefits in (4.23) are actuarially adjusted because they are divided by $(h - t_i)$, the length of the retirement subperiod. However, in contradiction to capital funding, contributions of period 1 are not augmented by the market interest rate, which implies $p = 0$ in equation (4.6). Altogether, we assume that the traditional PAYG system of old-age provision pays benefits equal to[141]

$$b^{PAYG} = b^{Beveridge} + b^{Bismarck} \tag{4.24}$$

The two parts of this benefit rule (4.24) focus on different socioeconomic goals. While the Beveridge part of the pension system provides a basic pension in order to prevent poverty, the Bismarck part pays earnings-related benefits in order to sustain the workers' previous standard of living (Jensen, Lau, and Poutvaara (2004)). In brief, the flat pension $b^{Beveridge}$ represents the level of minimum pension, which is also granted to those workers who have never stayed inside the labor market.

### 4.4.1 Retirement and Training with Traditional PAYG

Starting from the first-order conditions (4.15) and (4.16) of the general decision problem with a pension system in Section 4.3.3, the system of traditional PAYG yields the following optimality functions:

$$\begin{aligned} t_i &= R^{PAYG}(e_i) = (1 - (1 - \alpha)\tau)\frac{\beta\theta_i}{\gamma}(1 + e_i) - \frac{k}{\gamma} && (4.25) \\ e_i &= R^{PAYG}(t_i) = \delta(1 - (1 - \alpha)\tau)\beta\theta_i t_i && (4.26) \end{aligned}$$

Compared to the situation with laissez-faire in (4.11), the optimality function $t_i = R^{PAYG}(e_i)$ is affected twice. First, it is twisted downward because

[141] According to the propositions by Boskin, Kotlikoff, Puffert, and Shoven (1986), this two-pillar structure may be referred to as the "Boskin Proposal" (Huggett and Ventura (1999)).

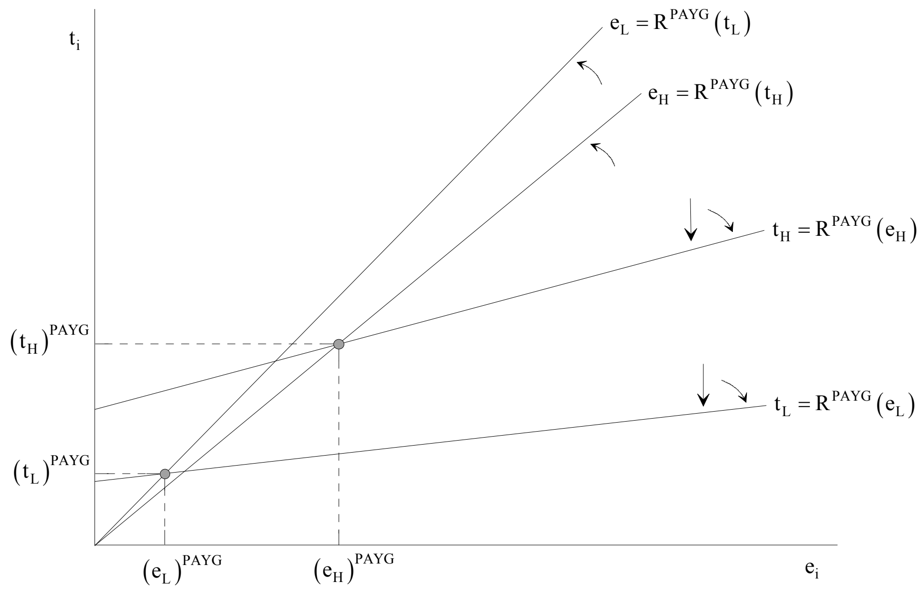


Figure 4.4: The Workers with Traditional PAYG

the marginal benefits of labor supply in period 2 are reduced by the factor $(1-(1-\alpha)\tau)$, that part of the retirement benefits which is not adjusted in an actuarially fair way. Obviously, the increase in the level of benefits depends on $\alpha$, the strength of the tax-benefit link. And second, $t_i$ is shifted downward because the Beveridge part of the benefits is independent of the worker's past contributions. Furthermore, the optimality function $e_i = R^{PAYG}(t_i)$ is twisted upward compared to (4.12) because the marginal benefits of training are scaled down. These modifications of the optimality functions are illustrated in Figure 4.4.

**Proposition 4.3** *With traditional PAYG, the desired retirement age and the optimal training intensity of a worker with ability $\theta_i$ are equal to*

$$(t_i)^{PAYG} = \frac{(1-(1-\alpha)\tau)\beta\theta_i - k}{\gamma - \delta\left[(1-(1-\alpha)\tau)\beta\theta_i\right]^2} \tag{4.27}$$

$$(e_i)^{PAYG} = \delta(1-(1-\alpha)\tau)\beta\theta_i(t_i)^{PAYG} \tag{4.28}$$

As with laissez-faire, the worker's optimal values for retirement age and training increase with the bargaining power $\beta$, the ability $\theta_i$, and the discount factor $\delta$ while they decrease with $\gamma$, the effort costs of labor. Furthermore, $(t_i)^{PAYG}$ and $(e_i)^{PAYG}$ decrease with $\tau$ and $k$ because these parameters of the pension system reduce the returns to labor supply and education. Hence, cutting the payroll tax rate or the Beveridge type benefits will, ceteris paribus, postpone retirement. This result is in line with Martín (2003) who concludes that especially the retirement age of low-skilled workers is distorted by the minimum pension as represented by the Beveridge part of the pension system. As suggested by Jensen, Lau, and Poutvaara (2004), strengthening the tax-benefit link implies an increase in the parameter $\alpha$, which fosters human capital formation and increases the worker's retirement age.

With traditional PAYG, $(t_i)^{PAYG}$ and $(e_i)^{PAYG}$ are smaller than with laissez-faire because there is an implicit tax on extensive labor supply of older generations near retirement and on training intensity at the intensive margin:

$$(\tau_t^*)^{PAYG} = (1-\alpha)\,\tau + \frac{k}{\beta\left(1+(e_i)^{PAYG}\right)\theta_i} \tag{4.29}$$

$$(\tau_e^*)^{PAYG} = (1-\alpha)\,\tau \tag{4.30}$$

The implicit tax on labor supply in (4.29) is composed of two effects that are both generated by the pension system if the tax-benefit link is imperfect and the adjustment of benefits is actuarially unfair. First, for $\alpha < 1$ the tax-benefit link is imperfect because additional contributions do not fully translate into higher retirement benefits. And second, the Beveridge type component of the pension system implies that the individual labor income is reduced without proportionally increasing the level of pensions. Hence, each worker forgoes benefits $k$ by postponing retirement for another unit of time. In line with Crémer and Pestieau (2003), the implicit tax rate $(\tau_t^*)^{PAYG}$ increases with $k$ and decreases with $\theta_i$.

### 4.4.2 The Current Situation with Traditional PAYG

Both workers and firms decide how long they want to produce in the second period. Depending on the worker's training intensity, the optimal separation date of the firm is described in equation (4.21). Altogether, the retirement date $(t_i)^*$ for ability type $\theta_i$ is defined by the smaller of the two values in (4.21) and (4.27) because then one party decides to leave the market and production is terminated:

$$(t_i)^* = \min\left\{(t_i)^{PAYG}, (t_i)^F\right\} \tag{4.31}$$

In the following analysis, we assume that the two ability types $\theta_L$ and $\theta_H$ satisfy the following condition:

$$(1-\tau)\frac{\beta}{1-\beta}\frac{f}{\gamma} < \theta_L < \frac{\beta}{1-\beta}\frac{f}{\gamma} < \theta_H \tag{4.32}$$

**Proposition 4.4** *The assumption (4.32) implies that the retirement date $(t_H)^*$ of ability type H corresponds to the worker's desired retirement age in (4.27) (cf. Appendix G.1):*

$$(t_H)^* = (t_H)^{PAYG} = \frac{(1-(1-\alpha)\tau)\beta\theta_H - k}{\gamma - \delta\left[(1-(1-\alpha)\tau)\beta\theta_H\right]^2} \tag{4.33}$$

Whether the retirement date $(t_L)^*$ of ability type L is determined by the worker or by the firm depends on the strength of the tax-benefit link.

**Proposition 4.5** *In order to focus on the most interesting case in the context of pension reform (cf. Sections 4.5 and 4.6), we assume that $\alpha$ is sufficiently small so that also $(t_L)^*$ corresponds to the worker's desired retirement age (cf. Appendix G.1):*

$$(t_L)^* = (t_L)^{PAYG} = \frac{(1-(1-\alpha)\tau)\beta\theta_L - k}{\gamma - \delta\left[(1-(1-\alpha)\tau)\beta\theta_L\right]^2} \tag{4.34}$$

Subject to these assumptions, the traditional PAYG pension system is graphically illustrated in figure 4.5. Because the retirement date is chosen

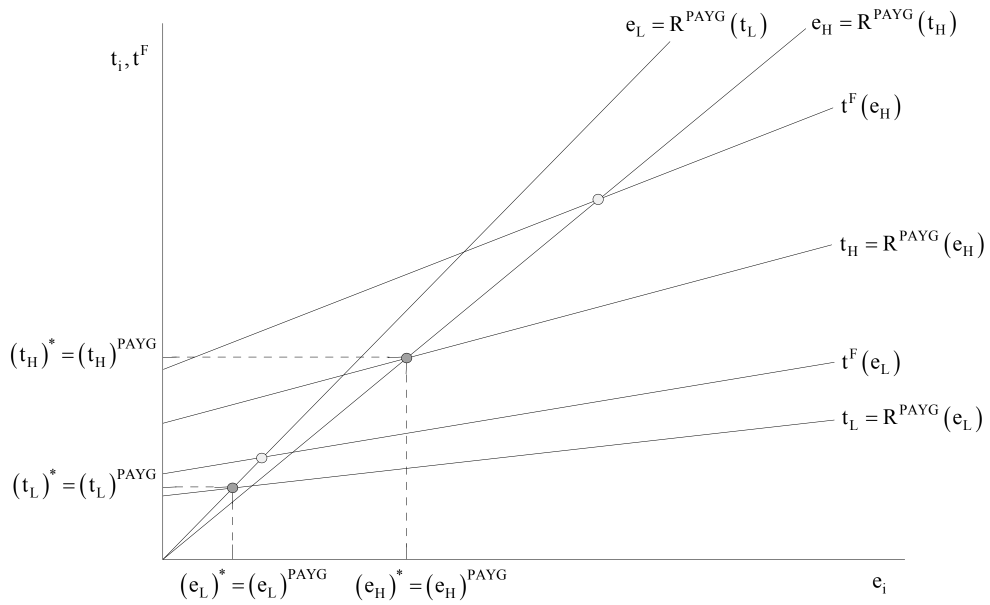


Figure 4.5: The Traditional PAYG Pension System

by the workers, also the training intensity of both ability types corresponds to the worker's optimal value in (4.28):

$$(e_H)^* = (e_H)^{PAYG} = \delta\left(1-(1-\alpha)\tau\right)\beta\theta_H\left(t_H\right)^{PAYG} \quad (4.35)$$

$$(e_L)^* = (e_L)^{PAYG} = \delta\left(1-(1-\alpha)\tau\right)\beta\theta_L\left(t_L\right)^{PAYG} \quad (4.36)$$

## 4.5 Pension Reform I: Individual Retirement Accounts

Policy discussions often propose to move from the traditional PAYG system to capital funding or at least to a mixed system comprising both a PAYG and a capital funded pillar. As discussed in Section 4.2, the capital funded system provides the fullest possible tax-benefit link and computes benefits in an actuarially fair way by definition:

$$b^{CF} = \frac{1}{h - t_i} \left[ R\tau\beta\theta_i + t_i\tau\beta \left(1 + e_i\right) \theta_i \right] \tag{4.37}$$

Capital funding implies that the sum of all contributions during the working life is refunded as annuities after the retirement. The contributions of period 1 are augmented by the factor $R = (1 + r)$.

Starting from the traditional PAYG system of old-age provision in Section 4.4, we analyze the implementation of individual retirement accounts and the impact on labor supply and human capital formation. Individual retirement accounts represent the capital funded pillar of the pension system according to (4.37). Suppose that $\omega$ is the proportion of benefits generated by individual retirement accounts and thus a measure for the degree of capital funding of the pension system. Hence, total benefits $b^{IRA}$ with individual retirement accounts (IRA) are composed of two parts that correspond to the different pillars of the pension system and are included by the weights $\omega$ and $(1 - \omega)$:[142]

$$\begin{aligned} b^{IRA} &= \omega b^{CF} + (1 - \omega)\, b^{PAYG} \\ b^{PAYG} &= k + \frac{\alpha}{h - t_i} \tau\beta\theta_i \left[1 + t_i \left(1 + e_i\right)\right] \end{aligned}$$

### 4.5.1 Retirement and Training with Individual Retirement Accounts

Similar to (4.25) and (4.26) with the traditional PAYG pension system, the worker's decision problem with individual retirement accounts yields the op-

[142] In fact, most PAYG systems determine the division of contributions rather than the division of benefits. In this case, $\omega$ is related to $\tau$ and the division of $b^{IRA}$ is endogenously determined by the different returns to the capital funded pillar and the PAYG pillar of the pension system. Because the return to capital funding corresponds to the market rate of return and thus exceeds the return to PAYG, the fraction of benefits generated by individual retirement accounts is larger than $\omega$ (Feldstein (2005a)).

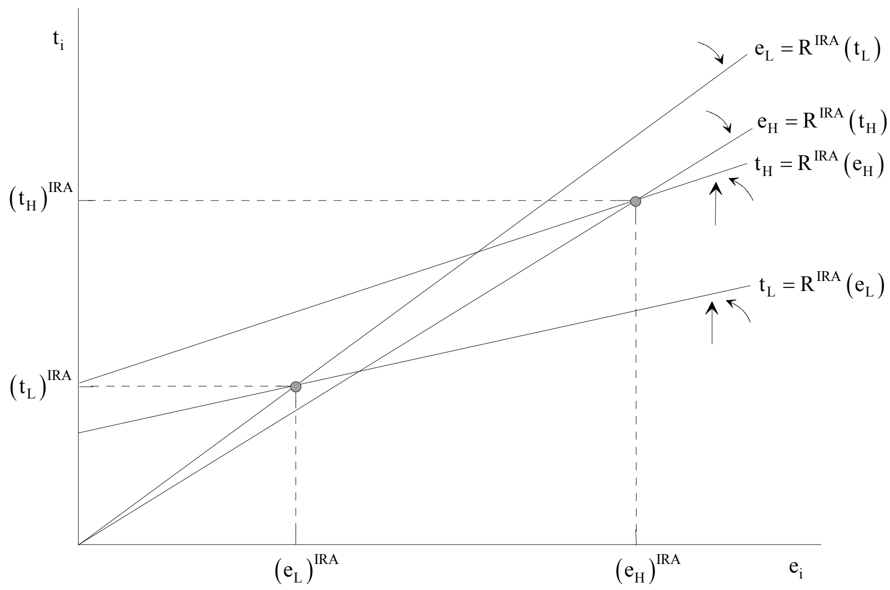


Figure 4.6: The Workers with Individual Retirement Accounts

timality functions

$$t_i = R^{IRA}(e_i) = [1-(1-\omega)(1-\alpha)\tau]\frac{\beta\theta_i}{\gamma}(1+e_i)-(1-\omega)\frac{k}{\gamma} \quad (4.38)$$

$$e_i = R^{IRA}(t_i) = \delta[1-(1-\omega)(1-\alpha)\tau]\beta\theta_i t_i \quad (4.39)$$

The optimality functions (4.38) and (4.39) are graphically illustrated in Figure 4.6. Compared to the situation with traditional PAYG in (4.25), the optimality function $t_i = R^{IRA}(e_i)$ is affected twice. It is twisted upward and shifted upward because both the tax-benefit link and the actuarial fairness are increased by the capital funded pillar of the pension system. Furthermore, $e_i = R^{IRA}(t_i)$ is twisted downward because the marginal benefits of training are raised.

**Proposition 4.6** *With individual retirement accounts, the desired retirement age and the optimal training intensity of a worker with ability $\theta_i$ are equal to*

$$(t_i)^{IRA} = \frac{\xi(\omega)\beta\theta_i - (1-\omega)k}{\gamma - \delta[\xi(\omega)\beta\theta_i]^2} \tag{4.40}$$

$$(e_i)^{IRA} = \delta\xi(\omega)\beta\theta_i(t_i)^{IRA} \tag{4.41}$$

*with* $\xi(\omega) \equiv 1 - (1-\omega)(1-\alpha)\tau$.

