Three Essays on the Taxonometrics of Labor Income

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Wirtschaftswissenschaftlichen Fakultät
der Eberhard-Karls-Universität zu Tübingen

vorgelegt von

Martin Frederik Weiss
aus Heidelberg

2010

Dekanin:	Professorin Dr. Kerstin Pull
Erstkorrektor:	Professor Dr. Dr. h.c. Franz W. Wagner
Zweitkorrektorin:	Professorin Dr. Kerstin Pull
Tag der mündlichen Prüfung:	14.7.2010

PREFACE

This thesis was accepted by the Faculty of Economics and Business Administration of the Eberhard-Karls-Universität Tübingen in July 2010. The contributions assembled here were written during my time as an assistant to Prof. Dr. Dr. h.c Franz W. Wagner from April 2006.

Issues concerning the taxation of labor income are a neglected area of research in Business Administration. One of the main pillars of the German tax system is the taxation of wages which consistently rakes in over 80% of German income tax revenue. Given this significance, the lack of research in this area is both surprising and astonishing. In this doctoral thesis, I provide three attempts to illuminate three aspects of labor taxation. As a "bonus", I also present a contribution to the taxation of capital income which was written in collaboration with my colleague Martin Jacob and Prof. Rainer Niemann.

I want to thank first and foremost Prof. Dr. Dr. h.c. Franz W. Wagner for his patience and encouragement along the way, and for the inspiration he provided. Special thanks are also due to Prof. Dr. Kerstin Pull for the second report on my thesis and for her course on scientific publishing during the Winter Semester 2006/2007. I also want to express my gratitude to my colleagues at the chair of Business Taxation: Martin Jacob, Michaela Ott, Andreas Pasedag and Jan Vossmerbäumer.

Contribution 2 profited from comments provided by participants at the fourth arqus conference at Bielefeld, July 2008. Contribution 3 owes much to comments by participants at the conference on “Empirical Research for the Tax Profession”, at the ZEW Mannheim, September 2009.

Last, but by no means least, I want to express my gratitude to my family and friends who have supported me along the way. Without them, this thesis would not have been written.

Tübingen, July 2010 Martin Frederik Weiss

OVERVIEW

List of Tables **I**

List of Figures **III**

Introduction **1**

1 Higher Tax Rates on Labor? Evidence from German Panel Data **9**

2 Training Decisions and the Income Tax **29**

3 Do Germans React to the Commuting Allowance? **63**

4 The Rich Demystified: A Reply to Bach, Corneo and Steiner (2008) **85**

Bibliography **127**

Index **141**

CONTENTS

List of Tables **I**

List of Figures **III**

Introduction **1**

1 Higher Tax Rates on Labor? Evidence from German Panel Data **9**

1.1 Introduction 10
1.2 Estimation of the Wage Equation 13
1.2.1 Estimation Strategy 13
1.2.2 Dataset 15
1.2.3 Estimation Results 17
1.3 Results 20
1.4 Conclusion 26

2 Training Decisions and the Income Tax **29**

2.1 Introduction 30
2.2 Training Decisions under an Income Tax 33
2.2.1 Tax Effects for Training Decisions 33
2.2.2 Relevant Provisions of the German Income Tax Code 2003 . 39
2.3 Empirical Returns to Training 41
2.3.1 Research Goal 41
2.3.2 Econometric Specification 42

2.3.3 The Dataset . . . 45
2.3.4 Estimation Results . . . 49
2.4 Tax Implications . . . 52
2.4.1 Description of the Approach . . . 52
2.4.2 Example for the Calculations . . . 54
2.4.3 Results . . . 55
2.5 Conclusion . . . 58
Appendix: The Duan Method . . . 60

3 Do Germans React to the Commuting Allowance? 63
3.1 Introduction . . . 64
3.2 The Commuting Allowance . . . 69
3.2.1 Legal Design of the Allowance . . . 69
3.2.2 Theoretical Research . . . 71
3.3 Estimation . . . 73
3.3.1 Estimation Strategy . . . 73
3.3.2 Dataset . . . 78
3.3.3 Results . . . 80
3.4 Conclusion . . . 82

4 The Rich Demystified: A Reply to Bach, Corneo and Steiner (2008) 85
4.1 Introduction . . . 86
4.2 Theoretical Examination of the Approach in Bach et al. (2008) . . . 88
4.2.1 The BCS Approach . . . 88
4.2.2 Effective Tax Rates . . . 89
4.2.3 Different Notions of Income . . . 91
4.2.4 The Relation between Economic Income and Taxable Income 93
4.2.5 The Tax Burden on Economic Activities . . . 95
4.3 Analysis of Effective Tax Rates on Top Incomes . . . 105
4.3.1 Description of Available Datasets . . . 105
4.3.2 “Adjusted Gross Income” Variable in BCS . . . 108
4.3.3 Adjustment for Taxes Paid . . . 113
4.3.4 Empirical Results . . . 116
4.3.5 Limitations . . . 117
4.4 Conclusion . . . 118

Appendix: Additional Tables 119

Bibliography **127**

Index **141**

LIST OF TABLES

1.1 Tax Rates in the Nordic Countries (in %) 10
1.2 Cohort Strengths in the Dataset 16
1.3 Conversion of Degrees into Imputed Educational Times S_i (in Years) 17
1.4 Estimation Results for Equation (1.1) 18
1.5 Required Marginal Tax Rates t_2 under the *Nielsen and Sørensen*-Model, for Varying Interest Rates i, Tax Rates t_1 and Thresholds Ω 26

2.1 Estimation Results for the OLS vs. Treatment Regression 50
2.2 Time Line for Example 55

3.1 Historical Development of the German Commuting Allowance 2001-2009 66
3.2 Wage Income and Deductions in German Income Tax Returns for the Year 2003 70
3.3 Sample Characteristics 80
3.4 Difference-in-Difference Estimation for the Commuting Distances . 81

4.1 Differing Income Definitions 91
4.2 Derivation of the Cash Flow 92
4.3 Comparison of Single Period ETRs 93
4.4 Calculation of Tax Liability for an Unincorporated Business 97

4.5 Calculation of Tax Liability for an Incorporated Business under the German Full-Imputation System and the German Half-Income System, Owners in the Highest German Tax Bracket 100
4.6 The Effect of Intertemporal Offsets on the ETR 103
4.7 Definition of Adjusted Gross Income 108
4.8 Adjustments to Recognized Tax Burden 113
4.9 Comparison of the Comprehensive Nominal Tax Rate and ETR^{BCS} for 2001 116
4.10 Adjusted Income Components as Percentage of Adjusted Sum of Incomes 120
4.11 Distribution of Adjusted Income Components among Quantiles . . . 121
4.12 Relevance of Intertemporal Loss Deductions 122
4.13 Reporting Frequency of Income Components 123

LIST OF FIGURES

1 Contributions of Income Types to the German Income Tax Base, Fiscal Year 2004 2

1.1 Prediction of Mean Gross Earnings for University and High School Graduates, Birth Cohort 1965-1969 22

1.2 Required Average Tax Rates for a University Graduate under a Dual Income Tax, for i=6%, Ω=25,000 € and $t_1 = \tau = 25\%$ 24

1.3 Illustration of the *Nielsen and Sørensen* Tax System for a Threshold Ω=26,000 € 25

2.1 Incidence of Occupational Training by Age Group and Gender . . . 31

2.2 Impact of Taxation on the Net Present Value, Interest Rate 5% . . . 36

2.3 Impact of Taxation on the Net Present Value, Interest Rate 3% . . . 37

2.4 Estimated Density of Monthly Wages before and after Intervention . 48

2.5 CDF of Effective Tax Rates for Sample Members 56

2.6 Estimated Density of Effective Tax Rates for Sample Members, Gross Interest Rate Varies 57

2.7 Estimated Density of Effective Tax Rates for Sample Members, Depreciation Rate Varies, Interest Rate 6% 58

2.8 Estimated Density of Effective Tax Rates for Sample Members, Gross Interest Rate Varying, Costs Inflated by 25% 59

3.1 Share of Income Types in German Income Tax Revenue 64

3.2 Itemized Deductions for “Wage Income” in German Income Tax Returns 2003 68

4.1 From Gross Economic Activity to Consumption 95

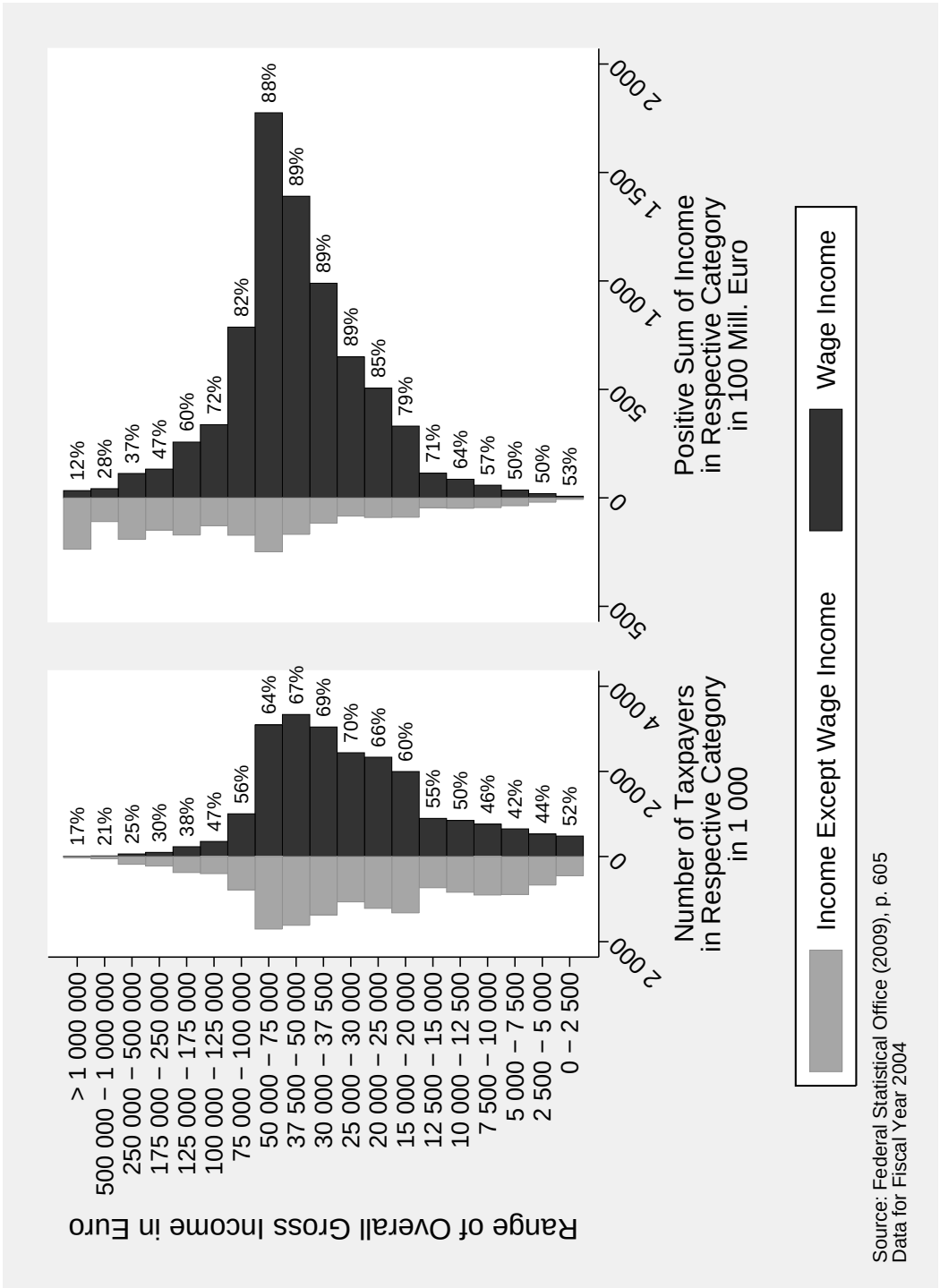
INTRODUCTION

The taxation of labor income and the effects that it generates on microeconomic decisions is a somewhat neglected area in tax research in Business Administration. *Wagner* (2008, p. 107) cites evidence from the recent survey article by *Hundsdoerfer, Kiesewetter and Sureth* (2008) that supports this assertion: Only around 1% of the literature cited there can be classified as true contributions to the taxation of labor income.

This conclusion does not quite chime with the observation that labor income has consistently provided the bulk of German income tax revenue. This fact, also emphasized in the introductions of contributions 2 and 3, is underscored by **figure 1** on the following page. It shows contributions to the tax base for the German income tax in the year 2004, the latest fiscal year for which figures are available (*Federal Statistical Office*, 2009, p. 605). In both panels, a distinction is made between the tax base for the income type[1] "Wage Income", and the other six income types combined. Along its y-axis, it additionally breaks down the overall income brackets that taxpayers belonged to in the relevant year 2004. To the right of the bars representing "Wage Income" sits its percentage share in the respective income brackets. The left panel concerns the *number* of taxpayers reporting labor / any other income – where any taxpayer can of course have taxable income both in the non-wage *and* wage category. The right panel shows the *sum of reported positive income* for both categories.

[1] Cf. figure 3.1 on page 64 for a disaggregated view onto the seven income types recognized under German income tax law.

FIGURE 1: Contributions of Income Types to the German Income Tax Base, Fiscal Year 2004



Source: Federal Statistical Office (2009), p. 605
Data for Fiscal Year 2004

As is obvious from figure 1, the dominance of "Wage Income" is indisputable. It seems to form the backbone of asset incomes for the middle class since its prevalence is most pronounced in the 20,000 to 75,000 € range. It was this "inescapable" picture which motivated my first foray into the realm of taxation of labor income in my diploma thesis (*Weiss*, 2007).

***Weiss* (2007)**

There, I researched the effects caused by the differences in tax bases between capital and labor income. Among the most salient results were the following:

- The massive body of human capital that modern economies rely on is indeed "as good as gold" (*Judd*, 2000) for the individuals endowed with it: Rates of return for an additional year of schooling between 7% and 10% were obtained from the Mincer regressions (*Mincer*, 1974) whose results are assembled in *Weiss* (2007, sect. B.3). To put these results into perspective, it must be borne in mind that "later" interventions in the labor market, such as job creation schemes, often measure their success in terms of a few percentage points of increases in employment probability (*Fitzenberger and Völter*, 2007). At the same time, these returns approach and even exceed those observed in stock markets over long time horizons (*DeLong and Magin*, 2009) – with the difference being that the vast majority of society is endowed with at least some human capital while stocks are held by a comparatively small minority.

- From the point of view of taxation, I built on the contributions by *Kaplow* (1996) and *Wagner* (2000) who showed that there was a substantial time effect favoring labor over capital income in an after tax perspective. In *Weiss* (2007, sect. 3.5), I calculated the effective tax rates (ETR) that were implied by the age-income-profiles I had derived. Depending on tax and interest rates, these ETRs were substantially lower than the nominal tax rates facing the worker, and in some cases well below zero. I then went on to ask how the differences between capital and labor income could be reconciled, and arrived at labor income tax rates substantially above prevailing capital income tax rates (*Weiss*, 2007, sect. 3.6). Since the publication of *Weiss* (2007), the introduction of a final withholding tax on capital income in Germany – featuring a tax rate of 25% – has proved the practical feasibility – in a mature democracy boasting one of the world's largest economies – of the theoretical insights following from my contribution.

In this doctoral thesis, I provide three contributions to the taxation of labor income. They form a chain of events in the worker‘s life which are related to human capital and encompass

- the *initial* impulse for human beings to be productive, i.e. a university education early in life (contribution 1)
- the *continuing* training necessary to keep up with developments and maintain productivity during the labor market participation (contribution 2)
- an example for the conditions under which the *existing* human capital stock can be exploited, the tax treatment of commuting costs (contribution 3).

Contribution 4 represents an unrelated analysis of more well-established results of capital income taxation.

Extant Human Capital Literature

One of the major insights gained while composing this doctoral thesis is the fact that there is a well-established and comprehensive strand of literature on human capital issues in the economics profession, starting with the seminal contribution by *Becker* (1964). This literature has already shed light on many aspects of human capital. A recent contribution by *Palacios-Huerta* (2003a) begins thus:

> Human capital resources are a crucial part of an individual's capital holdings and comprise much of the total aggregate wealth in the United States and other economically advanced nations. During the last few decades, much energy has been devoted to the analysis of human capital and its empirical regularities. The result has been the accumulation of a large amount of evidence supporting the importance of human capital to the structure and evolution of earnings, occupations, employment and unemployment, fertility, and economic growth and development.

In recent years, a tendency to apply well-rehearsed concepts borrowed from the finance literature has become apparent:

- The famous Capital Asset Pricing Model (*Sharpe*, 1964; *Lintner*, 1965) has been enriched by the addition of human capital payoffs, yielding new insights, as in *Palacios-Huerta* (2003b).

- The equity return puzzle (*DeLong and Magin*, 2009) has found its counterpart in the human capital return puzzle (*Constantinides and Duffie*, 1996; *Judd*, 2000; *Palacios-Huerta*, 2001).
- The fundamental schism between equity and debt financing at the firm level has been taken up and adapted to the specific properties of human capital by *Jacobs and van Wijnbergen* (2007), *Vandenberghe and Debande* (2008) and *Cigno and Luporini* (2009).
- Capital mobility observed for financial markets also manifests itself on markets for human capital so that a strand of literature has developed researching the tax burden on highly skilled – and hence mobile – labor (*Elschner and Schwager*, 2007).
- Finally, the strands of finance and human capital literature have recently[2] been joined in *Pantzalis and Park* (2009).

Contribution 1

I have aimed to add to the human capital literature in the following ways: Contribution 1, published as *Weiss* (2009), tries to blend the insights gained in Business Administration – and comprehensively explained in my above cited book *Weiss* (2007) – with the wider economic literature which arrived at strikingly similar conclusions with different reasoning. The article by *Nielsen and Sørensen* (1997) is taken as the basis of an investigation that is overdue, given that Dual Income Taxes are discussed publicly in many countries (*Genser and Reutter*, 2007): How big a spread between capital and labor income tax rates should be applied by legislators to account for the fact that labor income is privileged under a traditional income tax? As it turns out, the spread is quite substantial and much wider than most legislators allow for. Reassuringly, the conclusions that I draw based on the article *Nielsen and Sørensen* (1997) are strikingly similar to the ones arrived at in *Weiss* (2007).

To put the results into perspective, it should be borne in mind that the focus of contribution 1 is substantially narrower than the one adopted in *Weiss* (2007). I restrict my analysis to the student population, instead of a cross-section of society, and trace their earnings history over time. A comparison with the group of workers who could have gone to university, but chose not to, yields insights into the opportunity costs of a university degree which consist mainly of forgone wages. At the same

[2] Earlier contributions investigating the effect of the death of the CEO on firm value (*Slovin and Sushka*, 1993) may be viewed as precursors to this development.

time, in a refinement concerning my methodology, I employ the relatively rarely[3] used Hausman-Taylor estimator (*Hausman and Taylor*, 1981). The improvement over the workhorse OLS estimator used in *Weiss* (2007) lies in the more advanced treatment of the apparent endogeneity of schooling choices with an instrumental variables strategy. The use of panel data – as opposed to cross-sectional data – gives the estimates more stability.

Contribution 2 Contribution 2 then traces the development of human capital further, by taking the education phase as finished and asking how an additional investment into human capital in the form of training decisions is treated under tax law. Given the huge body of literature dealing with training decisions (*Leuven*, 2005), the lack of contributions on the tax influence is somewhat astonishing. I fill this gap by estimating the return to training measures adopted in the period 2001 to 2003. My estimation results find these measures to be highly profitable. On top of that, tax law favors them in a similar manner as the – considerably larger – investments into human capital that are the subject of contribution 1. For the much shorter durations over which payoffs are received, the time effect does not assert itself as forcefully, but is still appreciable.

Contribution 3 While the first two contributions focus on the investment into human capital, contribution 3 investigates the taxation of the commuting kilometers that taxpayers ply on their way to work. Given that there is a human capital stock to exploit, taxpayers in many cases need to appear physically in the workplace to allow the transformation of their accumulated human capital stock into payoffs, and hence consumption, to go ahead. This final step has been a political hot potato for a long time and has consequently witnessed its fair share of volatility in recent years. This volatility is harnessed in contribution 3 to extract behavioral reactions to different treatments of the commuting kilometers.

The innovation provided by contribution 3 consists of the fact that it represents one of the few attempts to dig into the rather intractable combination of *tax base effects* and *labor income*. Also, the treatment evaluation literature (*Cameron and Trivedi*, 2005) has not been exploited as much as one might hope for in tax research in Germany. A huge body of literature exists, for instance, harnessing the Tax Reform

[3] The only recent application in the German context seems to be *Schneider* (2005).

Act 1986 in the United States to gain insights into the tax influence on microeconomic decisions (*Auerbach and Slemrod*, 1997; *Kubik*, 2004; *Kumar*, 2008). For Germany, there are a few contributions that exploit policy changes with regard to the corporation tax system in 2001, such as *Edwards, Lang, Maydew and Shackelford* (2004) and *Blasch and Weichenrieder* (2007). I provide an extension of these efforts to labor income where the drop in the deductible amount for commuting activities between the fiscal years 2003 and 2004 provides a natural experiment to identify the tax influence on commuting behavior. As it turns out, Germans react rather inelastically to these changes. Contribution 3 concludes the main part of this doctoral thesis.

Contribution 4

In contribution 4, published as *Jacob, Niemann and Weiss* (2008), Prof. Rainer Niemann, my colleague Martin Jacob and myself subject the working paper by *Bach, Corneo and Steiner* (2008) to a rigorous analysis. They claim that "the rich" do not adequately contribute to income tax revenue in Germany, and provide proof by employing a dataset merged from official tax statistics and the German Socio-Economic Panel (GSOEP). From this combined source, they derive an "effective average income tax rate" of around 32% for the richest 0.0001% quantile of taxpayers in Germany in the year 2002. Similar results are reported for other years and quantiles.

We refute these claims by broadening the research question to all major tax burdens weighing on the gross income generated by "the rich". One of the more obvious omissions is the *local trade tax burden* weighing on business income which "the rich" largely rely on. This reliance is proved by publicly available data (*Federal Statistical Office*, 2006, p. 13) which shows that over 96% of the taxpayers who generated an overall sum of income in excess of 5,000,000 € in 2002 reported business income. This income accounted for over 82% of their overall sum of income.

Other issues that biased the rates derived by *Bach et al.* (2008) include the use of cross-sectional data that, by their very nature, are not capable of allowing inference with regard to the treatment of tax losses or the role of accruals in the reported business income. These intertemporal considerations can only be evaluated in a panel context, but the necessary data collection efforts are still in their infancy. The tax rates that we end up deriving in the contribution conclusively upend the notion that "the rich" shirk their tax responsibilities. At the same time, we argue that the entire setup chosen by *Bach et al.* is bound to fail and that the existing data cannot tell us

much about the questions they sought to address.

Datasets Lastly, the datasets employed in this doctoral thesis should be mentioned briefly – much more comprehensive descriptions can be found in section 1.2.2, section 2.3.3 and section 3.3.2 for the GSOEP, and section 4.3.1 for FAST 2001. The dearth of datasets available to the tax research community is a major obstacle standing in the way of more empirical research in this area. Compared to other subjects, such as finance, which have very good *and* recent data sources at their disposal, the demands placed by *taxpayers* on tax authorities in terms of confidentiality see to it that access to official datasets carrying information pertaining to tax issues is severely restricted.

FAST 2001 The FAST dataset, compiled by the Federal Statistical Office of Germany and used in contribution 4, goes back all the way to the year 2001, when macroeconomic conditions were considerably different, and, in particular, tax legislation was about to undergo several influential transformations, such as a radical overhaul of the corporation tax system or the introduction of a final withholding tax on capital income later on. Furthermore, it is a cross-sectional dataset, and, as argued in contribution 4, this property severely constrains researchers in the kind of questions that they can actively pursue. On a more positive note, the *sample* that researchers are allowed to use in a remote location represents 10% of all tax returns filed for the fiscal year 2001, which amounts to approximately 3,000,000 entries. For the purposes of contribution 4, the full *population* of tax returns could be used, which is an extremely rare occurrence in empirical research.

GSOEP The GSOEP, on the other hand, provided by the DIW, Berlin, is a longitudinal study that covers a vast array of sociological subjects, with taxation being only one of many. Some degree of vagueness is thus inevitable when making inference on tax questions out of this dataset. However, the GSOEP makes up for these deficits with its two major advantages: It is topical, i.e. last year‘s data are normally available the next autumn, and it provides a true panel dataset, allowing the tax analyst to trace subjects over several years, and thus to make inference much more robust (*Baltagi*, 2008, section 1.2). Indeed, contribution 1 makes considerable use of this property. Quite apart from the ability to account for fixed effects over time, changes in behavior can, by definition, only be observed in a panel context, and contribution 3 owes its existence to this fact.

CONTRIBUTION 1

HIGHER TAX RATES ON LABOR? EVIDENCE FROM GERMAN PANEL DATA

(Published as: FinanzArchiv: Public Finance Analysis 2009, pp. 73-92)

1.1 Introduction

The adoption of Dual Income Taxes in the Nordic countries during the 1990s and more recent moves in this direction in other European countries have elicited a sizeable number of academic contributions. Among them, the article by *Nielsen and Sørensen* (1997) stands out as a comprehensive attempt to buttress the case for the Dual Income Tax (DIT). The theoretical idea expounded in their contribution states that the traditional taxation of labor income on the basis of cash-flows and the taxation of returns from physical capital on an accrual basis may lead to distortions in investment behavior which in turn harm economic efficiency. As a traditional income tax lumps together these two streams of payoffs and applies a common tax schedule to the resulting sum, the fundamental distinction between them – the difference in the determination of their respective tax bases – is blurred. The DIT, on the other hand, separates the two income types and typically applies a constant marginal tax rate to capital income while labor income is taxed progressively.
This setup inevitably prompts one to investigate the "correct" spread between the two tax rates. Empirically, this difference varies from country to country, as recently shown in *Genser and Reutter* (2007, table 2), a part of which is reproduced in table 1.1 for convenience.

TABLE 1.1: Tax Rates in the Nordic Countries (in %)

Country	**Norway**	**Finland**	**Sweden**	**Denmark**
Implementation of DIT	1992	1993	1991	1987
Income Tax Rate for				
Capital Income	28	28	30	28/43
Earned Income	28-40	26.5-50	31.6-56.6	38.8-47.9

Source: *Genser and Reutter* (2007), Table 2

Table 1.1 highlights the fact that the approach towards the relationship between capital and labor ("Earned") income tax rates is nonuniform across the four Nordic countries. While the Norwegian tax system lets "labor income taxation begin where capital income taxation ends", the Finnish system sets the lowest labor income tax rate below the rate for capital income while the Swedish legislator reverses this relationship. With regard to the highest applicable marginal rate on labor income, the degree of the spread between the rates for top earners differs widely as well. While these observations are certainly incomplete in the sense that the determination of the

respective tax bases exerts a strong influence on the effective tax burden weighing on the economic activities of the taxpayer, they do show a rather surprising degree of variation even though the Nordic countries are relatively homogeneous in other economic aspects[1].

In Germany, the introduction of a final withholding tax for capital income in the year 2009, coupled with the continuing progressive taxation of labor income, can be viewed as a step toward the Dual Income Tax in all but name. The German capital income tax rate has been set at 25%, while labor income is taxed at marginal rates ranging from 15% to 45%. German tax law also allows taxpayers to elect to have their capital income taxed at their individual marginal tax rate if it is lower than 25%.

The goal of this contribution is the investigation of the "correct" spread between labor and capital income tax rates under a Dual Income Tax under the premises of the *Nielsen and Sørensen* (1997)-Model. This agenda is thus fairly narrowly focused on an empirical investigation and does not seek to provide a new reasoning for the results established by *Nielsen and Sørensen* (op. cit.). *Nielsen and Sørensen* build an "overlapping generations model where consumers face a trade-off between investment in human capital and investment in non-human capital" (p. 311, op. cit.). Set in a small open economy, with perfect foresight and perfect competition, the legislator has already committed himself to tax capital income at a constant marginal rate and now has to determine the appropriate taxation of labor income[2]. The "cash-flow treatment of human capital investment" is responsible for a distortion under a conventional income tax that slaps equal marginal tax rates on labor and capital income. *Nielsen and Sørensen* (op. cit.) argue that a justification for differential tax treatment of labor and capital income lies in the fact that the social and private after-tax rates of return coincide in the case of labor income while the taxation of capital income drives the private rate of return below the social rate of return. As their equation (5) shows[3], agents' optimization behavior leads to the usual prescription of equal marginal after-tax returns on human and non-human capital investment, which in turn must equal the rate of time preference. As taxation does

[1] Cf. *Elschner and Schwager* (2007) who classify the Scandinavian countries as uniformly high tax with regard to labor taxation.

[2] The *Nielsen and Sørensen*-Model thus conveys a "... typical second-best argument" (*Nielsen and Sørensen*, 1997, p. 313) in that the commitment to tax capital income is taken as given. I thank an anonymous referee for stressing this point.

[3] *Nielsen and Sørensen* (1997, p. 317).

not take a bite out of the return on human capital investment, overinvestment in human capital results and a surcharge is applied to the labor income tax rate above a certain threshold to counteract this effect.

Applying this model to real world data necessitates adaptions that may be open to criticism:

- Workers are not as homogeneous with regard to their ability to accumulate human capital as the *Nielsen and Sørensen*-Model envisions. The human capital production function $g(E)$ (*Nielsen and Sørensen*, 1997, p. 316) should be indexed $g_i(E)$, with i representing differently gifted brackets of society.

- The model is set in an OLG context, i.e. lives off the contrast between a young and old generation. The dividing line in the data between these generations is drawn at the time individuals enter the labor market after their formal education ends. In practice, the dividing line is blurred by the fact that lifelong learning is commonplace so that a formal end to education cannot be reliably determined[4].

- Costs of education in the *Nielsen and Sørensen*-Model come exclusively as opportunity costs during the education phase. Estimating these for differently gifted workers is challenging. It requires the specification of a counterfactual which constitutes another source of uncertainty.

The rest of the paper is organized as follows: In section 1.2, I describe my dataset and discuss the estimation strategy. The eight waves from 2000 to 2007 of the German Socio-Economic Panel are employed to estimate a Mincer-type wage equation. From the estimation results, the empirical age-earnings profiles for different education brackets of the German population are deduced. In section 1.3, the necessary surcharges to the labor income tax rate are computed under different constellations of parameters for the simple tax system envisioned by *Nielsen and Sørensen* (1997). Section 1.4 concludes.

[4] In the Mincer earnings equation, labor market experience and its square as well as tenure account for the effects of general and firm-specific human capital investments. I thank an anonymous referee for alerting me to this point.

1.2 Estimation of the Wage Equation

1.2.1 Estimation Strategy

I estimate a standard wage equation in a panel context based on the contribution by *Mincer* (1974),

$$\log(Y_{it}) = \alpha + \beta_1 S_i + \beta_2 EX_{it} + \beta_3 EX_{it}^2 + \beta_4 TE_{it} + \sum_{c=2}^{C} \gamma_c B_{ic} + W_{it}\psi + u_i + \varepsilon_{it} \quad (1.1)$$

where $i = 1, ..., N$ stands for individuals and $t = 1, ..., T$ denotes time periods. Y_{it} stands for labor income while B_{ic} represents cohort dummies for birth cohorts c, S_i stands for the (time-invariant) years of schooling, TE_{it} denotes the tenure with the current firm, EX_{it} represents (actual) labor market experience, u_i stands for a unit-specific effect and ε_{it} represents the usual idiosyncratic shock. The matrix W_{it} gathers further characteristics of the individuals, such as marital status, the blue-collar/white-collar distinction, health status or workplace autonomy.

The literature on the returns to education – as embodied in the coefficient β_1 in equation (1.1) – has evolved in several waves, with cross-sectional analyses employing OLS estimators dominating the first one. As the return to education is likely to be driven by unobservable characteristics of the individual, such as innate ability or stamina, the schooling variable is not orthogonal to the error term and the exogeneity assumption underlying the OLS estimator is violated: the resulting estimates are not even consistent. To account for this problem, the second wave of contributions proposed instrumental variable (IV) strategies, as surveyed by *Card* (2001). These strategies suffered from two major shortcomings, namely the lack of suitable instruments and the weak correlation of the existing instruments with the endogenous regressors (cf. *Ichino and Winter-Ebmer*, 1999).

The existence of suitable panel data allows one to get around the problem. Fixed effects (FE) regressions allow correlation between the regressors and the individual effect u_i, eliminating the bias inherent in OLS estimations of the wage equation, but cannot provide point estimates of time-invariant factors (such as length of schooling once an individual leaves school). Random effects (RE) models deliver such estimates, yet assume that the covariates are uncorrelated with the individual effects u_i and the idiosyncratic shocks ε_{it}. The polar cases of FE and RE estimation entail

drawbacks in terms of estimation output or in terms of almost untenable assumptions, with the choice between them often hinging on the Hausman specification test (*Hausman*, 1978).

The "all-or-nothing decision" (*Baltagi, Bresson and Pirotte*, 2003, p. 361) implied by an adoption of the FE or RE estimator was enriched by a third alternative formulated in *Hausman and Taylor* (1981). *Hausman and Taylor* developed the Efficient Generalized Instrumental Variables (EGIV) estimator which allows one to control for the correlation between individual effects and the regressors and at the same time to obtain estimates for the coefficients of time-invariant covariates[5]. Departing from the model

$$\log(Y_{it}) = x'_{it}\beta + z'_i\alpha + \eta_{it} \tag{1.2}$$

where x_{it} denotes the matrix of time-varying covariates and z_i the matrix of time-invariant regressors, these matrices are split, with the part x'_{1it} containing the exogenous, time-varying regressors, x'_{2it} the endogenous, time-varying regressors, z'_{1i} the exogenous, time-invariant regressors, z'_{2i} the endogenous, time-invariant regressors. The error term η_{it} is decomposed into an individual-specific part u_i and an idiosyncratic part ε_{it} so that equation (1.2) becomes

$$\log(Y_{it}) = x'_{1it}\beta_1 + x'_{2it}\beta_2 + z'_{1i}\alpha_1 + z'_{2i}\alpha_2 + u_i + \varepsilon_{it} \tag{1.3}$$

with the components of the error term independently and identically distributed $(0, \sigma_u^2)$ and $(0, \sigma_\varepsilon^2)$, respectively, and their conditional variance
$Var\left(\varepsilon_{it} + u_i | \, x_{it}, z_i\right) = \sigma_\varepsilon^2 + \sigma_u^2 = \sigma_\eta^2$.

The further technical implementation is extensively described in *Greene* (2008, chap. 12.8). Crucially, in a situation where the analyst suspects correlation between covariates and the individual effects u_i, the HT estimator relieves the analyst of the duty to find external instruments[6]. Instead, it constructs instruments from within the model, as it employs the time varying exogenous variables gathered in x'_{1it} "... twice, once as averages and another time as deviations from these averages" (*Baltagi*, 2008, p. 134).

[5] The contribution *Hausman and Taylor* (1981) explicitly concerns the estimation of a wage equation based on panel data.

[6] Note that a suspected correlation with the idiosyncratic part of the error term ϵ_{it} would call for external instruments.