In line with Feldstein (2005b), the worker's desired retirement age $(t_i)^{IRA}$ increases with $\omega$ because the incentives for early retirement are reduced by raising the degree of capital funding of the pension system. The reason is that the implicit tax rates are decreased by the introduction of individual retirement accounts:

$$(\tau_t^*)^{IRA} = (1-\omega)\left[(1-\alpha)\tau + \frac{k}{\beta\left(1+(e_i)^{IRA}\right)\theta_i}\right] \tag{4.42}$$

$$(\tau_e^*)^{IRA} = (1-\omega)(1-\alpha)\tau \tag{4.43}$$

If there are no individual retirement accounts (i.e. $\omega = 0$), $(t_i)^{IRA}$ and $(e_i)^{IRA}$ are equal to the optimal values of the pure PAYG system according to (4.27) and (4.28). In contrast, if the pension system is completely capital funded (i.e. $\omega = 1$), there are no distortions compared to the situation with laissez-faire because the implicit tax rates are equal to zero. In this case, the pension system is referred to as neutral or actuarially fair (Cremer, Lozachmeur, and Pestieau (2004)). The worker has the same incentives as with laissez-faire and chooses his optimal values for retirement age and training intensity according to (4.13) and (4.14).

Note that even the traditional PAYG system manages to mimic the capital funded system of old-age provision. As suggested by Lindbeck and Persson (2003), the implicit tax on labor supply can be reduced by strengthening the tax-benefit link. The implicit tax rates (4.29) and (4.30) are equal to zero if the pension system involves the parameters $\alpha = 1$ and $k = 0$. This result is in conformity with Lau and Poutvaara (2001b) who suggest that the best PAYG pension system is of the Bismarck type with an actuarially fair adjustment of benefits. Hence, instead of analyzing the implementation of individual retirement accounts, we could investigate an increase in $\alpha$. Indeed, the qualitative conclusions are nearly the same. However, we focus on the introduction of a capital funded pillar because this reform proposal is the most prevalent in current policy debates (cf. Section 4.2).

### 4.5.2 The Effects of Individual Retirement Accounts

Similar to (4.31) for the traditional PAYG system, the termination of production in period 2 for ability type $\theta_i$ is defined by

$$(t_i)^* = \min\left\{(t_i)^{IRA}, (t_i)^F\right\} \tag{4.44}$$

**Proposition 4.7** *With individual retirement accounts, the retirement date $(t_H)^*$ of ability type H corresponds to the worker's desired retirement age in (4.40) (Appendix G.2):*

$$(t_H)^* = (t_H)^{IRA} = \frac{\xi(\omega)\beta\theta_H - (1-\omega)k}{\gamma - \delta\left[\xi(\omega)\beta\theta_H\right]^2} \tag{4.45}$$

Concerning ability type L, the degree of capital funding ($\omega$) determines the retirement date in the second period. If the capital funded pillar is weak (i.e. $\omega$ is small), the retirement date of ability type L also corresponds to the worker's desired retirement age in (4.40). However, if the capital funded pillar is strong (i.e. $\omega$ is large), it is the firm that determines the termination of production in period 2 because the optimality function $t_L = R^{IRA}(e_L)$ moves upward. Hence, there must be a critical level $\bar{\omega}$ such that for $\omega > \bar{\omega}$ the firm's optimal separation date $(t_L)^F$ falls below the worker's desired retirement age $(t_L)^{IRA}$. In this case, $(t_L)^* = (t_L)^F$ is obtained by substituting the optimality function (4.39) into equation (4.21) and solving for $t$.

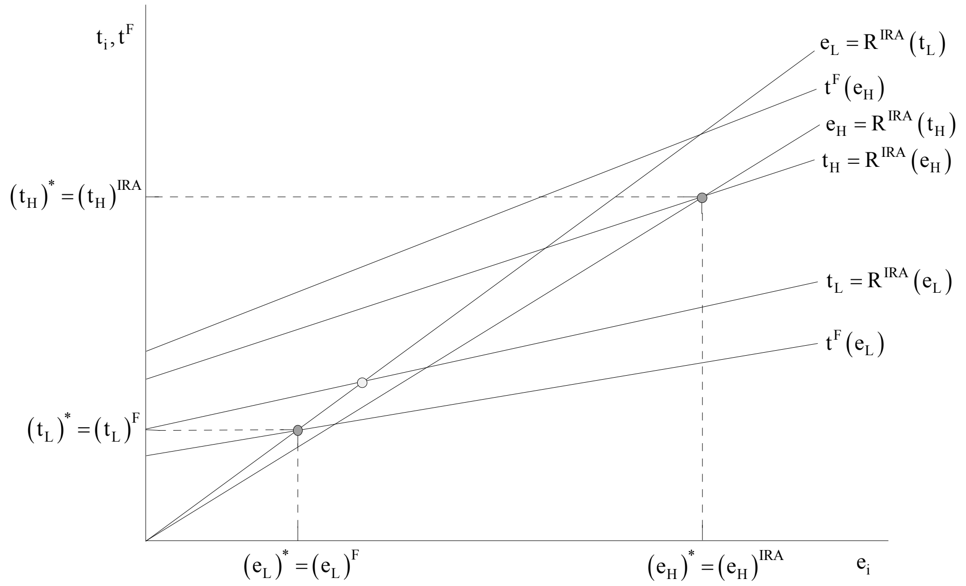


Figure 4.7: The Effecs of Individual Retirement Accounts

**Proposition 4.8** *Depending on $\omega$, the retirement date $(t_L)^*$ with individual retirement accounts is equal to (cf. Appendix G.2)*

$$(t_L)^* = \begin{cases} (t_L)^{IRA} = \frac{\xi(\omega)\beta\theta_L-(1-\omega)k}{\gamma-\delta[\xi(\omega)\beta\theta_L]^2} & \text{if } \omega \leq \bar{\omega} \\ (t_L)^F = \frac{(1-\beta)(\theta_L)^2}{f-\delta\xi(\omega)\beta(1-\beta)(\theta_L)^3} & \text{if } \omega > \bar{\omega} \end{cases} \tag{4.46}$$

For $\omega > \bar{\omega}$, the implementation of individual retirement accounts is graphically illustrated in Figure 4.7. Furthermore, the training intensity of both ability types is determined by substituting the retirement dates (4.45) and (4.46) into the optimality function (4.39), respectively. Hence, the level of training with individual retirement accounts is equal to

$$(e_H)^* = (e_H)^{IRA} = \delta\xi(\omega)\beta\theta_H(t_H)^{IRA} \tag{4.47}$$

$$(e_L)^* = \begin{cases} (e_L)^{IRA} = \delta\xi(\omega)\beta\theta_L(t_L)^{IRA} & \text{if } \omega \leq \bar{\omega} \\ (e_L)^F = \delta\xi(\omega)\beta\theta_L(t_L)^F & \text{if } \omega > \bar{\omega} \end{cases} \tag{4.48}$$

### 4.5.3 Interpretation

The implementation of individual retirement accounts implies that the implicit tax on labor supply and training is reduced. This can be seen by comparing the implicit tax rates with traditional PAYG in (4.29) and (4.30) and with individual retirement accounts in (4.42) and (4.43). As a consequence, human capital formation is increased and retirement is postponed. These results are in line with Lau and Poutvaara (2001a). However, concentrating on the worker's retirement and training decision in (4.40) and (4.41) and neglecting the demand side of the labor market may be misleading in evaluating the welfare gains of individual retirement accounts.

The effects of pension reforms strongly depend on the employment prospects of older workers near retirement age. These labor market prospects positively depend on the worker's ability. While high-ability workers indeed react to the introduction of individual retirement accounts according to (4.45) and (4.47), low-ability workers may be restricted by the firm's decision to terminate production. If the capital funded pillar of the pension system becomes strong, low-ability workers may not attain their desired retirement age because firms refuse to employ them any longer. For $\omega > \bar{\omega}$, the retirement date of ability type L is determined by the optimal separation date of the firm. This implies that the actual retirement date of low-ability workers increases to a lower extent than suggested by the worker's desired retirement age.

**Proposition 4.9** *Depending on the level of $\omega$, the effects of strengthening the capital funded pillar for low-ability workers are given by*

$$\frac{\partial \left(t_L\right)^*}{\partial \omega} = \begin{cases} \frac{\partial (t_L)^{IRA}}{\partial \omega} & \text{if } \omega \leq \bar{\omega} \\ \frac{\partial (t_L)^{F}}{\partial \omega} & \text{if } \omega > \bar{\omega} \end{cases}$$

*with (cf. Appendix H)*

$$0 < \frac{\partial \left(t_L\right)^F}{\partial \omega} < \frac{\partial \left(t_L\right)^{IRA}}{\partial \omega}$$

The retirement date $(t_L)^*$ unambiguously increases with $\omega$ because also the firm's optimal separation date $(t_L)^F$ positively depends on $\omega$. The reason is that an increase in $\omega$ fosters human capital formation according to (4.39) and so the firm can increase its profits by extending the time of production in period 2. However, for $\omega > \bar{\omega}$, there is a gap between the desired retirement age of low-ability workers and the actual retirement date due to the firm's decision to terminate production. This gap increases with $\omega$ and has its maximum for $\omega = 1$. As a consequence, the ratio of retirees per active worker is reduced by less than suggested by the worker's incentives to go on pension.

For $\omega = 1$, the pension system is completely moved towards capital funding, which implies that the worker has the same incentives as in the case of laissez-faire and chooses his optimal values for retirement age and training intensity according to (4.13) and (4.14). However, these values are not feasible once the separation decision of the firms is controlled for.

Note that the extent to which the retirement date for $\omega > \bar{\omega}$ responds to the introduction of individual retirement accounts strongly depends on the sensitivity of human capital formation. The less the worker's training intensity is affected by the implementation of individual retirement accounts, the lower is the impact on the actual retirement date. This implies that our conclusions will be even strengthened if training is exogenous and thus excluded from the analysis. In this case, the firm's optimal separation date in (4.21) simplifies to $\left.(t_L)^F\right|_{e=0} = \frac{1-\beta}{f}(\theta_L)^2$. It is completely independent of the degree of capital funding and only depends on constant parameters that are independent of the pension system. Hence, the firm's optimal separation date for ability type L remains unchanged and the implementation of individual retirement accounts has no effect at all on the retirement date of low-ability workers:

$$\frac{\partial \left.(t_L)^*\right|_{e=0}}{\partial \omega} = \frac{\partial \left.(t_L)^F\right|_{e=0}}{\partial \omega} = 0 \qquad \text{if } \omega > \bar{\omega}$$

Therefore, moving to a capital funded system of old-age provision does not necessarily have positive effects on labor force participation. The superiority of capital funding as postulated by Feldstein (2005a) is lowered and may even vanish in the presence of labor market imperfections.

## 4.6 Pension Reform II: Minimum Retirement Age

Postponed retirement is generally considered to be a key policy response to population aging because the ratio of retirees per active worker is reduced without changing the nature of the PAYG system (Lindbeck and Persson (2003)). In this context, a study of the EuropeanCommission (2001) concludes that increasing the effective retirement age to 65 will significantly limit the necessary increase in the social security contributions in the period between 2000 and 2050 (from 16% to 20.5% instead of 27%). For this reason, a second proposal for pension reform refers to an increase in the minimum retirement age (MRA) (for example Sayan and Kiraci (2001) and Gruber and Wise (2005)).

### 4.6.1 The Situation with Minimum Retirement Age

In our model, the minimum retirement age represents the lowest age of eligibility. Hence, workers are not allowed to choose a retirement age below the minimum retirement age $\bar{t}$. This lower bound $\bar{t}$ is identical for all workers and thus independent of their individual ability. In a nutshell, the workers' decision problem remains the same as with traditional PAYG but the minimum retirement age implies that the workers face an additional restriction. While the worker's retirement decision is not affected for low values of $\bar{t}$, the minimum retirement age is binding for $\bar{t} \geq (t_i)^{PAYG}$.

**Proposition 4.10** *With a minimum retirement age, the retirement date for both ability types depends on the level of $\bar{t}$:*

$$(t_i)^* = \begin{cases} (t_i)^{PAYG} & \text{if } \ \bar{t} < (t_i)^{PAYG} \\ \bar{t} & \text{if } \ (t_i)^{PAYG} \leq \bar{t} < (t_i)^F \\ (t_i)^F & \text{if } \ (t_i)^F \leq \bar{t} \end{cases} \tag{4.49}$$

In the following, we focus on the most interesting case where the minimum retirement age $\bar{t}$ lies between the firm's optimal separation dates for ability type L and H, i.e. $(t_L)^F < \bar{t} < (t_H)^F$. Hence, the situation for the two ability types is characterized by:

$$\begin{aligned} (t_L)^{PAYG} &< (t_L)^F < \bar{t} \\ (t_H)^{PAYG} &< \bar{t} < (t_H)^F \end{aligned}$$

In this case, the termination of production $(t_H)^*$ for ability type H is defined by the minimum retirement age:

$$(t_H)^* = \bar{t}$$

For the low-ability workers, the retirement date $(t_L)^*$ is determined by substituting the optimality function (4.39) into equation (4.21) and solving for $t$:

$$(t_L)^* = (t_L)^F = \frac{(1-\beta)(\theta_L)^2}{f - \delta(1-(1-\alpha)\tau)\beta(1-\beta)(\theta_L)^3}$$

This situation with $(t_L)^F < \bar{t} < (t_H)^F$ is graphically illustrated in Figure 4.8. According to (4.26), the worker's training intensity is equal to

$$\begin{aligned} (e_H)^* &= \delta(1-(1-\alpha)\tau)\beta\theta_i\bar{t} \\ (e_L)^* &= \delta(1-(1-\alpha)\tau)\beta\theta_L(t_L)^F \end{aligned}$$

### 4.6.2 The Effects of Increasing the Minimum Retirement Age

Increasing the minimum retirement age aims at augmenting the labor force participation at the extensive margin. However, the effective implications of

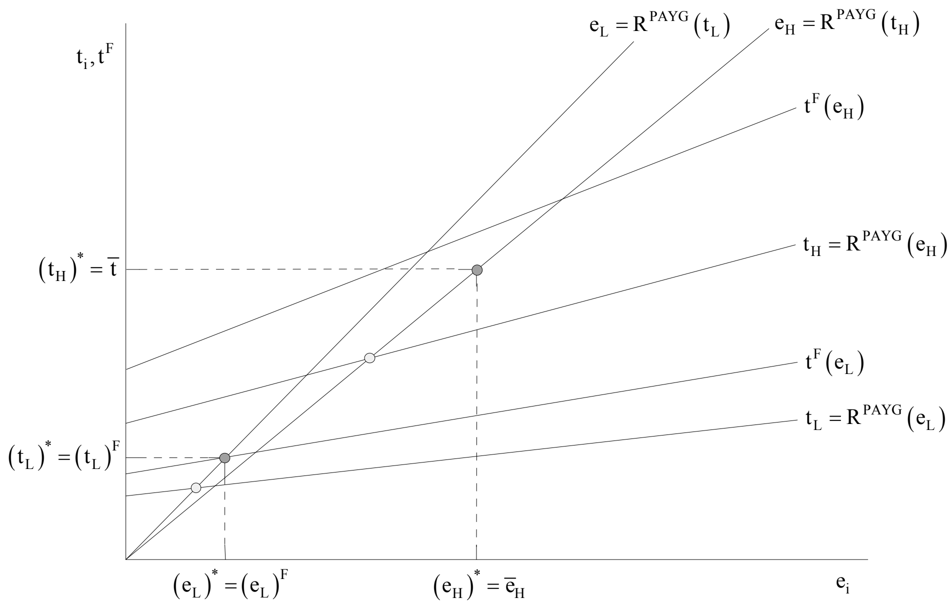


Figure 4.8: The Effects of Increasing the Minimum Retirement Age

an increase in $\bar{t}$ are obtained by analyzing the comparative statics of (4.49):

$$\frac{\partial (t_i)^*}{\partial \bar{t}} = \begin{cases} 0 & \text{if } \bar{t} < (t_i)^{PAYG} \\ 1 & \text{if } (t_i)^{PAYG} \leq \bar{t} < (t_i)^F \\ 0 & \text{if } \bar{t} \geq (t_i)^F \end{cases}$$

These comparative statics show that the retirement date $(t_i)^*$ is not affected if the minimum retirement age exceeds the optimal separation date of the firm. Hence, the effects of increasing the minimum retirement age are limited by the employment prospects of older workers.

**Proposition 4.11** *As illustrated in Figure 4.8, the situation* $(t_L)^F < \bar{t} < (t_H)^F$ *implies that only the retirement date of high-ability workers is increased by raising* $\bar{t}$*:*

$$\frac{\partial (t_H)^*}{\partial \bar{t}} = 1$$
$$\frac{\partial (t_L)^*}{\partial \bar{t}} = 0$$

In this case, high-ability workers are affected because they are forced to stay inside the labor market until they have reached the minimum retirement age. This is in line with the simulation results of Gruber and Wise (2002). By contrast, there is no impact on the retirement age of low-ability workers because it is the firm that decides to separate before the minimum retirement age is achieved.

Compared to the results of pension reform in Section 4.5, the effect of the firm's employment decision is even stronger because the increase in the actual retirement date due to the implementation of individual retirement accounts is only reduced but still existent, at least if human capital formation is endogenous. However, concentrating on the worker's retirement and training decision and neglecting the demand side of the labor market may again be misleading in evaluating the welfare gains of pension reforms.

## 4.7 Conclusion

The declining labor force participation in industrialized countries poses challenges to the pension system, which is essentially organized according to the PAYG principle in most countries. Chapter 4 of my PhD thesis presents a two-period partial-equilibrium model that systematically compares the implications different pension reforms for the retirement behavior at the extensive margin. Starting from the current system of traditional PAYG, two proposals for pension reforms are analyzed, which both aim at compensating the adverse effects of demographic change on the relative size of the active population.

Our formal analysis is based on recent literature on old-age provision, but the model manages to explain endogenously both the date of retirement and the formation of human capital. It is important to incorporate the training decision into the analysis of pension reforms because extending the working life increases the return to education and thus fosters the workers' incentives to acquire skills. Furthermore, the model incorporates worker heterogeneity in ability, which allows analyzing the implications of the pension system for different groups of workers. In line with Cremer, Lozachmeur, and Pestieau (2004) and Martín (2003), the desired retirement age positively depends on initial ability and the level of training.

Compared to the situation with laissez-faire, the traditional PAYG pension system with an imperfect tax-benefit link creates an implicit tax on labor supply at the extensive margin and training intensity at the intensive margin. The weaker the link between contributions and benefits, the larger the implicit tax rates and the more harmful are labor market distortions. In contrast, the capital funded system implies zero implicit tax rates because it constitutes a perfect substitute for private savings.

The first reform proposal suggests to introduce individual retirement accounts in order to (partly) move from the current PAYG towards a capital funded system of old-age provision. Indeed, increasing the size of the capital funded pillar of the pension system reduces the implicit tax rates at both margins. Hence, in conformity with Crémer and Pestieau (2003) and Lau and Poutvaara (2001a), this pension reform provides incentives for older workers to postpone retirement because capital funding rewards late retirement with an actuarially fair increase in subsequent pension benefits. A second benefit follows from the labor market and refers to human capital formation. In line with Trostel (1993), postponed retirement lengthens the time period over which individuals can appropriate the returns to education. In aggregate, there is a double benefit from the implementation of individual retirement accounts: reductions in labor market distortions at the extensive margin (retirement age) as well as increased human capital formation at the intensive margin (training intensity).

According to empirical estimates of Hernoes, Sollie, and Strom (2000), education is an important determinant of the retirement age. While early retirement is low among high-skilled workers, low-skilled workers tend to leave the labor force much earlier. One explanation could be that the retirement decision is essentially distorted for low-skilled workers (for example due to minimum pensions), which implies that the implementation of individual retirement accounts mainly affects the retirement decision of high-skilled workers (Martín (2003)).

However, we refer to another explanation by pointing out that the reduced impact of capital funding on low-ability workers stems from the demand side of the labor market. It is important to include the employment decision of the firms because the economic benefits of pension reforms strongly depend on the employment prospects of older workers near retirement age. In line with Lazear (1979), we find that the firm's optimal separation date increases with the worker's educational achievement. Hence, if the capital funded pillar of the pension system becomes strong and workers want to significantly postpone retirement, low-skilled workers may not attain their desired retirement age because firms refuse to employ them any longer. Depending on the degree of capital funding, the benefits of this pension reform mainly accrue for high-ability workers while the benefits for low-ability workers are reduced once their employment prospects near retirement age are controlled for.

The extent to which the retirement date responds to the implementation of individual retirement accounts strongly depends on the sensitivity of human capital formation. By completely focusing on retirement and taking training as exogenous, there is indeed no effect at all on the actual retirement date of low-ability workers. Therefore, moving to a capital funded system of old-age provision does not necessarily have positive effects on labor force participation. The superiority of capital funding as postulated by Feldstein (2005a) is lowered and may even vanish in the presence of labor market imperfections. As suggested by Lindbeck and Persson (2003), these qualitative conclusions also hold for a pension reform that aims at strengthening the tax-benefit link.

The second reform proposal suggests to increase the relative size of the active population by raising the minimum retirement age. Again, only high-ability workers may be affected because they are forced to stay inside the labor market until they have reached the minimum retirement age. In contrast, there may be no impact on low-ability workers if their retirement date is determined by the firms before the minimum retirement age is achieved. This is in line with simulation results of Gruber and Wise (2002).

In this model, retirement and human capital formation are endogenously determined and depend on the individual ability of workers. Nevertheless, the model has been kept simple for expositional and calculational reasons. The theoretical results of our stylized model only allow for qualitative conclusions concerning the evaluation of pension reforms. In order to assess the quantitative magnitude of these effects, we would have to estimate the elasticities of the labor market responses of both workers and firms. However, the underlying insights into the model presented here are robust to various types of generalization. Hence, they constitute a promising basis for policy recommendations and future research.

# Appendix A

# The Pivotal Productivities

With respect to the labor market decision of firms, the three pivotal productivities are the following (cf. Sections 2.3.4 and 2.4.1):

$$
\begin{aligned}
\theta^{LF} &= \frac{w_A}{\chi - (1-\beta) + \delta\,(1-\rho)\,(1-\beta)\,\alpha} \\
\theta_A^{PC} &= \frac{w_A - T}{\chi - (1-\beta) + \delta\,(1-\rho)\,(1-\beta)\,\alpha} \\
\theta_U^{PC} &= \frac{T}{1-\beta}
\end{aligned}
$$

Obviously, the implementation of penalty charges lowers the pivotal productivity between apprenticeship training and regular work:

$$\theta_A^{PC} \leq \theta^{LF}$$

Concerning the two pivotal productivities with penalty charges, the following relationship holds:

$$
\begin{array}{lcll}
\theta_A^{PC} & \geq & \theta_U^{PC} & \text{if } T \leq \bar{T} \\
\theta_A^{PC} & < & \theta_U^{PC} & \text{if } T > \bar{T}
\end{array}
$$

with

$$\bar{T} \equiv \frac{(1-\beta)\,w_A}{\chi + \delta\,(1-\rho)\,(1-\beta)\,\alpha}$$

Altogether, the three pivotal productivities show the following relationship:

$$\begin{array}{ll} \theta^{LF} \geq \theta_A^{PC} \geq \theta_U^{PC} & \text{if } 0 \leq T \leq \bar{T} \\ \theta^{LF} \geq \theta_U^{PC} > \theta_A^{PC} \;\; \textit{or} \;\; \theta_U^{PC} > \theta^{LF} > \theta_A^{PC} & \text{if } T > \bar{T} \end{array}$$

# Appendix B

# The Optimal Penalty Charges

## B.1 Calculation

The optimization problem is described in Section 2.4.2. Substituting the pivotal productivities $\theta_A^{PC}$ and $\theta_U^{PC}$ into the first-order condition (2.28) yields

$$\begin{aligned}\frac{\partial W^{PC}(T)}{\partial T} &= 0\\ (\chi - 1 + \delta\alpha)\,\theta_A^{PC}\left(-\frac{\partial \theta_A^{PC}}{\partial T}\right) &= (1+\delta\sigma)\,\theta_U^{PC}\frac{\partial \theta_U^{PC}}{\partial T} + cT\\ (\chi - 1 + \delta\alpha)\,\frac{w_A - T}{(v_1)^2} &= (1+\delta\sigma)\,\frac{T}{(1-\beta)^2} + cT\end{aligned}$$

with $v_1 \equiv \chi - (1-\beta) + \delta(1-\rho)(1-\beta)\,\alpha$. From this we obtain

$$\begin{aligned}\left(\frac{\chi - 1 + \delta\alpha}{(v_1)^2} + \frac{1+\delta\sigma}{(1-\beta)^2} + c\right)T &= \frac{\chi - 1 + \delta\alpha}{(v_1)^2}w_A\\ \frac{(1-\beta)^2(\chi - 1 + \delta\alpha) + (1+\delta\sigma)(v_1)^2 + (1-\beta)^2(v_1)^2 c}{(1-\beta)^2(v_1)^2}T &= \frac{\chi - 1 + \delta\alpha}{(v_1)^2}w_A\end{aligned}$$

Solving for $T$ yields the optimal penalty charges

$$T^* = \frac{(1-\beta)^2 (\chi - 1 + \delta\alpha)}{(1-\beta)^2 (\chi - 1 + \delta\alpha) + (1+\delta\sigma)(v_1)^2 + (1-\beta)^2 (v_1)^2 c} w_A$$

$$T^* = \frac{(1-\beta)^2 (\chi - 1 + \delta\alpha)}{(1-\beta)^2 (\chi - 1 + \delta\alpha) + (v_1)^2 v_2} w_A$$

with $v_2 \equiv 1 + (1-\beta)^2 c + \delta\sigma$.

## B.2 The Productivity-Enhancement

For the case $0 \leq T \leq \bar{T} \equiv \frac{(1-\beta) w_A}{\chi + \delta(1-\rho)(1-\beta)\alpha} = \frac{(1-\beta) w_A}{1-\beta+v_1}$, the optimal penalty charges are given by $T^*$ according to (2.29). Hence, $T^*$ has to satisfy the following two conditions in order to lie within the required interval:

$$T^* \geq 0 \quad \Leftrightarrow \quad \alpha \geq \frac{1-\chi}{\delta}$$

and

$$\begin{aligned} T^* &\leq \frac{(1-\beta) w_A}{1-\beta+v_1} \\ \frac{(1-\beta)(\chi-1+\delta\alpha)}{(1-\beta)^2(\chi-1+\delta\alpha)+(v_1)^2 v_2} &\leq \frac{1}{1-\beta+v_1} \\ (1-\beta)(\chi-1+\delta\alpha) &\leq v_1 v_2 \\ (1-\beta)(\chi-1+\delta\alpha) &\leq (\chi-(1-\beta)) v_2 + \delta(1-\rho)(1-\beta)\alpha v_2 \\ \delta(1-\beta)\alpha[1-(1-\rho)v_2] &\leq (1-\beta)(1-x) + (\chi-(1-\beta)) v_2 \quad \text{(B.1)} \end{aligned}$$

We assume that the administration costs exceed some lower bound

$$c \geq \bar{c} \equiv \frac{1-(1-\rho)(1+\delta\sigma)}{(1-\beta)^2(1-\rho)}$$

which implies that the inequality (B.1) is satisfied because $\alpha \geq 0$.

Altogether, the necessary condition for $0 \leq T^* \leq \bar{T}$ is the following:

$$\alpha \geq \frac{1-\chi}{\delta}$$

Note that the necessary condition for $T^* > \bar{T}$ is

$$c < \bar{c} \ \wedge \ \alpha > \frac{(1-\beta)(1-x) + (\chi - (1-\beta)) v_2}{\delta (1-\beta) [1 - (1-\rho) v_2]}$$

In this case, the optimal penalty charges would be equal to

$$T^* = \bar{T}$$

## B.3 Without the Participation Constraint

The participation constraint is not satisfied for high-ability workers if the productivity-enhancement of apprenticeship training is too low, i.e. if $\alpha < \frac{1}{\delta}$ (cf. Section 2.3.4). Concerning the optimal penalty charges, there are two cases that have to be considered:

(a) $\theta^W \geq \theta^{LF} \Leftrightarrow \alpha_0 \equiv \frac{1-\chi}{\delta(1-\rho(1-\beta))} \leq \alpha < \frac{1}{\delta}$

In case (a), the overall welfare is equal to

$$\begin{aligned} W^{(a)} &= \frac{1}{2}(\chi + \delta(1+\alpha)) - \frac{1}{2}(\chi - 1 + \delta\alpha)(\theta_A^{PC})^2 \\ &\quad - \frac{1}{2}(\chi - 1 + \delta\alpha)[1 - (\theta^W)^2] - \frac{1}{2}(1 + \delta\sigma)(\theta_U^{PC})^2 - \frac{c}{2}T^2 \end{aligned}$$

Because $\theta^W$ is independent of $T$, the optimal penalty charges are the same as in (2.29):

$$T^{(a)} = T^* = \frac{(1-\beta)^2 (\chi - 1 + \delta\alpha)}{(1-\beta)^2 (\chi - 1 + \delta\alpha) + (v_1)^2 v_2} w_A$$

Although workers with high ability refuse apprenticeship training, the overall welfare with $T^{(a)}$ is higher than in the laissez-faire equilibrium (cf. equation (2.21) for $\alpha_0 \leq \alpha < \frac{1}{\delta}$) because

$$\begin{aligned} (W^{(a)})^* &= \frac{1}{2}(1+\delta) + \frac{1}{2}(\chi - 1 + \delta\alpha)(\theta^W)^2 \\ &\quad - \frac{1}{8}\frac{\beta^2\chi^2(\chi - 1 + \delta\alpha) v_2}{(1-\beta)^2(\chi - 1 + \delta\alpha) + (v_1)^2 v_2} \\ &> \frac{1}{2}(1+\delta) + \frac{1}{2}(\chi - 1 + \delta\alpha)\left[(\theta^W)^2 - (\theta^{LF})^2\right] = W^{LF} \end{aligned}$$

if $\chi - 1 + \delta\alpha > 0 \;\Leftrightarrow\; \alpha > \frac{1-\chi}{\delta}$, which is implied by $\alpha_0 \geq \frac{1-\chi}{\delta}$.

(b) $\theta^W < \theta^{LF} \;\Leftrightarrow\; \alpha < \alpha_0$

In case (b), the overall welfare is equal to

$$W^{(b)} = \frac{1}{2}(1+\delta) - \frac{1}{2}(1+\delta\sigma)(\theta_U^{PC})^2 - \frac{c}{2}T^2$$

Because $W^{(b)}$ decreases with $T$, the optimal penalty charges are equal to zero (i.e. $T^{(b)} = 0$), which implies (cf. equation (2.21) for $\alpha < \alpha_0$)

$$(W^{(b)})^* = \frac{1}{2}(1+\delta) = W^{LF}$$

Altogether, the optimal penalty charges for $\alpha < \frac{1}{\delta}$ are equal to

$$T^{opt} = \begin{cases} 0 & \text{if} \quad \alpha < \alpha_0 \equiv \frac{1-\chi}{\delta(1-\rho(1-\beta))} \\ T^* & \text{if} \quad \alpha_0 \leq \alpha < \frac{1}{\delta} \end{cases}$$

## B.4 Comparative Statics

The comparative statics of $T^*$ with respect to productivity-enhancement, depreciation rate, separation probability, administration costs, and training wage are as follows:

$$\frac{\partial T^*}{\partial \alpha} = \frac{\delta(1-\beta)^2 v_1 v_2 \left[v_1 - 2(1-\beta)(1-\rho)(\chi - 1 + \delta\alpha)\right]}{\left[(1-\beta)^2(\chi - 1 + \delta\alpha) + (v_1)^2 v_2\right]^2} w_A > 0$$

$$\rightarrow \text{ if } \quad \alpha < \bar{\alpha} \equiv \frac{2(1-\chi)}{\delta} + \frac{\chi - (1-\beta)}{\delta(1-\rho)(1-\beta)}$$

$$\frac{\partial T^*}{\partial \sigma} = \frac{-\delta(1-\beta)^2(\chi - 1 + \delta\alpha)(v_1)^2}{\left[(1-\beta)^2(\chi - 1 + \delta\alpha) + (v_1)^2 v_2\right]^2} w_A < 0$$

$$\frac{\partial T^*}{\partial \rho} = \frac{2\delta(1-\beta)^3 \alpha(\chi - 1 + \delta\alpha) v_1 v_2}{\left[(1-\beta)^2(\chi - 1 + \delta\alpha) + (v_1)^2 v_2\right]^2} w_A > 0$$

$$\frac{\partial T^*}{\partial c} = \frac{-(1-\beta)^4(\chi - 1 + \delta\alpha)(v_1)^2}{\left[(1-\beta)^2(\chi - 1 + \delta\alpha) + (v_1)^2 v_2\right]^2} w_A < 0$$

$$\frac{\partial T^*}{\partial w_A} = \frac{(1-\beta)^2(\chi - 1 + \delta\alpha)}{(1-\beta)^2(\chi - 1 + \delta\alpha) + (v_1)^2 v_2} > 0$$