The quality of the instruments thus obtained must be subjected to the usual test of overidentifying restrictions as the above mentioned problem of weak correlation can bite here as well. A common approach in the literature employs the statistic developed by *Sargan* (1958) to prove the legitimacy of the instruments, as in, for instance, *Baltagi* (2008, p. 135) and *Kalwij* (2000, p. 66). The Sargan test statistic is distributed χ^2 under the null, with degrees of freedom – in the case of the Hausman-Taylor estimator – equal to the difference between the number of time-varying exogenous covariates (x'_{1it} in equation (1.3)) and the number of time-invariant endogenous covariates (z'_{2i})[7].

1.2.2 Dataset

The dataset comes[8] from the last eight waves of the German Socio-Economic Panel, conducted by the DIW, Berlin[9]. The estimation period ranges from 2000 to 2007. I restrict my estimation to German male dependent workers aged 18 to 65 years. To abstract from unemployment, I further restrict the estimation sample to workers who worked full time and have no missing data during the entire estimation period. This yields 1,629 valid observations for each of the eight years. The prevailing cohort strengths can be gauged from table 1.2.

The dependent variable $log(Y_{it})$ in equation (1.1) is the (natural) logarithm of gross labor earnings in the respective year t, unadjusted for inflation, which was exceptionally low during the estimation period in Germany, averaging about 1.6%. I include Christmas bonuses, holiday and performance pay in the calculation of gross compensation.

The covariates are

- Education time S_i which is calculated according to the standard durations for German school degrees, as contained in table 1.3: A high school degree holder with a completed university degree would thus be credited with 18

[7] As a corollary, the test can be calculated only in the overidentified case.

[8] The data used in this paper were extracted using the Add-On package PanelWhiz v2.0 (Nov 2007) for Stata. PanelWhiz was written by Dr. John P. Haisken-DeNew (john@panelwhiz.eu). The PanelWhiz generated DO file to retrieve the SOEP data used here and any Panelwhiz Plugins are available upon request. Any data or computational errors in this paper are my own. *Haisken-DeNew and Hahn* (2006) describe PanelWhiz in detail.

[9] For comprehensive information on the GSOEP, cf. *Haisken-DeNew and Frick* (2005).

TABLE 1.2: Cohort Strengths in the Dataset

Birth Cohort	No.	Percent	Cumul. Percent
1940-1944	59	3.62	3.62
1945-1949	136	8.35	11.97
1950-1954	233	14.30	26.27
1955-1959	301	18.48	44.75
1960-1964	323	19.83	64.58
1965-1969	299	18.35	82.93
1970-1974	176	10.80	93.74
1975-1979	96	5.89	99.63
1980-1984	6	0.37	100.00
Total	1,629	100	

Source: German Male Full-Time Workers, drawn from the GSOEP 2000-2007

years of education, 13 for his high school degree and five for completing university[10].

- Actual labor market experience EX_{it} is obtained from the GSOEP data which explicitly deliver the lengths of employment and unemployment spells.
- Tenure with the current employer, as reported in the GSOEP data, and the CNEF One Digit Industry Code[11].
- A set of cohort dummies, bundling five cohorts each. To avoid collinearity problems, the cohort 1940 to 1944 is omitted from the estimation.
- Dummy variables denoting marital status, bad health, blue-collar work, and a dummy for public servants.

I allocate these covariates into the vectors employed in equation (1.3) as follows:

- Tenure, actual labor market experience and its square enter as time varying endogenous variables gathered in x'_{2it}
- Cohort dummies and the public service dummy enter as time-invariant exogenous variables (z'_{1i})

[10] These standard durations are supplied by the data provider, the DIW in Berlin (*Haisken-DeNew and Frick*, 2005, p. 69).

[11] Cf. *Haisken-DeNew and Frick* (2005, p. 72).

TABLE 1.3: Conversion of Degrees into Imputed Educational Times S_i (in Years)

Schooling		$\mathbf{S_i}$
No Degree	-	7
Lower School Degree	Hauptschule	9
Intermediary School	Realschule	10
Degree For A Professional College	Fachhochschulreife	12
High School Degree	Hochschulreife	13
Additional Occupational Training		$\mathbf{S_i}$
Apprenticeship	Lehre	1.5
University Degree	Universität	5

Degree names are additionally given in German for ease of interpretation.

Source: *Haisken-DeNew and Frick* (2005, p. 69)

- The schooling duration S_i enters as a time-invariant endogenous variable (z'_{2i})

The remaining covariates are collected in x'_{1it}. I discuss this choice in conjunction with my results below.

1.2.3 Estimation Results

Table 1.4 gives the results of the Hausman-Taylor estimation of equation (1.1). The return to schooling is estimated at 10.50% per additional year of education, with a 95% confidence interval for this coefficient ranging from 7.56% to 13.44%.
These results can be compared against recent findings for German panel data. *Boockmann and Steiner* (2006) conduct a *random effects estimation* and put particular emphasis on cohort effects, thus addressing a set of questions different from the ones tackled in this paper[12]. Their dataset consists of GSOEP data from 1984 to 1997, and a digest of their findings translated into rates of return can be found in *Boockmann and Steiner* (2006, p. 1150). At least the finding of a return for German men with a university degree of 9.77% is roughly in line with the results in this paper. In terms of the tenure variable, the insignificance found in *Boockmann and Steiner* is also present in table 1.4, as well as the magnitude of the coefficients for labor market experience and its square.

[12] I did experiment with the cohort effects estimated by *Boockmann and Steiner* (2006), and found no significant influence. A joint Wald test of interaction terms between the education variables and the eight cohort dummies came in at 12.02 against a critical value at the 95% confidence level of $\chi^2_8 = 15.507$.

TABLE 1.4: Estimation Results for Equation (1.1)

Hausman-Taylor Estimator		
Bad Health	-0.0069	(0.0040)*
Blue-Collar Worker	-0.0221	(0.0095)**
Marital Status	-0.0118	(0.0050)**
CNEF Industry Code	-0.0023	(0.0027)
Experience	0.0365	(0.0026)***
Experience Squared	-0.0006	(0.0001)***
Tenure	0.0015	(0.0010)
Cohort 1945-1949	0.0746	(0.1729)
Cohort 1950-1954	0.1109	(0.1670)
Cohort 1955-1959	0.1800	(0.1627)
Cohort 1960-1964	0.3695	(0.1673)**
Cohort 1965-1969	0.4102	(0.1633)***
Cohort 1970-1974	0.4826	(0.1629)***
Cohort 1975-1979	0.5384	(0.1709)***
Cohort 1980-1984	0.5346	(0.2593)**
Civil Service	0.0339	(0.0457)
Amount of Education	0.1050	(0.0150)***
Constant	8.0603	(0.6292)***
σ_u	0.352	
σ_e	0.162	
ρ	0.825	
Sargan df	3	
Sargan Critical Value	7.815	
Sargan χ^2	2.503	
Sargan p-Value	0.475	

Analytic standard errors in parentheses

* p $<$ 0.10, ** p $<$ 0.05, *** p $<$ 0.01

Omitted Category: Cohort 1940-1944

Marital Status coded 1 for "married",

0 for "not married"

ρ denotes the fraction of the estimated

residual variance attributable to σ_u^2

Source: German Male Full-Time Workers,

drawn from the GSOEP 2000-2007

Examples of the adoption of the *Hausman-Taylor approach* for other countries are provided in *Kalwij* (2000) who employs panel data to measure the return to schooling in the Netherlands in the 1980s while *García-Mainar and Montuenga-Gómez* (2005) utilize the European Community Household Panel to estimate returns to education in Spain and Portugal in the 1990s. The significant differences between the OLS and EGIV results are presented in Table 2 of *Kalwij* (2000) and Tables 3 and 4 of *García-Mainar and Montuenga-Gómez* (2005) where the return to schooling increases from an OLS estimate of 6.9% and 2.4% to an EGIV estimate of 14.8% and 9.5%, respectively. An earlier study by *Baltagi and Khanti-Akom* (1990) which compared several variants of the Hausman-Taylor estimator with FE results using the "Panel Study of Income Dynamics" and was conducted using data featuring a comparable time dimension (1976-1982), yielded a range of coefficients for the education variable of 13.72% – 14.18%, depending on whether the original Hausman-Taylor estimator or the later enhancement by *Amemiya and MaCurdy* (1986) or the one by *Breusch, Mizon and Schmidt* (1989) was used. Thus, the point estimate presented in this paper finds itself within reach of the results established by comparable estimation strategies.

With regard to the precision of the estimates, the standard error of the coefficient for the schooling variable is remarkably low in comparison to, for instance, the ones reported in *Hausman and Taylor* (1981, p. 1392) or *Baltagi and Khanti-Akom* (1990, p. 404). The results reported in the latter contribution result in a confidence interval from 9.63% to 17.94%. The radical data selection strategy described in section 1.2.2 certainly plays a major role here: Abstracting from unemployment, which was particularly high during the estimation period, makes for low standard errors in the estimation output. On the other hand, the groups germane to the further discussion in section 1.3, i.e. university graduates and high school degree holders, tend to have lower than average unemployment rates.

Concerning the validity of the estimation strategy, a first check is a Hausman test of the within estimator of equation (1.1) against the random effects estimator which returned a test statistic of 48.93 which is distributed χ^2_7 under the null, with the critical value for the 95% confidence level at 14.067. The test rejects the additional orthogonality assumptions made by the RE estimator, i.e. the lack of correlation between the regressors and the individual effects, and suggests the application of an IV strategy such as the Hausman-Taylor estimator.

With regard to instrument quality for the purposes of this estimator, the Sargan

statistic[13] is reported at the bottom of table 1.4. The test statistic comes in at 2.503, far below the critical value of $\chi^2_3 = 7.814$. The null cannot be rejected which supports the case for the chosen IV strategy.

Other classifications into exogenous and endogenous variables, i.e. movements between x'_{it} and z'_i in equation (1.2), as conducted, for instance, in *Baltagi and Khanti-Akom* (1990, pp. 402-404), brought about worse results: The removal of labor market experience, its square and tenure from the set of endogenous variables, for instance, led to a Sargan statistic of 19.137, with the critical value[14] at $\chi^2_6 = 12.592$. Other tentative moves between the classifications for the covariates did not improve on the constellation reported in table 1.4, either.

1.3 Results

As the *Nielsen and Sørensen*-Model "postulates that the opportunity cost of education and training consists solely of foregone labor income" (*Nielsen and Sørensen*, 1997, p. 322), the calculation of rates of return on human capital investments requires the determination of opportunity costs, i.e. labor income earned without attending a certain educational program. To this end, I select the group of university graduates in the sample and define the earnings increment attributable to their university degree as the excess return over the group of high school degree holders who chose not to attend university. Setting the counterfactual in this fashion, I pit two relatively homogeneous groups against each other to minimize the selection bias that manifests itself in the choice of a university education: More able students are more likely to attend university than their less able peers. I experimented with treatment evaluation models in the spirit of *Cameron and Trivedi* (2005, chap. 25) and found that they rejected the notion of systematic differences between the two groups[15]. Furthermore, this setup enables a clear-cut vision of the kind of "... trade-off between accumulation of financial capital and accumulation of human capital..." (*Nielsen and Sørensen*, 1997, p. 314) envisioned in the *Nielsen*

[13] All estimation results are obtained from Stata 10.1. The overidentification test was conducted in the package xtoverid (*Schaffer and Stillman*, 2006).

[14] The removal of time-varying covariates from the vector of endogenous regressors adds them to the exogenous ones and thus leads to the increase in degrees of freedom for the Sargan statistic.

[15] *Collins and Davies* (2004) follow a cruder approach to calculate rates of return for the Canadian education system during the 1990s: They compare the earnings of university graduates to the median earnings of high school graduates.

and Sørensen-Model because the university degree is a particularly well-defined and important educational program.

On the other hand, the particular results derived on the back of this setup are primarily relevant to the comparison at hand, i.e. between high school and university graduates. If the comparison was conducted between two other groups, the outcome would certainly be different: The required spread is contingent on the comparison undertaken. Yet the thrust of the *Nielsen and Sørensen*-Model, the requirement for higher tax rates on labor income from an efficiency standpoint, can be confirmed. At the same time, the elusiveness of a definitive result for the required surcharge becomes apparent. On top of that, the result hinges on assumptions about long-term interest rates and several parameters of the *Nielsen and Sørensen* tax system, as well. I carry out the relevant sensitivity analyses below[16].

I predict the conditional mean of the earnings in year t from the receipt of the final educational degree onwards on the basis of my results in table 1.4 as

$$E_t = \exp\left(\hat{\alpha} + \hat{\beta}_1 S + \hat{\beta}_2 EX_t + \hat{\beta}_3 EX_t^2 + \sum_{c=2}^{C} \hat{\gamma}_c B_c + \frac{1}{2}\left(\hat{\sigma}_u^2 + \hat{\sigma}_e^2\right)\right) \tag{1.4}$$

where a hat stands for the point estimates from table 1.4, E_t denotes the mean of the earnings distribution, conditional on the covariates, and I assume that the error term in equation (1.1) is approximately normally distributed[17]. I omit the tenure variable from the calculations due to insignificance. To illustrate the result, figure 1.1 plots the conditional means of the earnings for university and high school graduates against age, up to the assumed retirement age of 65 years, for the birth cohort 1965 - 1969.

Subsequently, I calculate the public/social rate of return for a university education as the solution r_{pu}^* to

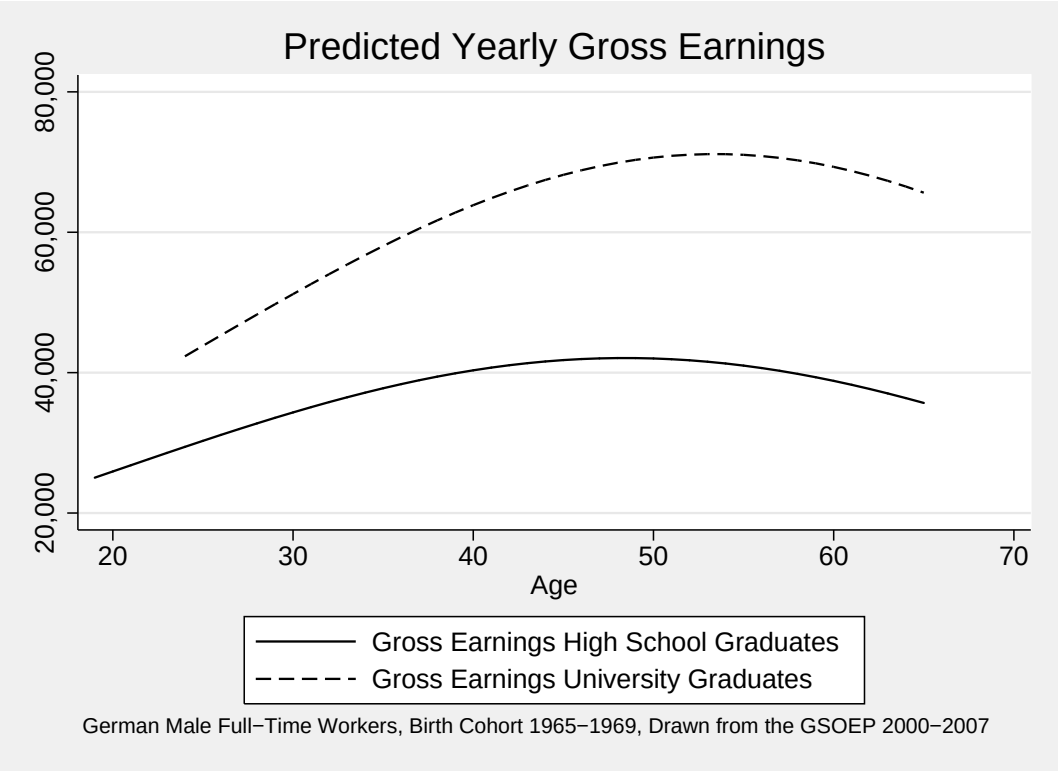
$$\sum_{t=19}^{23} \frac{E_t^{HS}}{(1+r_{pu})^t} \stackrel{!}{=} \sum_{t=24}^{65} \frac{E_t^{U} - E_t^{HS}}{(1+r_{pu})^t} \tag{1.5}$$

where the LHS represents the present value of foregone earnings of the individuals choosing a university education, i.e. the earnings of a worker with a high

[16] See table 1.5.

[17] A Shapiro-Wilk normality test on the errors from the regression in table 1.4 did not reject the null of normal distribution (p=0.176).

FIGURE 1.1: Prediction of Mean Gross Earnings for University and High School Graduates, Birth Cohort 1965-1969



school degree E_t^{HS}, and the RHS stands for the earnings increments of the university graduates over their peers, calculated over the remaining 41 years of their labor market participation. Equation (1.5) is designed to mimic the calculation of the public/social rate of return to human capital investment that lies at the heart of the *Nielsen and Sørensen*-Model.

To derive the private rate of return, I introduce the simple tax system of *Nielsen and Sørensen* (1997, p. 316), with a marginal tax rate t_1 applicable to labor income below a threshold Ω, and a rate τ for capital income. I constrain the lower labor income tax rate and the capital income tax rate to be equal, $t_1 = \tau$, in accordance with *Nielsen and Sørensen* (1997)[18]. A rate $t_2 > t_1$ kicks in for labor income above Ω. In the comparison of university graduates and high school degree holders, Ω is initially set equal to 25,000 €. I vary this rate below to check for the sensitivity to this assumption. The private rate of return r_{pr}^* is then computed as the solution to

[18] *Nielsen and Sørensen* (1997) show in a technical appendix that their result holds even in the absence of this condition.

$$\sum_{t=19}^{23} \frac{(1-t_1)\times \min\left(E_t^{HS};25,000\right)+(1-t_2)\times \max\left(E_t^{HS}-25,000;0\right)}{(1+r_{pr})^t} \stackrel{!}{=}$$
$$\sum_{t=24}^{65} \frac{(1-t_1)\times \min\left(E_t^{U};25,000\right)+(1-t_2)\times \max\left(E_t^{U}-25,000;0\right)}{(1+r_{pr})^t} -$$
$$\sum_{t=24}^{65} \frac{(1-t_1)\times \min\left(E_t^{HS};25,000\right)+(1-t_2)\times \max\left(E_t^{HS}-25,000;0\right)}{(1+r_{pr})^t} \tag{1.6}$$

The first row of equation (1.6) shows the after-tax opportunity costs for a university student: The wages obtainable in the labor market for a high school graduate are cut by the lower rate t_1 as long as they remain below $\Omega = 25,000$ €, while above that level the more onerous tax rate t_2 bites. The second row shows the after-tax values of the earnings of the university graduate. In the third row, the after-tax earnings of the high school graduate are subtracted to arrive at the excess earnings for the university graduate from the receipt of his degree to retirement.

Under the tax system described above, the private rate of return to investment in financial assets equals $i \times (1-\tau)$, where i denotes the prevailing gross interest rate. The idea at the core of *Nielsen and Sørensen* (1997) is that, given the (distortionary) taxation of capital income, this rate must coincide with the private rate of return on human capital investment r_{pr}^* to avoid distortions in investment behavior. To this end, the rate t_2 in equation (1.6) is raised until r_{pr}^* hits $i \times (1-\tau)$.

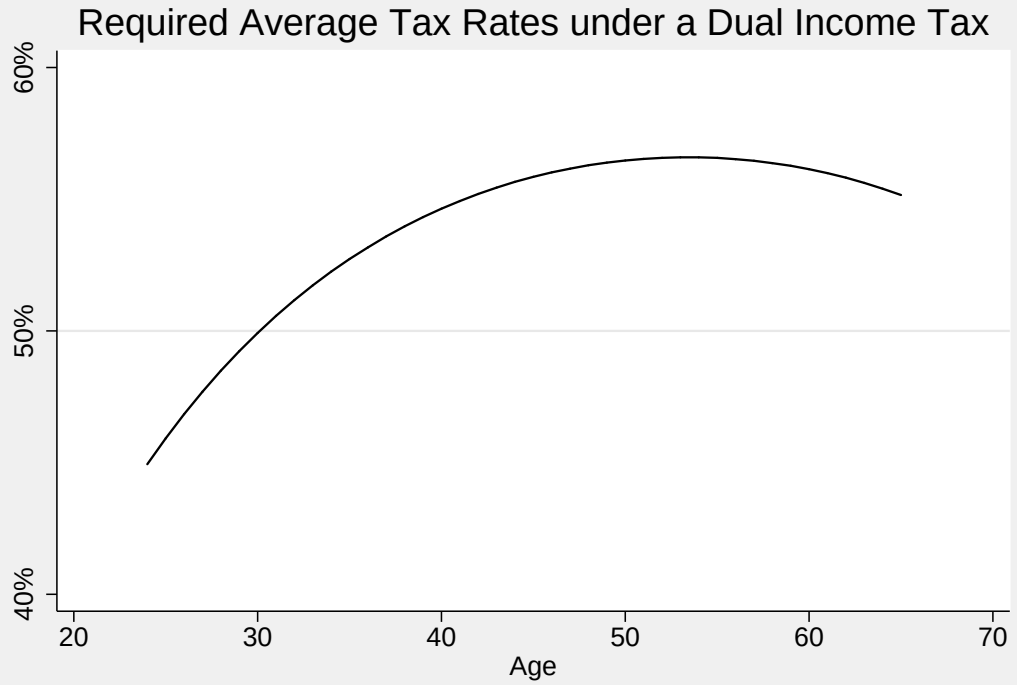
For the calculations, I continue the example above and compute r_{pu}^* and r_{pr}^* according to equation (1.5) and equation (1.6), respectively. The tax rate $t_1 = \tau$ is set at 25%. Assuming a gross interest rate of 6%, the goal is to lower the private rate of return on human capital to $6\% \times (1-25\%) = 4.5\%$. I iteratively raise t_2 in equation (1.6) until r_{pr}^* hits this value. This procedure yields a t_2 of 73.71%. Subsequently, one can derive the *average* tax rate on the labor income t^a of the university graduate in figure 1.1 resulting from the tax system outlined above according to

$$t^a = \frac{t_1 \times \min\left(E_t^{U};25,000\right)+t_2\times \max\left(E_t^{U}-25,000;0\right)}{E_t^{U}} \tag{1.7}$$

Figure 1.2 shows the development of the required average tax rates on labor income under the assumptions made above. To align the private return to his human capital

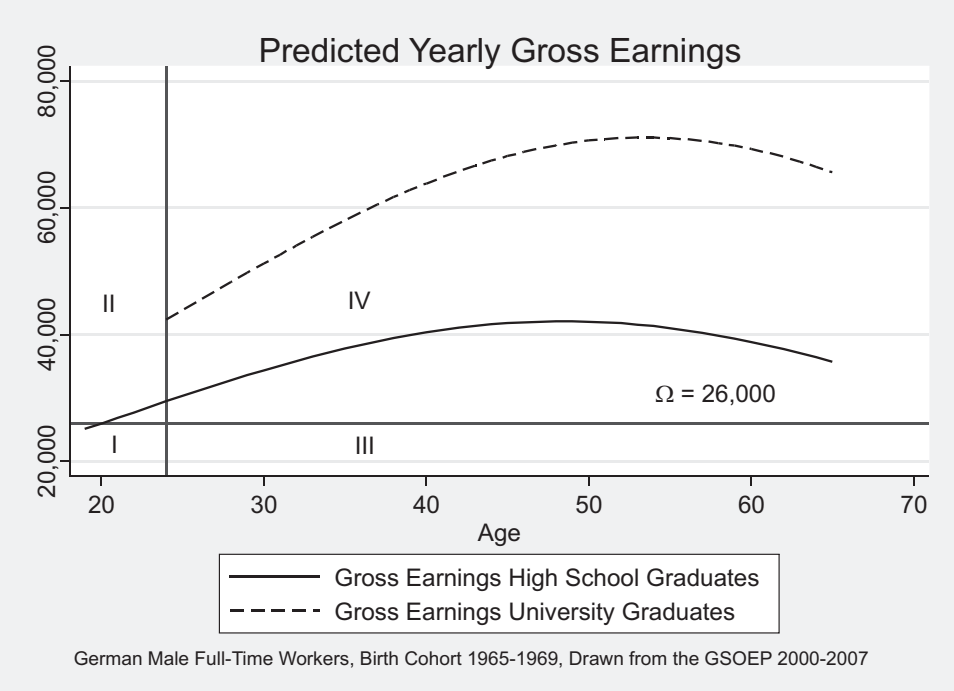
investment, the average tax rate for the university degree holder must unambiguously exceed the – assumed – tax rate of 25% for financial capital. The tax rate is age dependent because marginal labor income above the – immovable – threshold of $\Omega = 25,000$ € is subject to the higher tax rate t_2. Increased earnings thus drive up the weighted average rate in equation (1.7). The required tax rates end up mirroring the development of the gross earnings in figure 1.1.

FIGURE 1.2: Required Average Tax Rates for a University Graduate under a Dual Income Tax, for i=6%, Ω=25,000 € and $t_1 = \tau = 25\%$



To check for the sensitivity of the result with regard to the structure of the tax system, I vary the threshold Ω in the following and report the resulting t_2 rate that aligns the private rates of return on human and financial capital investment. To this end, consider the situation depicted in figure 1.3, where Ω is set equal to 26,000 €. In the left part of the graph – areas I and II –, the after-tax opportunity costs of a university degree, the after-tax wages obtainable between the age of 19 and 24, are depicted. In the after-tax perspective, the bulk of those costs is cut by the low tax rate t_1 (area I), raising their net burden, while only the small fraction in area II is

FIGURE 1.3: Illustration of the *Nielsen and Sørensen* Tax System for a Threshold Ω=26,000 €



cut by t_2. In the right part of the graph – areas III and IV –, both E_t^U and E_t^{HS} exhaust the maximum allowance for t_1. Their relative position is thus unchanged, and t_2 equals 73.22%. Raising the threshold to 28,000 €, the fraction of lower taxed opportunity costs in area I rises, while the situation in areas III and IV remains unchanged. More burdensome opportunity costs translate into a lower private rate of return which must be pushed up by a milder t_2, inducing a drop from 73.22% to 72.69%. A further hike in Ω to 30,000 € increases the net opportunity costs to their maximum, as now E_t^{HS} remains below the threshold in areas I and II and t_2 is lowered to 72.44%. From this point onwards, a rising threshold does not impact opportunity costs anymore, but merely subjects an ever larger part of the difference $E_t^U - E_t^{HS}$ to the lower tax rate t_1, raising the private rate of return. At Ω =35,000 €, for instance, t_2 must be raised to 74.83% to counteract this effect.

It is now straightforward to predict the behavior of the system for a changing t_1. A higher marginal rate for low incomes, at low thresholds Ω, will tend to lower

opportunity costs and leave the benefits of education unchanged, altogether driving up the private rate of return and triggering a higher t_2. At higher thresholds, a marginal increase in t_1 reduces both opportunity costs and benefits. The latter are discounted more heavily in equation (1.5), and their impact outweighs that of the benefits. In the same vein, a higher interest rate i would raise the required return on human capital investment r^*_{pr}, which calls for a lower tax rate t_2. Table 1.5 summarizes the behavior of t_2 just described for discreet changes in the exogenous parameters.

TABLE 1.5: Required Marginal Tax Rates t_2 under the *Nielsen and Sørensen*-Model, for Varying Interest Rates i, Tax Rates t_1 and Thresholds Ω

i				**4%**	
$\mathbf{t_1 = \tau}$		**10%**	**20%**	**30%**	**40%**
Ω	**10,000**	87.30%	89.63%	91.90%	93.69%
	20,000	77.83%	81.79%	85.50%	88.69%
	30,000	72.65%	77.72%	82.42%	85.75%
	40,000	83.30%	86.92%	89.65%	92.52%
i				**6%**	
$\mathbf{t_1 = \tau}$		**10%**	**20%**	**30%**	**40%**
Ω	**10,000**	79.84%	84.78%	88.53%	91.48%
	20,000	67.66%	74.17%	80.20%	85.50%
	30,000	61.71%	69.69%	75.78%	81.77%
	40,000	72.20%	78.96%	84.14%	88.72%
i				**8%**	
$\mathbf{t_1 = \tau}$		**10%**	**20%**	**30%**	**40%**
Ω	**10,000**	69.38%	77.08%	83.68%	88.56%
	20,000	53.61%	64.50%	73.39%	80.70%
	30,000	47.23%	59.10%	69.13%	76.59%
	40,000	57.27%	68.76%	76.82%	83.84%

Source: Own Calculations

1.4 Conclusion

This paper has attempted to estimate the extent of the required divergence between the tax rates on capital and labor income emanating from the model of *Nielsen and Sørensen* (1997). To this end, a Mincer-type earnings equation was estimated for a

panel of German workers in the years 2000 to 2007, employing the EGIV estimator proposed by *Hausman and Taylor* (1981). This approach takes into account the correlation between the regressors and the individual effects, entailing substantial advantages over competing IV-approaches which suffer from lack of suitable instruments. The *Hausman and Taylor* approach also trumps the competing fixed effects model in that it explicitly enables the user to estimate the impact of time-invariant endogenous covariates – such as schooling. It thus appears ideally suited for the task at hand.

The estimate of a return of 10.50% for every year of schooling finds itself within the range of results established in the literature, both in a national and international comparison. It is utilized to predict the conditional mean of the wages of differently educated brackets of German workers. Among the available groups, the relatively homogeneous one of High School Graduates is selected to emulate the setup of the *Nielsen and Sørensen*-Model. The marginal costs and benefits of a university graduation translate into a private rate of return which would equal the public rate of return if a proportional tax rate was applied. The taxation of the return to financial capital assumed in the *Nielsen and Sørensen*-Model necessitates measures to lower the private rate of return to human capital investments to the level of the competing financial investment. To this end, a surcharge is added to the labor tax rate above a certain threshold to bring the private rate of return on human capital investments into line with the private rate of return on financial capital. A special, yet relatively frequent educational biography – that of a university graduate – thus proxies for the stylized world spawned by the *Nielsen and Sørensen*-Model. Although the results cannot be readily generalized to the overall population, they do yield clues as to the extent of the necessary correction envisioned by *Nielsen and Sørensen*.

The extent of the surcharge for the case of university graduates is substantial, depending on the design of the tax system, i.e. the boundary between the higher and the lower labor income bracket, the tax rate for the lower labor income bracket and the interest rate. For realistic cases, i.e. a kink in the marginal tax rates at the threshold of 20,000 €[19], a lower marginal rate of 25% and a gross interest rate of 6%, the

[19] Checking back with the original *Nielsen and Sørensen*-Model, the threshold Ω is supposed to distinguish between "...'low' income from (unskilled) labor..." and earnings above that for "...skilled workers..." (*Nielsen and Sørensen*, 1997, p. 316). Taking a cue from table 1.3 and table 1.4, the unskilled can be represented by the "No Degree" group (credited with 7 years of schooling) who earn their keep primarily from manual labor. Their labor earnings are predicted between 13,000 € and 25,000 € so that a threshold of 20,000 € seems sensible.

necessary average tax rates range between 48% and 66% and thus by far exceed those applicable under German tax law which stipulates a top *marginal* income tax rate of 45% for labor income in excess of 250.000 €.
Importantly, this contribution has been written under the assumptions of the basic *Nielsen and Sørensen*-Model, in particular the assumption of an exogenous choice of leisure. *Nielsen and Sørensen* explore the case of an endogenous labor/leisure decision in their section 8, and find ambiguous results. Yet under plausible assumptions, the case for a surcharge to the tax rate on higher labor income remains intact.

Stepping out from the *Nielsen and Sørensen*-Model, it is important to realize that their model, and its empirical implementation above, is exclusively focused on time effects in the relationship between labor and capital income taxation. Quite apart from higher marginal tax rates, the effect of social security and pension contributions impacts labor's effective tax burden. Some of these contributions can be classified as taxes, e.g. on the basis of their actuarial unfairness, and the analysis presented here is thus complementary to a recent contribution by *Elschner and Schwager* (2007) who examine effective tax rates on the labor income of highly qualified employees, a group that certainly overlaps with the university graduates analyzed here.
Furthermore, tax base effects, such as the deductibility of commuting costs, exert a huge influence on the tax burden weighing on labor. Recent moves by the German legislator in this area have seen deductions for these costs abolished for the first twenty kilometers of the journey to the workplace. A judgement by the German Constitutional Court subsequently restored them. The episode shows that for the group of taxpayers who fully or partly exhaust the allowance for commuting costs, the legislator tried to use a tax base effect to increase income tax receipts. Higher tax rates are a more universal tool for the same purpose, though, as they hit taxpayers regardless of their distance to the workplace.
In conclusion, it appears that the *Nielsen and Sørensen*-Model should be subjected to further empirical examination in future research. In particular, additional insights could be gained with regard to the necessary tax rate spread in the lower tail of the earnings distribution.

CONTRIBUTION 2

TRAINING DECISIONS AND THE INCOME TAX

2.1 Introduction

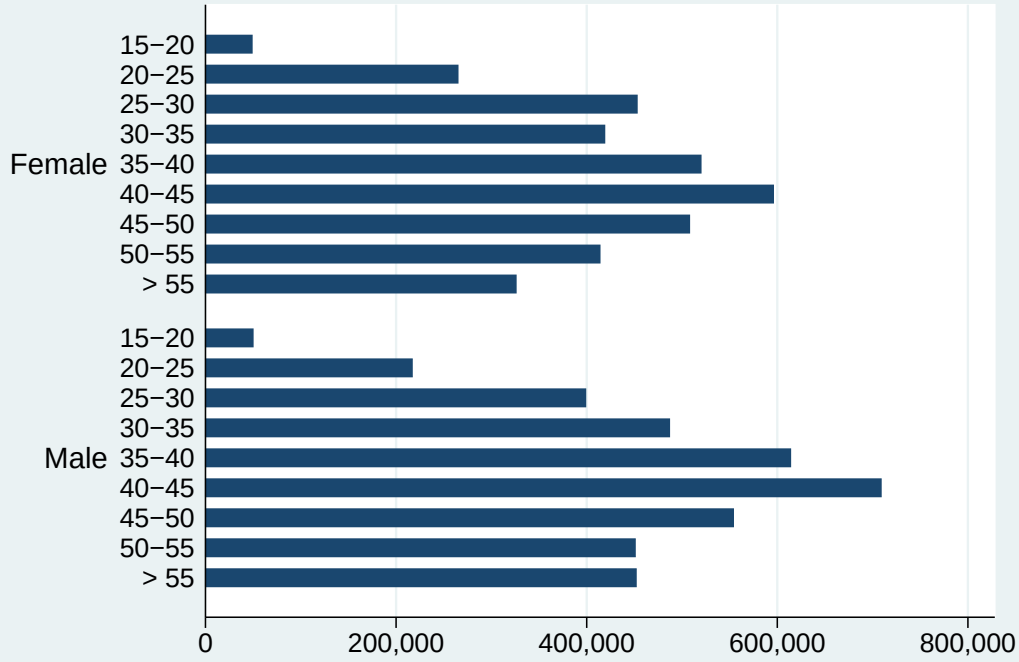
The importance of labor income for the consumption streams of most households is well documented, as their asset portfolios are heavily tilted toward human capital: Roughly 65% of Germany‘s national income has consistently been distributed as wages over time (*Federal Statistical Office*, 2008b, p. 621). This preponderance is also reflected in the overwhelming share of the income type “Employment” in taxable income in Germany over the last two decades, as reported in *Müller* (2004, p. 77): While income from employment, as recognized by § 19 of the German Income Tax Code, made up close to 80% of the tax base of the German income tax in 1989, this proportion had risen to over 85% in 1995. In 2004, its share still hovered around 81% (*Federal Statistical Office*, 2008a, table 1).

The labor market rewards workers for being productive. One major contribution to labor‘s productivity – which, in turn, is one of the major determinants of economic growth (*Cohen and Soto*, 2007) – is given by mandatory schooling, which households engage in during their youth. While these activities are confined to the phase *before* entering the labor market, technological change, which brings about obsolescence of acquired knowledge, and depreciation, i.e. forgetting, usually force agents to refresh and expand their knowledge and thus to engage in further education activities *during* their labor market participation. These activities are commonly called “training decisions” in the literature (*Leuven and Oosterbeek*, 2004). Apart from securing continued labor market attractiveness, training decisions can also be taken to increase earnings over and above the current level. Evidence of the profitability of these training investments for both workers and firms is abundant: *Almeida and Carneiro* (2009, p. 98), for instance, report an estimate for the rate of return of 8.6%, which approaches the long-term return on equities (*DeLong and Magin*, 2009).

These training activities exhibit all the hallmarks of investment decisions: Consumption possibilities are forgone today in order to obtain more consumption tomorrow, adjusted for some rate of time preference. Training decisions fit seamlessly into this fundamental framework. In fact, the literature has aptly described them as just another "...vehicle for carrying consumption into the future..." (*Rosen*, 1982, p. 443). There is abundant empirical evidence that highlights the dimensions of the resulting activities. For Germany, they can be gauged from *Federal Statistical Office* (2008b, p. 153): As **figure 2.1** on the facing page shows, over 6.8 million

courses were taken in Germany in 2007. The incidence of occupational training peaks between the age of 40 and 45.

FIGURE 2.1: Incidence of Occupational Training by Age Group and Gender



Source: Federal Statistical Office (2008), page 153
Data for Germany, 2007

An interesting angle onto training decisions is provided by the famous Mincer equation (*Mincer*, 1974) that relates the conditional expected (logarithmic) wage to years of schooling and the length of (potential) labor market participation and its square. Why does the labor market accord higher earnings to more experienced workers? The derivation of the Mincer equation, as shown in *Franz* (2009, p. 93), explicitly incorporates the assumption that a fraction of the worker's time and thus of his potential contemporaneous output is devoted to investment into his human capital. This fraction is assumed to be 100% during the formal schooling period and to decline linearly afterwards to hit 0% at the age of (planned) retirement. Hence, returns estimated in empirical research employing the Mincer equation constitute a hash of returns to schooling and returns to some kind of further training assumed to be undertaken, although this fact is seldom articulated in the vast literature on the Mincer equation. *Kuckulenz and Zwick* (2003, p. 3) support this argumentation

when they remark that "...postschool human capital investments are proxied ... by work experience or, in other words, left as a black box".

There is a vast strand of literature estimating rates of return to training, with recent contributions for Germany by *Muehler, Beckman and Schauenberg* (2007) and *Kuckulenz* (2007). From the point of view of *taxation*, though, there is a remarkable dearth of publications on the tax consequences of training measures - which is rather surprising given that the early contribution by *Rosen* (1982) had already established an influence of income taxation on training decisions. The scarcity of research in this area is also underscored by the lack of publications cited in the survey article *Hundsdoerfer et al.* (2008) and the lament in *Sørensen* (2004, p. 28).

Comprehensive recent German contributions to the topic, such as *Kuckulenz and Maier* (2006) and the longish article *Lechner and Wunsch* (2009), do not mention taxes. *Hungerbühler* (2007) and *Tremblay* (2009) constitute fairly recent *theoretical* contributions that do attempt treatments of these subjects. *Empirical* evidence regarding the interrelation between taxation on the one hand and returns to training on the other hand is still lacking, though, and is provided in this paper.

Since labor taxation does not recognize the notion of an “asset”, but is conducted along the lines of cash-flows flowing to the employee (*Kaplow*, 1996; *Nielsen and Sørensen*, 1997), there is a parallel to an active and closely related strand of literature that evaluates the accounting and tax consequences of investment decisions that are immediately expensed under current tax accounting rules, dubbed “internally developed intangible assets” in the seminal contribution *Robinson and Sansing* (2008). For instance, the fact that accounting rules explicitly provide for immediate expensing of research and development (R&D) expenditures has been researched with regard to the influence on share prices by *Chan, Lakonishok and Sougiannis* (2001), and with regard to earnings management in *Osma and Young* (2009). The connection to human capital issues has become more apparent in the recent contribution *Pantzalis and Park* (2009).

The rest of the paper is organized as follows: We review the tax effects in play and describe the relevant provisions of the German Tax Code in the evaluation period 2000 to 2003 in section 2.2. Section 2.3.1 specifies the research goal while section 2.3.2 introduces the necessary econometric methods. Section 2.3.3 describes the dataset, section 2.3.4 provides the estimation results, and the tax implications following from the estimation results are presented in section 2.4. Section 2.5 concludes. Appendix 2.5 provides additional methodological details.

2.2 Training Decisions under an Income Tax

2.2.1 Tax Effects for Training Decisions

The customary distinction between tax base-, time- and tax rate effects (*Hundsdoerfer et al.*, 2008) can be harnessed for the training investments considered here:

- Tax base effects, in the form of the deductibility of training costs: § 4 (5) of the German Income Tax Code does curtail deductibility for certain expenditures or provides lump-sum deductions. Training costs are, however, not constrained in their deductibility[1].
- Time effects, which manifest themselves in the full deductibility of the initial cost of the investment and the full taxability of the wages flowing back to the employee.

We provide a standard treatment of time effects in training decisions here. Let the tax treatment of human capital payoffs be marked by full deductibility of pecuniary costs and the full taxability of payoffs. Further assume that investments in human capital are evaluated in a two-period framework. This is a view of the world typical of the overlapping generations literature where the two periods are often interpreted as "youth" and "working age", as in *Nielsen and Sørensen* (1997) or *Jacobs and Bovenberg* (2009). Then the after tax rate of return[2] r_τ is related to the pre tax rate of return r as follows:

$$r_\tau = \frac{-(1-\tau)C + (1-\tau)(1+r)C}{(1-\tau)C} = r \tag{2.1}$$

with C denoting pecuniary costs and τ describing a time-invariant tax rate. As equation (2.1) shows, full deductibility and full taxability cause pre and post tax rates of return to coincide. Alternatively, the legislator could exempt "normal" returns r from taxable income entirely. Both payouts and payoffs would be irrelevant for taxable income then. Equation (2.1) highlights the effects of a cash-flow tax.
Yet while acquisition costs for human capital are fully tax deductible in many countries and the wages flowing back to workers are taxed in their entirety, this fact

[1] *Fuest and Huber* (2001) model this tax base effect theoretically.
[2] In a two period framework, none of the drawbacks of the internal rate of return bite so that an evaluation of investment with rates of return is permissible here.

in itself does not establish a cash-flow tax, as might be surmised following equation (2.1). The important point to grasp is that investment decisions are taken in a *relative world*. A whole set of investment possibilities with differing rates of returns constantly competes for investors' attention, and none of them is "desirable" or "undesirable" in itself. Indeed, investment theory is imbued by the notion of *relative* advantage, as expressed most succinctly by net present value calculations.

The standard alternative investment in the form of a savings account is usually not accorded the treatment set out in equation (2.1). Indeed, payments into the saving account are not tax deductible and only the return from the account is subject to tax. With human capital investments and the savings account going head to head in the investor's mind, the relation between them is unlikely to remain stable in an after-tax comparison. No matter whether a calculation before tax yields an advantage for human capital or the savings account, the order between them can be distorted by taxation.

This point was made in *Kaplow* (1996) and later applied in *Weiss* (2007) who discussed the tax treatment of human capital payoffs in relation to those for physical capital for differently educated German workers over their entire lifetime. The main result was the finding that, for reasonable constellations of average tax and interest rates, a major tax advantage in favor of human capital exists – a conclusion which is also supported in elaborate economic models such as *Nielsen and Sørensen* (1997).

One of the open questions is caused by the concept of economic income (*Wenger*, 1999) which forms the basis of this strand of literature. It concerns the treatment of economic rents accruing to workers: When should the rent first be taxed, i.e. the first appreciation be turned into taxable income (*Diller and Grottke*, 2010, section 3.2)? *Birth itself* endows most human beings with considerable earnings potential that can be transformed into wages through manual labor, without requiring much in terms of education. Thus, the question does arise how tax law should treat the resulting windfall. The present paper does *not grapple* with this issue as economic rents for training decisions can only accrue through a conscious decision by the worker or the firm, so that the time of the payout for the training measure taken can be regarded as a *natural* starting point for taxation.

Translating the above into a handy formula, the relative advantage of an investment into human capital compared to the savings account can be expressed succinctly in net present value terms as

$$NPV = \sum_{t=0}^{T} \frac{W_t - C_t}{(1+i)^t} \tag{2.2}$$

where W_t stands for the wage rate and t is a time index. The obvious after-tax counterpart NPV_τ to equation (2.2) is given by

$$NPV_\tau = \sum_{t=0}^{T} \frac{-(1-\tau_l) C_t + (1-\tau_l) W_t}{(1+i(1-\tau_c))^t} = (1-\tau_l) \sum_{t=0}^{T} \frac{W_t - C_t}{(1+i(1-\tau_c))^t} \tag{2.3}$$

where τ_l denotes the labor income tax rate, τ_c describes the capital income tax rate, and a full-blown traditional income tax represents the obvious special case $\tau_l = \tau_c$. A simple example may serve to clarify the full meaning of equation (2.2) and equation (2.3). Assume an investor facing a perfect capital market with the interest rate at 5%, and tax rates $\tau = \tau_l = \tau_c$. Let the costs for a training course at $t = 0$ be 1,000 monetary units and his payoff in $t = 1$ be 1,050 monetary units. Elementary computations show that the investor is indifferent before tax between investing in human capital and depositing the 1,000 in his savings account, facing a marginal investment opportunity.

Figure 2.2 on the next page shows the effect of taxation on the net present value. With a 0% tax rate obviously representing the no-tax case, the NPV must start off from 0. It rises to a peak of 12.20 at a tax rate of 50.6%, dropping back to 0 for a tax rate of 100% .

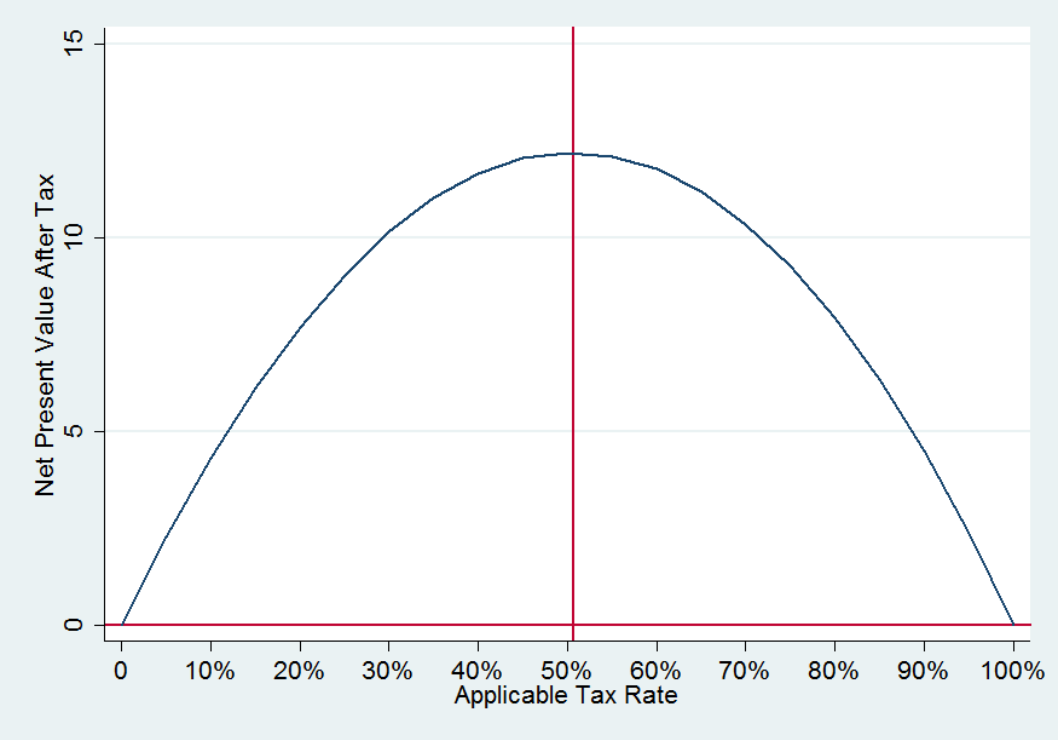
As the NPV after tax constantly hovers above the pre-tax NPV, tax law hands an unambiguous advantage to the human capital investment. The difference between the NPVs is represented by the NPV after tax. Transforming it,

$$\begin{aligned} NPV_\tau &= -1,000\,(1-\tau) + \frac{1,050\,(1-\tau)}{1+0.05\,(1-\tau)} \\ &= (1-\tau)\,1,000 \left(\frac{1.05}{1+0.05\,(1-\tau)} - 1 \right) \end{aligned} \tag{2.4}$$

Equation (2.4) shows that the curvature in figure 2.2 on the next page can be traced back to two countervailing forces:

1. The first part of equation (2.4) gives the *effective* payout for the investor; as his payout of 1,000 is tax deductible, he merely pays $(1-\tau)1,000$ for the right to the gross cashflow of 1,050 in $t = 1$.

FIGURE 2.2: Impact of Taxation on the Net Present Value, Interest Rate 5%



2. On this effective payout of $(1-\tau)1,000$, the investor earns a tax-induced "excess return" that is made up of the return to his human capital – untouched by taxation as in equation (2.1) – in relation to the return from his savings account which is cut by taxation.

Raising the tax rate τ causes the effective payout to decline and the tax induced excess return to rise. The phenomenon of an NPV rising with the tax rate has been described as the tax paradox in the literature before. The difference in equation (2.4) is easily identified as that between actual and economic depreciation charges, scaled by the tax rate τ. In the particular case of equation (2.4), the actual depreciation amounts to 1,000 at $t = 0$ while economic depreciation would be 1,000 at $t = 1$. This translates into

$$\tau\left(1,000 - \frac{1,000}{1+0.05\,(1-\tau)}\right) = \frac{50\tau\,(1-\tau)}{1+0.05\,(1-\tau)}$$

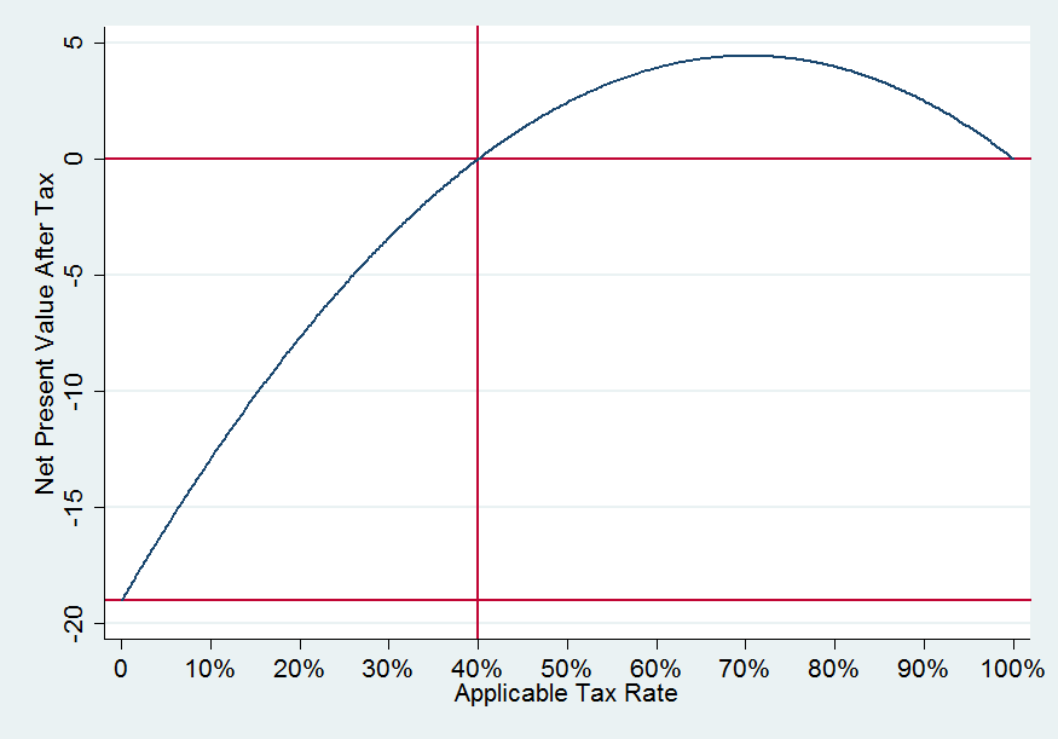
which allows one to calculate the top of the divergence in figure 2.2 as

$$\frac{\partial NPV_\tau}{\partial \tau} \stackrel{!}{=} 0 \Leftrightarrow \tau^* \approx 50.6\% \tag{2.5}$$

As it turns out, the unequal treatment of the two competing investments causes them to diverge after tax. In this particular example of a marginal investment, the investor's indifference before tax is turned into an incentive to take the course after tax.

Lowering the return to 3% and leaving the market interest rate untouched at 5%, the investment should not be undertaken, as the pre-tax NPV stands at -19.05. Yet under taxation, **figure 2.3** shows that the picture changes. The part of the panel to the left of 40% is a "comfort zone" where no tax-induced change in investment behavior is observed. To the right of 40%, the NPV switches sign, leading to a "danger zone" of tax-induced distortions.

FIGURE 2.3: Impact of Taxation on the Net Present Value, Interest Rate 3%



In a more formal treatment, let the return on the investment in human capital r diverge from the market interest rate i, implying possibly profitable investments in human capital, and allow for a divergence of the tax rates τ_l and τ_c. Sticking to the

two-period framework,

$$NPV = \frac{(1+r)\,C}{1+i} - C \tag{2.6}$$

while

$$NPV_\tau = \frac{(1+r)\,(1-\tau_l)\,C}{1+i\,(1-\tau_c)} - (1-\tau_l)\,C \tag{2.7}$$

so that the difference between the NPVs and thus the tax induced distortion between human capital investment and the standard alternative becomes

$$\begin{aligned} &\Delta NPV = NPV_\tau - NPV \\ &= \left(\frac{(1+r)\,(1-\tau_l)\,C}{1+i\,(1-\tau_c)} - (1-\tau_l)\,C\right) - \left(\frac{(1+r)\,C}{1+i} - C\right) \\ &= \left((1-\tau_l)\,\frac{r-i\,(1-\tau_c)}{1+i\,(1-\tau_c)} - \frac{r-i}{1+i}\right) C \end{aligned} \tag{2.8}$$

As equation (2.8) makes clear, the difference between before and after-tax NPV is a rather complex expression while under a pure cash-flow tax the difference should be as easy as

$$\begin{aligned} &\Delta NPV = NPV_\tau - NPV = (1-\tau)\,NPV - NPV \\ &= -\tau NPV = -\tau \frac{r-i}{1+i} C \end{aligned} \tag{2.9}$$

Reassuringly, equation (2.8) collapses to equation (2.9) by setting τ_c to zero.
From an economywide perspective, a harmful situation arises once the order of investment projects is distorted by taxation. If taxation induces changes in economic behavior, it imposes excess burdens that add to the economic costs of the tax systems. A non-neutral tax system may harm the economy in two respects:

1. It induces agents to expend planning costs that a neutral tax system avoids. If the ranking after tax can differ from the one established before tax, it may prove worthwhile for the investor to incur planning costs.

2. It may suggest unattractive investment options to an investor. Once tax planning has been carried out, a project that should have been discarded on the basis of the pre-tax NPV can reassert itself as attractive after tax. While this choice is rational from the investor‘s perspective, it is harmful to the overall

economy.

The analogies between physical and human capital are blurred by the following considerations, though: In a perfect capital market, it is customarily assumed that *any* project with a positive NPV is implemented, if necessary by borrowing the initial costs. This assumption is unlikely to be tenable for human capital investment. While in the case of physical capital investment, investments can be extended almost limitlessly, a human being is constrained by the time available to it and the attention spans it can devote to any one task. To put it differently, if two training opportunities should open up, both promising huge returns exceeding those from alternatives, the individual may only be able to take up one of them and thus "waste" value. This "waste", though, is forced upon him by his natural constraints.
This leads to the insight that a mere transfer of customary principles from physical to human capital may prove elusive. It also ties in with the conjectures in *Almeida and Carneiro* (2009, p. 98) on the question why training measures are not extended even further than the level observed in figure 2.1 on page 31. In particular, the persistence of economic rents from human capital investments may be explained with the above mentioned arguments. Apart from this attention restriction, there is a talent restriction because the natural distribution of abilities cannot be overcome.
A second line of argumentation rests on the fact that payoffs from human capital investments are highly correlated with the overall wealth of the worker: An accident that leaves the worker unable to work (and thus to utilize his human capital in the labor market) also entails significant costs for his health care, i.e. human capital does not pay off in states of the world when payoffs are most needed. This positive correlation calls for a higher required rate of return and goes some way to explaining the "human capital premium puzzle" (*Christiansen, Joensen and Nielsen*, 2007).

2.2.2 Relevant Provisions of the German Income Tax Code 2003

To evaluate the influence of taxes on the training decisions made by German workers, the relevant provisions of the German Income Tax Code for the period 2001 to 2003 are presented here. The determination of the tax liability in the German Income Tax Code is mostly governed by its first four sections, comprising §§ 1 to 34. While section two governs the tax base, section four contains the tax rate rules. The determination of taxable income proceeds in several distinct steps, enumerated in the "Income Tax Guidelines" (Einkommensteuerrichtlinien), number 3. German

Income Tax law recognizes seven income types, detailed in § 2 (1), among them the "Income from employment" that we focus on in this contribution. "Income" in the case of "employment" is calculated according to §§ 8 and 9 of the income tax code, with the latter stipulating the deductibility of training costs from the gross wages received in a certain year.

The income from the seven income types is determined according to the provisions of the second section of the code. It is added up to arrive at the sum of income (Summe der Einkünfte). Importantly, in the years 2001 to 2003 interest income counted towards the sum of income and was subject to the normal tax schedule. But for minor items, the sum of income equals the overall amount of income (Gesamtbetrag der Einkünfte). The next step involves the deduction of special items, such as special allowances (Sonderausgaben) and extraordinary burdens (außergewöhnliche Belastungen) born by the taxpayer in the relevant year. After these deductions, the income is determined, which, reduced by deductions for children, equals the taxable income.

For the case of special allowances, § 10 of the German Income Tax Code 2003 left the taxpayer with an option to itemize deductions or to go for a lump-sum deduction. Itemized deductions include those for contributions to health-, unemployment- and accident insurance or the church tax payed. While it would be difficult and in many cases arbitrary to try to calculate the itemized deductions, the lump-sum option, detailed in § 10c, can easily be inferred from the wages received by the individual.

Having calculated and deducted this special allowance, the taxable income thus calculated is then processed by the piecewise function that made up the German Income Tax Rate function for the year 2003[3]. During the years under observation in this paper, namely from 2001 to 2003, these tax rates remained in force during the entire period. The tax rate function consisted of four zones, with the first zone featuring average and marginal tax rates constant at 0%. After the first kink at a taxable income of 7,236 €, taxpayers faced an initial marginal tax rate of 19.9%, which rose to a top marginal rate of 48.5% for incomes above 55,008 €.

[3] § 32a (2003) stipulates minor complications like rounding the taxable income to the nearest number dividable by 36, and increasing it by 18, whose inclusion would not add to the insights desired in this paper.

2.3 Empirical Returns to Training

2.3.1 Research Goal

Within the microeconometrics literature, the quantities data analysts usually want to extract when evaluating treatment effects are the average treatment effect (ATE), the average treatment effect on the treated (ATT) and the local average treatment effect (LATE) (*Caliendo and Kopeinig*, 2008). In the absence of a truly experimental dataset, i.e. one where treatment has been assigned based on a random mechanism, the impact of a treatment is difficult to disentangle from the selection mechanism that caused subjects to take part in the first place. We expand this insight below in section 2.3.2.

For the purposes of this contribution, the desired quantity is the ATT, i.e. the effect that the treatment had on the participants. Note that interest here rests in the attribution of an economic rent to the payout for the training measure, while much of the training evaluation literature (*Fitzenberger and Völter*, 2007) is intended to answer the question whether a certain program "made sense", i.e. improved post-treatment outcomes in terms of the probability of employment.

Defining the treatment as T and the outcome variable as Y, the interesting parameter is

$$ATT = E\left(Y_1 | \, T = 1\right) - E\left(Y_0 | \, T = 1\right) \tag{2.10}$$

where $Y_{0,1}$ denote the outcome for untreated and treated cases and $E\left(\bullet\right)$ denotes the expectation operator. (2.10) thus says that the ATT denotes the difference in outcomes for the cases that chose to be treated. As usual, only one of the two outcomes is observed, the counterfactual is not.

If the only difference between $E\left(Y_1 | \, T = 1\right)$ and $E\left(Y_0 | \, T = 1\right)$ was the treatment assignment, one could estimate the population version of (2.10) from the sample moments for treated and untreated cases. Yet the non-random assignment of T frustrates this strategy, as outcomes *would have been different* even in the absence of treatment, i.e. iff

$$E\left(Y_0 | \, T = 1\right) - E\left(Y_0 | \, T = 0\right) \neq 0 \tag{2.11}$$

2.3.2 Econometric Specification

As argued above, the econometric specification employed to determine the rent obtained by training participants must account for the fact that the outcome of the training intervention is observed only for those individuals that take part in the training exercise. The possible outcome of training for an individual who – given the opportunity to participate – chose not to be treated – the counterfactual – can never be observed. If individuals *randomly* self-selected into training, the easiest and most obvious modeling strategy, including a dummy variable indicating treatment (=participation in training in 2001-2003) and non-treatment, would be justified. Yet the estimated treatment effect as embodied in the OLS coefficient of the dummy would be unlikely to be an unbiased estimator of the true impact. It is well known that individuals will self-select into treatment or non-treatment on the basis of the perceived benefits and costs that participation will entail. Only when this assessment leads to a positive expectation, i.e. the expectation of an economic rent, will agents participate in the treatment. Those who voluntarily submit to the treatment are likely to be endowed with more innate abilities to start with so that high propensity to engage in training will be correlated with the wage outcome.

The dataset does not contain information concerning the abilities of individuals. The correlation thus induced between the treatment dummy and the error term leads to a failure of the fundamental OLS assumption that the covariates are (conditionally) independent from the error term (*Greene*, 2008, table 2.1). Econometric methods are thus called for to circumvent the problem. The microeconometric literature has coined the term "treatment evaluation" for this kind of estimation problem. *Cameron and Trivedi* (2005, ch. 25) and *Greene* (2008, ch. 24.5) describe the parametric models applicable in this context.

The decision to engage in training activities can be observed from the available data, yet the calculus behind it cannot. From the above, one can surmise that potential candidates for participation in training activities will form an expectation about the economic rent obtainable from the program. The rent constitutes a *latent variable*. Denoting this rent by R, participation is conditional on R exceeding zero, where R is impacted by **z**, a set of covariates. Thus the **participation equation** for the i_{th} individual models his participation p_i as

$$p_i = \begin{cases} 1 & \text{if } R > 0 \\ 0 & \text{if } R \leqslant 0 \end{cases} \tag{2.12}$$

Letting the perceived benefit R depend on covariates $\mathbf{z}$ and an error term u

$$R = z\gamma + u \tag{2.13}$$

equation (2.12) turns into

$$p_i = \begin{cases} 1 & \text{if } z\gamma + u > 0 \\ 0 & \text{if } z\gamma + u \leqslant 0 \end{cases} \tag{2.14}$$

The **outcome equation** then models the resulting wage as

$$w_i = x\beta + \delta p_i + \varepsilon \tag{2.15}$$

where $\mathbf{x}$ stands for covariates in the outcome equation, and $\boldsymbol{\beta}$ denotes the associated parameter vector, while δ is the (scalar) parameter for the participation dummy. The errors of the participation and outcome equation are distributed according to the bivariate normal,

$$\begin{bmatrix} \varepsilon_i \\ u_i \end{bmatrix} \sim N \left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\varepsilon & \rho \\ \rho & 1 \end{bmatrix} \right] \tag{2.16}$$

The interesting output from equation (2.15) is the difference between (conditional) expected earnings for the treated and untreated groups which is given by *Greene* (2008, p. 890): The wage expectation of the trained $(p_i = 1)$, conditional on covariates x_i and z_i, is

$$E\left[w_i | \, p_i = 1, \; x_i, \; z_i\right] = x_i\beta + \delta + \rho\sigma_\varepsilon \lambda\left(-z\gamma\right) \tag{2.17}$$

where

$$\lambda\left(z\right) = \frac{\phi\left(z\right)}{1 - \Phi\left(z\right)} \tag{2.18}$$

is the inverse Mills ratio and ϕ and Φ denote the density and distribution function of the standard normal, respectively. The counterpart for the untrained $(p_i = 0)$ to

equation (2.17) is then

$$E\left[w_i|\, p_i = 0,\ x_i,\ z_i\right] = x_i\beta + \rho\sigma_\varepsilon \left[\frac{-\phi\left(z\gamma\right)}{1-\Phi\left(z\gamma\right)}\right] \tag{2.19}$$

The expectation of the difference in wages between trained and untrained workers is then given by

$$\begin{aligned} &E\left[w_i|\, p_i = 1,\ x_i,\ z_i\right] - E\left[w_i|\, p_i = 0,\ x_i,\ z_i\right] \\ &= \delta + \rho\sigma_\varepsilon \left[\frac{\phi\left(-z\gamma\right)}{1-\Phi\left(-z\gamma\right)} - \frac{-\phi\left(z\gamma\right)}{1-\Phi\left(z\gamma\right)}\right] \\ &= \delta + \rho\sigma_\varepsilon \left[\frac{\phi\left(z\gamma\right)}{\left(1-\Phi\left(z\gamma\right)\right)\Phi\left(z\gamma\right)}\right] \end{aligned} \tag{2.20}$$

This shows that a naive application of OLS techniques would wrongly attribute the second part of equation (2.20) to the coefficient for δ while in reality it is down to the correlation in the joint distribution of $\begin{pmatrix} \varepsilon_i & u_i \end{pmatrix}'$ as in equation (2.16). OLS estimates for the dummy are thus biased upwards if $\rho > 0$ is assumed.

The correlation coefficient in the joint error distribution ρ thus assumes critical importance. Setting $\rho = 0$, the second part of equation (2.20) drops out and the OLS estimator is back in business. The assessment of the merits of a special treatment regression and its simpler OLS counterpart thus critically hinges on the confidence interval for ρ emerging from the estimation.

As for the estimation of the model specified in equation (2.14) and equation (2.15), *Maddala* (1983, pp. 117-122) shows the derivation of the modeling strategy. Both maximum likelihood and twostep estimation are possible here. Under **full information maximum likelihood**, which is used exclusively in this paper, the necessary joint (log-)likelihood function of the random variables specified by equation (2.14) and equation (2.15) is derived in *Greene* (2006, p. 715):

$$\ln L = \begin{cases} \ln\Phi\left[\frac{z\gamma+(y-x\beta-\delta)\rho/\sigma}{\sqrt{1-\rho^2}}\right] - \frac{1}{2}\left[\frac{y-x\beta-\delta}{\sigma}\right] - \ln\left[\sqrt{2\pi}\sigma\right] & \text{if } p_i = 1 \\ \ln\Phi\left[\frac{-z\gamma-(y-x\beta)\rho/\sigma}{\sqrt{1-\rho^2}}\right] - \frac{1}{2}\left[\frac{y-x\beta}{\sigma}\right] - \ln\left[\sqrt{2\pi}\sigma\right] & \text{if } p_i = 0 \end{cases} \tag{2.21}$$

Estimation of $\hat{\beta}$ and $\hat{\gamma}$ – where a hat denotes estimated parameters – in equation (2.21) proceeds by maximizing the log-likelihood with respect to these parameters iteratively. The properties of maximum likelihood estimates are conveniently assembled in *Greene* (2008, p. 487). Recent applications of the method employed in this paper include *Jackson, Liu and Cecchini* (2009), who research the interaction between capital expenditure decisions and the depreciation method chosen, and have to account for the apparent endogeneity of this choice on the firm level, and *Kim, Chung and Firth* (2003).

2.3.3 The Dataset

Description of the Dataset

The holy grail for the analyses conducted in this contribution would be a panel dataset that lets the data analyst trace the employment and training history, alongside wage data, sociological indicators and detailed information on the tax status of both the individual and the spouse. The dataset employed in *Dearden, Reed and van Reenen* (2006) – for the British case – constitutes a rare example where these demands are to some extent met. For Germany, the data availability is rather limited, so that compromises are inevitable.

The dataset[4] employed for the estimation is a subsample from the 2000 and 2004 waves of the Socio-Economic Panel (GSOEP), compiled by the DIW, Berlin[5]. The survey is comprehensively described in *Haisken-DeNew and Frick* (2005). A topical module "Further Education, Training, Labor Market" was included in the GSOEP questionnaires in the years 1989, 1993, 2000 and 2004. The additional questions asked in this module pertain to the training efforts that respondents have undertaken in the previous three years, the kind of financial assistance that they could avail themselves of, the motivation for participation in a course and the self-assessed results (*Muehler et al.*, 2007). The subsample of the data used in this paper was drawn in the following fashion:

[4] The data used in this paper were extracted using the Add-On package PanelWhiz v2.0 (Nov 2007) for Stata. PanelWhiz was written by Dr. John P. Haisken-DeNew (john@panelwhiz.eu). The PanelWhiz generated DO file to retrieve the GSOEP data used here and any Panelwhiz Plugins are available upon request. Any data or computational errors in this paper are my own. *Haisken-DeNew and Hahn* (2006) describe PanelWhiz in detail.

[5] The data used in this publication were made available to me by the German Socio-Economic Panel Study (GSOEP) at the German Institute for Economic Research (DIW), Berlin.

We obtained information from the 2000 and 2004 waves of the GSOEP on a large number of socio-economic variables and initially restricted the sample to all adult respondents who provided information in both 2000 and 2004. Subsequently, we divided the sample into an *estimation* sample and an *auxiliary* sample. The estimation sample served to estimate the regressions in section 2.3.4 while the auxiliary sample provided information necessary for the calculation of the marginal tax rates for spouses. Inclusion in the estimation sample was determined as follows:

- The *estimation* sample was restricted to German workers who reported positive gross labor income in both 2000 and 2004.
- As the emphasis is placed on training decisions (which require future employment to pay off), the sample was confined to workers below the age of 61 and above the age of 19.
- The sample was limited to dependent full and part-time workers, thus dropping the self-employed whose inclusion would be unlikely to contribute to the question at hand.

As the final outcome variable of interest, we define a variable "Gross labor income in the year 2004". The GSOEP provides detailed information on the monthly gross labor income and the number of months during which the respondent drew this salary. The following components are added into the measure of gross yearly compensation:

- Any 13^{th} or 14^{th} month pay.
- Christmas and vacation bonuses.
- Other bonuses reported in the GSOEP[6].

The dataset was enriched by the addition of a variable on taxable income. In the case of a single household, the sum of income equals the sum of the seven income types mentioned in section 2.2.2. Of these seven, the income types reported in a more or less precise fashion in the dataset are the following:

1. Income from employment

[6] While the definition of this variable is a little hazy, it only averages about 93 € over the sample and is thus not quantitatively important.

2. Income from self-employment, which is of minor importance.

3. Income from capital, which lumps together interest and dividends. As a rough approximation, the ratio of interest to dividend income is assumed to be 4:1[7].