## B.5 The Overall Welfare

For $\alpha \geq \alpha_0$, the overall welfare with optimal penalty charges is determined by substituting the optimal penalty charges $T^*$ into equation (2.27):

$$\begin{aligned}
(W^{PC})^* &= \frac{1}{2}(\chi + \delta(1+\alpha)) - \frac{1}{2}(\chi - 1 + \delta\alpha)(\theta_A^{PC})^2 \\
&\quad - \frac{1}{2}(1+\delta\sigma)(\theta_U^{PC})^2 - \frac{c}{2}(T^*)^2 \\
&= \frac{1}{2}(\chi + \delta(1+\alpha)) - \frac{1}{2}(\chi - 1 + \delta\alpha)\left(\frac{w_A - T^*}{v_1}\right)^2 \\
&\quad - \frac{1}{2}(1+\delta\sigma)\left(\frac{T^*}{1-\beta}\right)^2 - \frac{c}{2}(T^*)^2 \\
&= \frac{1}{2}(\chi + \delta(1+\alpha)) - \frac{1}{2}\frac{(1-\beta)^2(\chi - 1 + \delta\alpha) + (v_1)^2 v_2}{(1-\beta)^2 (v_1)^2}(T^*)^2 \\
&\quad + \frac{\chi - 1 + \delta\alpha}{(v_1)^2} w_A T^* - \frac{1}{2}\frac{\chi - 1 + \delta\alpha}{(v_1)^2}(w_A)^2 \\
&= \frac{1}{2}(\chi + \delta(1+\alpha)) - \frac{1}{2}\frac{(1-\beta)^2(\chi - 1 + \delta\alpha) + (v_1)^2 v_2}{(1-\beta)^2 (v_1)^2} \\
&\quad \cdot \left(\frac{(1-\beta)^2(\chi - 1 + \delta\alpha)}{(1-\beta)^2(\chi - 1 + \delta\alpha) + (v_1)^2 v_2} w_A\right)^2 + \frac{\chi - 1 + \delta\alpha}{(v_1)^2} w_A \\
&\quad \cdot \left(\frac{(1-\beta)^2(\chi - 1 + \delta\alpha)}{(1-\beta)^2(\chi - 1 + \delta\alpha) + (v_1)^2 v_2} w_A\right) - \frac{1}{2}\frac{\chi - 1 + \delta\alpha}{(v_1)^2}(w_A)^2 \\
&= \frac{1}{2}(\chi + \delta(1+\alpha)) \\
&\quad + \frac{1}{2}\frac{(1-\beta)^2(\chi - 1 + \delta\alpha)^2}{(v_1)^2\left[(1-\beta)^2(\chi - 1 + \delta\alpha) + (v_1)^2 v_2\right]}(w_A)^2 \\
&\quad - \frac{1}{2}\frac{\chi - 1 + \delta\alpha}{(v_1)^2}(w_A)^2
\end{aligned}$$

Simplifying yields

$$(W^{PC})^* = \frac{1}{2}(\chi + \delta(1+\alpha)) - \frac{1}{2}\frac{(\chi - 1 + \delta\alpha) v_2}{(1-\beta)^2(\chi - 1 + \delta\alpha) + (v_1)^2 v_2}(w_A)^2 \quad \text{(B.2)}$$

## B.6 Comparative Statics of the Overall Welfare

For $\alpha \geq \alpha_0$, the comparative statics of the optimal overall welfare with respect to productivity-enhancement, depreciation rate, separating probability, administration costs, and training wage are as follows:

$$\begin{aligned}
\frac{\partial (W^{PC})^*}{\partial \alpha} &= \frac{1}{2}\delta \frac{(1-\beta)^4(\chi-1+\delta\alpha)^2+2(1-\beta)(\chi-1+\delta\alpha)v_1v_2\left[(1-\beta)v_1+(1-\rho)v_2(w_A)^2\right]}{\left[(1-\beta)^2(\chi-1+\delta\alpha)+(v_1)^2v_2\right]^2} \\
&\quad +\frac{1}{2}\delta\frac{(v_1)^2(v_2)^2\left[(v_1)^2-(w_A)^2\right]}{\left[(1-\beta)^2(\chi-1+\delta\alpha)+(v_1)^2v_2\right]^2} > 0 \\
\frac{\partial (W^{PC})^*}{\partial \sigma} &= -\frac{1}{2}\frac{\delta\left(1-\beta\right)^2\left(\chi-1+\delta\alpha\right)^2}{\left[(1-\beta)^2\left(\chi-1+\delta\alpha\right)+\left(v_1\right)^2 v_2\right]^2}(w_A)^2 < 0 \\
\frac{\partial (W^{PC})^*}{\partial \rho} &= -\frac{\delta\left(1-\beta\right)\left(\chi-1+\delta\alpha\right)\alpha v_1\left(v_2\right)^2}{\left[(1-\beta)^2\left(\chi-1+\delta\alpha\right)+\left(v_1\right)^2 v_2\right]^2}(w_A)^2 < 0 \\
\frac{\partial (W^{PC})^*}{\partial c} &= -\frac{1}{2}\frac{\left(1-\beta\right)^4\left(\chi-1+\delta\alpha\right)^2}{\left[(1-\beta)^2\left(\chi-1+\delta\alpha\right)+\left(v_1\right)^2 v_2\right]^2}(w_A)^2 < 0 \\
\frac{\partial (W^{PC})^*}{\partial w_A} &= -\frac{\left(\chi-1+\delta\alpha\right)v_2}{(1-\beta)^2\left(\chi-1+\delta\alpha\right)+\left(v_1\right)^2 v_2}w_A < 0
\end{aligned}$$

The comparative statics operate via two channels. First, there is a direct effect on the overall welfare which is identified by equation (2.27). Additionally, there is an indirect effect because the optimal penalty charges are altered according to Proposition 2.7. Altogether, the overall welfare with optimal penalty charges increases with $\alpha$ because the productivity of trained workers becomes larger. On the other hand, it decreases with $\sigma$ because previously unemployed workers have lower productivities in the second period. Furthermore, $(W^{PC})^*$ is negatively affected by the separating probability $\rho$ and the training wage $w_A$ because both higher fluctuation and higher wage costs of apprentices lead to increased inefficiencies in the provision of apprenticeship training positions (cf. Section 2.3.4). Finally, the overall welfare with optimal penalty charges decreases with the parameter $c$, which determines the administration costs of penalty charges.

# Appendix C

# The Pivotal Abilities

With respect to the labor market decision of workers, the four pivotal abilities are the following (cf. Sections 3.4.4, 3.4.5, and 3.5.1):

$$\begin{aligned}
\theta^{LF} &= \frac{e}{\delta\alpha} \\
\theta^{SA} &= z \\
\theta_A^{TC} &= \frac{e+\delta y}{\delta(\alpha+s)} \\
\theta_U^{TC} &= \frac{z-y}{1-s}
\end{aligned}$$

As discussed in Section 3.4.5, the assumption $z \leq \theta^{FB}$ implies

$$\begin{aligned}
\frac{e}{\delta\alpha} &\geq z \\
\theta^{LF} &\geq \theta^{SA} \qquad \text{(C.1)}
\end{aligned}$$

Furthermore, the assumption $\bar{\theta} \geq \theta^{LF}$ implies

$$\begin{aligned}
\frac{y}{s} &\geq \frac{e}{\delta\alpha} \\
\delta\alpha y &\geq se \\
\alpha(e+\delta y) &\geq (\alpha+s)e \\
\frac{e+\delta y}{\alpha+s} &\geq \frac{e}{\alpha} \\
\theta_A^{TC} &\geq \theta^{LF} \qquad \text{(C.2)}
\end{aligned}$$

and

$$
\begin{aligned}
\frac{y}{s} &\geq \frac{e}{\delta\alpha} \geq z \\
y &\geq sz \\
(1-s)\,z &\geq z-y \\
z &\geq \frac{z-y}{1-s} \\
\theta^{SA} &\geq \theta_U^{TC} \qquad \text{(C.3)}
\end{aligned}
$$

Taken together (C.1), (C.2), and (C.3), the four pivotal abilities show the following relationship:

$$\theta_A^{TC} \geq \theta^{LF} \geq \theta^{SA} \geq \theta_U^{TC}$$

# Appendix D

# The Optimal Basic Transfer

## D.1 Calculation

Substituting the pivotal abilities $\theta_U^{TC}$ and $\theta_A^{TC}$ into the first-order condition (3.22) implies

$$\begin{aligned}
W^{TC}(y,s) &= 0 \\
\delta\theta_U^{TC}\left(-\frac{\partial\theta_U^{TC}}{\partial y}\right) &= \delta\alpha\theta_A^{TC}\frac{\partial\theta_A^{TC}}{\partial y} - \frac{\partial\theta_A^{TC}}{\partial y}e \\
\delta\frac{z-y}{(1-s^*)^2} &= \alpha\frac{e+\delta y}{(\alpha+s^*)^2} - \frac{e}{\alpha+s^*} \\
\delta\left(\alpha+s^*\right)^2(z-y) &= (1-s^*)^2\,\alpha\,(e+\delta y) - (1-s^*)^2\,(\alpha+s^*)\,e \\
\delta[(\alpha+s^*)^2 + (1-s^*)^2\,\alpha]y &= \delta\left(\alpha+s^*\right)^2 z + (1-s^*)^2\,s^* e
\end{aligned}$$

Solving for $y$ yields the optimal basic transfer

$$y^* = \frac{(\alpha+s^*)^2}{(1+\alpha)\left(\alpha+(s^*)^2\right)}z + \frac{(1-s^*)^2\,s^*}{\delta\,(1+\alpha)\left(\alpha+(s^*)^2\right)}e$$

## D.2 Magnitude

The optimal basic transfer (3.24) is positive because

$$\begin{aligned} y^* &> 0 \\ \delta\left(\alpha+s^*\right)^2 z+\left(1-s^*\right)^2 s^* e &> 0 \end{aligned}$$

Furthermore, the optimal basic transfer falls below the level of social assistance if

$$\begin{aligned} y^* &< z \\ \frac{\delta\left(\alpha+s^*\right)^2 z+\left(1-s^*\right)^2 s^* e}{\delta[\left(\alpha+s^*\right)^2+\left(1-s^*\right)^2 \alpha]} &< z \\ \left(1-s^*\right)^2 s^* e &< \delta\left(1-s^*\right)^2 \alpha z \\ s^* e &< \delta \alpha z \end{aligned}$$

This inequality is satisfied if the following condition holds:

$$\alpha > \frac{e}{\delta z} s^*$$

# Appendix E

# Retirement and Training with LF

As shown in Section 4.3.2, the first-order conditions with laissez-faire are the following:

$$\delta\beta(1+e_i)\theta_i = \delta\gamma t_i \tag{E.1}$$

$$\delta t_i\beta\theta_i = e_i \qquad \Leftrightarrow \qquad t_i = \frac{e_i}{\delta\beta\theta_i} \tag{E.2}$$

Solving (E.2) for $t_i$ and substituting the result into (E.1) implies

$$\begin{aligned} \delta\beta(1+e_i)\theta_i &= \frac{\gamma e_i}{\beta\theta_i} \\ \delta(1+e_i)\left(\beta\theta_i\right)^2 &= \gamma e_i \\ \delta\left(\beta\theta_i\right)^2 &= \left[\gamma - \delta\left(\beta\theta_i\right)^2\right] e_i \end{aligned}$$

Solving for $e_i$ yields

$$(e_i)^{LF} = \frac{\delta\left(\beta\theta_i\right)^2}{\gamma - \delta\left(\beta\theta_i\right)^2}$$

By substituting $(e_i)^{LF}$ into equation (E.2) we obtain

$$(t_i)^{LF} = \frac{\beta\theta_i}{\gamma - \delta\left(\beta\theta_i\right)^2}$$

# Appendix F

# Tax-Deductibility of Training Costs

On the one hand, the tax-deductibility of training costs implies that the worker's education costs in period 1 are reduced compared to equation (4.2):

$$u_{1i} = (1-\tau)\,\beta\theta_i - s_i - \frac{1}{2}(1-\tau)\,(e_i)^2$$

On the other hand, the worker's benefits during the retirement subperiod are lowered because the level of benefits is linked to one's own contributions in the past:

$$b' = k + \frac{\alpha}{h-t_i}\left[\tau\beta\theta_i + t_i\tau\beta\,(1+e_i)\,\theta_i - \frac{1}{2}\tau(e_i)^2\right]$$

By proceeding in the same way as in Section 4.4.2, we obtain the following first-order conditions:

$$\begin{aligned}(1-(1-\alpha)\,\tau)\,\beta\theta_i\,(1+e_i) &= \gamma t_i + k \\ \delta\,(1-(1-\alpha)\,\tau)\,\beta\theta_i t_i &= (1-(1-\delta\alpha)\,\tau)\,e_i\end{aligned}$$

Combining the first-order conditions yields

$$(t_i)' = \frac{(1-(1-\alpha)\,\tau)\,\beta\theta_i - k}{(1-(1-\delta\alpha)\,\tau)\,\gamma - \delta\left[(1-(1-\alpha)\,\tau)\,\beta\theta_i\right]^2} \quad \text{(F.1)}$$

$$(e_i)' = \frac{\delta\,(1-(1-\alpha)\,\tau)}{1-(1-\delta\alpha)\,\tau}\beta\theta_i\,(t_i)' \quad \text{(F.2)}$$

Hence, the implicit tax rates on labor supply at the extensive margin and training intensity at the intensive margin are the following:

$$(\tau_t^*)' = (1-\alpha)\tau + \frac{k}{\beta\left(1+(e_i)'\right)\theta_i} \tag{F.3}$$

$$(\tau_e^*)' = \tau\left[1-\alpha\left(1-\frac{(e_i)'}{(t_i)'\beta\theta_i}\right)\right] \tag{F.4}$$

Both values $(t_i)'$ and $(e_i)'$ increase with the strength of the tax-benefit link if the implicit tax rates decrease with $\alpha$. The negative relation between $(\tau_t^*)'$, $(\tau_e^*)'$, and $\alpha$ is proved in the following:

$$\begin{aligned} \delta &\leq 1 \\ \delta\left(1-(1-\alpha)\tau\right) &\leq 1-(1-\delta\alpha)\tau \\ \frac{\delta\left(1-(1-\alpha)\tau\right)}{1-(1-\delta\alpha)\tau}\beta\theta_i(t_i)' &\leq \beta\theta_i(t_i)' \end{aligned}$$

Substituting in the definition of $(e_i)'$ in (F.2) yields

$$\begin{aligned} (e_i)' &\leq \beta\theta_i(t_i)' \\ \tau\left[-1+\frac{(e_i)'}{\beta\theta_i(t_i)'}\right] &\leq 0 \\ \frac{\partial(\tau_e^*)'}{\partial\alpha} &\leq 0 \end{aligned}$$

Because the implicit tax $(\tau_e^*)'$ in (F.4) is reduced by an increase in $\alpha$, the desired training intensity $(e_i)'$ in (F.2) goes up. Hence, also the implicit tax rate $(\tau_t^*)'$ in (F.3) is reduced because it negatively depends on the level of training. As a consequence, the desired retirement age $(t_i)'$ in (F.1) is increased.

# Appendix G

# The Retirement Date

## G.1 Traditional PAYG

It is the worker (and not the firm) who defines the end of production in period 2 if

$$\begin{aligned} (t_i)^{PAYG} &\leq (t_i)^F \\ \frac{(1-(1-\alpha)\tau)\beta\theta_i - k}{\gamma - \delta\left[(1-(1-\alpha)\tau)\beta\theta_i\right]^2} &\leq \frac{1-\beta}{f}(1+e_i)(\theta_i)^2 \\ (1-(1-\alpha)\tau) - \frac{1-\beta}{\beta}\frac{\gamma}{f}\theta_i &\leq \frac{k}{\beta(1+e_i)\theta_i} \end{aligned} \tag{G.1}$$

Subject to the assumption

$$(1-\tau)\frac{\beta}{1-\beta}\frac{f}{\gamma} < \theta_L < \frac{\beta}{1-\beta}\frac{f}{\gamma} < \theta_H \tag{G.2}$$

the retirement date of ability type H is determined by the worker. This is proved by the following lines:

$$\begin{aligned}
\frac{\beta}{1-\beta}\frac{f}{\gamma} &< \theta_H \\
(1-(1-\alpha)\tau)\frac{\beta}{1-\beta}\frac{f}{\gamma} &< \theta_H \\
(1-(1-\alpha)\tau) &< \frac{1-\beta}{\beta}\frac{\gamma}{f}\theta_H \\
(1-(1-\alpha)\tau)-\frac{1-\beta}{\beta}\frac{\gamma}{f}\theta_H &< 0 \\
(1-(1-\alpha)\tau)-\frac{1-\beta}{\beta}\frac{\gamma}{f}\theta_H &< \frac{k}{\beta(1+e_H)\theta_H} \\
(t_H)^{PAYG} &< (t_H)^F
\end{aligned}$$

Whether the retirement date $(t_L)^*$ of ability type L is determined by the worker or by the firm depends on the strength of the tax-benefit link. For example, the condition $(t_L)^{PAYG} \leq (t_L)^F$ is satisfied if there is no tax-benefit link (i.e. for $\alpha = 0$):

$$\begin{aligned}
(1-\tau)\frac{\beta}{1-\beta}\frac{f}{\gamma} &< \theta_L \\
(1-\tau) &< \frac{1-\beta}{\beta}\frac{\gamma}{f}\theta_L \\
(1-\tau)-\frac{1-\beta}{\beta}\frac{\gamma}{f}\theta_L &< 0 \\
(1-\tau)-\frac{1-\beta}{\beta}\frac{\gamma}{f}\theta_L &< \frac{k}{\beta(1+e_L)\theta_L} \\
\left.(t_L)^{PAYG}\right|_{\alpha=0} &< \left.(t_L)^F\right|_{\alpha=0}
\end{aligned}$$

However, the retirement date is determined by the firm if there is a perfect link between contributions and benefits (i.e. for $\alpha = 1$), at least for small values of $k$:

$$
\begin{aligned}
\theta_L &< \frac{\beta}{1-\beta}\frac{f}{\gamma} \\
0 &< 1 - \frac{1-\beta}{\beta}\frac{\gamma}{f}\theta_L \\
\left.(t_L)^F\right|_{\alpha=1,k=0} &< \left.(t_L)^{PAYG}\right|_{\alpha=1,k=0}
\end{aligned}
$$

Altogether, if the tax-benefit link is weak (i.e. $\alpha$ is small), condition (G.1) is also satisfied for ability type L. However, if the tax-benefit link is strong (i.e. $\alpha$ is large), it may be the firm that determines the termination of production in period 2 because the optimality function $t_L = R^{PAYG}(e_L)$ moves upward such that the firm's optimal separation date $(t_L)^F$ falls below the worker's desired retirement age $(t_L)^{PAYG}$.