4. Income from rent and lease.

Thus, only four income types can be ascertained from the data. Yet, having restricted the cases to dependent workers, the loss of information is minor. At the very least, one may hope that the deficiencies of the dataset balance out so that the taxable income derived can at least serve as an unbiased estimator of the real taxable income.
As the self-employed were eliminated from the dataset, most persons had little income apart from income from employment. For married couples, though, the complication arose that the marginal tax rate depends on the sum of income of both spouses. Fortunately, these income streams can be matched according to a unique partner number in the GSOEP. From the sum of income of the spouses, a marginal tax rate applicable to both of them can be inferred.

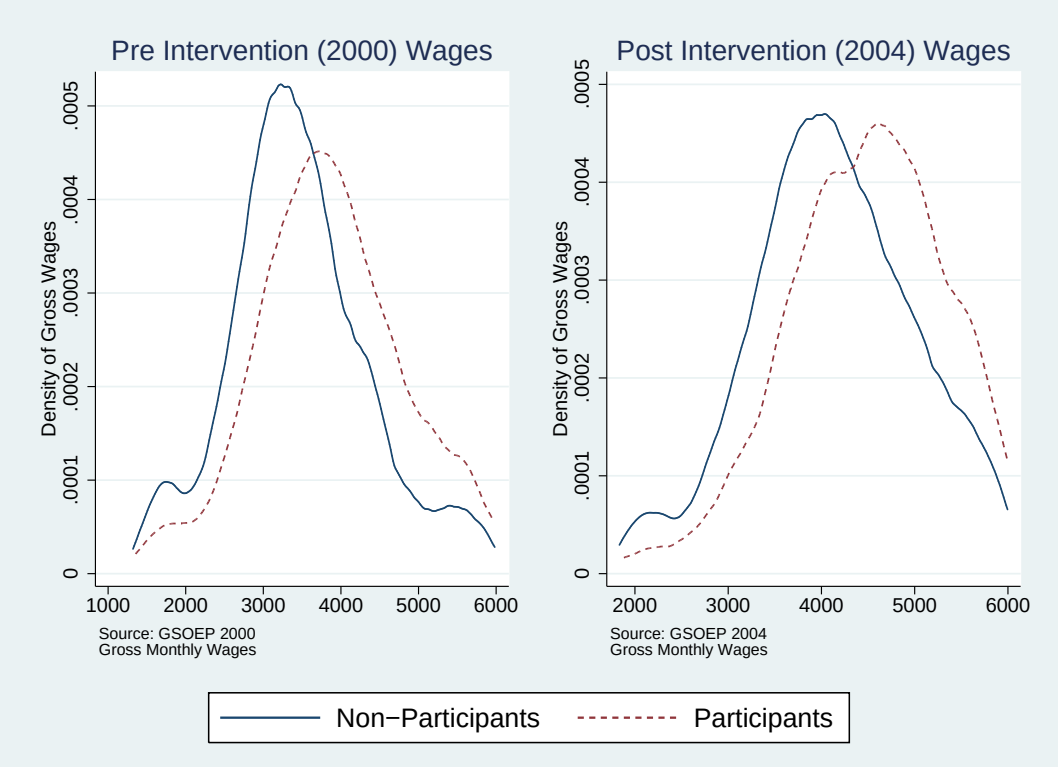
Descriptive Statistics

After the necessary deletions enumerated in section 2.3.3, the sample size dropped to 8,749, 2,994 of whom belong to the estimation sample and 5,755 do not. As of 2000, the median age is 39 and the mean age is 39.37. Within the estimation sample, 1,915 respondents did not take a single course in 2001 to 2003 while the remaining 1,079 did - figures that are comparable to those reported in *Muehler et al.* (2007).
Figure 2.4 on the next page provides the first evidence that training measures do indeed impact wages. It shows the pre- and postintervention wages, i.e. in 2000 and 2004, for participants and non-participants in the estimation sample. On the one hand, there are marked differences between the two groups before the intervention even took place, i.e. in the year 2000. Indeed, the difference in their mean wages was 880€, with a 95% confidence interval (CI) ranging from 829.28€ to 931.14€. After the intervention, in 2004, the difference amounts to 954.38€, with a 95% CI from 890.62€ to 1018.15€. Note that these are the "crude" differences emerging

[7] This is important for the 2004 wave, as dividends were taxed according to the half-income principle while interest was subject to income tax in its entirety.

from a descriptive analysis of the estimation sample, and that so far, no correction for the endogeneity of selection into training has been applied. Macroeconomic developments, though, such as the overall state of the labor market, are constant as the *difference* between the two groups is modeled in section 2.4.3 below.

FIGURE 2.4: Estimated Density of Monthly Wages before and after Intervention



The GSOEP asks respondents for information on the last three courses they attended[8]. With regard to the particular circumstances of these courses and the costs borne by respondents, this paper focuses on the most recent course. For one, this course is fresh in respondents' memories, lowering the infamous misreporting which can seriously harm the validity of econometric results. Furthermore, the after-tax calculus described in section 2.4.2 will feature the tax provisions in force during the years 2001 to 2003. It is most likely that the most recent course was taken with these provisions in mind, while for earlier courses, the decision to participate might have been made prior to 2001. This difference matters because the

[8] The original questionnaire is accessible at www.diw.de, with the relevant questions on pages 26-28.

tax rate structure in force in 2000 stipulated a top marginal tax rate of 51% instead of 48.5% in the years 2001-2003, for instance.
The goals pursued by respondents in taking courses in 2001-2003 were also queried for the GSOEP, and by far the most important one is the "adjustment to new demands in the job". This result must be viewed against the background of the restriction of the dataset to the working male population which suggest that "retraining" and introduction to a new job should not feature high on the priority list. As for the costs of the most recent course, the mean of the distribution amounts to 841.45€, on a standard deviation of 1,449.66.

2.3.4 Estimation Results

The treatment regression of section 2.3.2 is run to estimate (2.20). To this end, we introduce

- As joint covariates shared between the selection equation (2.14) and outcome equation (2.15):
  - The time span spent at the current firm by the training participant (tenure)
  - Marital status
  - Length of education which proxies for the highest school degree obtained
  - Age in the year 2004
  - The federal state of residence and the industry sector according to the NACE classification as controls
- As identifying information for the participation equation:
  - Firm size, based on the notion that it impacts the availability of training opportunities to workers
  - Workplace autonomy which plausibly impacts the need for training measures as more autonomous workplaces require more of it while not necessarily impacting wages.

We also use the *log* of the post treatment wage – as seen (in levels) in the right panel of figure 2.4 – to reduce the usual skew in the wage distribution, as is customary in this strand of literature (*Muehler et al.*, 2007).

TABLE 2.1: Estimation Results for the OLS vs. Treatment Regression

	OLS	Treatment Regression	
		Outcome	Selection
Current Year Age	-0.004***	-0.002***	0.004**
	[-0.004,-0.003]	[-0.003,-0.001]	[0.000,0.008]
Length of Education in Years	0.021***	0.021***	-0.056***
	[0.019,0.023]	[0.018,0.023]	[-0.070,-0.043]
Marital Status	-0.001	0.003	-0.063***
	[-0.007,0.005]	[-0.003,0.009]	[-0.099,-0.027]
Firm Tenure in Years	0.020***	0.018***	0.015
	[0.019,0.021]	[0.017,0.019]	[-0.006,0.036]
Training Measure in 2001-2003	0.171***	0.232***	
	[0.145,0.197]	[0.210,0.254]	
Firm Size			0.131***
			[0.106,0.156]
Occupational Autonomy			0.571***
			[0.480,0.662]
Constant	7.544***	7.248***	-2.646***
	[7.493,7.594]	[7.198,7.298]	[-3.273,-2.019]
Set of Controls for			
Federal State	Yes	Yes	
Sector (NACE)	Yes	Yes	
ρ		-0.819	
		[-0.8343, -0.8034]	
χ^2_1		1,116.735	
p-value		0.000	
σ		0.326	
		[0.3209, 0.33145]	
λ		-0.267	
		[-0.2748, -0.2600]	

Source: GSOEP 2004

Dependent Variable: Logarithmic Wage (OLS/Outcome equation),
Any Participation in Courses 2001-2003 (Selection equation)

95% confidence intervals in brackets

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

ρ refers to the estimated correlation between errors in selection and outcome equation

χ^2 and p-value refer to two-sided test of ρ being equal to zero

σ denotes the estimated standard deviation of the error in the outcome equation

λ denotes product of ρ and σ

Table shows maximum likelihood version of treatment regression

Table 2.1 on the facing page shows the results from an OLS regression in its left column and a treatment regression in the two columns on the right. The treatment regression results are unstacked into two subcolumns, one for the outcome and one for the selection equation. As for the latter, the identifying information from the "firm size" and "occupational autonomy" variables is significant at the 1% level, clearly signalling that it is relevant for the question whether taxpayers attended any courses. The estimated ρ at the bottom of the output and its associated confidence interval show that modeling the selection jointly with the outcome is vital, i.e. the second part of (2.20) must be taken into account.

Interest rests primarily in the coefficient for the training dummy which in the case of the OLS regression amounts to 17.1%, with the CI stretching from 14.5% to 19.7%, and for the treatment regression yields a return of 23.2%, with a CI from 21% to 25.2%. Note that the dependent variable is in logs and the results must be read as percentage changes in the dependent variable for a one-unit change in the independent variable.

To put the result into perspective, it must be born in mind that this rate represents the compounded wage gain over the entire training period, i.e. from the payout for the training course to 2004, attributable to the measure. Assuming a payout in 2002, the point estimate of the rate *per year* drops to

$$(1 + 0.232)^{0.5} - 1 = 0.11$$

i.e. 11%, and thus to a level that puts it above, but within reach of the long-term return on equity (*DeLong and Magin*, 2009). It is also consistent with the strand of literature dealing with the "human capital return puzzle", as represented by, for instance, *Panteghini* (2001), where excess returns for human capital are traced back to more severe frictions in human capital markets, such as short-sale constraints and a constricted set of insurance options. We provide more discussion and exemplary calculations in section 2.4.2 and 2.4.3 below.

2.4 Tax Implications

2.4.1 Description of the Approach

Bearing in mind the theoretical notions expounded in section 2.2.1, we evaluate the tax treatment of the costs and returns for the training courses taken by sample members in a multi-period context. The literature surveyed in section 2.1 seems to be largely set in a framework where the rates of return to a training measure are expressed as a percentage gain in the worker's wage, i.e. without reference to either opportunity costs due to the passage of time or monetary layouts for the measure (*Muehler et al.*, 2007, table 4). The same sentiment is expressed succinctly in *Almeida and Carneiro* (2009, p. 97): "...the literature estimating the effects of training on productivity has little or no mention of the costs of training".
The setup in the bulk of the literature renders a rate of return largely meaningless as its desirability can only be evaluated with a *reference to alternatives*: At least two rates are required. Indeed, *Almeida and Carneiro* (2009, p. 104) do remark that "...company job training is a sound investment for firms that do train, possibly yielding comparable returns to either investments in physical capital or investments in schooling." This quote shows that a comparison between competing investment opportunities is required, although no numerical attempt is undertaken to provide estimates for the returns of the alternatives.
The bulk of the literature insists on a single rate of return which shows the wage gain of the post-intervention wage over the pre-intervention wage. However, the question poses itself how fast the benefits of the training measures

1. materialize after the intervention
2. peter out once they have materialized.

Regarding the first point, there is not much literature to go by. The difficulties are compounded by the fact that the exact date of the payout for the costs of the training measure are not obtainable from the data. We assume a payout at the beginning of the year 2002 in the calculations in section 2.4.3 which seems plausible for the most recent training measure.
Regarding the second point, note that we do not attempt to employ a panel estimator to investigate the further fate of participants' wages – although the GSOEP data may allow this. The panel estimation of treatment effects seems not to have been

explored sufficiently in the theoretical econometrics literature (*Kyriazidou*, 1997; *Vella and Verbeek*, 1999). The goal is thus to estimate (2.2) and (2.3) with the available estimation results from table 2.1, and to make reasonable assumptions about the fate of the economic rents accruing to training participants after the year 2004.

A related strand of literature that investigates the relationship between research and development spending and stock market valuation of (R&D intensive) firms has established conventions which can be harnessed here. *Lev and Sougiannis* (1996) set out to estimate statistically reliable amortization rates for R&D spending, given that accounting rules prescribe immediate expensing of these costs. As *Lev and Sougiannis* (1996, p. 131) make clear, they let a period of 6 months expire after the end of the fiscal year during which the R&D costs were incurred, and then measure stock (excess-) returns for the subsequent 12 months. *Chan et al.* (2001), on the other hand, simply *assume* a depreciation rate of 20%, so that the benefits die out linearly over 5 years. In a recent contribution, *Hall and Oriani* (2006, p. 982) assume a depreciation rate of 15%.

Almeida and Carneiro (2009, p. 100), on the other hand, distinguish between two types of depreciation. Workers tend to forget existing knowledge, or it is rendered obsolete by technological progress, and part of the existing knowledge evaporates from the firm when workers quit their labor contracts and new workers establish new contracts. For the former effect, *Almeida and Carneiro* (2009, p. 101) employ a depreciation rate of 17%, while the combined effect amounts to 25%. Note that only the first of these problems affects the current discussion as the focus is on the worker, not on the knowledge stock within the firm.

To this end, we utilize the calculation of the marginal tax rate, as described in section 2.2.2, to determine $\tau_l = \tau_c = \tau_t$. Note that we evaluate the net present value against a joint marginal tax rate for labor and capital income, as this approach coincides with the state of the German income tax law in force at the time. The later adoption of a final withholding tax will thus not be discussed. Further note that a time-dependent tax rate τ_t that is evaluated for the time period under consideration should be used, i.e. according to the income tax schedule in force in the years 2001-2003. As the evaluation of the benefits from the training measure stretches out for five years, the calculation could plausibly incorporate the tax schedules for the years 2005 to 2008 as well. Yet the investment decision, i.e. the decision to conduct a training measure, was taken under the expectation of the tax rates *prevailing* in

2001/2003, so that it would be inappropriate to evaluate it against the *realized* tax rates in 2005 to 2008.
To put the above into a nutshell, the calculation consists of

1. obtaining the difference in wages attributable to the training measure, as in equation (2.20), for the year 2004: To establish the tax implications of the results in section 2.3.4, predictions along the lines of equation (2.20) must be obtained. Given the semilogarithmic nature of the estimation results in table 2.1, we use the "smearing method" proposed by *Duan* (1983), which is further described in section 2.5, to account for the influence of the error term on the predictions.

2. getting the present value of the benefits defined in 1. by letting them die out according to a linear depreciation schedule over the years 2004-2008

3. evaluating these benefits against the costs for the course obtained from the dataset, i.e. discounting them back to the date of the payout of the costs

4. repeating 1. to 3. both in a before and after tax perspective

5. calculating *the effective tax rate*[9] on the return from the training measure according to $\frac{NPV-NPV_\tau}{NPV}$, with the definitions for the acronyms given in equation (2.2) and equation (2.3) on page 35

6. varying assumptions to test for the sensitivity of the results.

2.4.2 Example for the Calculations

To illustrate the above, assume that the following constellation prevails for one of the sample members: An initial payout of 850€ – the mean payout in the sample – for a course in 2002, leading to a wage gain of 23.2% – obtained from the coefficient for "Training Measure in 2001-2003" in table 2.1 on page 50, which elevates his pre-intervention wage of 2,800€ to 3,449.60€ in 2004, for a gain of 649.6€. Adopting the benchmark case of a 20% linear depreciation rate, the time line of the human capital investment looks as follows:

[9] See the treatment of effective tax rates on human capital in *Weiss* (2007, subsection 3.4.2) for more comprehensive information.

TABLE 2.2: Time Line for Example

2002	2003	2004	2005	2006	2007	2008	2009
-850	0	649.6	519.68	389.76	259.84	129.92	0

Assuming an interest rate of 10% and a tax rate of 30%, the pre-tax values yield an NPV of 578.19€, and the NPV_τ comes in at 497.55€, yielding an ETR of

$$\frac{578.19 - 497.55}{578.19} = 13.9\%$$

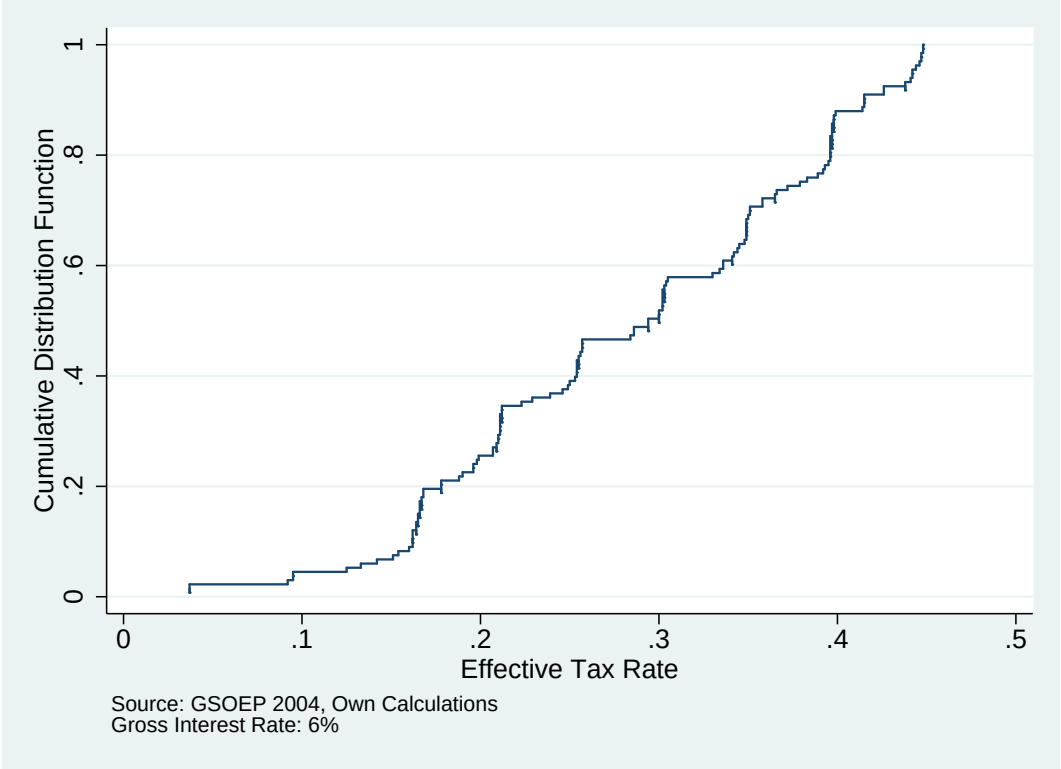
2.4.3 Results

We present the results of the calculations described above for the values of the members of the dataset, i.e. using the original cost data, and the original marginal tax rates as described in section 2.3.3, provided by respondents. Subsequently, we carry out sensitivity checks in terms of interest rates and the depreciation rate of the benefits.

As this exercise involves the evaluation of a time effect, the gross interest rate must be fixed, which we take to be 6% initially. We also restrict the sample information to sample members who report courses requiring more than a 100€ initial outlay. The questionnaire used to gather the data explicitly requests information regarding "further professional education", which prevents leisure activities from being reported and ensures that (primarily) *investment calculus* led to the participation decision, not the consumption value of the course activity. Further note that the restriction lets the number of available course participants drop to 133, which should be taken into account for the purpose of drawing conclusions from the results. We report the cumulative distribution function of the resulting effective tax rates in **figure 2.5** on the next page.

As can be inferred from figure 2.5, the effective tax rates emanating at the assumed gross interest rate of 6% are almost uniformly distributed between 20% and 45% across the population of course participants. In comparison with the much more dramatic results in *Weiss* (2007, sect. 3.4.2), the shorter duration of the benefits of five years asserts itself and prevents the ETRs from dropping precipitously. The distance to the nominal marginal tax rates for the sample members *averages 6.3 percentage points*, i.e. the effective tax rate is on average 6.3 percentage points

FIGURE 2.5: CDF of Effective Tax Rates for Sample Members



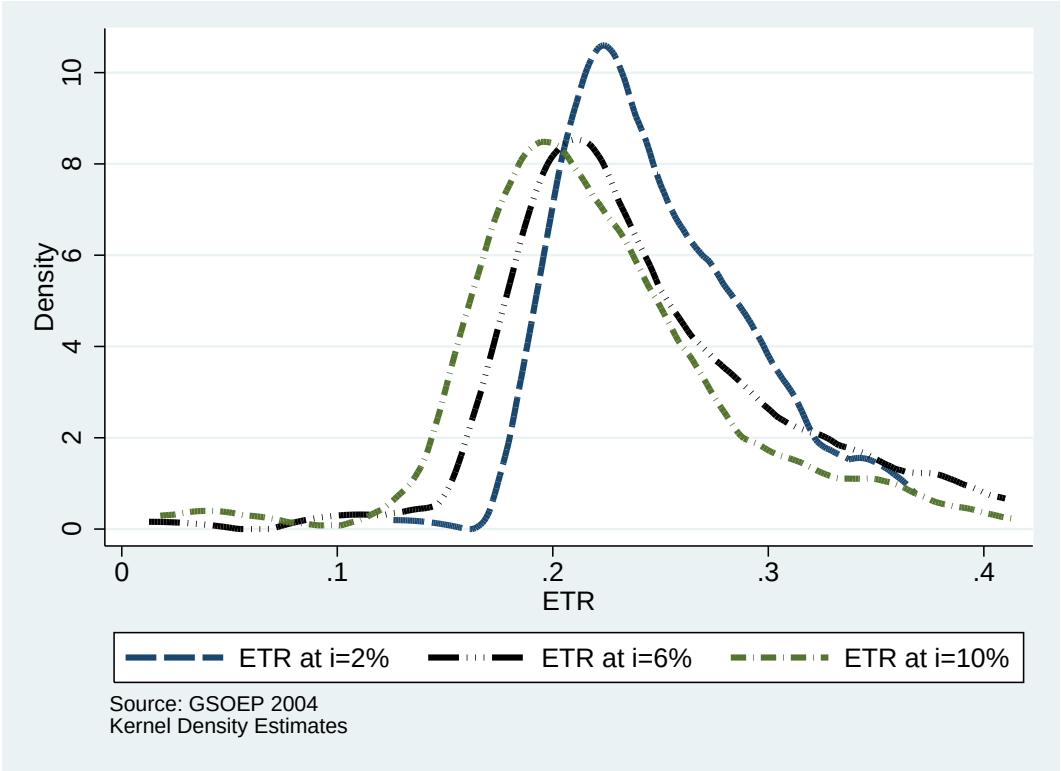
lower than the nominal rate facing the taxpayer.

Varying the gross interest rate between discrete values of 2%, 6% and 10%, **figure 2.6** on the facing page shows the time effect described in section 2.2.1 at work. The mode of the estimated densities of the effective tax rates shifts markedly to the left, the higher the interest rate rises, in accordance with intuition. The average deviation from the nominal tax rates is *4.5 percentage points* at an interest rate of 2%, and *9.8 percentage points* at 10%.

Figure 2.7 on page 58 sets the interest rate back to 6%, the standard value assumed here, and instead varies the depreciation rate of the wage benefits from the training measure, in line with the above mentioned contributions of *Chan et al.* (2001) and *Hall and Oriani* (2006). Note that the time period in years that elapses until the benefits have fully died out is given by

$$ceil\left(\frac{1}{depr}\right)$$

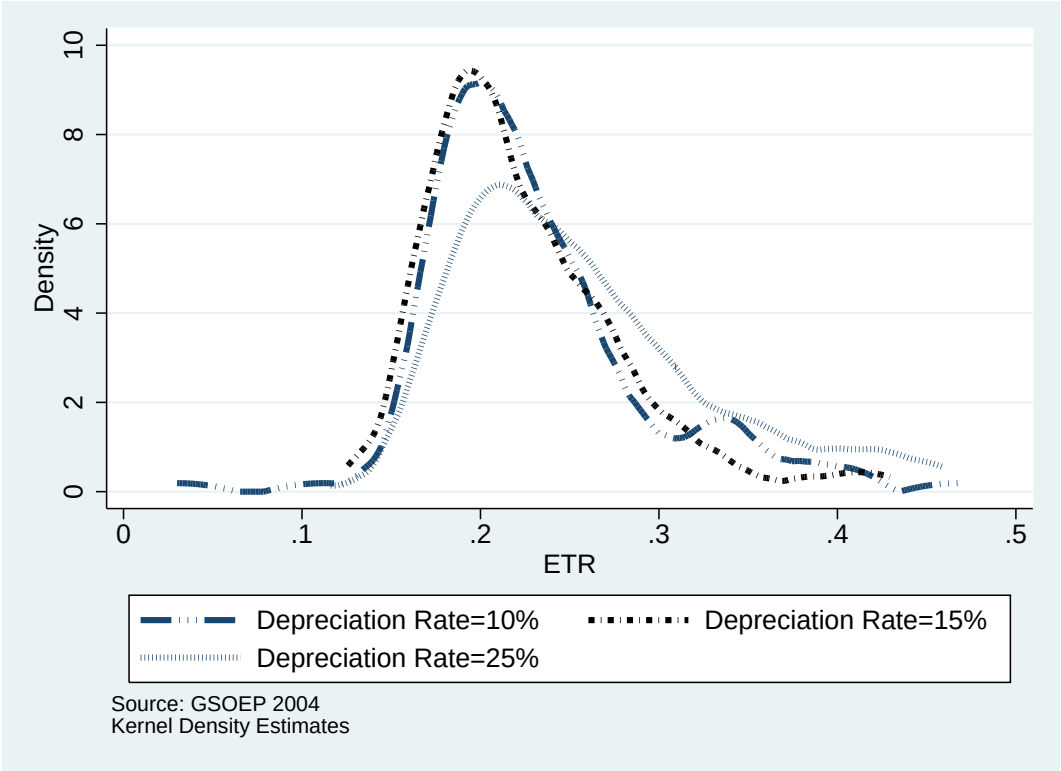
FIGURE 2.6: Estimated Density of Effective Tax Rates for Sample Members, Gross Interest Rate Varies



where $depr$ stands for the depreciation rate of the benefits, and the ceiling operator $ceil()$ gives the highest integer bigger than or equal to its argument. For the case of a 15% depreciation rate as in *Hall and Oriani* (2006), $ceil\left(\frac{1}{0.15}\right) = 7$ results. Higher depreciation rates translate into lower payback periods, so that the density estimates in figure 2.7 slide to the left the lower the depreciation rate is set. The average distance to the nominal tax rates is *4.7 percentage points* for the 25% depreciation rate and *7.3/8.4 percentage points* for 10% and 15%, respectively.

As an additional consistency check, we take a cue from the results in *Almeida and Carneiro* (2009) and augment the direct costs – obtainable from the dataset – with the forgone productivity costs that *Almeida and Carneiro* (2009, p. 103) estimate at 25% of the total cost, i.e. one third of the direct outlays. This estimate is well established in the literature and also used by *Trostel* (1993, table 1). Note that this cost element cannot be directly obtained from the dataset, as it would require more detailed information on working hours spent during the training measure. **Figure 2.8**

FIGURE 2.7: Estimated Density of Effective Tax Rates for Sample Members, Depreciation Rate Varies, Interest Rate 6%



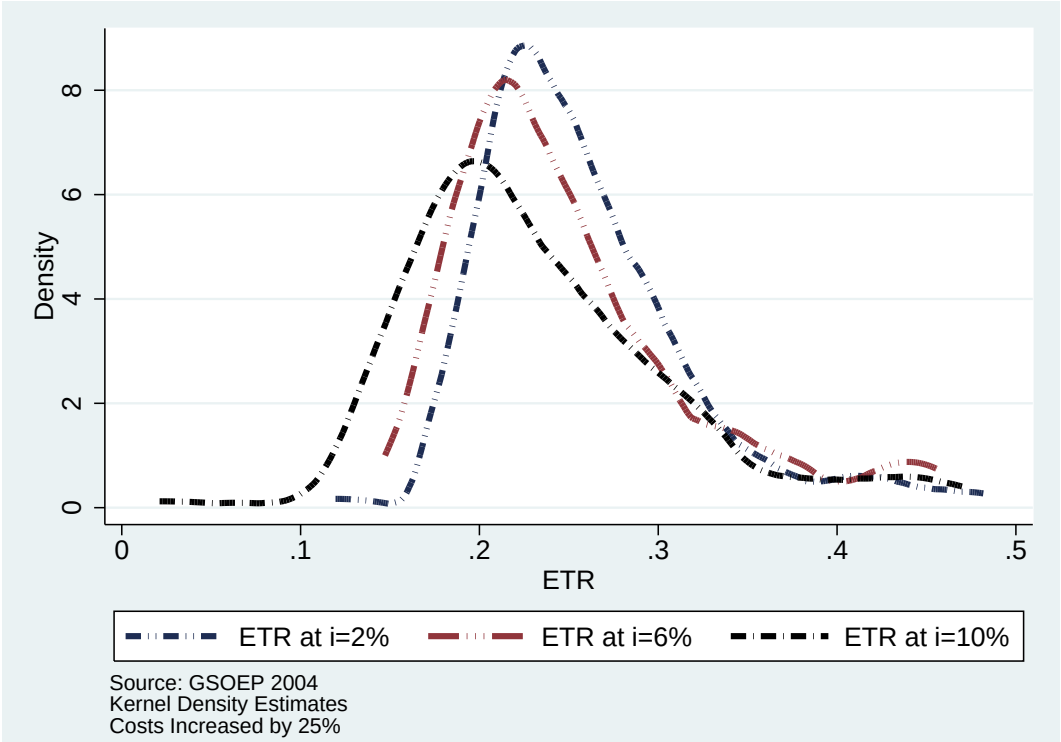
on the facing page contains the results of this experiment, depicting a slight decline in the ETRs when compared to figure 2.6, with the average deviation from nominal rates at *6.1 percentage points* at a gross interest rate of 6%.

2.5 Conclusion

In this contribution, the returns from training have been subjected to an analysis that encompasses the income tax consequences of both the returns from the training measure and the deductibility of the costs. The literature has so far failed to take adequate account of the tax implications that these measures entail, even though clearly they are bigger and more important by a huge multiple than the implications of capital income taxation can plausibly claim to be.

Taxation induces a split of economic rents generated by training measures between the worker, the firm and the tax authority. From the available dataset, only the

FIGURE 2.8: Estimated Density of Effective Tax Rates for Sample Members, Gross Interest Rate Varying, Costs Inflated by 25%



relationship between the tax authority and the worker could be inferred. In this respect, the results are incomplete as they do not incorporate the "whole" picture. The dataset that would be necessary to conduct research into this three way relationship would have to include *matched* information about both the employer and employee and their employment relationship while at the same time providing socio-economic information that allows the analyst to assess the tax implications of the training measure at hand. Importantly, tax information on the employee and the tax status of the employer would be indispensable to arrive at *after tax* results. There seems to be little such information in the literature researching these effects in a pre-tax perspective, such as *Ballot, Fakhfakh and Taymaz* (2006) for French and Swedish, *Kuckulenz* (2007) for German and *Dearden et al.* (2006) for British data.

The results in this paper are in this important respect "conditional" on the assumption that the third party, i.e. the employer, does not exert any pressure one way or the other to change the results materially. This assumption may be dubious on

account of several factors, the most salient of which would be that a labor relationship does not allow either side to appropriate the entire benefits resulting from the training decision. *Conti* (2005) and *Dearden et al.* (2006) do indeed contain evidence pointing in this direction. Other avenues pursued by this strand of literature include the effect of unionization on the frequency of training measures (*Dustmann and Schönberg*, 2009) or a more explicit focus on the perspective of the firm instead of the worker (*Almeida and Carneiro*, 2009).

Still, the results presented in section 2.4.3 do give a picture of the tax benefits that investments in human capital confer upon those who choose to subject themselves to them. The more dramatic results in *Weiss* (2007) are put into perspective by a focus on a smallish investment opportunity, where the payouts by course participants only average 850€. Also, the period of benefits derived from the measures is short, so that the time effect discussed in section 2.2.1 does not impact results too severely. Given that millions of courses are taken per year, as shown in figure 2.1 on page 31, the resulting tax savings for these investments compared to traditional savings products are substantial even if the individual case appears small.

Appendix: The Duan Method

The prediction of the level of wages from the results of a semilogarithmic equation – as used in section 2.3.4 and for the predictions in section 2.4.3– according to *Duan* (1983) assumes i.i.d. errors u_i. From a simplified semilog specification

$$\log(y) = X'\beta + u \tag{2.22}$$

the transformation to the conditional mean is given by

$$E(y|X) = \exp(X'\beta + u) = \exp(X'\beta)\exp(u) \tag{2.23}$$

The goal is to get a consistent estimate of the contribution of the errors to the exponentiated mean. To this end, the Duan method assumes that a consistent estimator of

$$E(\exp(u)) \tag{2.24}$$

is given by the sample average of the exponentiated errors, i.e.

$$\frac{1}{N}\sum_{j=1}^{N}\exp\left(\hat{u}\right) \tag{2.25}$$

where $\hat{u}$ denotes the estimated error term.

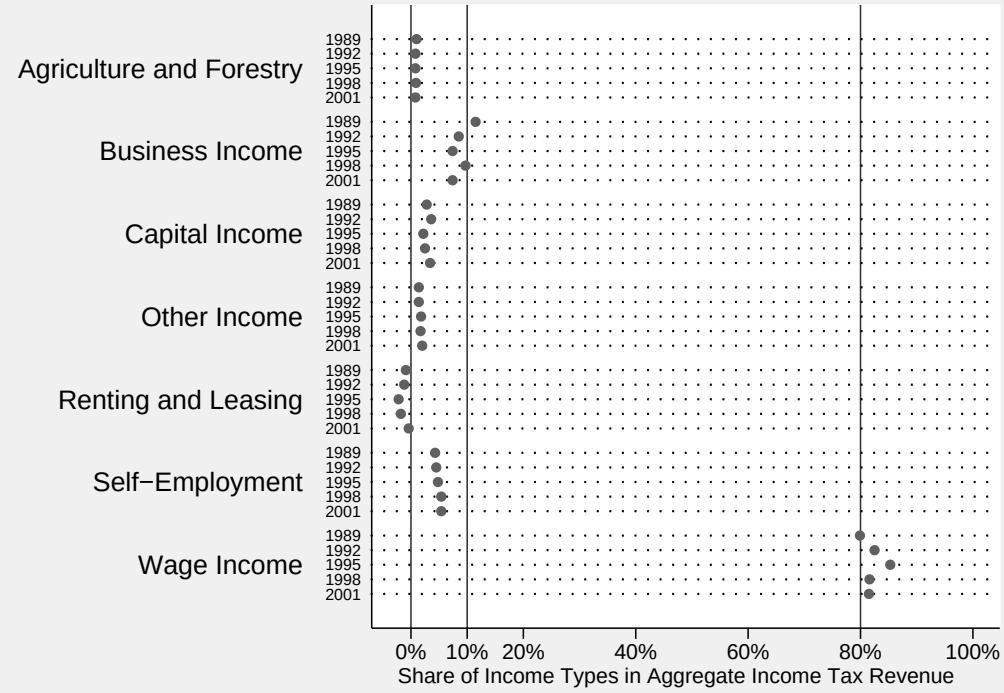
CONTRIBUTION 3

DO GERMANS REACT TO THE COMMUTING ALLOWANCE?

3.1 Introduction

"Human capital is the most important determinant of wealth and income for most individuals." This quote from *Judd* (1998) is borne out by both national income and income tax statistics for Germany. The former regularly shows a fraction of 65% of national income being distributed as wage income to private households (*Federal Statistical Office*, 2008b, p. 621), while the latter demonstrates that wage taxation contributes in excess of 90% of the aggregate income tax base (*Müller*, 2004). Figure 3.1 shows the historical development of contributions by the seven income types recognized in German income tax law over the last two decades, reinforcing the view that taxation of wage income – combined with the value added tax – has been responsible for the lion's share of tax revenue – and its stability – in Germany.