## G.2 Individual Retirement Accounts

It is the worker (and not the firm) who defines the end of production in period 2 if

$$
\begin{aligned}
(t_i)^{IRA} &\leq (t_i)^F \\
\xi(\omega)\frac{\beta\theta_i}{\gamma}(1+e_i) - (1-\omega)\frac{k}{\gamma} &\leq \frac{1-\beta}{f}(1+e_i)(\theta_i)^2 \\
\xi(\omega) - \frac{1-\beta}{\beta}\frac{\gamma}{f}\theta_i &\leq \frac{(1-\omega)k}{\beta(1+e_i)\theta_i}
\end{aligned}
\tag{G.3}
$$

Subject to assumption (G.2), the retirement date of ability type H is determined by the worker. The proof is the same as in Appendix G.1. Whether the retirement date $(t_L)^*$ of ability type L is determined by the worker or by the firm depends on the degree of capital funding. The case $\omega = 0$ corresponds to the situation with traditional PAYG and thus implies $\left.(t_L)^{IRA}\right|_{\omega=0} < \left.(t_L)^F\right|_{\omega=0}$, at least for small values of $\alpha$ (cf. Appendix G.1).

However, the opposite is true for $\omega = 1$:

$$\begin{aligned} \theta_L &< \frac{\beta}{1-\beta}\frac{f}{\gamma} \\ 0 &< 1 - \frac{1-\beta}{\beta}\frac{\gamma}{f}\theta_L \\ (t_L)^F\Big|_{\omega=1} &< (t_L)^{IRA}\Big|_{\omega=1} \end{aligned}$$

Taken together, there must be a critical degree of capital funding $\bar{\omega}$ such that for $\omega > \bar{\omega}$ the firm's optimal separation date $(t_L)^F$ falls below the worker's desired retirement age $(t_L)^{IRA}$:

$$(t_L)^* = \begin{cases} (t_L)^{IRA} & \text{if } \omega \leq \bar{\omega} \\ (t_L)^F & \text{if } \omega > \bar{\omega} \end{cases}$$

# Appendix H

# Comparative Statics with IRA

For $\omega > \bar{\omega}$, $(t_L)^{IRA}$ increases with $\omega$ more rapidly than $(t_L)^F$:

$$\frac{\partial (t_L)^F}{\partial \omega} < \frac{\partial (t_L)^{IRA}}{\partial \omega}$$

$$\frac{\delta \tau (1-\alpha) \beta (1-\beta)^2 (\theta_L)^5}{\left[f - \delta \xi (\omega) \beta (1-\beta) (\theta_L)^3\right]^2} < \frac{\tau (1-\mu) \beta \theta_L \left[\gamma + \delta \left(\xi (\omega) \beta \theta_L\right)^2\right]}{\left[\gamma - \delta \left(\xi (\omega) \beta \theta_L\right)^2\right]^2}$$

By applying the definitions of $(t_L)^{IRA}$ and $(t_L)^F$ in (4.46), this inequality simplifies to

$$\delta \left[(t_L)^F\right]^2 < \delta \left[(t_L)^{IRA}\right]^2 + \left[\gamma - \delta \left(\xi (\omega) \beta \theta_L\right)^2\right]^2 \gamma$$

This condition is unambiguously satisfied because $(t_L)^F < (t_L)^{IRA}$ for $\omega > \bar{\omega}$.

# Bibliography


Aaronson, D. (1998): "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes," *The Journal of Human Resources*, 33(4), 915–946.

Acemoglu, D. (1997): "Training and Innovation in an Imperfect Labour Market," *The Review of Economic Studies*, 64(3), 445–464.

——— (2002): "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, 40(1), 7–72.

Acemoglu, D., and J.-S. Pischke (1998a): "Beyond Becker: Training in Imperfect Labor Markets," *NBER Working Paper*, 6740.

——— (1998b): "Why Do Firms Train? Theory and Evidence," *The Quarterly Journal of Economics*, 113(1), 79–119.

——— (1999a): "Minimum Wages and On-the-Job Training," *NBER Working Paper*, 7184.

——— (1999b): "The Structure of Wages and Investment in General Training," *The Journal of Political Economy*, 107(3), 539–572.

Acemoglu, D., and R. Shimer (1999): "Holdups and Efficiency with Search Frictions," *International Economic Review*, 40(4), 827–849.

Akyol, A., and K. Athreya (2005): "Risky Higher Education and Subsidies," *Journal of Economic Dynamics and Control*, 29(6), 979–1023.

ALTONJI, J. G., AND T. A. DUNN (1996): "The Effects of Family Characteristics on the Return to Education," *The Review of Economics and Statistics*, 78(4), 692–704.

ANDERBERG, D., AND F. ANDERSSON (2003): "Investments in Human Capital, Wage Uncertainty, and Public Policy," *Journal of Public Economics*, 87(7-8), 1521–1537.

ANDERSSON, F., AND K. A. KONRAD (2001): "Globalization and Human Capital Formation," *IZA Discussion Paper*, 245.

——— (2003): "Human Capital Investment and Globalization in Extortionary States," *Journal of Public Economics*, 87(7-8), 1539–1555.

ANGRIST, J. D., AND A. B. KRUEGER (1991): "Does Compulsory School Attendance Affect Schooling and Earnings?," *The Quarterly Journal of Economics*, 106(4), 979–1014.

ARCIDIACONO, P. (2004): "Ability Sorting and the Returns to College Major," *Journal of Econometrics*, 121(1-2), 343–375.

ARCIDIACONO, P., AND S. NICHOLSON (2005): "Peer Effects in Medical School," *Journal of Public Economics*, 89(2), 327–350.

ARIAS, O., AND W. W. MCMAHON (2001): "Dynamic Rates of Return to Education in the U.S.," *Economics of Education Review*, 20(1), 121–138.

AROZAMENA, L., AND M. CENTENO (2006): "Tenure, Business Cycle and the Wage-Setting Process," *European Economic Review*, 50(2), 401–424.

ARUM, R. (1996): "Do Private Schools Force Public Schools to Compete?," *American Sociological Review*, 61, 29–46.

ASHENFELTER, O., C. HARMON, AND H. OOSTERBEEK (1999): "A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias," *Labour Economics*, 6(4), 453–470.

Ashenfelter, O., and A. Krueger (1994): "Estimates of the Economic Return to Schooling from a New Sample of Twins," *The American Economic Review*, 84(5), 1157–1173.

Astone, N. M., and S. S. McLanahan (1991): "Family Structure, Parental Practices and High School Completion," *American Sociological Review*, 56(3), 309–320.

Atkinson, A. B. (1973): *How Progressive Should Income Tax Be?* pp. 90–109. Longman.

Autor, D. H., L. F. Katz, and M. S. Kearney (2005): "Rising Wage Inequality: The Role of Composition and Prices," *NBER Working Paper*, 11628.

Ballou, D. (1999): "Do Public Schools Hire the Best Teachers?," *Quarterly Journal of Economics*, 111(1), 97–133.

Ballou, D., and M. Podgursky (1998): "Teacher Recruitment and Retention in Public and Private Schools," *Journal of Policy Analysis and Management*, 17(3), 393–418.

Balmaceda, F. (2001): "Firm-Sponsored General Training in a Frictionless Labor Market," *ILADES-Georgetown University Working Paper*, 134.

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

Bartel, A. P., and G. J. Borjas (1977): "Specific Training and Its Effects on the Human Capital Investment Profile," *Southern Economic Journal*, 44(2), 333–341.

Bartel, A. P., and N. Sicherman (1993): "Technological Change and Retirement Decisions of Older Workers," *Journal of Labor Economics*, 11(1), 162–183.

Bauer, T. K. (2002): "Educational Mismatch and Wages: A Panel Analysis," *Economics of Education Review*, 21(3), 221–229.

Becker, G. S. (1962): "Investment in Human Capital: A Theoretical Analysis," *The Journal of Political Economy*, 70(5), 9–49.

——— (1964): *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. Columbia University Press, New York.

——— (1992): "Human Capital and the Economy," *Proceedings of the American Philosophical Society*, 136(1), 85–92.

Bedard, K., and C. Ferrall (2003): "Wage and Test Score Dispersion: Some International Evidence," *Economics of Education Review*, 22(1), 31–43.

Belot, M., J. Boone, and J. V. Ours (2002): "Welfare Effects of Employment Protection," *CEPR Discussion Paper*, 3396.

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

Bénabou, R. (2002): "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 70(2), 481–517.

Bertola, G. (2001): "Aggregate and Disaggregated Aspects of Employment and Unemployment," *European University*.

Bingley, P., and G. Lanot (2004): "Employer Pay Policies, Public Transfers and the Retirement Decisions of Men and Women in Denmark," *European Economic Review*, 48, 181–200.

Bishop, J. H. (1997): "What We Know About Employer-Provided Training: A Review of the Literature," *Research in Labor Economics*, 16, 19–87.

Bishop, J. H., and L. Woessmann (2002): "Institutional Effects in a Simple Model of Educational Production," *IZA Discussion Paper*, 484.

Bittner, S., A. Hollederer, B. Kaltenborn, H. Rudolph, A. Vanselow, and C. Weinkopf (2001): *Ein Jahr Erfahrungen mit Dem Arbeitsmarktpolitischen Sonderprogramm CAST* vol. 290. Bundesministerium für Arbeit und Sozialordnung, Bonn.

Blackburn, M. L. (2004): "The Role of Test Scores in Explaining Race and Gender Differences in Wages," *Economics of Education Review*, 23(6), 555–576.

Blau, D. M. (1994): "Labor Force Dynamics of Older Men," *Econometrica*, 62(1), 117–156.

Blau, F. D. (1998): "Trends in the Well-Being of American Women, 1970-1995," *Journal of Economic Literature*, 36, 112–165.

Blinder, A., R. Gordon, and D. Wise (1981): *Life Cycle Savings and Bequest: Cross-Sectional Estimates of the Life Cycle Model* Macmillan, London and New York.

Blinder, A. S., and Y. Weiss (1976): "Human Capital and Labor Supply: A Synthesis," *The Journal of Political Economy*, 84(3), 449–472.

Bloch, F. E., and S. P. Smith (1977): "Human Capital and Labor Market Employment," *The Journal of Human Resources*, 12(4), 550–560.

——— (1979): "Human Capital and Labor Market Employment: Errata and Extension," *The Journal of Human Resources*, 14(2), 267–269.

Blundell, R. (2005): "Earned Income Tax Credit Policies: Impact and Optimality," *The Adam Smith Lecture*.

Blundell, R., L. Dearden, A. Goodman, and H. Reed (1997): *Higher Education, Employment and Earnings in Britain.* Institute for Fiscal Studies, London.

BLUNDELL, R., L. DEARDEN, C. MEGHIR, AND B. SIANESI (1999): "Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy," *Fiscal Studies*, 20(1), 1–23.

BLUNDELL, R., AND H. HOYNES (2001): "Has "In-Work" Benefit Reform Helped the Labour Market?," *NBER Working Paper*, 8546.

BLUNDELL, R. W., AND T. MACURDY (1999): *Labor Supply:: A Review of Alternative Approaches*vol. 3A. Elsevier Science B.V., Amsterdam.

BMBF (2004): *Berufsbildungsbericht*. Bundesministerium für Bildung und Forschung, Bonn.

BOADWAY, R., N. MARCEAU, AND M. MARCHAND (1996): "Investment in Education and the Time Inconsistency of Redistributive Tax Policy," *Economica*, 63(250), 171–189.

BOLDRIN, M., J. J. DOLADO, J. F. JIMENO, F. PERACCHI, F. BREYER, AND R. FERNANDEZ (1999): "The Future of Pensions in Europe," *Economic Policy*, 14(29), 287–320.

BOONE, J., AND L. BOVENBERG (2004): "The Optimal Taxation of Unskilled Labor with Job Search and Social Assistance," *Journal of Public Economics*, 88, 2227–2258.

——— (2006): "Optimal Welfare and in-Work Benefits with Search Unemployment and Observable Abilities," *Journal of Economic Theory*, 126, 165–193.

BOONE, J., AND R. A. DE MOOIJ (2003): "Tax Policy in a Matching Model with Training," *Oxford Economic Papers*, 55(1), 121–147.

BOOTH, A. L., AND S. E. SATCHELL (1994): "Apprenticeships and Job Tenure," *Oxford Economic Papers*, 46(4), 676–695.

Börsch-Supan, A. (2000): "Incentive Effects of Social Security on Labor Force Participation: Evidence in Germany and Across Europe," *Journal of Public Economics*, 78, 25–49.

——— (2003): "Labor Market Effects of Population Aging," *Review of Labor Economics and Industrial Relations*, 17, 5–44.

Börsch-Supan, A., I. Duzgun, and M. Weiss (2005): "Altern und Produktivität: Zum Stand der Forschung," *MEA Discussion Paper*, 05073.

Boskin, M., L. Kotlikoff, D. Puffert, and J. Shoven (1986): "Social Security: A Financial Appraisal Across and Within Generations," *NBER Working Paper*, 1891.

Bottazzi, R., T. Jappelli, and M. Padula (2006): "Retirement Expectations, Pension Reforms, and their Impact on Private Wealth Accumulation," *Journal of Public Economics*, 90(12), 2187–2212.

Boucekkine, R., D. de la Croix, and O. Licandro (2002): "Vintage Human Capital, Demographic Trends, and Endogenous Growth," *Journal of Economic Theory*, 104, 340–375.

Bougheas, S., and Y. Georgellis (2004): "Early Career Mobility and Earning Profiles of German Apprentices: Theory and Empirical Evidence," *Labour*, 18(2), 233–263.

Bovenberg, A. L., and B. Jacobs (2005): "Redistribution and Education Subsidies are Siamese Twins," *Journal of Public Economics*, 89(11), 2005–2035.

Bovenberg, A. L., and T. Knaap (2005): "Ageing, Funded Pensions and the Dutch Economy," *CESifo Working Paper*, 1403.

Bovenberg, A. L., and P. B. Sorensen (2004): "Improving the Equity-Efficiency Trade-Off: Mandatory Savings Accounts for Social Insurance," *International Tax and Public Finance*, 11, 507–529.

Bowles, S., H. Gintis, and M. Osborne (2001): "The Determinants of Earnings: A Behavioral Approach," *Journal of Economic Literature*, 39(4), 1137–1176.

Brooks-Gunn, J., G. J. Duncan, P. K. Klebanov, and N. Sealand (1993): "Do Neighborhoods Influence Child and Adolescent Development?," *American Journal of Sociology*, 99(2), 353–395.

Brown, C. (1976): "A Model of Optimal Human-Capital Accumulation and the Wages of Young High School Graduates," *The Journal of Political Economy*, 84(2), 299–316.

Brunello, G. (2002): "Is Training More Frequent When Wage Compression is Higher? Evidence from 11 European Countries," *CESifo Working Paper*, 637.

Brunello, G., and D. Checchi (2005): "School Quality and Family Background in Italy," *Economics of Education Review*, 24, 563–577.

Burtless, G. (1986): "Social Security, Unanticipated Benefit Increases, and the Timing of Retirement," *The Review of Economic Studies*, 53(5), 781–805.

Burtless, G., and R. A. Moffitt (1985): "The Joint Choice of Retirement Age and Postretirement Hours of Work," *Journal of Labor Economics*, 3(3), 209–236.

Butcher, K. F., and A. Case (1994): "The Effect of Sibling Composition on Women's Education and Earnings," *The Quarterly Journal of Economics*, 109(3), 531–563.

Cahuc, P., and A. Zylberberg (2004): *Labor Economics*. MIT Press, Cambridge.

Card, D., and A. B. Krueger (1992): “Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States,” *The Journal of Political Economy*, 100(1), 1–40.

Caucutt, E. (2001): “Peer Group Effects in Applied General Equilibrium,” *Economic Theory*, 17(1), 25–51.

Caucutt, E. M. (2002): “Educational Vouchers When There are Peer Group Effects - Size Matters,” *International Economic Review*, 43(1), 195–222.

Chan, S., and A. H. Stevens (2001): “Job Loss and Employment Patters of Older Workers,” *Journal of Labor Economics*, 19(2), 484–521.

Chang, C., and Y. Wang (1996): “Human Capital Investment under Asymmetric Information: The Pigovian Conjecture Revisited,” *Journal of Labor Economics*, 14(3), 505–519.

Charlot, O., B. Decreuse, and P. Granier (2005): “Adaptability, Productivity, and Educational Incentives in a Matching Model,” *European Economic Review*, 49(4), 1007–1032.

Cingano, F. (2003): “Returns to Specific Skills in Industrial Districts,” *Labour Economics*, 10(2), 149–164.

Cogan, J. F. (1981): “Fixed Costs and Labor Supply,” *Econometrica*, 49, 945–963.

Coile, C. C., and P. B. Levine (2006): “Labor Market Shocks and Retirement: Do Government Programs Matter?,” *NBER Working Paper*, W12559.

Coleman, J. S. (1966): *Equality of Educational Opportunity*. U.S. Government Printing Office, Washington DC.

Collins, K. A., and J. B. Davies (2003): "Measuring Effective Tax Rates on Human Capital: Methodology and an Application to Canada," *CESifoWorking Paper*, 965.

Conde-Ruiz, J. I., and V. Galasso (2004): "The Macroeconomics of Early Retirement," *Journal of Public Economics*, 88(9-10), 1849–1869.

CongressionalBudgetOffice (1993): *Displaced Workers: Trends in the 1980s and Implications for the Future.* Government Printing Office, Washington, DC.

Cooke, L. P. (2003): "A Comparison of Initial and Early Life Course Earnings of the German Secondary Education and Training System," *Economics of Education Review*, 22, 79–88.

Corcoran, M., R. Gordon, D. Laren, and G. Solon (1992): "The Association Between Men's Economic Status and Their Family and Community Origins," *The Journal of Human Resources*, 27(4), 53–79.

Corneo, G., and M. Marquardt (2000): "Public Pensions, Unemployment Insurance, and Growth," *Journal of Public Economics*, 75, 293–311.

Cremer, H., J.-M. Lozachmeur, and P. Pestieau (2004): "Social Security, Retirement Age and Optimal Income Taxation," *Journal of Public Economics*, 88, 2259–2281.

Cremer, H., and P. Pestieau (2000): "Reforming our Pension System: Is It a Demographic, Financial or Political Problem?," *European Economic Review*, 44, 974–983.

Crémer, H., and P. Pestieau (2003): "The Double Dividend of Postponing Retirement," *International Tax and Public Finance*, 10, 419–434.

Crépon, B., N. Deniau, and S. Pérez-Duarte (2002): "Wages, Productivity, and Workers Characteristics: A French Perspective," *mimeo.*

Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov (2005): *Interpreting the Evidence on Life Cycle Skill Formation*Handbook of the Economics of Education. North Holland, Amsterdam.

Cutler, D. M., J. M. Poterba, L. M. Sheiner, and L. H. Summers (1990): "An Ageing Society: Opportunity or Challenge?," *Brookings Papers on Economic Activity*, 1, 1–73.

Daniel, K., and J. S. Heywood (2007): "The Determinants of Hiring Older Workers: UK Evidence," *Labour Economics*, 14(1), 35–51.

Dann, S., A. Kirchmann, A. Spermann, and J. Volkert (2002): *Einstiegsgeld in Baden-Württemberg*. Sozialministerium Baden-Württemberg, Tübingen.

Datcher, L. P. (1982): "Effects of Community and Family Background on Achievement," *The Review of Economics and Statistics*, 64(1), 32–41.

de la Croix, D., and O. Licandro (1999a): "Life Expectancy and Endogenous Growth," *Economic Letters*, 65, 255–263.

——— (1999b): "Life Expectancy and Endogenous Growth," *Economic Letters*, 65, 255–263.

Dearden, L. (1998): "Ability, Families, Education and Earnings in Britain," *Institute for Fiscal Studies*, 98/14.

DeCicca, P. (2007): "Does Full-Day Kindergarten Matter? Evidence from the First Two Years of Schooling," *Economics of Education Review*, 26(1), 67–82.

DeMeza, D., and D. C. Webb (2001): "Advantageous Selection in Insurance Markets," *RAND Journal of Economics*, 32(2), 249–262.

Demmel, R., and C. Keuschnigg (2000): "Funded Pensions and Unemployment," *Finanzarchiv*, 57, 22–38.

Devereaux, P. J. (2001): "The Cyclicality of Real Wages Within Employer-Employee Matches," *Industrial and Labor Relations Review*, 54(4), 835–850.

Diamond, P., and J. Gruber (1997): "Social Security and Retirement in the US," *NBER Working Paper*, 6097.

Diamond, P. A., and J. A. Hausman (1984): *The Retirement and Unemployment Behavior of Older Men*pp. 97–132. Brookings Institution, Washington, DC.

Diamond, P. A., and P. R. Orszag (2005): "Saving Social Security," *Journal of Economic Perspectives*, 19(2), 11–32.

Disney, R. (2000): "Crises in Public Pension Programmes in OECD: What are the Reform Options?," *The Economic Journal*, 110(461), F1–F23.

Donaldson, D., and B. C. Eaton (1976): "Firm-Specific Human Capital: A Shared Investment or Optimal Entrapment?," *The Canadian Journal of Economics*, 9(3), 462–472.

——— (1977): "Specific Human Capital and Shared Investments Again," *The Canadian Journal of Economics*, 10(3), 474–475.

Driffill, E. J., and H. S. Rosen (1983): "Taxation and Excess Burden: A Life Cycle Perspective," *International Economic Review*, 24(3), 671–683.

Duncan, G. J., and S. D. Hoffman (1981): "The Incidence and Wage Effects of Overeducation," *Economics of Education Review*, 1(1), 75–86.

Dur, R., and C. N. Teulings (2003): "Are Education Subsidies an Efficient Redistributive Device," *CEE Discussion Paper*, 0030.

Dustmann, C., N. Rajah, and A. V. Soest (2003): "Class Size, Education, and Wages," *The Economic Journal*, 113(485), 99–120.

EATON, J., AND H. S. ROSEN (1980): "Taxation, Human Capital, and Uncertainty," *The American Economic Review*, 70(4), 705–715.

ECHEVARRÍA, C. A. (2003): "Life Expectancy, Retirement and Endogenous Growth," *Economic Modelling*, 21, 147–174.

ECHEVARRÍA, C. A. (2004): "Life Expectancy, Schooling Time, Retirement, and Growth," *Economic Inquiry*, 42, 602–617.

ECHEVARRÍA, C. A., AND A. IZA (2005): "Life Expectancy, Human Capital, Social Security and Growth," *DFAEII Working Paper*, 200517.

ECKAUS, R. S. (1963): "Investment in Human Capital: Comment," *The Journal of Political Economy*, 71(5), 501–504.

EISSA, N., AND H. HOYNES (2005): "Behavioral Responses to Taxes: Lessons from the EITC and Labor Supply," *NBER Working Paper*, 11729.

EISSA, N., AND H. W. HOYNES (2004): "Taxes and the Labor Market Participation of Married Couples: The Earned Income Tax Credit," *Journal of Public Economics*, 88, 1931–1958.

EISSA, N., H. J. KLEVEN, AND C. T. KREINER (2004): "Evaluation of Four Tax Reforms in the United States: Labor Supply and Welfare Effects for Single Mothers," *NBER Working Paper*, 10935.

——— (2006): *Welfare Efects of Tax Reform, and Labor Supply at the Intensive and Extensive Margins*MIT Press, Cambridge.

EISSA, N., AND J. B. LIEBMAN (1996): "Labor Supply Response to the Earned Income Tax Credit," *The Quarterly Journal of Economics*, 111(2), 605–637.

ELMESKOV, J., J. MARINT, AND S. SCARPETTA (1998): "Key Lessons for Labour Market Reforms: Evidence from OECD Countries' Experiences," *Swedish Economic Policy Review*, 5(2), 205–252.

England, P. (1982): "The Failure of Human Capital Theory to Explain Occupational Sex Segregation," *The Journal of Human Resources*, 17(3), 358–370.

Epple, D., and R. E. Romano (1998): "Competition Between Private and Public Schools, Vouchers, and Peer-Group Effects," *The American Economic Review*, 88(2), 33–62.

Erosa, A., L. Fuster, and D. Restuccia (2002): "Fertility Decisions and Gender Differences in Labor Turnover, Employment, and Wages," *Review of Economic Dynamics*, 5(4), 856–89.

EuropeanCommission (2001): "Reforms of Pension Systems in the EU - An Analysis of the Policy Options," *European Economy*, 73, 171–222.

Evans, W., W. Oates, and R. Schwab (1992): "Measuring Peer Group Effects: A Study of Teenage Behavior," *The Journal of Political Economy*, 100(5), 966–991.

Evans, W. N., and R. M. Schwab (1995): "Finishing High School and Starting College: Do Catholic Schools Make a Difference?," *Quarterly Journal of Economics*, 110, 941–974.

Farber, H. S. (1997): "The Changing Face of Job Loss in the United States, 1981-1995," *Brookings Papers on Economic Activity: Microeconomics*, pp. 55–142.

Fehr, H. (2000): "Pension Reform During the Demographic Transition," *Scandinavian Journal of Economics*, 102, 419–443.

Fehr, H., S. Jokisch, and L. J. Kotlikoff (2004): "The Role of Immigration in Dealing with the Developed World's Demographic Transition," *Finanzarchiv*, 60, 296–324.

FELDSTEIN, M. (1974): "Social Security, Induced Retirement, and Aggregate Capital Accumulation," *The Journal of Political Economy*, 82(5), 905–926.

——— (1982): "Social Security and Private Saving: Reply," *Journal of Political Economy*, 90, 630–642.

——— (1985): "The Optimal Level of Social Security Benefits," *The Quarterly Journal of Economics*, 100(2), 303–320.

——— (1996): "The Missing Piece in Policy Analysis: Social Security Reform," *The American Economic Review*, 86(2), 1–14.

——— (2005a): "Rethinking Social Insurance," *The American Economic Review*, 95(1), 1–24.

——— (2005b): "Structural Reform of Social Security," *Journal of Economic Perspectives*, 19(2), 33–55.

FELDSTEIN, M., AND A. SAMWICK (2002): *Potential Paths of Social Security Reform*chap. 16, pp. 181–224. MIT Press, Cambridge.

FELDSTEIN, M. S., AND J. LIEBMAN (2002): *Social Security*North-Holland, Amsterdam.

FELDSTEIN, M. S., AND A. SAMWICK (1999): "Maintaining Social Security Benefits and Tax Rates Through Personal Retirement Accounts: An Update Based on the 1998 Social Security Trustees Report," *NBER Working Paper*, 6540.

FELLA, G. (2005): "Termination Restrictions and Investment in General Training," *European Economic Review*, 49(6), 1479–1499.

FELLI, L., AND C. HARRIS (1996): "Learning, Wage Dynamics, and Firm-Specific Human Capital," *The Journal of Political Economy*, 104(4), 838–868.

Fenge, R., and M. Werding (2003): "Ageing and the Tax Implied in Public Pension Schems: Simulations for Selected OECD Countries," *CESifo Working Paper*, 841.

Fertig, M. (2003): "Educational Production, Endogeneous Peer Group Formation and Class Composition - Evidence from the PISA 2000 Study," *IZA Discussion Paper*, 714.

Fields, G. S., and O. S. Mitchell (1984): "Economic Determinants of the Optimal Retirement Age: An Empirical Investigation," *The Journal of Human Resources*, 19(2), 245–262.

Fisher, W. H., and C. Keuschnigg (2007): "Parametric Pension Reform and Labor Market Incentives," *Economics Series from Institute for Advanced Studies*, 208.

Freeman, R. B., and J. L. Medoff (1984): *What Do Unions Do?* Basic Books, New York.

Galasso, V., and P. Profeta (2002): "The Political Economy of Social Security: A Survey," *European Journal of Political Economy*, 18, 1–29.

Gentry, W. M., and R. G. Hubbard (2002): "The Effects of Progressive Income Taxation on Job Turnover," *NBER Working Paper*, 9226.

Goldhaber, D. D. (1996): "Public and Private High Schools: Is School Choice an Answer to the Productivity Problem?," *Economics of Education Review*, 15, 93–109.

Goode, R. B. (1959): "Adding to the Stock of Physical and Human Capital," *The American Economic Review*, 49(2), 147–155.

Greenwald, R., L. V. Hedges, and R. D. Laine (1996a): "The Effect of School Resources on Student Achievement," *Review of Educational Research*, 66(3), 391–396.

——— (1996b): "Interpreting Research on School Resources and Student Achievement: A Rejoinder to Hanushek," *Review of Educational Research*, 66(3), 411–416.

Gruber, J., and D. A. Wise (2002): "Social Security Programs and Retirement Around the World: Micro Estimation," *NBER Working Paper*, 9407.

——— (2005): "Social Security Programs and Retirement Around the World: Fiscal Implications," *NBER Working Paper*, 11290.

Hakola, T., and R. Uusitalo (2005): "Not so Voluntary Retirement Decisions? Evidence from a Pension Reform," *Journal of Public Economics*, 89, 2121–2136.

Haley, W. J. (1973): "Human Capital: The Choice Between Investment and Income," *The American Economic Review*, 63(5), 929–944.

——— (1976): "Estimation of the Earnings Profile from Optimal Human Capital Accumulation," *Econometrica*, 44(6), 1223–1238.

Hall, A., and T. R. Johnson (1980): "The Determinants of Planned Retirement Age," *Industrial and Labor Relations Review*, 33(2), 241–254.

Haltiwanger, J. C., J. I. Lane, and J. R. Spletzer (1999): "Productivity Differences Across Employers: The Roles of Employer Size, Age, Human Capital," *The American Economic Review*, 89(2), 94–98.

Hamilton, G. (1996): "The Market for Montreal Apprentices: Contract Length and Information," *Explorations in Economic History*, 33(4), 496–523.

Hanoch, G. (1967): "An Economic Analysis of Earning and Schooling," *The Journal of Human Resources*, 2(3), 310–329.

Hanushek, E. A. (1971): "Teacher Charachteristics and Gains in Student Achievement: Estimation Using Micro Data," *The American Economic Review*, 61(2), 280–288.

——— (1979): "Conceptual and Empirical Issues in the Estimation of Educational Production Functions," *The Journal of Human Resources*, 14(3), 351–388.

——— (1986): "The Economics of Schooling: Production and Efficiency in Public Schools," *The Journal of Economic Literature*, 24(3), 1141–1177.

——— (1989): "The Impact of Differential Expenditures on School Performance," *Educational Researcher*, 18(4), 45–51+62.

——— (1992): "The Trade-Off Netween Child Quantitiy and Quality," *The Journal of Political Economy*, 100(1), 84–117.

——— (1996a): "Measuring Investment in Education," *The Journal of Economic Perspectives*, 10(4), 9–30.

——— (1996b): "A More Complete Picture of School Resource Policies," *Review of Educational Research*, 66(3), 397–409.

——— (1997): "Assessing the Effects of School Resources on Student Performance: An Update," *Educational Evaluation and Policy Analysis*, 19(2), 141–164.

——— (1999): "Some Findings from an Independent Investigation of the Tennessee STAR Experiment and from Other Investigations of Class Size Effects," *Educational Evaluation and Policy Analysis*, 21(2), 143–163.

——— (2006): "Alternative School Policies and the Benefits of General Cognitive Skills," *Economics of Education Review*, 25(4), 447–462.

Hanushek, E. A., J. Kain, and S. Rivkin (1998): "Teachers, Schools, and Academic Achievement," *NBER Working Paper*, 6691.

Hanushek, E. A., C. K. Y. Leung, and K. Yilmaz (2003): "Redistribution Through Education and Other Transfer Mechanisms," *Journal of Monetary Economics*, 50, 1719–1750.

Hanushek, E. A., and J. A. Luque (2003): "Efficiency and Equity in Schools Around the World," *Economics of Education Review*, 22(5), 481–502.