FIGURE 3.1: Share of Income Types in German Income Tax Revenue



Source: Müller (2004)

This overwhelming importance of human capital – as an input factor in a knowledge based economy and as a major source of factor incomes for most individuals – is matched by the increasingly sophisticated research methods applied to human

capital issues. Part of this literature is concerned with the tax influence on the acquisition of human capital through schooling or – after the end of formal schooling – on-the-job training (*Nielsen and Sørensen*, 1997; *Jacobs and van Wijnbergen*, 2007; *Hungerbühler*, 2007). Quite another chunk of literature deals with the conditions under which the *existing* human capital stock can be exploited. For instance, the deductibility of work-related expenses has attracted some attention recently, as in *Baake, Borck and Löffler* (2004) and *Richter* (2006).

A significant aspect of this exploitation process is the fact that, given modern means of transportation, taxpayers have the opportunity to live far from their place of work, and consequently have to ply the route to work on a more or less regular basis. Zooming in on this process, the income tax treatment of the costs expended by taxpayers to commute between their homes and places of work is controversial: Huge volatility of tax rules governing the deduction of commuting costs is observable both in a cross-sectional/international and longitudinal/national context.

In an international comparison, the rules governing the deductibility of commuting costs are quite diverse[1], and a common underlying ideology is hard to discern: *Great Britain, Spain, Ireland, Canada* and the *United States* simply disallow the deductions. *Italy* grants a high blanket deduction of 4,500 €, but disallows an itemized deduction of commuting costs. *Switzerland* and *Finland* allow a deduction of the corresponding amount that the taxpayer would have to expend to use public transportation to cover the same distance. *Belgium* allows a deduction of 75% of costs if the taxpayer credibly itemizes her travel. Otherwise, an allowance of 15 eurocents is permissible. *Luxembourg* allows a standard deduction of 396 € for distances of less than 4 kilometers which rises to a maximum of 2,970 € in step with the distance traveled. *Japan* applies income dependent allowances instead of itemized deductions so that there is no dependence on the actual distance to the workplace.

In Germany, deductibility of commuting costs has been allowed via a commuting allowance for the last couple of decades. *Steenken* (2002, p. 146) traces the historic development of the German commuting deduction from the first beginnings in Prussia to the state of the tax law around the year 2000. The German legislator initially granted a relatively high allowance per kilometer of approximately 25 eurocents in the later part of the 1950s, but required the taxpayer to prove that his costs were

[1] Cf. *Steenken* (2002, pp. 183-187) and *Bach* (2003, p. 607). All descriptions refer to the year 2000.

"necessary" (*Bach*, 2003, p. 603). The allowance was differentiated along the lines of the means of transport used, and a maximum for the distance traveled was fixed at 40 kilometers. The year 1967 witnessed a major reduction to approximately 18 eurocents which was motivated by the high volume of traffic hitting German roads. The maximum of 40 kilometers was eventually abolished in 1971. The 1990s then saw increases in the allowance to counteract the additional tax burden on fuel enacted at the same time.

The commuting allowance consists in a rate per kilometer of distance to the workplace. To arrive at the yearly deduction from the income tax base, this rate needs to be multiplied by the number of days of commuting and the distance in kilometers. The development of the rate during the years 2001 to 2009 in the German income tax code can be observed in table 3.1.

TABLE 3.1: Historical Development of the German Commuting Allowance 2001-2009

Year	**Allowance per km**			**Max**	**Blanket**
	0-10 km	**11-20 km**	**>20 km**		**Deduction**
2001	0.36	0.40	0.40	5,112	1,044
2002	0.36	0.40	0.40	5,112	1,044
2003	0.36	0.40	0.40	5,112	1,044
2004	0.30	0.30	0.30	4,500	920
2005	0.30	0.30	0.30	4,500	920
2006	0.30	0.30	0.30	4,500	920
2007	0.00 (0.30)*	0.00 (0.30)*	0.30	4,500	920
2008	0.00 (0.30)*	0.00 (0.30)*	0.30	4,500	920
2009	0.30	0.30	0.30	4,500	920

*: Changed to amount in brackets retrospectively by judgment of German Constitutional Court

Sources: § 9 (1) 4, (2) (2001-2006, 2009), § 9 (2) (2007-2008) of the German Income Tax Code

Column "Blanket Deduction" refers to § 9a of the German Income Tax Code

Thin horizontal lines denote policy changes

All amounts in €

As can be inferred from the table, the rules governing the deduction have been subject to quite dramatic changes in the last couple of years. Starting from a graduated scheme of a deduction of 0.36 € for the first 10 daily kilometers commuted, and

0.40 € thereafter, the deduction was lowered to 0.30 € regardless of the distance covered in 2004. This implies a reduction of 16.6% for the first 10 kilometer, and of 25% thereafter. The year 2007 then saw the abolishment of the deduction for the first twenty kilometers, while from the twenty-first kilometer onwards, it stayed in place. The German constitutional court later struck down the 2007 law[2], and the rules in force in 2006 were reinstated and retrospectively applied for the years 2007 and 2008[3].

To appreciate the practical importance of the commuting allowance in the determination of the tax due on labor incomes in Germany, figure 3.2 shows the distribution of itemized deductions for "wage income" in German income tax returns: Miscellaneous items comprise approximately 19% of these deductions, additional expenditures for food and dual households[4] another 6% and 5%, respectively. The remaining 70% of itemized deductions can be attributed to the commuting allowance.

Given the huge volume of the commuting allowance, several considerations shape the rules governing its application: On the one hand, the opportunity to raise more tax revenue on the part of the fiscal authorities is tempting. A "ballpark figure" for a total abolition would be around 6 billion € for the year 2001[5]. On the other hand, distributional implications of changes in the allowance are also a key driver. The 16 German federal states feature hugely different land masses which in turn entails a different distribution of commuting distances. The political economy of the commuting allowance, as modeled in *Borck and Wrede* (2005), consequently pits small German federal states, such as Bremen and Berlin, against large – and more populous – states[6]. The rules originally enacted for the fiscal year 2007, as seen in table 3.1, can be interpreted as a result of the political clout wielded by the larger German federal states[7]: For long distance commuters, the deduction was maintained.

I contribute to the literature in the following ways: The consequences of variations in the deductible amounts for commuting costs have not been subjected to thorough

[2] Cf. the judgement handed down by the German Constitutional Court dated 12/9/2008.

[3] Cf. *Federal Ministry of Finance* (2008).

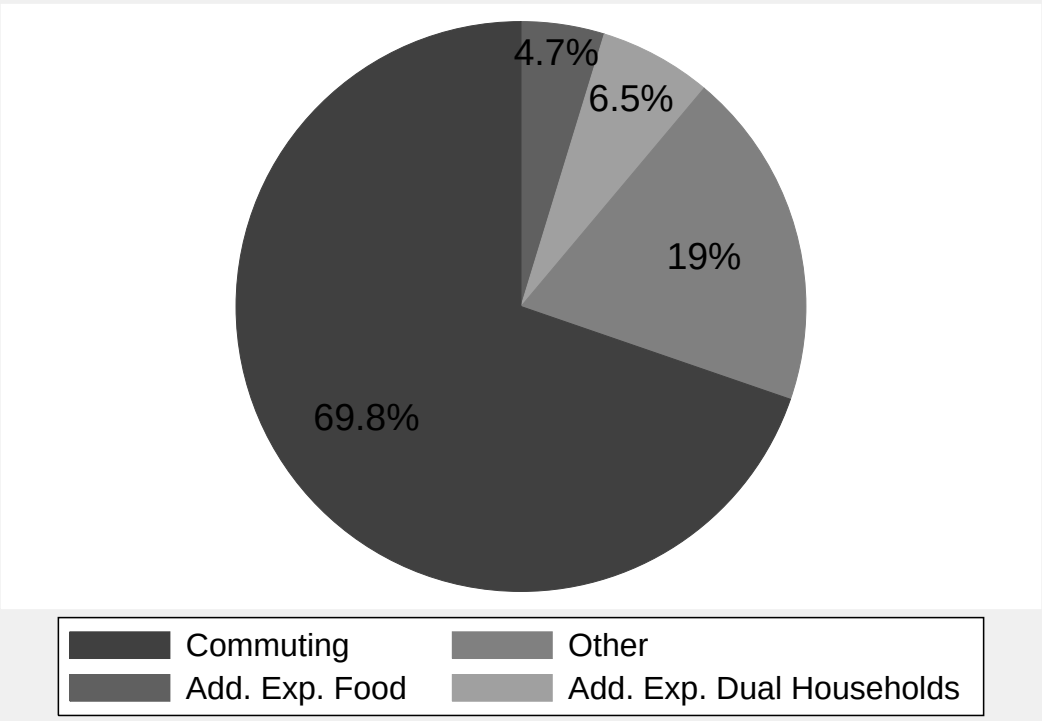
[4] Cf. § 4 (5) of the German income tax code for the legal rules governing these deductions.

[5] Cf. *Bach* (2003, p. 606) and *Federal Statistical Office* (2005, table 11).

[6] Official figures released in *Federal Statistical Office* (2005, table 8) show that almost 50% of taxpayers who chose to claim the commuting allowance in their 2001 tax returns hailed from the three federal states Baden-Wuerttemberg, Bavaria and North Rhine-Westphalia.

[7] Distributional aspects of the commuting allowance are covered in *Bach* (2003) and *Kloas and Kuhfeld* (2003).

FIGURE 3.2: Itemized Deductions for "Wage Income" in German Income Tax Returns 2003



Source: Federal Statistical Office 2007, p. 15

empirical research yet. From such an investigation, one can expect insights into the question how this deductibility shapes the behavior of taxpayers and whether more or less generous allowances are likely to trigger reactions from taxpayers. The legislator must take these reactions into account when changing the tax rules governing commuting costs. Should the taxpayers turn out to react inelastically to the deduction, the current rules could be regarded as waste, i.e. the subsidy implicit in the deduction squanders tax revenue without eliciting any behavioral changes. Any serious discussion of the environmental impact of commuting or the abatement of urban sprawl (*Su and DeSalvo*, 2008) must be based on solid insights into these reactions. Given the volume of tax receipts at stake, the state of the literature on this question is not satisfactory, a sentiment echoed by *Wrede* (2001, p. 80): "...the question is, with just a few exceptions, widely ignored in the theoretical economics literature".

To arrive at dependable results, I employ the insights of the well established treat-

ment evaluation literature, pitting the observed behavior of a *treatment and a control group* against each other before and after an *intervention*, which in this case consists in the changes observed in the commuting allowance between 2003 and 2004, which induced a steep drop in the deductible amounts for commuting. I utilize the well known difference-in-difference estimator to extract the elasticity of commuting kilometers with respect to the deduction.

There is a voluminous literature that can serve as examples of this strategy: The Tax Reform Act of 1986 in the United States represents a prominent example of just such an intervention, inducing a huge variation in tax rates that has helped researchers identify the tax influence on several microeconomic behavioral margins. Consequently, it has spawned a voluminous literature (*Kubik*, 2004; *Kumar*, 2008). *Gruber and Poterba* (1994), for instance, pursue a similar research strategy to the one in this paper, investigating the changed incentives for health insurance purchases. Other examples include the research conducted by *Leuven and Oosterbeek* (2004) into the tax effects of a *special deduction* for training expenditures, and *Jappelli and Pistaferri* (2003) who exploit changes in *marginal tax rates* to identify the tax influence on demand for life insurance assets in Italy in the 1990s.

The remainder of the paper is structured as follows: Section 3.2 describes the tax treatment of commuting costs for the relevant years under German Income Tax Law and introduces relevant research in this area. Section 3.3 provides the estimation strategy, the dataset and the results, while section 3.4 concludes.

3.2 The Commuting Allowance

3.2.1 Legal Design of the Allowance

German income tax law distinguishes seven income types, as already seen in figure 3.1. The sum of these income types is subjected to a common tax schedule[8]. Table 3.2 details the determination of this process for wage income in the year 2003 further. Row (1) gives the overall number of German taxpayers who filed tax returns in that year. Row (2) shows the number and percentage of those who reported gross income/deductions from wages: More than 85% of the tax return population do contain these items.

[8] A recent change in the law saw the introduction of a final withholding tax on capital income. The text refers to the German tax law in 2003/2004.

TABLE 3.2: Wage Income and Deductions in German Income Tax Returns for the Year 2003

	Gross Wages	**Deductions**
(1) Overall # of taxpayers filing returns	27,008,320	
(2) # of taxpayers reporting "Wage Income"	23,271,841	23,336,429
As Percent	86.17%	86.40%
(3) Volume (1,000)	827,995,421	57,684,777
Mean	35,579	2,472
Median	30,151	1,725
(4) # of taxpayers claiming blanket deduction		11,236,429
As Percent		48.15%
(5) # of itemizing taxpayers		12,100,000
As Percent		51.85%
(6) Volume of itemization (1,000)		41,800,000
(7) Thereof for Commuting Allowance (1,000)		29,176,400
As Percent		69.8%

All Amounts in €
Row (2) percentages: Fraction of all taxpayers in Row (1)
Row (4)/(5) percentages: Fraction of taxpayers in Row (2)
Row (7) percentage: Fraction of itemized deductions in Row (6)
Jointly filing couples are treated as one taxpayer
Source: *Federal Statistical Office* (2007, table 8)

Row (3) gives the overall volume of wage income and deductions reported, and means and medians for them. Given that the sum of income from all seven income types, as shown in figure 3.1, amounted to 938 billion € in 2003, the significance of labor income is reinforced. Rows (4) and (5) then highlight the split between those who claimed the blanket deduction and those itemizing their deductions. Row (6) gives the volume of itemization, amounting to over 41 billion € in 2003, of which almost 70% was claimed for commuting costs[9] (row (7)). The numbers reported in table 3.2 underscore both the importance of wage taxation within the German income tax system, and the preponderance of the commuting deduction within the itemized deductions allowed.

The commuting allowance, as stipulated in § 9 of the German Income Tax Code, gives commuters a – partial – right to deduct the costs associated with their commuting behavior from their taxable income in the category "wage income". The allowance is granted according to the kilometers for a one-way trip to the work-

[9] This information chimes with figure 3.2.

place, multiplied by the number of days of commuting. Note that there is only partial relief for the costs, as there is no dependence on the means of transport used by the taxpayer or the type of car driven. An upper ceiling, as seen in the fifth column of table 3.1, applies, unless the taxpayer commutes in a car.
Matters are rendered slightly more complicated by § 9a of the income tax code, which provided a *blanket deduction* of 1,044/920 € in the years 2003 and 2004, respectively, as seen in the sixth column of table 3.1. This deduction is available regardless of actual expenses and granted without any itemization. It thus provides a "floor" for the deduction.
As an example, consider a taxpayer who in 2003 commuted 17 kilometer on 180 days to his workplace. He calculated his deduction as

$$(10 \times 0.36 + 7 \times 0.40) \times 180 = 1,152$$

which exceeded the blanket allowance of 1,044 €. Given a marginal tax rate of 20% in the respective year, this deduction would be worth 1,152 € × 20%=230.40 € in tax refunds for the year.
The same taxpayer, given constant commuting behavior, would calculate his deduction in the fiscal year 2004 as

$$17 \times 0.30 \times 180 = 918$$

inducing him to claim the blanket deduction of 920 €.
Note that, while the commuting deduction is the most important one for wage earners in Germany, it is by no means the only one, as figure 3.2 has already made clear. Other itemized deductions can thus push the taxpayer over the blanket deduction as well, and the estimation strategy in section 3.3 must take account of this fact.

3.2.2 Theoretical Research

The subjects of scientific interest to tax research can be divided along the lines of time-, tax rate- and tax base- effects. While time effects concern the question whether certain parts of the tax base are to be taxed in time period 1 or 2, tax rate effects govern the question whether a higher or lower tax rate is to be applied. A recent example for the former effect is given by the temporary reintroduction of the accelerated depreciation of assets under the German income tax code: The

distribution over time of the tax base is altered while its overall sum stays the same. The latter effect is best thought of in terms of the recent introduction of a final withholding tax on capital income in Germany: A certain part of the tax base is subjected to a different tax rate.

Tax base effects – as the third category – concern the question which parts of the taxpayer‘s economic activities are subjected to tax at all: The fundamental decision to exempt home production from the tax code (*Sandmo*, 1990) provides a prominent example. The commuting allowance itself also falls into this category, as the costs associated with commuting are either deductible or not, depending on the rates shown in table 3.1.

Empirical research into tax base effects has been conducted, for instance, in the seminal contribution by *Gruber and Poterba* (1994) which looked into the tax incentives afforded by the deductibility of health insurance premia. Fifteen years later, *Selden* (2009) picked up the same topic with a new dataset. *Jappelli and Pistaferri* (2003) investigate how changes in the tax incentives provided by Italian income tax law changed the demand for life insurance assets. *Parry* (2002) researches the marginal welfare cost of tax base effects.

Within the strand of theoretical literature concerned with tax base effects, a smallish part is concerned with the deductibility of commuting costs from the income tax base, with early contributions by *Cogan* (1981) and *Parry and Bento* (2001). *Gutierrez-i-Puigarnau and van Ommeren* (2009) recently investigated the relationship between commuting and labor supply. Theoretical insights into workers' actual commuting costs are contained in *van Ommeren and van der Straaten* (2008) and *Ng* (2008).

Empirical research into commuting behavior and the costs associated with it is delivered in *van Ommeren and Fosgerau* (2009) for the Netherlands. A separate strand represented by *Baldry* (1998a,b) adds equity considerations to the discussion which will not be germane for this paper.

For the German case, *Wrede* (2000) provides a theoretical treatment of the interaction between deductibility of commuting costs and pre-existing distortions of the labor-leisure choice induced by wage taxation. Note that the actual treatment of commuting costs under income tax law via an *allowance* differs from the one assumed in most models in the literature – where the "costs", such as depreciation for the car used or fuel, are made deductible. This observation is particularly relevant for the case of *Wrede* (2000) where increased costs are shouldered by taxpayers

to decrease traveling time: German income tax law would be invariant to any such increase in expenditures, apart from the case where it leads to less *distance* covered. *Wrede* (2009) and *Borck and Wrede* (2009) investigate the relationship between commuting costs and taxation theoretically, as do *Richter* (2006) and *Richter and Söhn* (2008). Empirical evidence regarding commuting behavior for Germany is still sorely lacking, though, as only *Ismer, Kaul and Rath* (2008) provide a rare example of an attempt to empirically investigate the effects of the commuting allowance.

3.3 Estimation

3.3.1 Estimation Strategy

The final goal of the estimation strategy pursued is the extraction of the elasticity of the commuting distance with respect to the after-tax price paid for commuting one kilometer. This elasticity can give the legislator an idea of the behavioral changes to be expected from variation in the deductible amounts for commuting, as explained in section 3.1.
The changes in commuting distances investigated here can be achieved through two principal means: Taxpayers can choose to alter their place of residence or their place of work, or both simultaneously (*Wrede*, 2001). Apart from changes in the commuting distance, other microeconomic decisions could theoretically be impacted by the allowance and investigated as well. As mentioned above, a voluminous strand of literature concerned with the time-cost trade-off with regard to commuting (*Wrede*, 2000) would suggest research into, for instance, the decision to acquire a faster car to cover the distance to the workplace more quickly. Yet given the nature of the *allowance*, which ties deductibility solely to the distance commuted, the decision to purchase a different car does not influence the deductible amount, as it is determined with respect to the distance plied by the taxpayer alone.

To isolate the tax influence on commuting behavior, the treatment evaluation literature, as surveyed by *Cameron and Trivedi* (2005, sect. 25), is relevant: Since complete randomization of treatment assignment is unattainable in practice, the correlations between assignment of treatment and outcome must be accounted for. Variation in after-tax prices induced by policy changes is commonly exploited in

this literature, as already mentioned in section 3.2.2. With regard to the tax effects mentioned above, most of this literature exploits a tax-rate effect that induces changes in the after tax prices of the goods in question. In the present case, a tax-base effect provides the variation. Given that the income tax liability is calculated as the product of tax base and tax rate, this distinction does not preclude the application of the methods in this strand of literature to the problem at hand.

One of the candidate methods offered by this literature is the "regression discontinuity" design, as recently applied to a tax problem in *Leuven and Oosterbeek* (2004). To successfully apply this method, a discontinuity in the treatment assignment would be required. Given that the tax treatment of commuting kilometers changed discontinuously at the 10 kilometer mark between 2003 and 2004, one could imagine employing the commuters who used to commute 9 and 11 kilometer in 2003, respectively, as the two comparison groups. The main obstacle to the implementation of this strategy is given by the fact that a 10 kilometer commuting distance alone does not suffice to take taxpayers above the blanket deduction. As the exemplary calculations in section 3.2.1 show, a distance of at least 18 kilometer is required to achieve this. Taxpayers below this threshold potentially claim the blanket deduction, which means that their commuting behavior may well be unresponsive to changes in the commuting allowance. While additional deductions for other costs, contained in the "Other" category in figure 3.2, can achieve this, the dataset does not deliver the detailed information required for this conclusion. Note that this investigation is concerned with the *marginal* decision to commute one more or one less kilometer, and the certainty that the blanket deduction has been exceeded is necessary to be sure that the taxpayer does indeed make this decision according to the after-tax price induced by the commuting allowance.

Instead, I employ a difference-in-difference estimation strategy (*Cameron and Trivedi*, 2005, sect. 22.6.). The most obvious requirement for this approach is a longitudinal dataset, which I describe in more detail in section 3.3.2. While even a collection of separate cross-sections would be sufficient for this exercise[10], the SOEP data used here allow the analyst to trace sample members with a unique id across time periods, and thus provide proper panel data.

The main requirement imposed by this method with regard to the intervention is that it should be exogenous to the behavior in question. In the case researched here,

[10] Special conditions would in this case attach to the composition of the groups compared, though (*Cameron and Trivedi*, 2005, p. 770).

the change in the deductibility should not reflect an impression on the part of the legislator that there is too much commuting and that it should be discouraged via the tax system. From the discussions at the time, there seem to be no such indications, and the major concern seems to have been with the additional tax revenue raised by lowering the allowance.
Several factors must be accounted for before a conclusion on the tax influence should be made:

- Any time trends between the periods in question should be eliminated from the estimate, i.e. a trend towards more or less commuting that would have been observed even in the absence of the tax law change.

- Any fixed effects on the part of the individuals observed before and after the tax law change should not influence the conclusions about the tax change. For instance, the unobserved propensity to commute longer or shorter distances should be swept away during estimation. Indeed, the literature provides evidence for heterogeneous preferences for commuting, as recently demonstrated for the United States in *Small, Winston and Yan* (2005).

For the case of the commuting allowance, I exploit the change in the deductible amount per kilometer between the years 2003 and 2004 observable in table 3.1 as the intervention in question. Arguably, the policy change from 2006 to 2007, which abolished the deduction entirely for the first 20 kilometer of commuting distance, could also be used. This particular change was immediately challenged and ultimately undone in court, though, so that many taxpayers may have simply stuck with their existing commuting arrangement in anticipation of future policy reversals. It thus would be difficult to see how any genuine reaction to the change could be disentangled from the inevitable anticipation effects.
Regarding the treatment and control groups, several margins of identifying variation in the after-tax price of commuting can be exploited. *Gruber and Poterba* (1994, p. 711) enumerate three possible sources of variation, of which two are relevant for the case at hand. On the one hand, the fact that the reduction of the commuting allowance was uneven across the distance covered, with a kink at 10 kilometers, might enable one to construct a sample of long distance and short distance commuters that should be unequally affected by the change. A taxpayer that commuted more than 10 kilometer daily in 2003 would feel a more pronounced increase in

the after-tax cost of his commuting in 2004. On the other hand, the deductions are taken against the marginal tax rate, and high income taxpayers should feel the pinch from a reduction of the deductibility more forcefully than low-income taxpayers[11]. Given the above mentioned concern that taxpayers plying short distances might claim the blanket deduction – and thus be expected to be unresponsive to changes in the commuting allowance – a comparison between high and low income commuters, who commute enough kilometers to exceed the blanket deduction with this item alone, seems more promising. Denote the quantity of interest as

$$c_{it}\left[\tau_{i,t}^{m}\left(TI_{i,t}\right), D_{i,t}\right]$$

which gives the after-tax cost of a commuting kilometer to be borne by individual i in period of time t, given a marginal tax rate τ^m, determined by her taxable income TI in this period, and the commuting distance $D_{i,t}$ plied in the given year[12]. The marginal tax rate τ^m is determined by the taxable income in the respective year under the progressive income tax schedule stipulated in § 32a of the German income tax code.

Note how other literature mentioned above grapples with the issue of progressivity: *Leuven and Oosterbeek* (2004, p. 465) define their after-tax costs against the (Dutch) corporation tax which featured a *constant* marginal tax rate of 35% at the time. *Jappelli and Pistaferri* (2003, table 2) show the structure of the (Italian) income tax in the 1990s which exhibited constant marginal rates within tax brackets. Finally, *Gruber and Poterba* (1994, p. 724) pit "High-Income" and "Low-Income" individuals against each other.

I follow a similar strategy in this paper, taking into account the concerns regarding the influence of the blanket deduction mentioned above. I let the marginal tax rate applicable to the labor income of the household head proxy for the marginal tax rate τ^m, as in *Jappelli and Pistaferri* (2003, p. 1786), and proceed under the – admittedly – crude assumption that the middle class researched in this paper is mainly endowed with human capital whose sale to the labor market feeds its consumption streams. Additional capital income, which would drive up the marginal tax rate

[11] The strand of literature that deals with the (limited) loss offset possibilities under tax law (*Ahsan and Tsigaris*, 2009) is not relevant here. Losses for wage income are extremely rare, numbering 76,150 out of 22,860,534 cases in 2003 (*Federal Statistical Office*, 2007, table 5).

[12] Conditioning on the commuting distance becomes necessary due to the graduated structure of the commuting allowance in 2003.

under the income tax law in 2003/2004, is thus neglected. I do include robustness checks below to check for the validity of this assumption.

With treatment and control groups thus allocated, the difference-in-difference estimation proceeds as follows: Let the commuting kilometers $ck_{i,t}$ plied by a member of group $i \in [linc; hinc]$ in period $t \in [2003; 2004]$ be generated by the following data generating process:

$$ck_{i,t} = \alpha + \beta_t + \gamma D_{hinc} + \delta D_{hinc} D_{2004} + \varepsilon_{it} \tag{3.1}$$

where "linc" and "hinc" stand for low and high income, respectively, and $D_\bullet$ denotes a dummy variable for the subgroup/time period indicated in the subindex. A time trend, capturing changes in the demand for commuting kilometers common to both groups, is given by β_t. α and γ capture fixed effects for the groups in question.

Differencing (3.1) across time subsequently sweeps away the fixed effects α and γ:

$$\Delta^i = \beta_2 - \beta_1 + \delta + (\varepsilon_{i,2} - \varepsilon_{i,1}) \tag{3.2}$$

Taking the difference between the time differences in (3.2) eliminates the drift term β_t:

$$\Delta^{hinc} - \Delta^{linc} = \delta + (\varepsilon_{hinc,2} - \varepsilon_{hinc,1}) - (\varepsilon_{linc,2} - \varepsilon_{linc,1}) \tag{3.3}$$

Taking expectations on both sides of (3.3),

$$\hat{\delta} = \Delta^{hinc} - \Delta^{linc} \tag{3.4}$$

with $\hat{\delta}$ an unbiased estimator for the average treatment effect on the treated. To make (3.4) econometrically tractable, the estimator $\hat{\delta}$ can be obtained from the equivalent OLS regression of the stacked commuting distances on the interacted dummy for the postintervention period times the dummy for the treated group:

$$ck_t = \alpha + \beta D_{2004} + \gamma D_{hinc} + \delta D_{hinc} D_{2004} + \varepsilon_t \tag{3.5}$$

3.3.2 Dataset

The GSOEP[13] dataset used[14] in this study is one of the most popular longitudinal surveys of households for Germany. Since 1984, it has gathered data on a wide array of sociological indicators, among them an extensive section dealing with the labor market status and behavior of respondents.
The following demands are made for inclusion in my estimation sample:

- Only German born workers are considered.
- The subject must report labor income and commuting distances for both 2003 and 2004. One corollary is that sample members who discontinue their labor supply *due* to the deterioration in the commuting allowance (*Gutierrez-i-Puigarnau and van Ommeren*, 2009) do not enter the estimation below.
- I restrict the estimation to workers between the age of 25 and 55 years (in 2003). This restriction is intended to reduce the hazard that the commuting decision might be contaminated with education and early retirement decisions - or anticipation thereof.

My retrieval, with these conditions imposed, yields a preliminary sample of 2,339 respondents for each year.
Marginal tax rates ranged between 20% and 48.5% in 2003 and 16% to 45% in 2004[15]. I calculate the marginal tax rate for the sample members as follows: From § 32a of the German Income Tax Code 2003/2004, I obtain the formula for the determination of the income tax liability, taking into account the marital status of the respondent[16]. To arrive at the taxable amount that is fed to this formula, I calculate

[13] The data used in this publication were made available to me by the German Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW), Berlin. Cf. *Haisken-DeNew and Frick* (2005) and *Wagner, Frick and Schupp* (2007) for further information.

[14] The data used in this paper were extracted using the Add-On package PanelWhiz v2.0 (Nov 2007) for Stata. PanelWhiz was written by Dr. John P. Haisken-DeNew (john@panelwhiz.eu). The PanelWhiz generated DO file to retrieve the SOEP data used here and any Panelwhiz Plugins are available upon request. Any data or computational errors in this paper are my own. *Haisken-DeNew and Hahn* (2006) describe PanelWhiz in detail.

[15] The marginal rate can also be 0% for the taxpayers with less than 7,235/7,664 € of taxable income in 2003/2004, respectively. Commuting distances for this group rarely exceed the minimum threshold of 18 kilometer demanded below.

[16] Jointly filing couples are subject to lower tax rates if one spouse is the major "breadwinner". Note that the use of income deciles, as in *Ismer, Kaul and Rath* (2008, p. 62), would sweep this problem under the rug.

gross labor earnings for the years 2003 and 2004 from the available data. I also make allowance for the blanket deduction, as stipulated in § 9a of the German Income Tax Code 2003, of 1,044/920 €. Special allowances ("Sonderausgaben") are granted according to § 10c of the German Income Tax Code 2003[17].

Given the marginal rates thus calculated, I stratify the sample into "low marginal tax rate" and "high marginal tax rate" taxpayers. Taxpayers with marginal rates in 2003 between 20% and 28% are categorized as the former, while those featuring 42% and higher marginal tax rates are entered into the latter group. Correlation between the marginal rates in the two years is high at 0.8260, so that the classification is applied to both years.

This approach to the division of sample members should be preferable to the one used in *Gruber and Poterba* (1994, p. 724) where the choice is conditioned on income itself: Joint filers and single filers should not be lumped together, given the quite different marginal rates they face. Furthermore, the results reported below are remarkably stable with regard to small changes in the selection rule – such as an increase of the upper ceiling to 29% – which chimes with the observation by *Gruber and Poterba* (1994, footnote 24). A further demand is that the commuting distances must exceed 18 kilometer for inclusion in the estimation sample. As the exemplary calculations in section 3.2.1 have shown, this minimum is required for the taxpayer to exceed the blanket deduction no matter whether other deductions are itemized. Also, sample members reporting only weekly commuting behavior are excluded. This approach to the problem of the blanket deduction is extremely conservative, as it is likely that other deductions will be present in a considerable number of cases, and that the number of days should be higher for the daily commuters[18]. I also cut off the upper 1% of commuting distances which in this sample means those exceeding 150 kilometer. The extreme right tail of commuting distances is likely to bump into the upper limit shown in column 5 of table 3.1 anyway, so not much is lost owing to this restriction of the sample.

[17] The provisions in § 10c remained constant between 2003 and 2004.

[18] Official figures presented in *Federal Statistical Office* (2005, p. 18) drive home the point that a weighted average of the number of days across all (including weekly) taxpayers is 172.

3.3.3 Results

With these restrictions in place, the sample size drops to 630. A considerable loss of observations is thus the price for a sharp distinction between the low and high marginal tax rate groups used here. The distribution across years and groups and some sample characteristics are shown in table 3.3.

TABLE 3.3: Sample Characteristics

	2003	**2004**	**Total**
Low Income	132	132	264
	38.68	39.64	39.17
	14,242.81	16,935.18	15,620.30
	0.25	0.27	0.26
High Income	183	183	366
	41.39	42.40	41.90
	58,304.97	57,818.34	58,060.98
	0.47	0.46	0.46
Total	315	315	630
	40.28	41.24	40.77
	40,220.82	40,631.79	40,428.63
	0.38	0.38	0.38

Legend: Frequency
Average age
Average gross labor earnings
Average marginal tax rate
Source: GSOEP 2003/2004

As can be inferred from the table, the overall sample is more affluent than the average taxpayer, with a mean gross labor income over 40,000 € compared to the 35,500 € found in table 3.2, and the increased marginal tax rates[19] that this income entails. The two groups are roughly similar with regard to the age structure. The high income group is substantially larger, leading to lower standard errors for the means computed below.

The fact that the sample is not representative of the population of German taxpayers in terms of indicators like labor income and age does not invalidate the results: The estimation derives its information from the *differences* over time and groups that

[19] Note that the sample consists of both joint and single filers so that a weighted average results.

sweep away all the fixed effects affecting demand for commuting kilometers. All the time invariant factors, such as the propensity to commute, do not confound the conclusions.

The results of the difference-in-difference estimation are shown in table 3.4.

TABLE 3.4: Difference-in-Difference Estimation for the Commuting Distances

	Sample		
	Low Income	**High Income**	**Difference between Groups**
2003	35.34	38.95	3.61
	(1.62)	(1.78)	(2.41)
2004	34.56	39.57	5.01
	(1.60)	(1.86)	(2.46)
Difference across time	- 0.78	0.62	
	(2.28)	(2.58)	
		Difference in Difference	1.40
			(3.44)

Legend: Mean/Difference in Means
(Standard Error of Mean)/(SE of Difference in Means)
Source: GSOEP 2003/2004

The table makes clear the substantially higher commuting distances in the high income sample, with a difference that increased between 2003 and 2004. A t-test between the groups pooled across the two years yields a rejection at the 5% level against the two-sided alternative (p-value 0.0169). Surprisingly, the higher after-tax cost in 2004 for the high income commuters has not dampened their commuting behavior: Actually, the distances increase for them, while the lower income sample members reduce them. Finally, the difference-in-difference result shows the treatment effect to be positive instead of the expected negative one, with a large standard error[20]: No firm conclusion either way is apparent from these results, raising the specter of taxpayers being indifferent to the allowance in the short run.

Sensitivity tests are conducted to check the result for stability. The political economy of the commuting allowance, as discussed in section 3.1, would suggest a division of the sample into large and small states. Official statistics[21] show that the fed-

[20] All reported standard errors are robust to heteroscedasticity.

[21] Cf. *Federal Statistical Office* (2005, table 8).

eral states Baden-Wuerttemberg, Bavaria, Hesse, Lower Saxony and North-Rhine Westphalia accounted for over two thirds of claims for the commuting allowance in 2001. Restricting the analysis in table 3.4 to these five states – accounting for over 58% of Germany's land mass between them – yields a result remarkably stronger than the one in table 3.4. The difference-in-difference estimate now comes in at 4.11, on a standard error of 4.07. The increase in the difference between the groups over time becomes much more pronounced (in comparison with the last column of table 3.4), with the point estimate increasing from 3.11 in 2003 to 7.22 in 2004.