Hanushek, E. A., and J. M. Quigley (1985): "Life-Cycle Earning Capacity and the OJT Investment Model," *International Economic Review*, 26(2), 365–385.

Hanushek, E. A., S. G. Rivkin, and L. L. Taylor (1996): "Aggregation and the Estimated Effects of School Resources," *The Review of Economics and Statistics*, 78(4), 611–627.

Harhoff, D., and T. J. Kane (1995): *Financing Apprenticeship Training: Evidence from Germany*Office of Educational Research and Improvement (ED), Washington, DC.

——— (1997): "Is the German Apprenticeship System a Panacea for the U.S. Labor Market?," *Journal of Population Economics*, 10(2), 171 – 196.

Harmon, C., V. Hogan, and I. Walker (2003): "Dispersion in the Economic Return to Schooling," *Labour Economics*, 10(2), 205–214.

Harmon, C., and I. Walker (1995): "Estimates of the Economic Return to Schooling for the UK," *The American Economic Review*, 85(5), 1278–1286.

Hashimoto, M. (1981): "Firm-Specific Human Capital as a Shared Investment," *The American Economic Review*, 71(3), 475–482.

Hausman, J. A. (1985): *Taxes and Labor Supply*vol. I, chap. 4, pp. 213–263. North-Holland, Amsterdam.

HAVEMAN, R., AND B. WOLFE (1995): "The Determinants of Children's Attainments: A Review of Methods and Findings," *Journal of Economic Literature*, 33(4), 1829–1878.

HAWLEY, J. D. (2004): "Changing Returns to Education in Times of Prosperity and Crisis, Thailand 1985-1998," *Economics of Education Review*, 23(3), 273–286.

HECKMAN, J., A. LAYNE-FARRAR, AND P. TODD (1996): "Human Capital Pricing Equations with an Application to Estimating the Effect of Schooling Quality on Earnings," *The Review of Economics and Statistics*, 78(4), 562–610.

HECKMAN, J., AND D. M. MASTEROV (2004): "The Productivity Argument for Investing in Young Children," *Center for Economic Development Working Paper*, 5.

HECKMAN, J., AND E. VYTLACIL (2000): "Identifying the Role of Cognitive Ability in Explaining the Level of and Change in the Return to Schooling," *NBER Working Paper*, 7820.

HECKMAN, J. J. (1976): "A Life-Cycle Model of Earnings, Learning, and Consumption," *The Journal of Political Economy*, 84(4), 11–44.

——— (1993): "What Has Been Learned About Labor Supply in the Past Twenty Years?," *American Economic Review*, 83(2), 116–121.

——— (2000): "Policies to Foster Human Capital," *Research in Economics*, 54(1), 3–56.

——— (2002): "Flexibility and Job Creation: Lessons for Germany," *NBER Working Paper*, 9194.

HECKMAN, J. J., L. LOCHNER, AND R. COSSA (2002): "Learning-By-Doing Vs. On-the-Job Training: Using Variation Induced by the EITC to Distinguish Between Models of Skill Formation," *NBER Working Paper*, 9083.

HECKMAN, J. J., L. LOCHNER, AND C. TABER (1998a): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1(1), 1–58.

——— (1998b): "Tax Policy and Human-Capital Formation," *The American Economic Review*, 88(2), 293–297.

——— (1999a): "General Equilibrium Cost Benefit Analysis of Education and Tax Policies," *NBER Working Paper*, 6881.

——— (1999b): "Human Capital Formation and General Equilibrium Treatment Effects: A Study of Tax and Tuition Policy," *Fiscal Studies*, 20(1), 25–40.

HEDGES, L. V., R. D. LAINE, AND R. GREENWALD (1994): "Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes," *Educational Researcher*, 23(3), 5–14.

HENRY, G. T., AND D. K. RICKMAN (2007): "Do Peers Influence Children's Skill Development in Preschool?," *Economics of Education Review*, 26(1), 100–112.

HERNOES, E., M. SOLLIE, AND S. STROM (2000): "Early Retirement and Economic Incentives," *The Scandinavian Journal of Economics*, 102(3), 481–502.

HEYWOOD, J. S., L.-S. HO, AND X. WEI (1999): "The Determinants of Hiring Older Workers: Evidence from Hong Kong," *Industrial and Labor Relations Review*, 52(3), 444–459.

HIGASHI, Y. (2002): "Firm Specific Human Capital and Unemployment in a Growing Economy," *Japan and the World Economy*, 14(1), 35–44.

Hines, J. R., and T. Taylor (2005): "Shortfalls in the Long Run: Predictions About the Social Security Trust Fund," *Journal of Economic Perspectives*, 19(2), 3–9.

Horowitz, J. B., and L. Spector (2005): "Is There a Difference Between Private and Public Education on College Performance?," *Economics of Education Review*, 24, 189–195.

Hoxby, C. (2000): "Peer Effects in the Classroom: Learning from Gender and Race Variation," *NBER Working Paper*, 7867.

——— (2002a): "Would School Choice Change the Teaching Profession?," *Journal of Human Resources*, 37(4), 846–891.

Hoxby, C. M. (2002b): "Does Competition Among Public Schools Benefit Students and Taxpayers?," *The American Economic Review*, 90, 1209–1238.

Huggett, M., and G. Ventura (1999): "On the Distributional Effects of Social Security Reform," *Review of Economic Dynamics*, 2, 498–531.

Hungerbühler, M. (2007): "Tax Progression and Training in a Matching Framework," *Labour Economics*, 14(2), 185–200.

Hurd, M. D., and M. J. Boskin (1984): "The Effect of Social Security on Retirement in the Early 1970s," *The Quarterly Journal of Economics*, 99(4), 767–790.

Hutchens, R. (1999): "Social Security Benefits and Employer Bahavior: Evaluating Social Security Early Retirement Benefits as a Form of Unemployment Insurance," *International Economic Review*, 40(3), 659–678.

Hutchens, R. M. (1988): "Do Job Opportunities Decline with Age?," *Industrial and Labor Relations Review*, 42(1), 89–99.

Ill, P. B. S. (1984): "Federal Income Taxation and Human Capital," *Virginia Law Review*, 70(7), 1357–1427.

IMF (2005): *Consultation Concluding Statement of the Mission*, Germany - 2005 Article IV.

Immervoll, H., H. J. Kleven, C. T. Kreiner, and E. Saez (2005): "Welfare Reform in European Countries: A Microsimulation Analysis," *IZA Discussion Paper*, 1810.

Imrohoroglu, A., S. Imrohoroglu, and D. H. Joines (1995): "A Life Cycle Analysis of Social Security," *Economic Theory*, 6, 83–114.

Ishikawa, M., and D. Ryan (2002): "Schooling, Basic Skills and Economic Outcomes," *Economics of Education Review*, 21(3), 231–243.

Jensen, S. E. H., M. I. Lau, and P. Poutvaara (2004): "Efficiency and Equity Aspects of Alternative Social Security Rules," *Finanzarchiv*, 60, 325–358.

Johnson, R. W., and D. Neumark (1997): "Age Discrimination, Job Separations, and Employment Status of Older Workers," *The Journal of Human Resources*, 32(4), 779–811.

Jones, I. S. (1986): "Apprentice Training Costs in British Manufacturing Establishments: Some New Evidence," *British Journal of Industrial Relations*, 24(3), 333–362.

Judd, K. L. (1998): "Taxes, Uncertainty, and Human Capital," *The American Economic Review*, 88(2), 289–292.

Juhn, C., K. M. Murphy, and B. Pierce (1993): "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 101(3), 410–442.

JÜLICHER, P. (2002): *Kombilohn Als Instrument der Beschäftigungspolitik*, vol. 4 of *Bundesarbeitsblatt*. Kohlhammer, Stuttgart.

KALEMLI-OZCAN, S. (2002): "Does the Mortality Decline Promote Economic Growth?," *Journal of Economic Growth*, 7, 411–439.

KALEMLI-OZCAN, S., H. E. RYDER, AND D. N. WEIL (2000): "Mortality Decline, Human Capital Investment, and Economic Growth," *Journal of Development Economics*, 62, 1–23.

KALTENBORN, B. (2001): "Kombilöhne in Deutschland - Eine Systematische Übersicht," *IAB-Werkstattbericht*, 14.

——— (2003): "Kombilöhne: Stand und Perspektiven," *DIW Vierteljahreshefte zur Wirtschaftsforschung*, 72(1), 124–132.

KANAZAWA, S. (2005): "Is "discrimination" Necessary to Explain the Sex Gap in Earnings," *Journal of Economic Psychology*, 26, 269–287.

KAPLOW, L. (1994): "Human Capital under an Ideal Income Tax," *Virginia Law Review*, 80(7), 1477–1514.

——— (1996): "On the Divergence Between "Ideal" and Conventional Income-Tax Treatment of Human Capital," *The American Economic Review*, 86(2), 347–352.

KESSLER, A. S., AND C. LUELFESMANN (2002): "The Theory of Human Capital Revisited: On the Interaction of General and Specific Investments," *CESifo Working Paper*, 776.

KEUSCHNIGG, C. (2004): "Venture Capital Backed Growth," *Journal of Economic Growth*, 9, 239–261.

——— (2005): *Öffentliche Finanzen: Einnahmenpolitik*, Neue Ökonomische Grundrisse. Mohr Siebeck, Tübingen.

KEUSCHNIGG, C., AND M. KEUSCHNIGG (2004): "Aging, Labor Markets and Pension Reform in Austria," *Finanzarchiv*, 60, 359–392.

KIKER, B. F. (1966): "The Historical Roots of the Concept of Human Capital," *The Journal of Political Economy*, 74(5), 481–499.

KIRCHGÄSSNER, G. (2005): "Wirtschaftliche Probleme Einer Alternden Bevölkerung: Übersicht und Grundsätzliche Aspekte," *Teil 3 des Jahresberichts 2005 der Kommission für Konjunkturfragen.*

KOMMISSIONFÜRKONJUNKTURFRAGEN (2005): "Jahresbericht 2005. Wirtschaftliche Auswirkungen Einer Alternden Bevölkerung," *Beilage zur "Volkswirtschaft"*, 384.

KONRAD, K. A. (2001): "Privacy and Time-Consistent Optimal Labor Income Taxation," *Journal of Public Economics*, 79(3), 503–519.

KOTLIKOFF, L., AND J. GOKHALE (1992): "Estimating a Firm's Age-Productivity Profile Using the Present Value of Workers' Earnings," *Quarterly Journal of Economics*, 107(4), 1215–1242.

KOTLIKOFF, L., AND D. WISE (1989): *Employee Retirement and a Firm's Pension Plan*pp. 279–334. University of Chicago Press, Chicago.

KOTLIKOFF, L. J. (1979): "Testing the Theory of Social Security and Life Cycle Accumulation," *The American Economic Review*, 79, 396–410.

——— (1996): "Privatizing Social Security at Home and Abroad," *The American Economic Review*, 86(2), 368–372.

——— (2002): *Generational Policy*North-Holland, Amsterdam.

KOTLIKOFF, L. J., A. SPIVAK, AND L. H. SUMMERS (1982): "The Adequacy of Savings," *The American Economic Review*, 72(5), 1056–1069.

KREKEL, E. M. (2005): *Finanzierung der Ausbildung*Bundesinstitut für Berufsbildung, Bremen.

KRUEGER, A. B. (2003): "Economic Considerations and Class Size," *The Economic Journal*, 113(483), 34–63.

KRUEGER, A. B., AND M. LINDAHL (2001): "Education for Growth: Why and For Whom?," *Journal of Economic Literature*, 39(4), 1101–1136.

KUNZE, A. (2005): "The Evolution of the Gender Wage Gap," *Labour Economics*, 12, 73–97.

LACH, S., AND M. SCHANKERMAN (2003): "Incentives and Invention in Universities," *NBER Working Paper*, 9727.

LAM, D., AND R. F. SCHOENI (1993): "Effects of Family Background on Earnings and Returns to Schooling: Evidence from Brazil," *The Journal of Political Economy*, 101(4), 710–740.

LAU, M. I., AND P. POUTVAARA (2001a): "Social Security Incentives and Human Capital Investment," *CESifo Working Paper*, 438.

——— (2001b): "Social Security Rules, Labor Supply and Human Capital Formation," *CEBR Discussion Paper*, 2001-10.

LAYARD, R., AND S. NICKELL (1999): *Labor Market Institutions and Economic Performance*vol. 3C. North Holland, Amsterdam.

LAZEAR, E. P. (1979): "Why Is There Mandatory Retirement?," *The Journal of Political Economy*, 87(6), 1261–1284.

——— (1981): "Agency, Earnings Profiles, Productivity, and Hours Restrictions," *American Economic Review*, 71(4), 606–620.

——— (1986): "Salaries and Piece Rates," *The Journal of Business*, 59(3), 405–431.

——— (2000): "The Power of Incentives," *The American Economic Review*, 90(2), 410–414.

——— (2001): "Educational Production," *The Quarterly Journal of Economics*, 116(3), 777–803.

——— (2003): "Firm-Specific Human Capital: A Skill-Weights Approach," *NBER Working Paper*, 9679.

Lefgren, L. (2004): "Educational Peer Effects and the Chicago Public Schools," *Journal of Urban Economics*, 56(2), 169–191.

Leigh, D. E. (1984): "Why is There Mandatory Retirement? An Empirical Reexamination," *The Journal of Human Resources*, 19(4), 512–531.

Lemieux, T. (2006): "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *The American Economic Review*, 96(3), 461–498.

Lillard, L. A., and H. W. Tan (1992): "Private Sector Training: Who Gets It and What Are Its Effects?," *Research in Labor Economics*, 13, 1–62.

Lindbeck, A., and M. Persson (2003): "The Gains from Pension Reform," *Journal of Economic Literature*, 41, 74–112.

Lindsay, C. M. (1971): "Measuring Human Capital Returns," *The Journal of Political Economy*, 79(6), 1195–1215.

Ljungqvist, L. (1995): "Wage Structure as Implicit Insurance on Human Capital in Developed versus Underdeveloped Countries," *Journal of Development Economics*, 46(1), 35–50.

Loewenstein, M. A., and J. R. Spletzer (1997): "Delayed Formal On-the-Job Training," *Industrial and Labor Relations Review*, 51(1), 82–99.

——— (1998): "Deviding the Costs and Returns to General Training," *Journal of Labor Economics*, 16(1), 142–171.

——— (1999): "General and Specific Training: Evidence and Implications," *The Journal of Human Resources*, 34(4), 710–733.

Loury, L. D. (2006): "All in the Extended Family: Effects of Grandparents, Aunts, and Uncles on Educational Attainment," *The American Economic Review*, 96(2), 275–278.

Lumsdaine, R. L., and D. A. Wise (1990): "Aging and Labor Force Participation: A Review of Trends and Explanations," *NBER Working Paper*, 3240.

Lynch, L. (1991): "The Role of Off-the-Job Vs. On-the-Job Training for the Mobility of Women Workers," *American Economic Review*, 81(2), 151–156.

Malcomson, J. M., J. W. Maw, and B. McCormick (2003): "General Training by Firms, Apprentice Contracts, and Public Policy," *European Economic Review*, 47(2), 197–227.

Manski, C. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, 531–542.

Martín, A. R. S. (2003): "Endogenous Retirement and Public Pension System Reform in Spain," *Cebtro Atlántico de Estudios Ríos Pérez Working Paper*, 17.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995): *Microeconomic Theory*. Oxford University Press, New York.

Meyer, B. D. (2002): "Labor Supply at the Extensive and Intensive Margins: The EITC, Welfare, and Hours Worked," *The American Economic Review*, 92(2), 373–379.

Meyer, B. D., and D. T. Rosenbaum (2001): "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers," *Quarterly Journal of Economics*, 116(3), 1063–1114.

MILES, D., A. TIMMERMANN, J. DE HAAN, AND M. PAGANO (1999): "Risk Sharing and Transition Costs in the Reform of Pension Systems in Europe," *Economic Policy*, 14(29), 251–286.

MINCER, J. (1958): "Investment in Human Capital and Personal Income Distribution," *The Journal of Political Economy*, 6(4), 281–302.

——— (1962): "On-the-Job Training: Costs, Returns, and Some Implications," *The Journal of Political Economy*, 70(5), 50–79.

——— (1970): "The Distribution of Labor Incomes: A Survey With Special Reference to the Human Capital Approach," *Journal of Economic Literature*, 8(1), 1–26.

——— (1974): *Schooling, Earnings, and Experience*. Columbia University Press, New York.

——— (1989): "Human Capital and the Labor Market: A Review of Current Research," *Educational Research*, 18(4), 27–34.

——— (1997): "The Production of Human Capital and the Life Cycle of Earnings: Variation on a Theme," *Journal of Labor Economics*, 15(1), 26–47.

MINCER, J., AND H. OPEK (1982): "Interrupted Work Careers: Depreciation and Restoration of Human Capital," *The Journal of Human Resources*, 17(1), 3–24.