To check for the uncertainty associated with the marginal tax rates, I add information about other household income in 2003 to the dataset. Those with large non-labor income components could well be misclassified in the initial analysis, given that only labor income is employed to derive the marginal tax rate[22]. The SOEP data contain information about income from dividends and saving accounts. For the sensitivity check, I exclude sample members with such income in excess of 5,000 € from the estimation, resulting in the removal of only 12 of the 630 persons. The results are similar to the ones obtained in table 3.4: The difference-in-difference estimate is computed as 1.51, on a standard error of 3.46, confirming the conclusions above. Overall, the absence of any significant result suggests that the behavioral changes induced by changes in the commuting allowance are largely indeterminate in the short run. Indeed, although the point estimates for the treatment effects are positive, all reported results feature 95% confidence intervals that comprise both positive and negative values.

3.4 Conclusion

The effects of the commuting allowance on commuting behavior in Germany have been less than comprehensively investigated in the literature so far. This is surprising given the huge volume of income tax revenue at stake and the increased environmental concerns in connection with urban sprawl in public discussions. The results presented here show that taxpayers in Germany do not exhibit statistically significant short run behavioral changes when confronted with an increase in the after-tax cost of commuting. Given this relative inertia, the German legislator may reconsider the strategy attempted in 2007, i.e. a reduction in the commuting al-

[22] Cf. the description in section 3.3.1.

lowance, to raise more income tax revenue. An across-the-board cut to, say, 0.20 € per kilometer may well be simultaneously distributionally acceptable and unlikely to trigger much in terms of a reaction on the part of taxpayers or a challenge at the constitutional court. Indeed, the main reasoning provided by the court when it struck down the 2007 law was the asymmetry between the tax treatment for the first 20 kilometer commuted and the remainder[23].

It is of course arguable that the effect of altered commuting allowances only materializes over time, and that the research approach chosen here thus masks important results to be had from a long-term analysis. Such an analysis would have to include the variation in the taxation of car fuel, as an important input factor to enable commuting behavior, as well. This additional aspect was indeed a driving force behind *increases* in the allowance enacted in 1990[24].

Any serious discussion of the merits of granting commuting allowances also has to factor in the significant *compliance costs* involved. The fact that an *allowance* is granted instead of an obligation to itemize all costs connected with commuting relieves both the taxpayer and the tax authorities of some of the potential costs. Yet the common practice of ride sharing to work among colleagues or the additional deductions for accidents while commuting still give rise to many clashes that ultimately have to be resolved in court. The number of days a taxpayer commuted in a given year relies on self reporting, which is bound to lead to controversy. And the shortest available distance from the place of residence to the workplace – which the law requires taxpayers to choose for purposes of the commuting allowance – give rise to further complications. Only a total prohibition of the deduction, such as practiced in other developed countries, can obviate these costs. The estimated revenue gain reported above[25] of 6 billion € per year is thus a lower bound for the economic gains to be expected from a total abolition.

[23] Cf. footnote 2 and table 3.1.
[24] Cf. *Steenken* (2002, p. 154).
[25] Cf. footnote 5.

CONTRIBUTION 4

THE RICH DEMYSTIFIED: A REPLY TO BACH, CORNEO AND STEINER (2008)

(Joint with Rainer Niemann and Martin Jacob
Published as: CESifo Working Paper no. 2478)

4.1 Introduction

Discussions concerning the taxation of "the rich" have dominated the German political and social landscape for a long time. The perception that "the rich" do not adequately contribute their share to the public finances is deep-rooted, yet scientific tests of this presumption are quite rare. The contribution by *Bach, Corneo and Steiner* (2008)[1] constitutes an attempt to add to the public discourse by calculating the tax burden weighing on "the income rich" in the years 1992 to 2002. To this end, BCS employ a dataset merged from the German socio-economic panel (GSOEP), provided by the DIW, Berlin, and the whole population of income tax returns for Germany for the years 1992, 1995, 1998, 2001 and 2002. Their major finding is an "effective average income tax rate of the German economic elite – the top 0.001% quantile of the gross income distribution – ... (of) about 34 percent, which is well below the legislated tax rate." They also report significant variation over time, with the rate ranging from a low in 1995 of 31% to a high in 1998 of 45%. Similar results are reported for the top 0.0001% quantile where the average income tax rate drops to only 32.0% for 2002 (BCS, p. 17).

While this result seems shocking at first blush and might lead one to conclude that the superrich are not contributing adequately to the German tax base, we discuss possible reasons for it and provide appropriate remedies. BCS restrict their analyses to the income tax burden. Yet, while some economic activities, e.g. workers' labor supply, are subject to personal income taxation alone, this is not true for the case of income from business activity where both the local trade tax and, for incorporated businesses, the corporation tax reduce the sole destination of earnings, the consumption of a natural person. As it happens, the top 0.001% quantile[2] of the German taxpayers derive their *major* earnings stream from business activity (68.84% of adjusted gross income) while fully 94.96% report any earnings from this source.

BCS conduct their analysis with regard to a single year – necessitated by the lack of panel data sets covering taxation in Germany – and thus neglect the intertemporal aspect of economic activities that investment theory is imbued with. This point is made all the more poignant as the intertemporal aspect of taxation "creeps

[1] Henceforth BCS. As there are three almost identical versions of the paper, we cite the most recent and important discussion paper series where it has been published.

[2] Measured by adjusted gross income (defined in section 4.3.2). Note that jointly filing couples would count as one unit in this analysis.

in through the backdoor": German Income Tax Law allows taxpayers to conduct unlimited carry-forwards and limited carry-backs of tax losses, which BCS completely ignore. If carry-forwards and carry-backs lower the tax payments due in the year under examination but are disregarded for purposes of the income tax burden attributed to this year, the resulting number is biased toward an artificially low tax rate.

Furthermore, BCS interpret the deduction of capital expenses and accruals, i.e. the transformation of cash flows into tax bases, as "loopholes". The term "loophole" is also applied to the losses resulting from the income type "renting and leasing" which is subsequently truncated arbitrarily at –5.000 €. In the same vein, BCS attempt to explain the astonishingly low tax rates that they find with reference to terms such as "loophole", which appears five times throughout the text, "tax base erosion" (twelve appearances) and "tax avoidance" (nine appearances).

Additionally, BCS devote an entire section to the definition of an "adjusted gross income" which differs markedly from the legal definition supplied by the income tax code. It also diverges from the well-established and readily available concept of economic profits as a measure of economic income. Being derived from investment theory, the notion of economic profits is based on a multi-period concept. Similarly, the concept of "effective tax rates" is only rendered meaningful in an intertemporal perspective.

We thus recap the theory behind the academic concepts applied in BCS. Along the way, we highlight the instances where these multi-period concepts are not amenable to cross-sectional, i.e. single period, analyses. A frictionless transfer is not feasible, leading us to conclude that any tax rate computed on the back of the available data can only be interpreted as a nominal tax burden. As a prominent example, the apparent rule that faster depreciation today translates into less depreciation in the next year, serves to highlight the problematic nature of a single-period view: Eventually, today's lower tax base catches up with the taxpayer. We tentatively improve on the numbers provided by BCS, remedying several of the points of critique above and labeling the results as a "comprehensive nominal tax rate". Ironically, and in contrast to BCS, many publications concerning business taxation neglect the personal tax and solely integrate the corporate tax. We reestablish the *link between corporate and personal taxation* that is quite standard in modern economic thought. Our argumentation dissects their definition of an "effective tax rate", where we argue that ultimately, *the numerator, the denominator and the number itself* are misleading.

We provide an extensive analysis of "comprehensive nominal tax rates" on the population of German taxpayers in 2001. The spotlight put on top incomes by BCS is mirrored in our analysis. Our tentative measure of a tax rate that explicitly accounts for a comprehensive tax burden, encompassing a reasonable definition of pre-tax economic income and an enlarged set of taxes weighing on it, results in "comprehensive nominal tax rates" that are up to 12 percentage points higher than the tax rates calculated by BCS. As it turns out, the finding of a higher tax rate is robust across the entire income distribution and increases in income.

The rest of the paper is organized as follows: In section 4.2, we critically examine the theoretical approach adopted in BCS. In section 4.3, we provide an extensive analysis of "comprehensive nominal tax rates" on top incomes and show that a comprehensive, bird's eye view of the tax burden for the income rich leads to sharply different conclusions. Section 4.4 concludes. An Appendix provides additional statistical material.

4.2 Theoretical Examination of the Approach in Bach et al. (2008)

4.2.1 The BCS Approach

The BCS contribution is part of a broader literature that researches the issues of inequality and its reduction through government policies, in particular through taxation, and the degree of "effective progressivity" that the German tax system exhibits. BCS cite most of these contributions in their chapter 1. The authors themselves have contributed to the subject matter before, in *Bach, Corneo and Steiner* (2005), with a focus on tax issues, and *Bach, Corneo and Steiner* (2007), with a focus on the distribution of market incomes. These papers share several common traits that we subject to a critical examination below, in particular the same database[3] and the focus on the income tax burden alone[4]. The correction for presumed "tax avoidance strategies" or the – arbitrary – truncation of losses from renting and leasing[5] provide other examples.

[3] Cf. subsection 4.3.1.
[4] Cf. subsection 4.2.5.
[5] Cf. subsection 4.2.5.

Bach et al. (2005, sec. 4.3) also employ the concept of effective tax rates that we discuss in subsection 4.2.2. Table 11 in that chapter is quite similar in structure to table 4 of BCS: Both tables consistently find effective tax rates for the top 1% fractile of the German income tax population well below the legislated nominal tax rates, leading to the conclusion "Tax progression is real and strong, although definitely not as strong as the statutory tax rates would imply." (*Bach et al.*, 2005, p. 24).

4.2.2 Effective Tax Rates

The development of effective tax rates in the literature was instigated by *King and Fullerton* (1984) and extended by *Devereux and Griffith* (1999). This strand of the literature has witnessed huge growth since then (*Knirsch*, 2007). It is based on the notion that nominal tax rates as stipulated by tax codes do not adequately express the tax burden weighing on economic activities. Quite to the contrary, nominal tax rates are meaningless numbers because they are applied to a tax base that does not coincide with investors' economic goals. Investment theory is based on the notion of investors' goals defined as the *discounted* consumption possibilities created by an investment over and above those readily available from some standard investment, i.e. a savings account. If the definition of the tax base in the tax code does not hit these goals but anything other than them, then the nominal tax rate is no longer a valid expression of the reduction of the degree to which investor reach their goals after tax. A comparison of nominal tax rates, e.g. across countries, is consequently not meaningful.

A very general formulation of effective tax rates (ETR) that remedies the aforementioned deficiencies is thus (*Niemann, Bachmann and Knirsch*, 2003)

$$\text{ETR} = \frac{\text{Economic goal before tax} - \text{economic goal after tax}}{\text{Economic goal before tax}} = \frac{\text{Tax wedge}}{\text{Economic goal before tax}}. \tag{4.1}$$

Importantly, the concept of effective tax rates is inherently designed to address a *multi-period context* because it builds on investment theory. Departing from equation (4.1), the "economic goal" has been defined in various ways in the literature,

e.g. as net present values, final values or rates of return as defined by *Baldwin* (1959). Regardless of the specific figure employed to calculate equation (4.1) on the preceding page, they all address multi-period contexts. If this property of effective tax rates is disregarded, the connection to the literature is severed.

In contrast to the multi-period effective tax rate models, BCS, p. 10 derive their measure of effective taxation from a single period calculus. Single-period expressions, as they are typical of cost accounting or producer rents in microeconomic theory, are unlikely to cover the tax burden on multi-period investments such as real estate or shareholdings adequately. What would a single-period ETR be able to deliver in terms of information about tax effects, i.e. tax base, tax rate and time effects? As time effects are by the nature of the construction of a single-period ETR out of the equation, only tax base and tax rate effects can be detected.

BCS not only deliver one effective tax rate, but rather a whole vector of effective tax rates over points in time. They define[6] their "effective average tax rate", denoted here as ETR_t^{BCS}, as the assessed income tax liability (plus solidarity surcharge) paid in time t, denoted as PIT_t, divided by the *adjusted gross income* in t, denoted as AGI_t.

Consider AGI_t being the economic goal before tax and $AGI_t - PIT_t$ being the economic goal after tax, then ETR_t^{BCS} can be written as the well-known effective tax rate approach from equation (4.1).

$$ETR_t^{BCS} = \frac{AGI_t - (AGI_t - PIT_t)}{AGI_t} = \frac{PIT_t}{AGI_t}. \tag{4.2}$$

The interpretation of ETR_t^{BCS} cannot be the same as in a multi-period approach as one year's income is not an economic goal in itself. Therefore, ETR_t^{BCS} delivers a tax rate that tells us how much in taxes one investor paid in relation to an – imprecise – measure of income in *one* year, i.e. an adjusted nominal tax rate.

Furthermore, BCS establish entire time series of ETRs[7], which might leave the impression of an intertemporal perspective being adopted. Crucially, though, the available micro datasets for Germany do not allow one to track taxpayers' behaviour over time. Several cross-sectional datasets appended to each other cannot make up for this deficiency. They do not contain unique id variables which could establish a link between them. There is no mention of this problem in BCS, nor is an idea of

[6] BCS, p. 15.

[7] Cf. table 4 on page 17.

how to interpret the varying tax rates given.

4.2.3 Different Notions of Income

Quite apart from the doubts related to the interpretation of ETR_t^{BCS} as an ETR, the definition of the income employed to compute this number leads to yet another problem. Consider for example an individual who purchases 0.5% of the equity of an incorporated business[8] for 100 €. After holding the non-dividend paying stocks for 5 years, it is sold for a nominal capital gain. Discounting the proceeds of the share sale amounting to 133.82 € at the prevailing interest rate $i = 6\%$, the investment turns out to be marginal before tax.

Applying the *concept of economic profits*, the microeconomic view would consider the individual to have income each year due to an appreciation in each period: with each passing year, the final payment in $t = 5$ draws nearer. In a macroeconomic view, the *national accounts* would include this gain in $t = 5$. The notion of income espoused in the German tax code, the *taxable income* (fiscal view), would lead to zero income since the German income tax code of 2001 did not tax capital gains from minority shareholdings that had been realized after the expiration of the speculation period of one year.

TABLE 4.1: Differing Income Definitions

		Income according to		
t	Cashflow	Economic Profit	National Accounts	German Tax Law until 2008
0	–100.00	0.00	0.00	0.00
1	0.00	6.00	0.00	0.00
2	0.00	6.36	0.00	0.00
3	0.00	6.74	0.00	0.00
4	0.00	7.15	0.00	0.00
5	133.82	7.57	33.82	0.00

The three definitions result in periodical income (economic profit), income at realization (national accounts), or in no income at all, as shown in **table 4.1**. Adding to the confusion thus created, BCS applied a fourth definition of income, their *adjusted gross income* (AGI), that they derived from taxable income[9]. In microeco-

[8] Mainly AGs and GmbHs according to German company law.

[9] A detailed discussion of the derivation can be found in subsection 4.3.2 on page 108.

nomic approaches, however, the concept of the economic profit is well established and remains the yardstick for economic income for which the AGI does not even serve as a rough approximation.

We now provide a simple numerical example of the time-series behaviour of the single-period ETRs based on the AGI by BCS and on the economic profit. Given the prominence accorded to income from renting and leasing, we model a taxpayer over a period of five years and highlight the result of a narrow focus on data for one specific year.

Suppose an individual taxpayer invested 1 million € in real estate, where the Income Tax Code stipulates a straight-line depreciation rate of 2% per year. From the income tax statistic, we observe the income from renting and leasing. Multiplying this income with the investor's marginal tax rate gives the personal income tax in t, PIT_t, which is attributable to the real estate investment. Adding the depreciation allowances at each point in time, we can derive the cash flow CF_t. At the end of $t = 5$ the house is sold for 800,000 €, the taxpayer realizes a capital loss of 100,000 €[10]. The investment's economic profit, EP_t is defined as cash flow in t less economic depreciation. BCS' adjusted gross income equals the tax base except for losses from renting and leasing exceeding 5,000 €, which are truncated.

TABLE 4.2: Derivation of the Cash Flow

	t	1	2	3	4	5
(1)	CF_t	200.00	0.00	250.00	4.56	800.00
(2)	$Depr_t$	20.00	20.00	20,00	20.00	20.00
(3)	$TaxBase_t$	180.00	–20.00	230.00	–15.44	–120.00
(4)	PIT_t	75.60	–8.40	96.60	–7.78	–50.40
(5)	EP_t	60.00	51.60	54.69	42.98	45.28
(6)	AGI_t^{BCS}	180.00	–5.00	230.00	–5.00	–105.00

Given the economic profit and the adjusted gross income in each period, we compute the vectors of ETRs from equation (4.2):

The variation from –111% to 176% of the "ETR" based on the economic profit and equation (4.2) indicates the inaccuracy of ETR_t^{EP} as a tax burden measure. Applying the BCS definition of the *adjusted gross income* to this example results in an "ETR" ranging from 42% to 168%. Hence, even if one applies the universally

[10] We assume that he can fully offset the loss against other income.

TABLE 4.3: Comparison of Single Period ETRs

	t	1	2	3	4	5
(4)/(5)	ETR_t^{EP}	126.00%	–16.28%	176.61%	–15.09%	–111.30%
(4)/(6)	ETR_t^{BCS}	42.00%	168.00%	42.00%	155.60%	48.00%

accepted measure of the economic profit, single period ETRs lack any explanatory power. Any deviation from this concept, such as the *adjusted gross income* in BCS, does not render the measure meaningful, either.

By contrast, the ETR based on Baldwin rates of return[11] delivers an ETR of 44.48%, indicating that the real estate investment carries a higher tax burden than the alternative investment since the present value of the economic depreciation is higher than the present value of the deduction allowed by § 7 (4) of the German Income Tax Code. This interpretation would not have been possible with any of the single period "ETRs" as they cannot account for time effects.

4.2.4 The Relation between Economic Income and Taxable Income

The numerical example above has shown that the economic profit and the tax base differ at each point in time and in their time series behaviour. This section discusses the restrictions under which the economic profit can be derived from the taxable income in a single period context and how the resulting single period ETRs can be interpreted. If these restrictions held, then we could use the available cross-sectional data to compute the taxpayer's economic profit and interpret the resulting number.

Consider the earnings E_t^* as well as the expense A_t^* to be constant and to be different from the incoming payments E_t and the outgoing payments A_t respectively, then the cash flow and the tax base evolve as:

$$CF_t = E_t - A_t \tag{4.3}$$

$$TB_t = E_t^* - A_t^*. \tag{4.4}$$

Under the assumption that the cash flow, the tax base and their components are constant until infinity, the income from the microeconomic view, i.e. the economic

[11] The pre-tax net present value is zero in order to allow for any meaningful interpretation.

profit, is equal to the cash flow but not necessarily equal to the fiscal income. Consequently, the tax base has to be adjusted. The following relation between economic and taxable income has to hold to validate the interpretation of the outcome from cross-sectional data.

$$EP_t = CF_t = TB_t + [E_t - E_t^*] - [A_t - A_t^*] = const. \quad \forall \quad t \tag{4.5}$$

A prominent example for a necessary adjustment is given by the income from dividends under the half income system where both the earnings and the expenses have to be revised. At first, dividends are included with 50% of the gross distribution in the taxable income. Additionally, the gross dividend is lowered by corporate and local trade tax. Furthermore, consider this shareholding to be debt-financed, then only half of the interests paid can be deducted from the taxable income, while the cash flow was lowered by the full expenses. Consequently, the earnings side as well as the expense side has to be adjusted in order to deliver an accurate measure of the economic profit[12]. Therefore, the taxable dividend has to be adjusted on the earnings side .
The economic profit can only be derived from the taxable income observed in a cross-sectional dataset if the following restrictive assumptions hold:

- Infinite time horizon,
- Cash flow is constant over time,
- Tax base is constant over time,
- Incorporated businesses do not retain earnings,
- Economic depreciation is equal to deduction allowed by tax code,
- Accumulated tax burden is known for all income fractiles,
- Origin of foreign income, especially dividends, is known
- Capital gains are not realized, since tax burden on capital gains is not measurable.

[12] A full overview of adjustments for the individual income fractiles is presented in section 4.3.2.

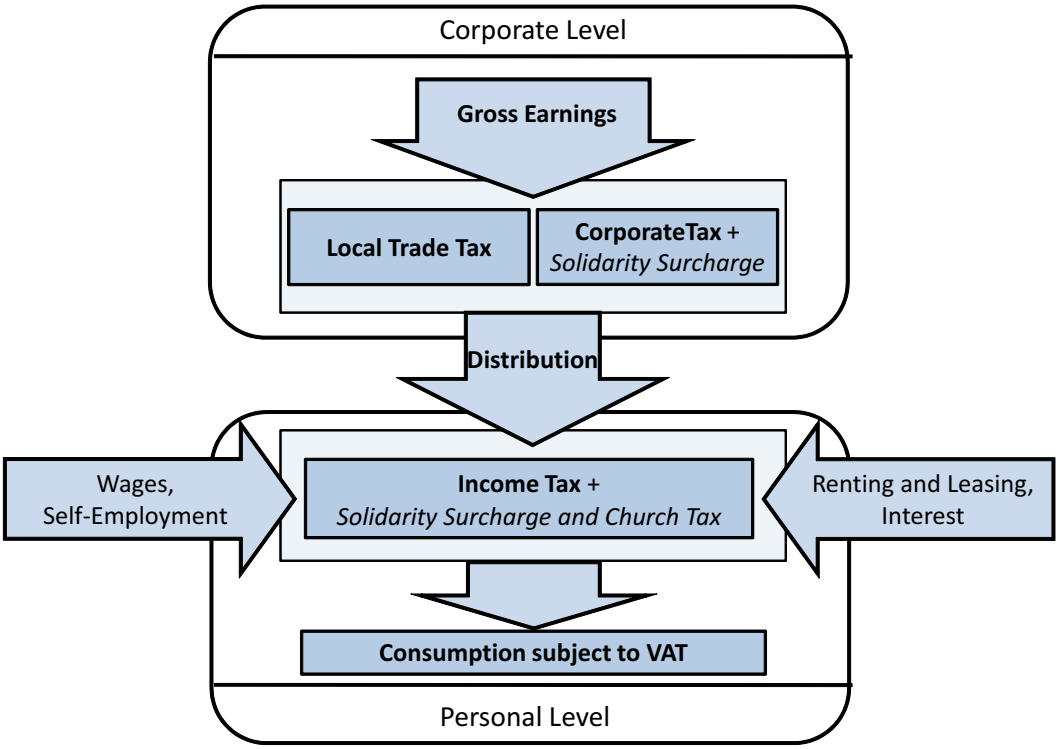
If these restrictions held, then the application of equation (4.2) would deliver a tax rate which could be interpreted as the comprehensive nominal tax burden of the economic profit in the examined year.

4.2.5 The Tax Burden on Economic Activities

Relevant Tax Rules for Incorporated and Unincorporated Business

With regard to the taxes constituting the tax burden on economic activities, **figure 4.1** highlights those central to the main thrust of our argumentation, emphasizing the importance of the integration of both the personal and the corporate level.

FIGURE 4.1: From Gross Economic Activity to Consumption



Each of these taxes comes with its own set of rules concerning the determination of the tax base and tax rates. A priori, it is not possible to discount any of these taxes as irrelevant for the measurement of effective taxation[13]. In particular, the inclusion of the tax burden on the corporate level is mandatory and largely uncontroversial:

[13] The universe of taxes possibly impacting the tax burden on economic activities exceeds the one in figure 4.1 and could conceivably include the Value Added Tax and property taxes.

While from a *legal* standpoint corporate and personal taxes are owed by different *legal* persons, economists trace profits from their origin on the corporate level to their sole final destination: The consumption of an individual or a household. Along the way, taxation reduces the consumption value repeatedly. For the distributional consequences, it does not matter what a tax is being called and whether it is collected on the corporate or on the individual level.

The refusal of BCS to take into account the enlarged universe of taxes depicted in figure 4.1 contradicts with the legislator himself: § 35 of the German income tax code establishes a link between the income tax and the local trade tax, providing a relief in the former for the burden represented by the latter[14]. The obvious corollary of this construction, i.e. that the legislator is aware of the accumulated tax burden established by *both* taxes and the possibility of a substitution between them, is turned on its head in BCS: The relief stipulated in § 35 is incorporated into their calculations, but the concomitant burden of the local trade tax is ignored.

BCS choose to include the Income Tax[15] and the Solidarity Surcharge on the "Personal Level" in their calculations[16]. There is no explicit explanation of this restriction, nor is this choice self-evident in the light of the foregoing discussion. Interestingly, the corporate level is barely touched upon: The gross dividend before corporation tax is imputed (BCS, p. 9) rather crudely, neglecting the local trade tax burden[17]. The more fundamental decision to omit the corporate level entirely is not made transparent at all and its effect will be the matter of the following paragraphs.

We show the dimension of the underestimation bias in BCS that can be traced back to the neglect of the local trade tax. This bias is prevalent for the rich as the top 1% quantile[18] generates 17.89% of their total income from business activity. The income from unincorporated business is for the top 0.1% quantile the most important income component with 37.83% of the gross income. This share rises to 68.84% for the top 0.001% quantile and to 78.66% for the top 0.0001% quantile respectively[19].

[14] See table 4.4 on the facing page.
[15] Abbreviated "PIT" in their paper.
[16] BCS, p. 15.
[17] BCS seem to suggest in their conclusion that a "move from the taxpayer level to the individual level – taking household composition into account – ... seems an interesting topic for future research".
[18] The quantiles reported in this section refer to the overall amount of income inferred from the almost 30,000,000 German tax returns in 2001.
[19] Cf. table 4.10.

With regard to this point, we provide a comparison of the comprehensive tax burden and the "narrow tax burden" by BCS for the income from unincorporated businesses in **table 4.4**. German income tax law allowed taxpayers to deduct the local trade tax from its own tax base and from the income tax base until 2007. Depending on the local municipal rate of the local trade tax, which we fix at 400%[20], the initial local trade tax burden, factoring in its deductibility from its own tax base (120), was 20 (*16.67%*). On the remainder (100), income tax – with 48.5% as the highest marginal rate in 2001 – and solidarity surcharge was due (51.17)[21]. Furthermore, § 35 of the German income tax code provided for additional lump-sum tax relief with regard to the local trade tax (9.50).

TABLE 4.4: Calculation of Tax Liability for an Unincorporated Business

		Our approach	BCS
(1)	Business Profit=Income	120	100
(2)	Local Trade Tax	–20	–
(3)	**Income from Business Activity**	**100**	100
(4)	Personal Income Tax[1]	–51.17	–51.17
(5)	Paragraph 35[1]	+9.50	+9.50
(6)	**After Tax**	**58.33**	58.33
	$CNTR/ETR^{BCS}$	51.39%	41.67%

[1] Solidarity Surcharge contained in Income Tax Payment as in BCS

From table 4.4, a tax burden on a business profit of 120 of 61.67 can be inferred, leading to a comprehensive nominal tax rate (CNTR) of 51.39%. BCS' narrow focus on the income tax burden alone – without accounting for the local trade tax paid, but factoring in the tax relief provided for it by § 35 – would calculate a tax burden of 51.17-9.50=41.67 and come up with an "effective tax rate" of 41.67%, almost ten percentage points lower. The "true" tax burden of unincorporated business that represents the most important income source for the rich and superrich is consequently underestimated by BCS. As the observed income tax liability – including the § 35 offset – constitutes the nominator of their ETR, the ETR itself is unambiguously biased downward. Worse yet, BCS account for the local trade tax offset in their

[20] This simplification is suspended in our calculations in section 4.3, in favor of a more detailed estimate of the applicable municipal rate: We distinguish between the former east and west. A more detailed analysis on the level of federal states was not allowed by the anonymisation laws.

[21] We omit the liability for church tax in all our calculations.

income tax measure without acknowledging the underlying local trade tax payment that the offset was supposed to mitigate in the first place.

In the same vein, the corporate and the local trade tax is omitted when measuring the tax burden on the income from incorporated business, i.e. income from dividends. The lower numerator – due to the omission – entails a lower effective tax rate. If we take into account the empirically observed concentration of shareholdings in the upper quantiles of the German income distribution, then this effect does not balance out over the entire distribution; instead it biases the results in favor of the impression that the rich do not pay taxes adequately as for the income of incorporated business[22].

With regard to this point, we compute the resulting bias for distributed profits[23]. We integrate the corporate and the personal level into an exemplary calculation for both the full imputation system (FIS) that was in effect until 2001 and the half income system (HIS) valid from 2001 to 2007[24]. To explain the structures, **table 4.5** on page 100 highlights the tax burden for an incorporated business in Germany for both tax systems, where the distribution, which can be directly derived from the ITR dataset, is exogenously set to 100 in row 4 in table 4.5 on page 100. The relevant income is given in row 1 with the business profit.

If an investor received 100 of dividends, the business profit, i.e. the economic income, was 175.57. One pays local trade tax (29.26) and corporation tax at a rate of 40% (plus solidarity surcharge) for retained earnings (61.74). When the profit is distributed, the corporate tax rate is lowered to 30% and a relief of 15.44 is granted in this case. The distribution (100) is subject to the personal income tax rate and solidarity surcharge at a tax base of 146.31 that is equal to the business profit after local trade tax. As the corporation tax is imputed at a rate of 30%, the individual increases his consumable amount by 71.44, implying a CNTR of 59.31%. Applying the narrow focus, one neglects the presence of the local trade tax such that the income is equal to the income tax base (146.31). The result of this narrow focus

[22] The top 1% quantile received 42.85% of the dividend and interest income in 2001. The income from dividends and interests represent the second largest income component for the top 0.1% (0.01%) quantile with 21.22% (27.71%) of gross income. Cf. tables 4.10 and 4.11.

[23] Consider that a corporation retains earnings. A narrow focus on the personal tax alone would imply these earnings to be temporarily tax-free, while they are evidently reduced by the local trade tax and the corporate tax.

[24] The dividends taxable in 2001 were taxed in 95.9% of the cases under the full imputation system, in 4.1% of the cases under the half income system. (Source: Income Tax Return (ITR) Data for 2001, own calculations).

is a CNTR (here ETR^{BCS}) of 51.17% which is well below the true tax burden on profits from incorporated businesses.

TABLE 4.5: Calculation of Tax Liability for an Incorporated Business under the German Full-Imputation System and the German Half-Income System, Owners in the Highest German Tax Bracket

		Our approach		BCS		
		FIS 2001	HIS 2002	FIS 2001	HIS 2002a	HIS 2002b
(1)	**Business Profit**	**175.57**	**162.99**	**146.31**	**133.33**	**66.67**
(2)	Local Trade Tax	29.26	27.16	-	-	-
(3)	Corporation Tax	61.74	35.82	-	-	-
(3a)	Distribution Tax Relief	–15.44	-	-	-	-
(4)	**Distribution**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
(5)	Tax Base PIT	146.31	50.00	146.31	50.00	50.00
(6)	Personal Income Tax	74.86	25.58	74.86	25.58	25.58
(7)	Imputation Credit	46.31	-	46.31	-	-
(8)	**After Tax**	**71.44**	**74.42**	**71.44**	**107.75**	**41.08**
(9)	CNTR/ETR^{BCS}	59.31%	54.34%	51.17%	19.19%	38.38%

Note: Solidarity Surcharge contained in income and corporation tax; municipal rate for local trade tax of 400%

For profits that are taxed under the half income system, one pays local trade tax (27.16) and corporation tax at a rate of 25% plus solidarity surcharge (35.82) on the tax base (135.82) where the local trade tax is deductible as a business expense and from its own tax base (162.99). The remainder (100) is distributed to owners immediately where half of the distribution is subject to personal income tax and solidarity surcharge (25.58). This procedure leaves owners with an after-tax increase in their wealth of 74.42 which translates into a CNTR of 54.34%.

For the sake of our argument, let us assume that we only focus on the income tax and adjust the observed income tax base of dividend recipients (row 4) by $^4/_3$, as in BCS, p. 9. On our adjusted gross income in row 1 we now slap the empirically observed income tax and solidarity surcharge (25.58) and calculate an ETR^{BCS} of 19.19% or 38.38%, depending on whether we adjust the distribution in row 4 (100) – as demonstrated in column 3 – or the tax base (50), which is half the distribution, as in column 4. Proceeding in this fashion, BCS underestimate the CNTR by over 8 percentage points for the full imputation system and by 35.15 or 15.97 percentage points for the half income system, respectively, depending on the adjustment.

In a nutshell, BCS severely underestimate the CNTR of over half of the income components for the top 0.1% quantile and over 94% for the 0.001% quantile. Furthermore, the bias increases in 2002 as the share of dividends taxed under the half income system will be higher than in 2001.

Tax Treatment of Losses

With regard to the tax treatment of losses, intertemporal and intra-temporal offset measures must be distinguished. With regard to intra-temporal offsets, German tax law restricted the amount of losses that could be offset between the seven income types mentioned in § 2 of the German Income Tax Code to 51,500 € for single filers and 103,000 € for joint filers. Losses above this level were subject to a limited offset. BCS[25] mention this source of complications but any adjustment mechanism remains undiscussed. The remaining losses should be factored into both the economic income and the tax liability. While the income adjustment is rather straightforward, the determination of tax liability assigned to the examined year cannot be conducted satisfactorily. One would need the present value of the tax liability reductions due to the remaining losses from the year under examination accruing in

[25] Cf. footnote 9.

the future. Potentially, this assessment could stretch out into a very distant future if taxable income for the offset is slow to come by or if the carry-forward is sizable. The restrictions only affect 0.05% of the taxpayers with a positive income but it should be noted that 0.55% of the top 1% quantile is affected with a mean of 3,374,415 €, showing that the flaw in the calculation is more severe for the rich.
Intertemporal offsets were limited to a carry-back to the previous year up to 511,500 € and unlimited carry-forwards. Bearing these legal rules in mind, the tax payments attributable to a single period must be adjusted. BCS' starting point for computing the adjusted gross income in t, the *overall amount of income* (Gesamtbetrag der Einkünfte), does not include the intertemporal offsets. However, their measure of the assessed income tax liability in t, i.e. the nominator of ETR_t^{BCS}, is lowered by the intertemporal loss deductions while the denominator – the AGI – remains constant. The reduction of the assessed income tax liability and the nominator, respectively, has to be attributed to the period when the loss occurred, i.e. to $t+1$ or $t-x$.
In a nutshell, the tax liability attributable to the examined period t is the assessed income tax liability that would have been due without any intertemporal loss offsets in t but – as argued above – lowered by the present value of tax liability reductions from losses that occurred in t but could not be offset in that period.
The following example gives an idea of the degree to which the omission of losses can bias the CNTR. Assume a (simplified) scenario with four single taxpayers where each investor has an overall amount of income of 5,000,000 € in 2001. To steer clear of the complications of taxes other than the income tax, assume that his overall amount of income consists of wage income only. Investor 1 has neither loss carry-back nor carry-forwards. Investor 2 suffers a loss of 1,000,000 € in 2002, such that he can carry back 511,500 € to 2001. Investor 3 has a loss carry-forward of 1,000,000 € which he fully offsets in 2001. Investor 4 features both a loss carry-forward of 1,000,000 € and a loss of 1,000,000 € in 2002 that is carried back to 2001 to the tune of 511,500 €. We now compute the assessed income tax liability for each investor in 2001 (row 5) and the tax payment assigned to 2001 (row 6). By dividing row 5 by the income in 2001, one arrives at ETR^{BCS}. Row 8 delivers the comprehensive nominal tax rate including the tax payment assigned to 2001 ($CNTR^{2001}$).
While each investor receives the same income in 2001, ETR^{BCS} varies from 35.28% to 50.75% while the $CNTR^{2001}$ is actually 50.75% for all four investors. If one

TABLE 4.6: The Effect of Intertemporal Offsets on the ETR

	Investor	1	2	3	4
(1)	Income in 2001	5,000,000	5,000,000	5,000,000	5,000,000
(2)	Carryback	–	511,500	–	511,500
(3)	Carryforward	–	–	1,000,000	1,000,000
(4)	Taxable Income	5,000,000	4,488,500	4,000,000	3,488,500
(5)	Assessed Income Tax Liability[1]	2,537,359	2,275,638	2,025,684	1,763,963
(6)	Tax Payment assigned to 2001[1]	2,537,359	2,537,359	2,537,359	2,537,359
(7)=(5)/(1)	$\mathbf{ETR^{BCS}}$	**50.75%**	**45.51%**	**40.51%**	**35.28%**
(8)=(6)/(1)	$\mathbf{CNTR^{2001}}$	**50.75%**	**50.75%**	**50.75%**	**50.75%**

[1] Solidarity Surcharge contained in income tax payment

raised the loss carryforward for investor 4, ETR^{BCS} would drop further. The intertemporal loss offset lowers the assessed income tax liability while the income of 2001 is not touched upon. But the reduction of the income tax liability cannot be assigned to 2001 since the losses and the tax reduction must be assigned to the period when the loss occurred. What this paragraph shows is that it is virtually impossible to disentangle taxpayers' tax affairs and focus on a single year to assess the cut that taxation takes out of their income.