MIRRLEES, J. A. (1971): "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies*, 38, 175–208.

MITCHELL, O., AND S. ZELDES (1996): "Social Security Privatization: A Structure for Analysis," *The American Economic Review*, 82(2), 363–367.

MITCHELL, O. S., AND G. S. FIELDS (1984): "The Economics of Retirement Behavior," *Journal of Labor Economics*, 2(1), 84–105.

Mitchell, O. S., and J. W. Phillips (2000): "Retirement Responses to Early Social Security Benefit Reductions," *NBER Working Paper*, 7963.

Murnane, R. J., J. B. Willet, Y. Duhaldeborde, and J. H. Tyler (2000): "How Important are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings?," *Journal of Policy Analysis and Management*, 19(4), 547–568.

Neal, D. (1995): "Industry-Specific Human Capital: Evidence from Displaced Workers," *Journal of Labor Economics*, 13(4), 653–677.

Nechyba, T. J. (1996): "Public School Finance in a General Equilibrium Tiebout World: Equalization Programs, Peer Effects and Private School Vouchers," *NBER Working Paper*, 5642.

Nerlove, M., A. Razin, E. Sadka, and R. K. V. Weizsäcker (1993): "Comprehensive Income Taxation, Investments in Human and Physical Capital, and Productivity," *Journal of Public Economics*, 50(3), 397–406.

Nickell, S. (1997): "Unemployment and Labor Market Rigidities: Europe Vs. North America," *Journal of Economic Perspectives*, 11(3), 55–74.

Nickell, S. (1998): "Unemployment: Questions and Some Answers," *Economic Journal*, 108, 802–816.

Nielsen, S. B., and P. B. Sorensen (1997): "On the Optimality of the Nordic System of Dual Income Taxation," *Journal of Public Economics*, 63(3), 311–329.

Ochel, W. (2005): "Hartz IV: Welfare to Work in Germany," *CESifo DICE Report*, 3(2), 18–25.

OECD (2004): *Benefits and Wages: OECD Indicators*. Paris.

——— (2005): *Employment Outlook*. Paris.

Orszag, J. M., and D. Snower (2003): "Designing Employment Subsidies," *Labour Economics*, 10(5), 557–572.

Ortigueira, S. (2006): "Skills, Search and the Persistence of High Unemployment," *Journal of Monetary Economics*, 53, 2165–2178.

Paglin, M., and A. M. Rufolo (1990): "Heterogeneous Human Capital, Occupational Choice, and Male-Female Earnings Differences," *Journal of Labor Economics*, 8(1), 123–144.

Parsons, D. A. (1972): "Specific Human Capital: An Application to Quit Rates and Layoff Rates," *The Journal of Political Economy*, 80(6), 1120–1143.

Parsons, D. O. (1974): "The Cost of School Time, Foregone Earnings, and Human Capital Formation," *The Journal of Political Economy*, 82(2), 251–266.

Pastore, F., and A. Verashchagina (2006): "Private Returns to Human Capital over Transition: A Case Study of Belarus," *Economics of Education Review*, 25, 91–107.

Perri, T. J. (2003): "The Cost of Specialized Human Capital," *Economics of Education Review*, 22(4), 433–438.

Perroni, C. (1995): "Assessing the Dynamic Efficiency Gains of Tax Reform When Human Capital is Endogeneous," *International Economic Review*, 36(4), 907–925.

Petrongolo, B., and C. A. Pissarides (2001): "Looking Into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39, 390–431.

Phelps, E. S. (1997): *Rewarding Work - How to Restore Participation and Self-Support to Free Enterprise*. Harvard University Press, Cambridge.

Piketty, T., and E. Saez (2003): "Income Inequality and in the United States," *Quarterly Journal of Economics*, 118(1), 1–39.

Pirenne, H. (1936): *Economic and Social History of Medieval Europe*. Harcourt Brace, New York.

Pissarides, C. A. (1992): "Loss of Skill During Unemployment and the Persistence of Unemployment Shocks," *The Quarterly Journal of Economics*, 107(4), 1371–1391.

Plotnick, R. D., and S. D. Hoffman (1999): "The Effect of Neighborhood Characteristics on Young Adult Outcomes: Alternative Estimates," *Social Science Quarterly*, 80(1), 1–18.

Polachek, S. (1981): "Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure," *The Review of Economics and Statistics*, 63(1), 60–69.

Poutvaara, P. (2001): "Alternative Tax Constitutions and Risky Education in a Federation," *Regional Science and Urban Economics*, 31(2), 355–377.

Prendergast, C. (1993): "The Role of Promotion in Inducing Specific Human Capital Acquisition," *The Quarterly Journal of Economics*, 108(2), 523–534.

Psacharopoulos, G. (1994): "Returns to Investment in Education: A Global Update," *World Development*, 22(9), 1325–1343.

Raffelhüschen, B. (2001): "Soziale Grundsicherung in der Zukunft: Eine Blaupause," *Diskussionsbeiträge des Instituts für Finanzwissenschaft der Albert-Ludwigs-Universität Freiburg*, 96.

Reimers, C., and M. Honig (1996): "Responses to Social Security by Men and Women: Myopic and Far-Sighted Bahavior," *The Journal of Human Resources*, 31(2), 359–382.

Reinberg, A., and M. Hummel (2002): "Qualifikationsspezifische Arbeitslosenquoten - Reale Entwicklung Oder Statistisches Artefakt," *IAB Werkstattbericht*, 4.

Rephann, T. J. (2002): "The Importance of Geographical Attributes in the Decision to Attend College," *Socio-Economic Planning Sciences*, 36(4), 291–307.

Ritzen, J. M., and D. R. Winkler (1977): "The Production of Human Capital Over Time," *The Review of Economics and Statistics*, 59(4), 427–437.

Rivkin, S. G. (2001): "Tiebout Sorting, Aggregation and the Estimation of Peer Group Effects," *Economics of Education Review*, 20(3), 201–209.

Roed, K., and S. Strom (2002): "Progressive Taxes and the Labour Market: Is the Trade-Off Between Equality and Efficiency Inevitable?," *Journal of Economic Surveys*, 16(1), 77–110.

Rose, H. (2006): "Do Gains in Test Scores Explain Labor Market Outcomes," *Economics of Education Review*, 25(4), 430–446.

Rosen, S. (1983): "Specialization and Human Capital," *Journal of Labor Economics*, 1(1), 43–49.

Rosenzweig, M. R., and K. I. Wolpin (1994): "Are There Increasing Returns to the Intergenerational Production of Human Capital? Maternal Schooling and Child Intellectual Achievement," *The Journal of Human Resources*, 29(2), 670–693.

Rosholm, M., and M. Svarer (2004): "Endogenous Wage Dispersion in a Search-Matching Model," *Labour Economics*, 11(5), 623–645.

Rouse, C. (1998): "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program," *Quarterly Journal of Economics*, 113, 553–602.

——— (1999): "Further Estimates of the Economic Return to Schooling from a New Sample of Twins," *Economics of Education Review*, 18(2), 149–157.

RUBB, S. (2003): "Overeducation in the Labor Market: A Comment and Re-Analysis of a Meta-Analysis," *Economics of Education Review*, 22(6), 621–629.

RUST, J., AND C. PHELAN (1997): "How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets," *Econometrica*, 65(4), 781–831.

RYAN, P. (1980): "The Costs of Job Training for a Transferable Skill," *British Journal of Industrial Relations*, 18(3), 334–352.

SACERDOTE, B. (2001): "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *The Quarterly Journal of Economics*, 116(2), 681–704.

SAEZ, E. (2002): "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 117, 1039–1073.

SAMWICK, A. A. (1998): "New Evidence on Pensions, Social Security, and the Timing of Retirement," *Journal of Public Economics*, 70, 207–236.

——— (2000): "Is Pension Reform Conducive to Higher Saving?," *The Review of Economics and Statistics*, 82(2), 264–272.

SAYAN, S., AND A. KIRACI (2001): "Parametric Pension Reform with Higher Retirement Ages: A Computational Investigation of Alternatives for a Pay-as-You-Go-Based Pension System," *Journal of Economic Dynamics and Control*, 25, 951–966.

Scarpetta, S. (1996): "Assessing the Role of Labour Market Policies and Institutional Settings on Unemployment: A Cross-Country Study," *OECD Econimic Studies*, 26, 43–98.

Schöb, R., and J. Weimann (2003): "Kombilohn: Die Magdeburger Alternative," *Perspektiven der Wirtschaftspolitik*, 4, 1–16.

Schultz, T. W. (1961): "Investment in Human Capital," *The American Economic Review*, 51(1), 1–17.

Schumann, P. L., D. A. Ahlburg, and C. B. Mahoney (1994): "The Effects of Human Capital and Job Characteristics on Pay," *The Journal of Human Resources*, 29(2), 481–503.

Scott, F. A., M. C. Berger, and J. E. Garen (1995): "Do Health Insurance and Pension Costs Reduce the Job Opportunities of Older Workers?," *Industrial and Labor Relations Review*, 48(4), 775–791.

Shaffer, H. G. (1961): "Investment in Human Capital: Comment," *The American Economic Review*, 51(5), 1026–1035.

Shaw, K. L. (1989): "Life-Cycle Labor Supply with Human Capital Accumulation," *International Economic Review*, 30(2), 431–456.

Shea, J. (2000): "Does Parents' Money Matter?," *Journal of Public Economics*, 77(2), 155–184.

Sheshinski, E. (1971): "On the Theory of Optimal Income Taxation," *Harvard Institute for Economic Research Discussion Paper*, 172.

Sheshinski, E., and Y. Weiss (1981): "Uncertainty and Optimal Social Security Systems," *The Quarterly Journal of Economics*, 96(2), 189–206.

Sinn, H.-W. (2002): "Die Höhle in der Eiger-Nordwand, eine Anmerkung Zum Mainzer Modell und Zum Wohlfahrtsstaat an Sich," *Ifo-Schnelldienst*, 55(3/2002), 20–25.

——— (2003): *Ist Deutschland Noch Zu Retten?* Econon, München.

——— (2004): "The Pay-as-You-Go Pension System as Fertility Insurance and an Enforcement Device," *Journal of Public Economics*, 88, 1335–1357.

Sinn, H.-W., C. Holzner, W. Meister, W. Ochel, and M. Werding (2006): *Welfare to Work in Germany: The Country's Current Agenda for Redesigning the Welfare State*. Edward Elgar, Cheltenham.

Snow, A., and R. S. Warren (1990): "Human Capital Investment and Labor Supply Under Uncertainty," *International Economic Review*, 31(1), 195–206.

Snower, D. (1994): "Converting Unemployment Benefits Into Employment Subsidies," *American Economic Review*, 84(2), 65–70.

Soares, R. R. (2005): "Mortality Reductions, Educational Attainment, and Fertility Choice," *The American Economic Review*, 95, 580–601.

Solow, R. M. (1965): "Human Capital," *The Journal of Political Economy*, 73(5), 552–553.

Soskice, D. (1994): *Reconciling Markets and Institutions: The German Apprenticeship System*University of Chicago Press, Chicago.

Stähler, N. (2006): "Employment Protection: Its Effects on Different Skill Groups and on the Incentive to Become Skilled," *Jahrbücher für Nationalökonomie und Statistik*, 226(1), 41–54.

Steuerle, C. E. (1990): "Policy Watch: Tax Credits for Low-Income Workers with Children," *The Journal of Economic Perspectives*, 4(3), 201–212.

——— (1996): "How Should Government Allocate Subsidies for Human Capital?," *The American Economic Review*, 86(2), 353–357.

Stock, J. H., and D. A. Wise (1990): "Pensions, the Option Value of Work, and Retirement," *Econometrica*, 58, 1151–1180.

Straka, J. W. (1992): *The Demand for Older Workers: The Neglected Side of a Labor Market*vol. 15 of *Studies in Income Distribution*. Social Security Administration, Washington, DC.

Temple, J. A., and A. J. Reynolds (2007): "Benefits and Costs of Investments in Preschool Education: Evidence from the Child-Parent Centers and Related Programs," *Economics of Education Review*, 26(1), 126–144.

Trostel, P. A. (1993): "The Effect of Taxation on Human Capital," *The Journal of Political Economy*, 101(2), 327–350.

Tyler, J. H. (2004): "Basic Skills and the Earnings of Dropouts," *Economics of Education Review*, 23(3), 221–235.

Vandenberghe, V., and S. Robin (2004): "Evaluating the Effectiveness of Private Education Across Countries: A Comparison of Methods," *Labour Economics*, 11(4), 487–506.

Verdugo, R. R., and N. T. Verdugo (1989): "The Impact of Surplus Schooling on Earnings: Some Additional Findings," *The Journal of Human Resources*, 24(4), 629–643.

Veum, J. R. (1999): "Training, Wages, and the Human Capital Model," *Southern Economic Journal*, 65(3), 526–538.

VonBardeleben, R., U. Beicht, and K. Fehér (1995): *Betriebliche Kosten und Nutzen der Ausbildung - Repräsentative Ergebnisse Aus Industrie, Handel und Handwerk*, Berichte Zur Beruflichen Bildung 187. Bertelsmann, Bielefeld.

Wallace, T. D., and L. A. Ihnen (1975): "Full-Time Schooling in Life-Cycle Models of Human Capital Accumulation," *The Journal of Political Economy*, 83(1), 137–156.

Wasmer, E. (2006): "General versus Specific Skills in Labor Markets with Search Frictions and Firing Costs," *The American Economic Review*, 96(3), 811–831.

Weisbrod, B. A. (1961): "The Valuation of Human Capital," *The Journal of Political Economy*, 69(5), 425–436.

——— (1966): "Investing in Human Capital," *The Journal of Human Resources*, 1(1), 5–21.

Weiss, Y. (1986): *The Determinants of Life Cycle Earnings: A Survey*vol. 1 of *Handbook of Labor Economics*, pp. 603–640. North Holland, Amsterdam.

Willis, R. J., and S. Rosen (1979): "Education and Self-Selection," *Journal of Political Economy*, 87, 7–36.

Wilson, K. (2001): "The Determinants of Educational Attainment: Modeling and Estimating the Human Capital Model and Education Production Functions," *Southern Economic Journal*, 67(3), 518–551.

——— (2002): "The Effects of School Quality on Income," *Economics of Education Review*, 21(6), 579–588.

Winter-Ebmer, R., and A. Wirz (2002): "Public Funding and Enrolment Into Higher Education in Europe," *IZA Discussion Paper*, 503.

Winter-Ebmer, R., and J. Zweimüller (1992): "Occupational Segregation and Career Advancement," *Economic Letters*, 39(2), 229–234.

Zabalza, A., C. Pissarides, and M. Barton (1980): "Social Security and the Choice Between Full-Time Work, Part-Time Work, and Retirement," *Journal of Public Economics*, 14, 245–276.

Zietz, J., and P. Joshi (2005): "Academic Choice Bahavior of High School Students: Economic Rationale and Empirical Evidence," *Economics of Education Review*, 24, 297–308.

ZUCKER, A. (1967): "A Note on the Declining Tendency with Age for Investment in Human Capital," *The Journal of Human Resources*, 2(4), 538–540.

ZWEIMÜLLER, J., AND R. WINTER-EBMER (2000): "Firm-Specific Training: Consequences for Job Mobility," *IZA Discussion Paper*, 138.

# Curriculum Vitae

## Particulars

| | |
|---|---|
| Date of birth | 02/13/1978 |
| Place of birth | Bochum, Germany |
| Nationality | German |
| Marital status | married |

## Education

| | |
|---|---|
| 05/2004 – 05/2007 | University of St. Gallen, PhD in Economics and Finance |
| 01/2003 – 05/2004 | Study Center Gerzensee, Program for Beginning Doctoral Students in Economics |
| 10/1998 – 10/2002 | University of St. Gallen, Studies in Economics, Specialization in Finance, Diploma in Economics (lic.oec.HSG) |
| 08/1988 – 07/1997 | Gymnasium Holthausen, Abitur |

## Work Experience

| | |
|---|---|
| 03/2006 – 08/2007 | Institute of Public Finance and Fiscal Law, Research Assistant |
| 04/2003 – 07/2006 | Institute of Public Finance and Fiscal Law, Teaching Assistant |
| 10/2002 – 02/2007 | Research Institute for Empirical Economics and Economic Policy, Teaching Assistant |
| 04/2002 – 07/2007 | University of St. Gallen, Department of Economics, Teaching Assistant |
| 10/2001 – 08/2007 | University of St. Gallen, Department of Economics, Research Assistant |
| 02/2000 – 03/2000 | DuPont, Internship |
| 02/1999 – 04/1999 | VFL Bochum, Internship |
| 07/1998 – 09/1998 | Volksbank, Internship |
| 07/1997 – 06/1998 | Diakonisches Werk, Civilian Service |