With regard to the importance of this issue, the intertemporal loss offset affects 1.38% of all taxpayers, this proportion rises to 9.38% (13.12% /13.31%) for the top 0.1% (0.01% /0.001%) quantile with a mean of 1,200,810 € (3,630,884 €/9,019,291 €)[26], implying that the underestimation bias is not uniformly distributed among the taxpayers. It rather serves to reinforce the impression that the rich are not paying taxes.

Income from renting and leasing

A hazy subject in BCS is the treatment of income from renting and leasing. According to BCS, "losses from renting and leasing exceeding some thresholds (are disregarded) since most of these losses are likely to arise from tax avoidance."[27] Even disregarding the status of the term "tax avoidance" (which is customarily used to denote *perfectly legal* minimization strategies of one's tax liability, as opposed to

[26] Cf. table 4.12.

[27] BCS, p. 23. The threshold is set equal to 2,500 € / 5,000 €.

criminal "tax evasion"), and the reasoning provided, the asymmetric truncation of losses from renting and leasing strikes one as arbitrary.

BCS dwell on this income type extensively and suspect major loopholes for the rich there. According to the calculations by *Müller* (2004, p. 77), income from renting and leasing plays a minor, yet special role in that aggregate income tax revenue from this income type has been consistently negative over the years. With regard to losses from this income type, *Müller* (2004, p. 93) goes on to argue that these are concentrated in the upper quantiles of the German income distribution. Our analyses show[28] that 18.05% of all taxpayers report earnings from renting and leasing. Among these taxpayers, 52.01% report losses. This explains the negative share of -0.14% of the income from renting and leasing in the overall amount of income. It has to be noted, though, that 1,652,865 taxpayers (32,48% of the reported cases) report losses exceeding 5,000 € such that the truncation envisioned by BCS becomes relevant. The effects of this – arbitrary – threshold of 5,000 € on the calculations in BCS are easily explained: the truncation cuts off the lower tail of the distribution of income from rent and leasing, which – artificially – increases the mean income from this income type from -687 € to 4.604 €. *Müller* (2004, p. 92, Fig. 2.2) draws attention to the extremely uneven distribution of losses from renting and leasing for the *years 1989, 1992* and *1995*. The 90 % of taxpayers with the lowest overall amount of income in the respective year declared less than 60 % of the entire amount of losses accrued in this income type, while the remaining right tail of the income distribution declared the missing 40 odd %.

Apart from the arbitrary truncation of losses (which we undo in our calculations in subsection 4.3.2), several problems aggravate the calculations envisioned by BCS. Firstly, it should be borne in mind that real estate management can also be conducted in an unincorporated or incorporated business, making any truncation impossible. Furthermore, any gains or losses from the divestment of real estate would be subject to corporate and personal taxation. Secondly, the data only record the current returns and deductions, but exclude any gains or losses realized upon sale of the real estate itself in the income from renting and leasing. These would show up in the income type "other income" only if the sale took place within the speculation period[29]. Since an investor includes both the returns from renting and the resale price in his calculus, it is impossible to separate the income from renting

[28] Cf. table 4.10.

[29] Cf. our example in table 4.2.

and leasing and the attendant gains and losses from sale. The datasets available for Germany do not allow one to make this connection, though.

With regard to the returns to renting and leasing, the state of the German real estate market is not easily ascertainable. As is true of any real estate market, transactions are infrequent and the terms of these transactions are rarely released. Recent years have witnessed the first concerted efforts toward the compilation of a price index for real estate in Germany. *Hoffmann and Lorenz* (2006, p. 30-36) report several indicators of price developments for German real estate. While the multitude of indicators reported there do not give a consistent picture, they certainly *refute* the notion that outsized gains could be realized in German real estate in the last two decades. Indeed, government had to stimulate real estate investments in East Germany to entice investors at the beginning of the 1990s, where faster depreciation for tax purposes was one of the main instruments. As the diagrams in *Hoffmann and Lorenz* (2006, p. 33) show, prices for East German real estate declined precipitously during the period 1995 to 2005. Given that large-scale pre-tax gains obviously did not materialise, the claim that income from renting and leasing has been a "a vast loophole for tax-saving activities in Germany for decades, *especially in the 1990s*[30]" rings hollow[31].

So far, we can conclude that not only the measure of effective taxation itself is disputable but also the definition of the income does not strictly follow any of the three concepts presented in section 4.2.3. As BCS also omitted the local trade tax, the corporate tax, the intertemporal effects of loss deduction in their computation and regarded economic losses from renting and leasing as "tax avoidance", the numerator – the taxes paid – as well as the denominator – the income – are strongly biased. Furthermore, the quotient of these (biased) numbers cannot be interpreted as a measurement of effective taxation, at the utmost as a nominal tax burden in one year.

4.3 Analysis of Effective Tax Rates on Top Incomes

4.3.1 Description of Available Datasets

BCS describe the components of their dataset and their matching procedure quite

[30] Emphasis added by the authors.

[31] We report empirical results for gains/losses from real estate transactions in table 4.13.

extensively in their chapter 3. The dataset used in BCS is a sample composed of matched information from the SOEP and the FAST 1992-2001: To overcome deficiencies in each of the two datasets, BCS lump them together to form their "integrated ITR-SOEP data set". The matching algorithm employed is described in *Bach et al.* (2007, Appendix 2).

Their first component, the FAST[32] is comprehensively described in *Merz, Vorgrimmler and Zwick* (2006). It is a scientific use file that is collected every three years and contains a 10 % random sample of tax returns filed in the respective year, resulting in approximately 3,000,000 cases out of a population of approximately 30,000,000 tax returns. So far – as of the time of writing – the data for the years 1992, 1995, 1998 and 2001 have been made available. As a stratified sample, this file contains very precise information for certain minority groups of the population whose data are consequently anonymised more strongly. In particular, the right tail of the income distribution is entirely present in the dataset. Overall, more than 600 variables are contained in the dataset.

The time structure suggested by the triennial compilation process for the FAST does not imply a panel structure, though. Crucially for the analysis in BCS and ours, the concatenation of the four yearly files does not allow one to track taxpayers across time. As an anonymised and randomised file, information cannot be linked across years. Assessments of tax burdens are thus only possible for a given individual taxpayer and year.

As an additional crucial hurdle, German Income tax law stipulates the taxation of accounting profits for certain income types, i.e. income from forestry and agriculture, income from business activity and self-employment. The process of the determination of taxable profit is not observable in the data whereas the resulting balance is reported. This contrasts with the other income types where the determination of their contribution to the tax base is observable in much greater depth. As a corollary, the presence or absence of accruals within the reported profits cannot be determined. Any attempt to undo their effects is thus doomed.

Regarding the second major data source for BCS, comprehensive information on the German Socio-Economic Panel, SOEP, compiled by the DIW, Berlin, can be found in *Haisken-DeNew and Frick* (2005, p. 16): "The SOEP was started in 1984

[32] Abbreviation of the German "Faktische Anonymisierung der Steuerstatistik" – De Facto Anonymisation of the Tax Statistics.

as a longitudinal survey of private households and persons in the Federal Republic of Germany. The central aim of this panel study is to collect representative micro-data on persons, households and families...". It is designed as a yearly sample of the same units (households and individuals) and thus *does allow* inference with panel methods. Compared with the FAST samples, the size of the dataset is more modest, varying between 10,000 and 25,000 due to the addition of several new subsamples[33]. On the other hand, the SOEP sample contains a broader cross-section of individuals as it also contains taxpayers who are not legally obliged to file tax returns. With regard to tax variables, it is quite apparent that the SOEP is not designed with the tax researcher in mind and that the tax environment of the respondents must be inferred in a rather piecemeal fashion and with a relatively high degree of uncertainty[34].

Admittedly, a match with the SOEP could theoretically provide a tentative link across years which might yield a panel structure. Yet the reasoning for the matching provided in BCS, p. 6/7 does not mention the panel structure, but is explicitly intended to more accurately reflect conditions in the lower tail of the income distribution.

We tried to replicate this particular dataset as far as possible. Our contract with the provider of the FAST sample did not allow us to match with another data source, though. Consequently, we could not work on the same data. Given the high degree of anonymisation in the right tail of the income distribution prevailing in the FAST dataset, we ran a controlled data retrieval under the supervision of the statistical office which gave us access to the entire population of almost 30,000,000 tax returns (ITR) in Germany in 2001 and 900 variables for each of those[35].

The deficiency of the missing matching is mitigated because we emphasize the conditions in the right tail of the German income distribution. As this particular subgroup, the "target group", is completely represented in our dataset, we can make inferences largely unperturbed by the missing cases for the left tail of the income distribution. Given that our (and BCS') main interest lies in the right tail, we can omit the matching from our analysis[36].

[33] Cf. *Haisken-DeNew and Frick* (2005, p. 26).

[34] Cf. *Wagenhals and Buck* (2007) for a possible solution.

[35] Data handling issues forced us to drop taxpayers with a reported overall amount of income around 0 €. This induced a drop in the sample size to approximately 28,150,000.

[36] In any case, the link thus established would be too flimsy as the small number of 25,000 units in the SOEP has to be matched to the approximately 3,000,000/30,000,000 units in the FAST/ITR. On average, each SOEP unit must account for well over 100/1,000 units in the FAST/ITR database.

4.3.2 "Adjusted Gross Income" Variable in BCS

BCS devote their section 4 and appendix 2 to the derivation of an "adjusted gross income". We argue above[37] that any attempt to measure the economic profit from the data available to researchers is severely restricted. Bearing the limitations in mind, we provide our digest for the derivation of an economically based single-period income. **Table 4.7** displays our adjustments compared to the adjustments by BCS, grouped by the seven income types recognized under German income tax law. We highlight major deviations from the methodology in BCS in bold face.

TABLE 4.7: Definition of Adjusted Gross Income

Income from	Adjustments made by BCS	Necessary adjustments of the taxable income
Wages	• Adjustment for employers' social security contributions	• Adjustment for employers' social security contributions
	• Allowable expenses excluded	• Allowable expenses **included**
	• Social security contributions for civil servants imputed	• Social security contributions for civil servants imputed
	• Tax exempt foreign wage income added	• Tax exempt foreign wage income added

[37] Cf. section 4.2.4 on page 93.

Income from	Adjustments made by BCS	Necessary adjustments of the taxable income
Business Activity	• Includes income from agriculture and forestry, from unincorporated business enterprise and from self-employed activities • Tax exempt profits from outbound business activities included • Tax subsidies added • Income should be adjusted for “tax expenditures" such as depreciation, but data is not available	• Includes income from agriculture and forestry, from unincorporated business enterprise and from self-employed activities • Tax exempt profits from outbound business activities included • Tax subsidies added • Since these “tax expenditures" are “capital expenditures", **an adjustment cannot be justified** • **Local trade tax added** for unincorporated business
Capital Gains	• Includes gains from disposal of enterprises or substantial shareholdings (Income from business activity) and of private investments (Speculation gains § 23)	• Includes gains from disposal of enterprises or substantial shareholdings (Income from business activity) and of private investments (Speculation gains § 23) • “Tax-Free” capital gains **are added**, as far as reported

Income from	Adjustments made by BCS	Necessary adjustments of the taxable income
Interest and Dividends	• Saver Allowance added to income • Recorded net dividend is multiplied by $\frac{4}{3}$	• Saver Allowance added to income • Recorded net dividend under half income system is **adjusted by corporate tax** and **by the local trade tax** • Dividends under full imputation systems **are adjusted by local trade tax**
Renting and Leasing	• Losses exceeding 5,000 € are ignored • Losses from shareholdings of closed property funds, etc. exceeding 2,500 € are ignored	• Losses are **fully included** in income
Transfer Income	• Corrected for the allowance for taxable pensions from employment • Non-taxable share of life annuity funds is added to income (70% of the pension) • Social assistance, housing benefits and other public transfers are taken from the SOEP	• Corrected for the allowance for taxable pensions from employment • Non-taxable share of life annuity funds is **not added** to income • Social assistance, housing benefits and other public transfers **are ignored** since top 1% are unlikely to receive public transfer income

Income from	Adjustments made by BCS	Necessary adjustments of the taxable income
Losses		• Remaining non-deducted losses occurred in 2001 **should be subtracted**.

We include the deductible expenses for income from **wages**, i.e. we do not add these expenses to the adjusted gross income since for example the costs for training or education dilute the taxpayer's economic income while also being recognized as tax deductible. Adding these costs would lead to an overestimation of the adjusted gross income and, consequently, to an underestimation of the nominal tax burden.
BCS describe a category "**business activity**" that bundles the profit income types recognized under German income tax law. Firstly, a line must be drawn between the income from unincorporated businesses and income from self-employment: The local trade tax burden is added to the taxable income as shown in table 4.4 for unincorporated business, while the self-employed are not subject to this tax. As BCS omit the local trade tax entirely, this problem does not crop up in their contribution. Secondly, any – putative – adjustments for "tax expenditures" cannot be justified. BCS term the provisions for pension reserves or depreciation charges "tax expenditures", without properly explaining the implications of this classification. From circumstantial evidence, it seems that any "accruals", i.e. charges that do not coincide with the associated cash flow, are regarded as "in need of correction". Adding these components to the taxable income would imply a fundamental flaw since the tax expenditures either have already been capital expenditures – as in the case of depreciation allowances – or will be wage expenditures in the future – in the case of pension reserves. As mentioned in subsection 4.3.1, the profit- and loss statements of the tax balance sheets underlying the income from unincorporated businesses is not available. Thus, the ability to correct for supposed "loopholes" is restricted by data availability.
The **capital gains** from the sale of an enterprise or a substantial shareholding – even if they are tax-exempted – are added to the speculation gains to form the income from capital gains. The gains realized from the disposal outside the speculative period are not included in the taxable income effectuating an underestimation of

the adjusted gross income from capital gains. These gains must be added to the adjusted gross income but unfortunately they are not observable[38]. It has to be noted that the "tax-free" capital gains are far away from being tax-free. Many studies proved that corporate and dividend taxes reduce the valuation of shares. Thus, the so called "tax-free" capital gains are already taxed at the corporate and the shareholder level[39].

The income from **interest and dividends** represents another major source of error in the BCS approach. Firstly, the bundling of this income stream ties together two heterogeneous sources of income, with dividends carrying a burden from the corporate level while interest is unencumbered by any prior taxation. While BCS multiply the recorded net dividend by the factor 1.333 in order to try to account for the corporation tax under the half income system, they ignore the local trade tax as shown in table 4.5 for both the half income system and the full imputation system. Therefore, we adjust the taxable dividend by adding the local trade tax. Additionally we add the corporation tax under the half income system. Consistent with BCS, we add the saver allowance to the income.

For the reasons discussed in subsection 4.2.5, we do not adjust the income from **renting and leasing**. The consistently negative income that renting and leasing has delivered over the last decades indicates that investors suffered economic losses. Furthermore, we do not go along with the practice in BCS of truncating losses for this particular income type at a threshold of 5,000 €.

We do not match the ITR data and the SOEP datasets, hence in the case of **transfer income**, we cannot account for social assistance, housing benefits or other public transfers. This biases the income generally downwards but the effect for the top incomes is negligible since they are not likely to receive any kind of transfer. In contrast to BCS, we do not add the non-taxable share of life-annuity funds which we regard as a payback of contributions made.

Finally, **losses** occurred in 2001 that could not have been offset against other (pos-

[38] Since the "tax-free" speculation gains are neither reported in the ITR data nor in the SOEP, any matching process does not help to uncover this flaw. With the introduction of the final withholding tax in 2009, capital gains from the disposal of shares will be taxed with 25% regardless of the holding period solving the information problem behind the "tax-free" capital gains.
But we assume the speculation gains to have minor impact on the income in 2001 and 2002, since the substantial interest for tax purposes was lowered to 1% in 1999. Furthermore, the stock market showed rather bad performance in 2001 where the TecDax crashed from over 9,500 points in 2000 to a minimum 683.82 in 2001 and the DAX fell to its 1998 level.

[39] See among others *Ball* (1984) and *Collins and Kemsley* (2000).

itive) sources of income are not integrated in the overall amount of income and thus, should be recognized in the computation of the economic income. However, as the present value of tax liability reduction from the remaining losses is not observable, one cannot correctly account for the intertemporal effect of the restriction to the intra-temporal loss offset. Therefore, we cannot adjust for these losses and *underestimate* the "true" CNTR.
Bearing in mind that the "correct" and representative figure of income can only be derived from any cross-sectional dataset under the restrictions discussed in section 4.2.4, we conclude that our definition of income differs from the *adjusted gross income* in BCS for every income type.

4.3.3 Adjustment for Taxes Paid

In addition to the adjustments enumerated in table 4.7, we also adjust the numerator in equation (4.2) and account for the most important taxes. We detail these corrections in **table 4.8** which itself mirrors table 4.7.

TABLE 4.8: Adjustments to Recognized Tax Burden

Income from	Necessary adjustments to Recognized Tax Burden
Wages	• **Accounted for possible taxes paid** on foreign wages
Business Activity	• **Taxes paid abroad** on tax exempted profits from outbound business activities **included** • Local Trade Tax is **added to the PIT**
Interest and Dividends	• Corporate tax **and** the local trade tax are **added to the PIT** for the half income system • Local trade tax **added** for the full imputation system

Income from	Necessary adjustments to Recognized Tax Burden
Losses	• **Loss carry-forwards** used in 2001 and **carry-backs** from 2002 lower recorded tax burden in 2001. Difference in income tax burden with and without carryforward and carryback is added. • Present value of **tax reduction from non-deducted losses** occurred in 2001 should lower tax liability.

The adjustment of the recognized tax burden follows a simple rule. If you increase the taxable income for any income type in order to derive economic income, you have to account for possible taxes paid on the respective levels. For example, BCS adjust for tax exempted foreign wages income. If these wages carry a tax burden from abroad then these taxes must be added to the PIT. The same problem results for tax exempted profits from outbound business activities with a potential accumulated tax burden. For the correct adjustment, information on the country of origin, the tax rate and the tax base would be necessary.

For the case of **capital gains**, the problem of inferring the accumulated tax burden proves insurmountable. It is neither observable nor measurable with any method as one would have to compare the pre-tax capital gain with the (observable) after-tax capital gain. However, several studies prove that capital gains carry a tax burden, as discussed in the last section.

By contrast, the local trade tax (τ^{LTT}) and the corporate tax (τ^{Corp}) paid on the corporate level can be derived from the information available from the ITR dataset. If an **incorporated business** paid out one Euro of dividends under the half income system, resp. full imputation system, then the profit on the corporate level was $\frac{1}{(1-\tau^{Corp})(1-\tau^{LTT})}$, resp. $\frac{1}{1-\tau^{LTT}}$. Since we adjusted the gross income, we have to adjust the recognized tax burden by the corporate tax and the local trade tax as well. The accumulated tax burden from dividend income on the corporate level (ATB_{Div}), which is added to the PIT and which is defined as the difference between the pre-corporate-tax profit and the pre-income-tax gross dividend, evolves under

the half income system as

$$ATB_{Div} = \left(\tau^{Corp} + \tau^{LTT} - \tau^{Corp} \cdot \tau^{LTT}\right) \frac{Div}{(1-\tau^{Corp})(1-\tau^{LTT})} \tag{4.6}$$

and under the full imputation system respectively as

$$ATB_{Div} = \tau^{LTT} \frac{Div}{1-\tau^{LTT}}. \tag{4.7}$$

Additionally, the local trade tax must be added to the recognized tax burden if a taxpayer has income from an **unincorporated business**. We can observe the lump-sum offset of the local trade tax (*Steuerermäßigung bei Einkünften aus Gewerbebetrieb*) from § 35 denoted as $O_{\S 35}$. Thus, we derive the taxable business income (*Gewerbesteuer-Messbetrag*) by dividing the offset by 1.8. The accumulated tax burden of the income from an unincorporated income ATB_{Uninc} evolves as the product of the taxable business income and the municipal collection rate (*Gewerbesteuer-Hebesatz*) CR_i[40].

$$ATB_{Uninc} = CR_i \frac{O_{\S 35}}{1.8} \tag{4.8}$$

Losses carried forward to 2001 from prior years and/or losses carried back from 2002 into 2001 impact the tax liability due in 2001 yet leave the adjusted gross income untouched. A recognition of the tax reduction solely in 2001 would unduly bias the resulting tax rates downward. For the reasons expounded in subsection 4.2.5, we account for this effect in the manner described at the bottom of table 4.8. The present value of the tax liability reduction from non-deducted losses occurred in 2001 should, but cannot be recognized for the reasons discussed in subsection 4.3.2.

Since taxable capital gains from the disposal of enterprises and substantial shareholdings, business activity and dividends represent the major part of the income of the rich and the super-rich, BCS' measure of "effective taxation" provides an incorrect, downward biased picture of the "tax burden" of the rich.

[40] The index i indicates that we account for regional differences in the collection rate in 2001. The federal state of the taxpayer is known from the dataset which allows us to apply the average collection rate for the individual federal state.

4.3.4 Empirical Results

In this section, we provide our results for the cross-section of income tax returns in Germany in 2001. **Table 4.9** shows the mean comprehensive nominal tax rates emanating from our adjustments, conditional on the quantiles of the income distribution, and the corresponding numbers from BCS, table 4. The last column gives the difference in percentage points between the two sets of results. Note that the adjustments in our approach and in BCS differ in the manner described in table 4.7 so that the allocation to quantiles may not coincide perfectly.

TABLE 4.9: Comparison of the Comprehensive Nominal Tax Rate and ETR^{BCS} for 2001

Quantile	**CNTR**	**ETR$^{\mathbf{BCS}}$**	**$\triangle$=CNTR–ETR$^{\mathbf{BCS}}$**
Decile 1-5	3.60%	2.90%	0.70%
Decile 6-9	12.87%	10.10%	2.77%
Decile 10	23.27%	21.90%	1.37%
Top 1%	37.21%	33.40%	3.81%
Top 0.1%	45.25%	38.20%	7.05%
Top 0.01%	48.05%	38.70%	9.35%
Top 0.001%	50.30%	38.10%	12.20%
Top 0.0001%	51.39%	36.00%	15.39%

Source: ITR 2001, own calculations

Apparently, there is a broad agreement between the results. Within the top 1% quantile, the results diverge markedly, with the differences ranging from 3.81 percentage points for the 1% quantile to 15.39 percentage points for the top 0.0001% quantile. Given the focus on the taxation of "the rich" in BCS, this development acquires a lot of significance. As an explanation for the differences, we provide **table 4.10** on page 120 which clearly shows the change in composition in the adjusted sum of income between the quantiles.

Taxpayers below the 90% quantile rely heavily on domestic wage income (84.54% for the deciles 1-5 and 91.81% for the deciles 6-9 respectively) which is taxed in a cash flow manner, i.e. in an uncontroversial way which requires very few adjustments. Within the upper 1% quantile, though, the importance of wage income declines precipitously from 44.22% for the top 1% quantile to 0.27% for the top

0.0001% while business and capital income assume greater weight. Taxpayers in the top 0.001% (0.0001%) quantile generate 94.24% (99.99%) of their income from capital or business income. These are the income types where our criticism of the BCS approach bites most heavily and where consequently the divergence is greatest.

These results cast grave doubt on the validity of the claim in BCS, p. 17 that "...average effective tax rates paid by the economic elite..." was 38.1% in 2001. The economic elite – defined in BCS as the upper 0.001% quantile – actually faced a CNTR of 50.30% which is over 12 percentage points higher. Bearing all the caveats of our prior analysis in mind, the CNTRs exceed the top statutory income tax rates by several percentage points across the board, highlighting the influence of the local trade tax and the corporation tax on the overall tax burden. On top of that, the relevance of intertemporal loss deductions, depicted in **table 4.12** on page 122, increases with the quantiles, so that the bias arising from the omission of the losses in BCS also increases.

4.3.5 Limitations

Concerning the methodology used, we describe additional[41] limitations in this section. As mentioned above, our main interest lies in the right tail of the income distribution. Note here that we must contend with the fact that the dataset only contains information pertaining to taxation. The definition of "income rich" should thus be read as "taxable income rich", subject to the modifications that we describe in table 4.7. Assume, in the fashion of table 4.1 on page 91, a taxpayer who makes spectacular gains in real estate and shares after the expiration of the speculation period of 10, resp. 1 year: if you further assume that this is his major source of income, he could make millions without the ITR dataset even noticing him. The particular definition of income dictated by German income tax law would effectively allow him to "fly below the radar". Similarly, if a taxpayer's assets consisted of a share in an incorporated business and the business retained its entire earnings for 2001, the taxpayer would be classified as "poor" in the 2001 cross-section because he does not have taxable income[42] in this year. Note how differently this issue would be

[41] Note that the main criticism of the BCS approach is contained in section 4.2.

[42] This is true for the purposes of income taxation. Corporation tax is due on the profit, anyway. There is no link, though, in the data that would enable us to establish a connection.

handled under economic income: if the investments enabled by the retention were profitable, an economic appreciation would be due, generating economic income in the year 2001. Given the inevitable departures of a cross-sectional income definition in the dataset from economic income stressed throughout our article, the classification of taxpayers into quantiles based on the "overall amount of income" (Gesamtbetrag der Einkünfte) is certainly problematic. These problems are shared with BCS and anyone who makes inferences from tax data.

4.4 Conclusion

We have provided a comprehensive examination of the most salient and controversial results in *Bach et al.* (2008). Our argumentation rests on several layers of critique:

Very fundamentally, we argue that a restriction on the income tax burden has lowered the "effective tax rates" calculated in BCS by up to 12 percentage points. We replace it with a comprehensive measure of tax burden and calculate tax rates that are well *above* the nominal tax rate. In particular, the omission of the local trade tax and the corporation tax from the calculations is bound to yield a skewed picture of the nominal tax burden. Given that income from business activity is concentrated in the upper quantiles of the income distribution, this omission does not cause a downward shift across all taxpayers, but an unambiguous bias toward lighter taxation of the rich. We also find that the concept of effective tax rates is not applicable, as the very notion of ETRs precludes their meaningful deployment in a single period calculus. Furthermore, we have questioned the role of accruals for tax purposes. By their very nature, these items must be interpreted in an intertemporal perspective, as they are designed to transform cash flows into tax bases over time. We also find substantial shortcomings with respect to the treatment of tax losses, as intertemporal aspects reassert themselves.

Reassuringly, the official tax statistics (*Federal Ministry of Finance*, 2001, p. 21) for 2001 provide a glimpse at the massive contribution that the income rich make toward income tax receipts: while the upper 50 % of German income taxpayers contributed over 90 % of income tax receipts in 2001, the upper ten percent contribute over 50 % and the upper five percent, the smallest quantile for which these numbers are available for 2001, chips in 40 odd %.

We have shown that the complaint in BCS, p. 20: "... effective tax progression stops at income levels within the top percentile, i.e. the effective tax rate is not monotonically increasing in gross income within the top percentile of the income distribution" is unsubstantiated and disregards elementary rules for measuring effective tax rates.

Appendix: Additional Tables

TABLE 4.10: Adjusted Income Components as Percentage of Adjusted Sum of Incomes

	Income from:										
Quantiles	**§13**[1]	**§15**[2]	*CG*[3]	**§18**[4]	*CG*[5]	**§19**[6]	**§20**[7]	**§21**[8]	**§22**[9]	*CG*[10]	**Total**
Losses	2.99	42.27	*1.34*	7.53	*0.37*	29.81	5.84	17.53	-5.97	*7.42*	100
Decile 1–5	0.66	3.77	*0.10*	1.67	*0.01*	84.54	3.16	–0.05	6.23	*–0.05*	100
Decile 6–9	0.65	3.91	*0.08*	1.60	*0.01*	91.81	1.45	–0.18	0.75	*–0.02*	100
Decile 10	0.73	8.29	*0.46*	9.14	*0.13*	78.03	3.66	–0.40	0.49	*–0.08*	100
Top 1%	0.82	17.89	*2.50*	24.63	*0.55*	44.22	12.31	–0.79	0.72	*–0.44*	100
Top 0.1%	0.75	37.83	*5.75*	12.75	*0.75*	25.09	21.22	0.63	1.46	*0.36*	100
Top 0.01%	0.36	58.34	*12.11*	2.81	*0.25*	10.79	27.71	–2.12	1.89	*–0.20*	100
Top 0.001%	0.00	68.84	*11.36*	1.64	*0.00*	3.04	25.40	0.51	0.58	*0.18*	100
Top 0.0001%	–0.02	78.66	*3.89*	0.06	*0.00*	0.27	21.33	–0.30	0.00	*–0.43*	100
Total	**0.66**	**4.23**	*0.13*	**2.37**	*0.02*	**85.89**	**2.50**	**–0.14**	**3.43**	*–0.04*	**100**

Source: ITR 2001, own calculations.
[1] Income from Agriculture and Forestry – [2] Income from Unincorporated Business – [3] Capital Gains from § 16 and 17 (included in [2]) – [4] Income from Self-Employment – [5] Capital Gains from § 18 (included in [4]) [6] Income from Wages – [7] Income from Capital Income – [8] Income from Renting and Leasing – [9] Other Income – [10] Speculation Gains (included in [9])

TABLE 4.11: Distribution of Adjusted Income Components among Quantiles

	Income from:										
Quantiles	**§13**[1]	**§15**[2]	*CG*[3]	**§18**[4]	*CG*[5]	**§19**[6]	**§20**[7]	**§21**[8]	**§22**[9]	*CG*[10]	**SoI**
Losses	–1.02	–4.49	*1.14*	–0.43	*0.50*	0.15	0.52	61.74	0.97	*4.15*	–0.32
Decile 1–5	24.80	11.73	*7.89*	5.73	*5.38*	23.35	19.02	–17.73*	68.37	*17.19*	22.43
Decile 6–9	43.83	24.95	*8.64*	16.16	*10.76*	51.36	19.91	42.33	20.39	*21.78*	46.36
Decile 10	32.38	67.81	*82.32*	78.54	*83.37*	25.13	60.55	13.66	10.28	*56.88*	31.53
Top 1%	9.88	44.09	*71.29*	39.49	*57.75*	4.15	42.85	–21.19*	4.04	*40.87*	9.94
Top 0.1%	3.08	29.42	*54.66*	6.95	*21.03*	0.89	25.43	–14.37*	1.76	*21.71*	4.07
Top 0.01%	0.49	15.90	*33.03*	0.69	*2.22*	0.16	11.23	–1.53*	0.77	*1.95*	1.69
Top 0.001%	–0.01	6.32	*9.23*	0.13	*0.02*	0.02	4.26	–0.58*	0.27	*–2.86**	0.62
Top 0.0001%	0.00	1.79	*0.89*	0.00	*0.00*	0.00	1.07	0.21	0.00	*1.25*	0.16
Total	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**

Source: ITR 2001, own calculations.
[1] Income from Agriculture and Forestry – [2] Income from Unincorporated Business – [3] Capital Gains from § 16 and 17 (included in [2]) – [4] Income from Self-Employment – [5] Capital Gains from § 18 (included in [4]) [6] Income from Wages – [7] Income from Capital Income – [8] Income from Renting and Leasing – [9] Other Income – [10] Speculation Gains (included in [9]) – * Negative percentage indicates positive income since overall income from renting and leasing and from speculation gains respectively is negative.

TABLE 4.12: Relevance of Intertemporal Loss Deductions

	RL §2[1]	**Intertemporal LD**[2]				**Total**[3]
Quantiles		**Total**	*Mean (€)*	**CB**	**CF**	
Losses	99.51%	0.09%	*19,894.51*	0.00%	0.09%	99.51%
Decile 1–5	0.04%	1.90%	*10,809.46*	0.22%	1.69%	4.01%
Decile 6–9	0.03%	0.71%	*37,038.94*	0.18%	0.53%	0.74%
Decile 10	0.16%	1.57%	*187,327.72*	0.44%	1.16%	1.06%
1%	0.55%	4.83%	*415,176.84*	1.34%	3.65%	5.23%
0.1%	0.71%	9.38%	*1,200,816.34*	2.84%	7.03%	9.86%
0.01%	0.72%	13.12%	*3,630,884.23*	4.02%	9.97%	13.66%
0.001%	0.00%	13.31%	*9,019,291.05*	3.60%	11.15%	13.67%
0.0001%	0.00%	22.22%	*9,092,094.67*	3.70%	18.52%	22.22%
Total	**1.09%**	**1.38%**	*36,062,25*	**0.23%**	**1.16%**	**2.46%**

[1] Percentage of taxpayers with remaining non-deducted intratemporal losses of 2001 (§ 2 German Tax Code) – [2] Percentage of taxpayers with intertemporal loss deductions in 2001. Total percentage and proportion of carrybacks from 2002 (CB) and all carryforwards from prior years reported –
[3] Cumulates the cases, where the CNTR is biased either due to a remaining non-deducted intratemporal loss or an intertemporal loss offset as percentage of all taxpayers.

TABLE 4.13: Reporting Frequency of Income Components

	Income from:						
Quantiles	**§13**[1]	**§15**[2]	**§18**[3]	**§19**[4]	**§20**[5]	**§21**[6]	**§22/23**[7]
Losses	3.97%	58.18%	11.10%	28.09%	12.42%	41.47%	17.68%
Decile 1–5	1.87%	12.49%	3.46%	82.09%	9.40%	12.73%	21.51%
Decile 6–9	2.27%	12.71%	5.39%	94.82%	9.46%	18.84%	8.38%
Decile 10	2.13%	25.28%	20.21%	89.32%	26.59%	42.11%	7.80%
Top 1%	2.80%	53.36%	38.96%	75.15%	65.09%	72.57%	15.79%
Top 0.1%	4.27%	75.45%	32.63%	69.62%	89.07%	82.65%	23.50%
Top 0.01%	6.99%	89.46%	32.52%	63.32%	94.69%	86.05%	27.50%
Top 0.001%	10.07%	94.96%	46.40%	58.99%	95.32%	87.05%	35.97%
Top 0.0001%	14.81%	96.30%	62.96%	62.96%	96.30%	85.19%	33.33%
Total	2.05%	13.84%	5.88%	87.84%	19.62%	18.05%	14.96%
Thereof Reported Losses							
Losses	62.15%	93.38%	77.90%	36.01%	12.17%	85.28%	0.62%
Decile 1–5	12.66%	30.41%	17.25%	0.36%	2.56%	43.03%	0.08%
Decile 6–9	9.65%	25.58%	14.79%	0.02%	5.17%	55.49%	0.42%
Decile 10	13.87%	27.27%	8.97%	0.05%	6.19%	59.65%	2.42%
Top 1%	24.06%	29.63%	5.59%	0.15%	4.54%	58.27%	0.00%
Top 0.1%	39.08%	22.20%	7.60%	0.21%	2.40%	52.08%	2.09%
Top 0.01%	56.92%	13.07%	8.38%	0.17%	2.01%	48.38%	1.96%
Top 0.001%	64.29%	11.36%	3.10%	0.00%	1.51%	46.69%	4.00%
Top 0.0001%	75.00%	7.69%	0.00%	0.00%	0.00%	60.87%	22.22%
Total	11.47%	28.09%	13.54%	0.18%	2.44%	52.01%	0.28%

Source: ITR 2001, own calculations.
[1] Income from Agriculture and Forestry – [2] Income from Unincorporated Business – [3] Income from Self- Employment – [4] Income from Wages – [5] Income from Capital Income – [6] Income from Renting and Leasing – [7] Other Income

CONCLUSION

In this doctoral thesis, I have provided three contributions to the taxation of labor income. Bearing in mind the fact that this is a relatively new and – so far – "fringe" subject in tax research in Business Administration, these contributions constitute an innovative new branch of investigation that should be pursued more vigorously. Given the massive concentration of theoretical research in this area in the hands of economists, this thesis is focused on empirical research. *Adding* new theoretical insights will be a steep uphill struggle, though. The perceived skills of the Business Administration profession in the area of accounting do note bite here, since labor taxation is conducted along the lines of cash-flows. All accounting knowledge that can be brought to bear for research into capital income issues becomes obsolete, and the complexity of microeconomic modeling creeps in.

The subject of labor taxation does promise to feed future generations of researchers, though. The tension between household production and participation in the official labor market, i.e. between taxed and untaxed deployment of human capital, represents one underresearched aspect. *Ferri, Molto and Uriel* (2009), for instance, have recently investigated the trade-off between home produced meals and those served in restaurants for Spain, albeit with a reference to the value-added tax, not the income tax that was researched in this doctoral thesis. Given that this strand of literature has been around since at least *Becker* (1965) and *Boskin* (1975), there should be pent-up demand for a greater emphasis on this gigantic tax base effect (*Freeman and Schettkat*, 2005; *Booth and Coles*, 2010).

Another vast but relatively unexplored area is tackled by *Jacobs and van Wijnbergen*

(2007), who explore an intriguing new avenue for research: Most students finance their studies using debt, be it in the form of state support or from private sources. The fact that the market for student financial support is dominated by debt positions is unusual when compared to firm behavior where both equity and debt positions are regularly sold. This restriction is forced upon workers by legal constraints, though: The prohibition of slavery sees to it that workers cannot commit to pay a residual claim on their wages to an equity holder. In contrast to the worker‘s inability to sell equity in his human capital, labor taxation *can* perform this function and thus complete the capital market for human capital. Indeed, by taxing the cash-flow flowing to the worker after deductions for costs, taxation does claim the residual of human capital payoffs that financiers cannot legally appropriate. This strand of literature is currently emerging and has spawned contributions from *Vandenberghe and Debande* (2008), *Cigno and Luporini* (2009) and *Del Rey and Racionero* (2010).
I sincerely hope that my contributions have alerted the research community to the promises that labor taxation holds for them.

BIBLIOGRAPHY

AHSAN, SYED M. AND TSIGARIS, PANAGIOTIS (2009), The Efficiency Loss of Capital Income Taxation under Imperfect Loss Offset Provisions, *Public Finance Review* 37 (6), pp. 710–731.

ALMEIDA, RITA AND CARNEIRO, PEDRO (2009), The Return to Firm Investments in Human Capital, *Labour Economics* 16 (1), pp. 97–106.

AMEMIYA, TAKESHI AND MACURDY, THOMAS E. (1986), Instrumental Variable Estimation of an Error-Components Model, *Econometrica* 54 (4), pp. 869–880.

AUERBACH, ALAN J. AND SLEMROD, JOEL (1997), The Economic Effects of the Tax Reform Act of 1986, *Journal of Economic Literature* 35 (2), pp. 589–632.

BAAKE, PIO, BORCK, RAINALD AND LÖFFLER, ANDREAS (2004), Complexity and Progressivity in Income Tax Design: Deductions for Work-Related Expenses, *International Tax and Public Finance* 11 (3), pp. 299–312.

BACH, STEFAN (2003), Entfernungspauschale: Kürzung gerechtfertigt, *Wochenbericht des DIW Nr. 40/2003*, pp. 602–608.

BACH, STEFAN, CORNEO, GIACOMO AND STEINER, VIKTOR (2005), Top Incomes and Top Taxes, DIW Discussion Paper 532.

BACH, STEFAN, CORNEO, GIACOMO AND STEINER, VIKTOR (2007), From Top to Bottom: The Entire Distribution of Market Income in Germany 1992-2001, DIW Discussion Paper 683.

BACH, STEFAN, CORNEO, GIACOMO AND STEINER, VIKTOR (2008), Effective Taxation of Top Incomes in Germany, 1992-2002, CESifo Working Paper 2233.

BALDRY, JONATHAN (1998a), Abolishing Income Tax Deductions for Work-Related Expenses, *Agenda* 5 (1), pp. 49–60.

BALDRY, JONATHAN (1998b), Income Tax Deductions for Work-Related Expenses: The Rationale Examined, *Australian Economic Papers* 37 (1), pp. 45–57.

BALDWIN, ROBERT H. (1959), How to Assess Investment Proposals, *Harvard Business Review* 37 (3), pp. 98–104.

BALL, RAY (1984), The Natural Taxation of Capital Gains and Losses when Income Is Taxed, *Journal of Banking and Finance* 8 (3), pp. 471–481.

BALLOT, GERARD, FAKHFAKH, FATHI AND TAYMAZ, EROL (2006), Who Benefits from Training and R&D, the Firm or the Workers?, *British Journal of Industrial Relations* 44 (3), pp. 473–495.

BALTAGI, BADI H. (2008), *Econometric Analysis of Panel Data*, John Wiley & Sons, 4th edition.

BALTAGI, BADI H., BRESSON, GEORGES AND PIROTTE, ALAIN (2003), Fixed Effects, Random Effects or Hausman-Taylor? A Pretest Estimator, *Economics Letters* 79 (3), pp. 361–369.

BALTAGI, BADI H. AND KHANTI-AKOM, SOPHON (1990), On Efficient Estimation with Panel Data: An Empirical Comparison of Instrumental Variables Estimators, *Journal of Applied Econometrics* 5 (4), pp. 401–406.

BECKER, GARY S. (1964), *Human Capital: A Theoretical and Empirical Analysis*, Columbia University Press, New York.

BECKER, GARY S. (1965), A Theory of the Allocation of Time, *The Economic Journal* 75 (299), pp. 493–517.

BLASCH, FRANK AND WEICHENRIEDER, ALFONS J. (2007), When Taxation Changes the Course of the Year: Fiscal-Year Adjustments and the German Tax Reform of 2000-01, *Fiscal Studies* 28 (3), pp. 367–377.

BOOCKMANN, BERNHARD AND STEINER, VIKTOR (2006), Cohort Effects and the Returns to Education in West Germany, *Applied Economics* 38 (10), pp. 1135–1152.

BOOTH, ALISON L. AND COLES, MELVYN G. (2010), Tax Policy and Returns to Education, *Labour Economics* 17 (1), pp. 291–301.

BORCK, RAINALD AND WREDE, MATTHIAS (2005), Political Economy of Commuting Subsidies, *Journal of Urban Economics* 57 (3), pp. 478–499.

BORCK, RAINALD AND WREDE, MATTHIAS (2009), Subsidies for Intracity and Intercity Commuting, *Journal of Urban Economics* 66 (1), pp. 25–32.

BOSKIN, MICHAEL J. (1975), Efficiency Aspects of the Differential Tax Treatment of Market and Household Economic Activity, *Journal of Public Economics* 4 (1), pp. 1–25.

BREUSCH, TREVOR S., MIZON, GRAYHAM E. AND SCHMIDT, PETER (1989), Efficient Estimation Using Panel Data, *Econometrica* 57 (3), pp. 695–700.

CALIENDO, MARCO AND KOPEINIG, SABINE (2008), Some Practical Guidance for the Implementation of Propensity Score Matching, *Journal of Economic Surveys* 22 (1), pp. 31–72.

CAMERON, A. COLIN AND TRIVEDI, PRAVIN K. (2005), *Microeconometrics. Methods and Applications*, Cambridge University Press.

CARD, DAVID (2001), Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems, *Econometrica* 69 (5), pp. 1127–1160.

CHAN, LOUIS K. C., LAKONISHOK, JOSEF AND SOUGIANNIS, THEODORE (2001), The Stock Market Valuation of Research and Development Expenditures, *The Journal of Finance* 56 (6), pp. 2431–2456.

CHRISTIANSEN, CHARLOTTE, JOENSEN, JUANNA SCHRÖTER AND NIELSEN, HELENA SKYT (2007), The Risk-Return Trade-Off in Human Capital Investment, *Labour Economics* 14 (6), pp. 971–986.

CIGNO, ALESSANDRO AND LUPORINI, ANNALISA (2009), Scholarships or Student Loans? Subsidizing Higher Education in the Presence of Moral Hazard, *Journal of Public Economic Theory* 11 (1), pp. 55–87.

COGAN, JOHN F. (1981), Fixed Costs and Labor Supply, *Econometrica* 49 (4), pp. 945–963.

COHEN, DANIEL AND SOTO, MARCELO (2007), Growth and Human Capital: Good Data, Good Results, *Journal of Economic Growth* 12 (1), pp. 51–76.

COLLINS, JULIE H. AND KEMSLEY, DEEN (2000), Capital Gains and Dividend Taxes in Firm Valuation: Evidence of Triple Taxation, *The Accounting Review* 75 (4), pp. 405–427.

COLLINS, KIRK A. AND DAVIES, JAMES B. (2004), Measuring Effective Tax Rates on Human Capital: Methodology and an Application to Canada, in: SØRENSEN, PETER BIRCH (Ed.), *Measuring the Tax Burden on Capital and Labor*, CESifo Seminar Series, chapter 6, MIT Press, pp. 171–211.

CONSTANTINIDES, GEORGE M. AND DUFFIE, DARRELL (1996), Asset Pricing with Heterogeneous Consumers, *The Journal of Political Economy* 104 (2), pp. 219–240.

CONTI, GABRIELLA (2005), Training, Productivity and Wages in Italy, *Labour Economics* 12 (4), pp. 557–576.

DEARDEN, LORRAINE, REED, HOWARD AND VAN REENEN, JOHN (2006), The Impact of Training on Productivity and Wages: Evidence from British Panel Data, *Oxford Bulletin of Economics and Statistics* 68 (4), pp. 361–421.

DEL REY, ELENA AND RACIONERO, MARÍA (2010), Financing Schemes for Higher Education, *European Journal of Political Economy* 26 (1), pp. 104–113.

DELONG, J. BRADFORD AND MAGIN, KONSTANTIN (2009), The U.S. Equity Return Premium: Past, Present, and Future, *Journal of Economic Perspectives* 23 (1), pp. 193–208.

DEVEREUX, MICHAEL P. AND GRIFFITH, RACHEL (1999), The Taxation of Discrete Investment Choices (Revision 2), Working Paper W98/16, The Institute for Fiscal Studies.

DILLER, MARKUS AND GROTTKE, MARKUS (2010), Grenzen und Erweiterungsmöglichkeiten der investitionsneutralen Besteuerung nach dem ökonomischen Gewinn, *Zeitschrift für Betriebswirtschaft* 80 (2), pp. 123–146.

DUAN, NAIHUA (1983), Smearing Estimate: A Nonparametric Transformation Method, *Journal of the American Statistical Association* 78 (383), pp. 605–610.

DUSTMANN, CHRISTIAN AND SCHÖNBERG, UTA (2009), Training and Union Wages, *Review of Economics and Statistics* 91 (2), pp. 363–376.

EDWARDS, COURTNEY, LANG, MARK, MAYDEW, EDWARD AND SHACKELFORD, DOUGLAS (2004), Germany's Repeal of the Corporate Capital Gains Tax: The Equity Market Response, *Journal of the American Taxation Association* 26 (suppl.), pp. 73–97.

ELSCHNER, CHRISTINA AND SCHWAGER, ROBERT (2007), A Simulation Method to Measure the Effective Tax Rate on Highly Skilled Labor, *FinanzArchiv: Public Finance Analysis* 63 (4), pp. 563–582.

FEDERAL MINISTRY OF FINANCE (2001), Datensammlung zur Steuerpolitik, downloadable at: www.starweb.hessen.de, retrieved 4/24/08.

FEDERAL MINISTRY OF FINANCE (2008), BVerfG-Urteil vom 9. Dezember 2008 (2 BvL 1/07 u.a.) zur Entfernungspauschale, 30.12.2008 - IV C 5 -S 2351 - 08/10005, Federal Tax Bulletin 2009 I, p. 28.

FEDERAL STATISTICAL OFFICE (2005), Finanzen und Steuern, Jährliche Einkommensteuerstatistik auf Basis der Geschäftsstatistik der Finanzverwaltung, Fachserie 14 / Reihe 7.1.1, Wiesbaden.

FEDERAL STATISTICAL OFFICE (2006), Finanzen und Steuern, Jährliche Einkommensteuerstatistik auf Basis der Geschäftsstatistik der Finanzverwaltung, Fachserie 14 / Reihe 7.1.1, Wiesbaden.

FEDERAL STATISTICAL OFFICE (2007), Finanzen und Steuern, Jährliche Einkommensteuerstatistik, Fachserie 14 / Reihe 7.1.1, Wiesbaden.

FEDERAL STATISTICAL OFFICE (2008a), Finanzen und Steuern, Jährliche Einkommensteuerstatistik, Fachserie 14 / Reihe 7.1.1, Wiesbaden.

FEDERAL STATISTICAL OFFICE (2008b), *Statistical Yearbook 2008 for the Federal Republic of Germany*, Wiesbaden.

FEDERAL STATISTICAL OFFICE (2009), *Statistical Yearbook 2009 for the Federal Republic of Germany*, Wiesbaden.

FERRI, JAVIER, MOLTO, MARIA LUISA AND URIEL, EZEQUIEL (2009), Time Use and Food Taxation in Spain, *FinanzArchiv: Public Finance Analysis* 65 (3), pp. 313–334.

FITZENBERGER, BERND AND VÖLTER, ROBERT (2007), Long-Run Effects of Training Programs for the Unemployed in East Germany, *Labour Economics* 14 (4), pp. 730–755.

FRANZ, WOLFGANG (2009), *Arbeitsmarktökonomik*, Springer, Berlin, Heidelberg, 7th edition.

FREEMAN, RICHARD B. AND SCHETTKAT, RONALD (2005), Marketization of Household Production and the EU-US Gap in Work, *Economic Policy* 20 (41), pp. 6–50.

FUEST, CLEMENS AND HUBER, BERND (2001), Tax Progression and Human Capital in Imperfect Labour Markets, *German Economic Review* 2 (1), pp. 1–18.

GARCÍA-MAINAR, IMMACULADA AND MONTUENGA-GÓMEZ, VICTOR M. (2005), Education Returns of Wage Earners and Self-Employed Workers: Portugal vs. Spain, *Economics of Education Review* 24 (2), pp. 161–170.

GENSER, BERND AND REUTTER, ANDREAS (2007), Moving Towards Dual Income Taxation in Europe, *FinanzArchiv: Public Finance Analysis* 63 (3), pp. 436–456.

GREENE, WILLIAM (2006), Censored Data and Truncated Distributions, in: MILLS, TERENCE C. AND PATTERSON, KERRY (Eds.), *Palgrave Handbook of Econometrics*, volume 1, Palgrave Macmillan, pp. 695–736.

GREENE, WILLIAM H. (2008), *Econometric Analysis*, Prentice Hall, Upper Saddle River, NJ, 6th edition.

GRUBER, JONATHAN AND POTERBA, JAMES (1994), Tax Incentives and the Decision to Purchase Health Insurance: Evidence from the Self-Employed, *The Quarterly Journal of Economics* 109 (3), pp. 701–733.

GUTIERREZ-I-PUIGARNAU, EVA AND VAN OMMEREN, JOS (2009), Labour Supply and Commuting, Mimeo, SOEPpapers, downloadable at www.diw.de, DIW Berlin, retrieved 12/15/2009.

HAISKEN-DENEW, JOHN P. AND FRICK, JOACHIM R. (2005), Desktop Companion to the German Socio-Economic Panel (SOEP), Working Paper, DIW Berlin, downloadable at www.diw.de, retrieved 1/10/2010.

HAISKEN-DENEW, JOHN P. AND HAHN, MARKUS (2006), PanelWhiz: A Flexible Modularized Stata Interface for Accessing Large Scale Panel Data Sets, mimeo, downloadable at www.panelwhiz.eu, retrieved 4/1/2010.

HALL, BRONWYN H. AND ORIANI, RAFFAELE (2006), Does the Market Value R&D Investment by European Firms? Evidence from a Panel of Manufacturing Firms in France, Germany, and Italy, *International Journal of Industrial Organization* 24 (5), pp. 971–993.

HAUSMAN, JERRY A. (1978), Specification Tests in Econometrics, *Econometrica* 46 (6), pp. 1251–1271.

HAUSMAN, JERRY A. AND TAYLOR, WILLIAM E. (1981), Panel Data and Unobservable Individual Effects, *Econometrica* 49 (6), pp. 1377–1398.

HOFFMANN, JOHANNES AND LORENZ, ANDREAS (2006), Real Estate Price Indices for Germany: Past, Present and Future, Paper for the OECD-IMF Workshop 7, OECD, Paris.

HUNDSDOERFER, JOCHEN, KIESEWETTER, DIRK AND SURETH, CAREN (2008), Forschungsergebnisse in der Betriebswirtschaftlichen Steuerlehre - eine Bestandsaufnahme, *Zeitschrift für Betriebswirtschaft* 78 (1), pp. 61–139.

HUNGERBÜHLER, MATHIAS (2007), Tax Progression and Training in a Matching Framework, *Labour Economics* 14 (2), pp. 185–200.

ICHINO, ANDREA AND WINTER-EBMER, RUDOLF (1999), Lower and Upper Bounds of Returns to Schooling: An Exercise in IV Estimation with Different Instruments, *European Economic Review* 43 (4-6), pp. 889–901.

ISMER, ROLAND, KAUL, ASHOK AND RATH, SILKE (2008), Der verfassungsrechtliche Streit um die Entfernungspauschale: eine ökonomisch-empirische Betrachtung, *Finanz-Rundschau* 2, pp. 58–67.

JACKSON, SCOTT B., LIU, XIAOTAO (KELVIN) AND CECCHINI, MARK (2009), Economic Consequences of Firms' Depreciation Method Choice: Evidence from Capital Investments, *Journal of Accounting and Economics* 48 (1), pp. 54–68.

JACOB, MARTIN, NIEMANN, RAINER AND WEISS, MARTIN (2008), The Rich Demystified - A Reply to Bach, Corneo, and Steiner (2008), CESifo Working Paper 2478.

JACOBS, BAS AND BOVENBERG, A. LANS (2009), Human Capital and Optimal Positive Taxation of Capital Income, *International Tax and Public Finance* 17, doi: 10.1007/s10797-009-9120-5.

JACOBS, BAS AND VAN WIJNBERGEN, SWEDER J. G. (2007), Capital-Market Failure, Adverse Selection and Equity Financing of Higher Education, *FinanzArchiv: Public Finance Analysis* 63 (1), pp. 1–32.

JAPPELLI, TULLIO AND PISTAFERRI, LUIGI (2003), Tax Incentives and the Demand for Life Insurance: Evidence from Italy, *Journal of Public Economics* 87 (7-8), pp. 1779–1799.

JUDD, KENNETH L. (1998), Taxes, Uncertainty, and Human Capital, *American Economic Review* 88 (2), pp. 289–292.

JUDD, KENNETH L. (2000), Is Education as Good as Gold? A Portfolio Analysis of Human Capital Investment, Mimeo, Hoover Institution, Stanford, CA.

KALWIJ, ADRIAAN (2000), Estimating the Economic Return to Schooling on the Basis of Panel Data, *Applied Economics* 32 (1), pp. 61–71.

KAPLOW, LOUIS (1996), On the Divergence between "Ideal" and Conventional Income-Tax Treatment of Human Capital, *American Economic Review* 86 (2), pp. 347–352.

KIM, JEONG-BON, CHUNG, RICHARD AND FIRTH, MICHAEL (2003), Auditor Conservatism, Asymmetric Monitoring, and Earnings Management, *Contemporary Accounting Research* 20 (2), pp. 323–359.

KING, MARVIN A. AND FULLERTON, DON (1984), *The Taxation of Income from Capital*, Chicago, London.

KLOAS, JUTTA AND KUHFELD, HARTMUT (2003), Entfernungspauschale: Bezieher hoher Einkommen begünstigt, *Wochenbericht des DIW Nr. 42/2003* , pp. 623–629.

KNIRSCH, DEBORAH (2007), Measuring Tax Distortions with Neutrality-Based Effective Tax Rates, *Review of Managerial Science* 1 (2), pp. 151–165.

KUBIK, JEFFREY D. (2004), The Incidence of Personal Income Taxation: Evidence from the Tax Reform Act of 1986, *Journal of Public Economics* 88 (7-8), pp. 1567–1588.

KUCKULENZ, ANJA (2007), Wage and Productivity Effect of Continuing Training in Germany: A Sectoral Analysis, ZEW Discussion Paper 06-025, Mannheim.

KUCKULENZ, ANJA AND MAIER, MICHAEL (2006), Heterogeneous Returns to Training: An Analysis with German Data Using Local Instrumental Variables, ZEW Discussion Paper 06-002, Mannheim.

KUCKULENZ, ANJA AND ZWICK, THOMAS (2003), The Impact of Training on Earnings – Differences between Participant Groups and Training Forms, ZEW Discussion Paper 03-57, Mannheim.

KUMAR, ANIL (2008), Labor Supply, Deadweight Loss and Tax Reform Act of 1986: A Nonparametric Evaluation Using Panel Data, *Journal of Public Economics* 92 (1-2), pp. 236–253.

KYRIAZIDOU, EKATERINI (1997), Estimation of a Panel Data Sample Selection Model, *Econometrica* 65 (6), pp. 1335–1364.

LECHNER, MICHAEL AND WUNSCH, CONNY (2009), Are Training Programs More Effective when Unemployment Is High?, *Journal of Labor Economics* 27 (4), pp. 653–692.

LEUVEN, EDWIN (2005), The Economics of Private Sector Training: A Survey of the Literature, *Journal of Economic Surveys* 19 (1), pp. 91–111.

LEUVEN, EDWIN AND OOSTERBEEK, HESSEL (2004), Evaluating the Effect of Tax Deductions on Training, *Journal of Labor Economics* 22 (2), pp. 461–488.

LEV, BARUCH AND SOUGIANNIS, THEODORE (1996), The Capitalization, Amortization, and Value-Relevance of R&D, *Journal of Accounting and Economics* 21 (1), pp. 107–138.

LINTNER, JOHN (1965), The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *The Review of Economics and Statistics* 47 (1), pp. 13–37.

MADDALA, GANGADHARRAO S. (1983), *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.

MERZ, JOACHIM, VORGRIMMLER, DANIEL AND ZWICK, MARKUS (2006), De Facto Anonymised Microdata File on Income Statistics 1998, *Journal of Applied Social Science Studies* 126 (2), pp. 313–327.

MINCER, JACOB (1974), *Schooling, Experience and Earnings*, National Bureau of Economic Research, New York.

MUEHLER, GRIT, BECKMAN, MICHAEL AND SCHAUENBERG, BERND (2007), The Returns to Continuous Training in Germany: New Evidence from Propensity Score Matching Estimators, *Review of Managerial Science* 1 (3), pp. 209–235.

MÜLLER, HEIKO (2004), *Das Aufkommen der Steuern vom Einkommen in Deutschland*, Deutscher Universitäts-Verlag, Wiesbaden.

NG, CHEN FENG (2008), Commuting Distances in a Household Location Choice Model with Amenities, *Journal of Urban Economics* 63 (1), pp. 116–129.

NIELSEN, SØREN BO AND SØRENSEN, PETER BIRCH (1997), On the Optimality of the Nordic System of Dual Income Taxation, *Journal of Public Economics* 63 (3), pp. 311–329.

NIEMANN, RAINER, BACHMANN, MARK AND KNIRSCH, DEBORAH (2003), Was leisten die Effektivsteuersätze des European Tax Analyzer?, *Die Betriebswirtschaft* 63 (2), pp. 123–137.

OSMA, BEATRIZ GARCÍA AND YOUNG, STEVEN (2009), R&D Expenditure and Earnings Targets, *European Accounting Review* 18 (1), pp. 7–32.

PALACIOS-HUERTA, IGNACIO (2001), The Human Capital of Stockholders and the International Diversification Puzzle, *Journal of International Economics* 54 (2), pp. 309–331.

PALACIOS-HUERTA, IGNACIO (2003a), An Empirical Analysis of the Risk Properties of Human Capital Returns, *The American Economic Review* 93 (3), pp. 948–964.

PALACIOS-HUERTA, IGNACIO (2003b), The Robustness of the Conditional CAPM with Human Capital, *Journal of Financial Econometrics* 1 (2), pp. 272–289.

PANTEGHINI, PAOLO M. (2001), Dual Income Taxation: The Choice of the Imputed Rate of Return, *Finnish Economic Papers* 14 (1), pp. 5–13.

PANTZALIS, CHRISTOS AND PARK, JUNG CHUL (2009), Equity Market Valuation of Human Capital and Stock Returns, *Journal of Banking & Finance* 33 (9), pp. 1610–1623.

PARRY, IAN W. H. (2002), Tax Deductions and the Marginal Welfare Cost of Taxation, *International Tax and Public Finance* 9 (5), pp. 531–552.

PARRY, IAN W. H. AND BENTO, ANTONIO (2001), Revenue Recycling and the Welfare Effects of Road Pricing, *The Scandinavian Journal of Economics* 103 (4), pp. 645–671.

RICHTER, WOLFRAM F. (2006), Efficiency Effects of Tax Deductions for Work-Related Expenses, *International Tax and Public Finance* 13 (6), pp. 685–699.

RICHTER, WOLFRAM F. AND SÖHN, H. (2008), Streichung der Entfernungspauschale - (verfassungs-) rechtliche und ökonomische Aspekte, *Steuer und Wirtschaft*, pp. 117–133.

ROBINSON, LESLIE A. AND SANSING, RICHARD (2008), The Effect of "Invisible" Tax Preferences on Investment and Tax Preference Measures, *Journal of Accounting and Economics* 46 (2-3), pp. 389–404.

ROSEN, HARVEY S. (1982), Taxation and On-The-Job-Training Decisions, *The Review of Economics and Statistics* 64 (3), pp. 442–449.

SANDMO, AGNAR (1990), Tax Distortions and Household Production, *Oxford Economic Papers* 42 (1), pp. 78–90.

SARGAN, J. DENIS (1958), The Estimation of Economic Relationships Using Instrumental Variables, *Econometrica* 26 (3), pp. 393–415.

SCHAFFER, MARK E. AND STILLMAN, STEVEN (2006), xtoverid: Stata Module to Calculate Tests of Overidentifying Restrictions after xtreg, xtivreg, xtivreg2 and xthtaylor, downloadable at www.ideas.repec.org, retrieved 1/2/2009.

SCHNEIDER, KERSTIN (2005), Union Wage Setting and Progressive Income Taxation with Heterogeneous Labor: Theory and Evidence from the German Income Tax Reforms 1986-1990, *Labour Economics* 12 (2), pp. 205–222.

SELDEN, THOMAS M. (2009), The Impact of Increased Tax Subsidies on the Insurance Coverage of Self-Employed Families: Evidence from the 1996-2004 Medical Expenditure Panel Survey, *Journal of Human Resources* 44 (1), pp. 115–139.

SHARPE, WILLIAM F. (1964), Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *The Journal of Finance* 19 (3), pp. 425–442.

SLOVIN, MYRON B. AND SUSHKA, MARIE E. (1993), Ownership Concentration, Corporate Control Activity, and Firm Value: Evidence from the Death of Inside Blockholders, *The Journal of Finance* 48 (4), pp. 1293–1321.

SMALL, KENNETH A., WINSTON, CLIFFORD AND YAN, JIA (2005), Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability, *Econometrica* 73 (4), pp. 1367–1382.

SØRENSEN, PETER BIRCH (2004), Measuring Taxes on Capital and Labor: An Overview of Methods and Issues, in: SØRENSEN, PETER BIRCH (Ed.), *Measuring the Tax Burden on Capital and Labour*, CESifo Seminar Series, chapter 1, MIT Press, pp. 1–34.

STEENKEN, ULF (2002), *Die Zulässigkeit gesetzlicher Pauschalierungen im Einkommensteuerrecht*, Juristische Reihe Tenea, Dissertation, Passau.

SU, QING AND DESALVO, JOSEPH S. (2008), The Effect of Transportation Subsidies on Urban Sprawl, *Journal of Regional Science* 48 (3), pp. 567–594.

TREMBLAY, JEAN-FRANCOIS (2009), Taxation and Skills Investment in Frictional Labor Markets, *International Tax and Public Finance* 17 (1), pp. 52–66.

TROSTEL, PHILIP A. (1993), The Effect of Taxation on Human Capital, *Journal of Political Economy* 101 (2), pp. 327–350.

VAN OMMEREN, JOS AND FOSGERAU, MOGENS (2009), Workers' Marginal Costs of Commuting, *Journal of Urban Economics* 65 (1), pp. 38–47.

VAN OMMEREN, JOS AND VAN DER STRAATEN, J. WILLEMIJN (2008), The Effect of Search Imperfections on Commuting Behaviour: Evidence from Employed and Self-Employed Workers, *Regional Science and Urban Economics* 38 (2), pp. 127–147.

VANDENBERGHE, VINCENT AND DEBANDE, OLIVIER (2008), Refinancing Europe's Higher Education through Deferred and Income-Contingent Fees: An Empirical Assessment Using Belgian, German & UK data, *European Journal of Political Economy* 24 (2), pp. 364–386.

VELLA, FRANCIS AND VERBEEK, MARNO (1999), Two-Step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias, *Journal of Econometrics* 90 (2), pp. 239–263.

WAGENHALS, GERHARD AND BUCK, JÜRGEN (2007), Möglichkeiten und Grenzen der Datenfusion für Mikrosimulationsmodelle, in: ZWICK, MARKUS AND MERZ, JOACHIM (Eds.), *MITAX – Mikroanalysen und Steuerpolitik*, volume 7 of *Statistik und Wissenschaft*, Federal Statistical Office, pp. 84–105.

WAGNER, FRANZ W. (2000), Korrektur des Einkünftedualismus durch Tarifdualismus – Zum Konstruktionsprinzip der Dual Income Taxation, *Steuer und Wirtschaft*, pp. 431–441.

WAGNER, FRANZ W. (2008), Steuerforschung: Welche Probleme finden Ökonomen interessant, und welche sind relevant?, *Steuer und Wirtschaft*, pp. 97–116.

WAGNER, GERT G., FRICK, JOACHIM R. AND SCHUPP, JÜRGEN (2007), The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements, *Journal of Applied Social Science Studies* 127 (1), pp. 129–169.

WEISS, MARTIN (2007), *Labor Taxation – A Hidden Privilege*, Shaker Verlag, Aachen.

WEISS, MARTIN (2009), Higher Tax Rates on Labor? Evidence from German Panel Data, *FinanzArchiv: Public Finance Analysis* 65 (1), pp. 73–92.

WENGER, EKKEHARD (1999), Warum die Finanzwissenschaft mit der Suche nach einem theoretischen Ideal der Einkommensteuer scheitern musste, in: SMEKAL, CHRISTIAN, SENDLHOFER, RUPERT AND WINNER, HANNES (Eds.), *Einkommen versus Konsum - Ansatzpunkte zur Steuerreformdiskussion*, Physica-Verlag, Heidelberg, pp. 37–63.

WREDE, MATTHIAS (2000), Tax Deductibility of Commuting Expenses and Leisure: On the Tax Treatment of Time-Saving Expenditure, *FinanzArchiv: Public Finance Analysis* 57 (2), pp. 216–224.

WREDE, MATTHIAS (2001), Should Commuting Expenses Be Tax Deductible? A Welfare Analysis, *Journal of Urban Economics* 49 (1), pp. 80–99.

WREDE, MATTHIAS (2009), A Distortive Wage Tax and a Countervailing Commuting Subsidy, *Journal of Public Economic Theory* 11 (2), pp. 297–310.

INDEX

Ahsan and Tsigaris (2009), 76
Almeida and Carneiro (2009), 30, 39, 52, 53, 57, 60
Amemiya and MaCurdy (1986), 19
Auerbach and Slemrod (1997), 7
Baake et al. (2004), 65
Bach et al. (2005), 88, 89
Bach et al. (2007), 88, 106
Bach et al. (2008), 7, 86–93, 96, 97, 101–119
Bach (2003), 65–67
Baldry (1998a), 72
Baldry (1998b), 72
Baldwin (1959), 90
Ballot et al. (2006), 59
Ball (1984), 112
Baltagi and Khanti-Akom (1990), 19, 20
Baltagi et al. (2003), 14
Baltagi (2008), 8, 14, 15
Becker (1964), 4
Becker (1965), 125
Blasch and Weichenrieder (2007), 7
Boockmann and Steiner (2006), 17
Booth and Coles (2010), 125
Borck and Wrede (2005), 67
Borck and Wrede (2009), 73
Boskin (1975), 125
Breusch et al. (1989), 19
Caliendo and Kopeinig (2008), 41
Cameron and Trivedi (2005), 6, 20, 42, 73, 74
Card (2001), 13
Chan et al. (2001), 32, 53, 56
Christiansen et al. (2007), 39
Cigno and Luporini (2009), 5, 126
Cogan (1981), 72
Cohen and Soto (2007), 30
Collins and Davies (2004), 20
Collins and Kemsley (2000), 112
Constantinides and Duffie (1996), 5
Conti (2005), 60
DeLong and Magin (2009), 3, 5, 30, 51
Dearden et al. (2006), 45, 59, 60
Del Rey and Racionero (2010), 126
Devereux and Griffith (1999), 89
Diller and Grottke (2010), 34
Duan (1983), 54, 60
Dustmann and Schönberg (2009), 60

Edwards et al. (2004), 7
Elschner and Schwager (2007), 5, 11, 28
Federal Ministry of Finance (2001), 118
Federal Ministry of Finance (2008), 67
Federal Statistical Office (2005), 67, 79, 81
Federal Statistical Office (2006), 7
Federal Statistical Office (2007), 70, 76
Federal Statistical Office (2008a), 30
Federal Statistical Office (2008b), 30, 64
Federal Statistical Office (2009), 1
Ferri et al. (2009), 125
Fitzenberger and Völter (2007), 3, 41
Franz (2009), 31
Freeman and Schettkat (2005), 125
Fuest and Huber (2001), 33
García-Mainar and Montuenga-Gómez (2005), 19
Genser and Reutter (2007), 5, 10
Greene (2006), 44
Greene (2008), 14, 42, 43, 45
Gruber and Poterba (1994), 69, 72, 75, 76, 79
Gutierrez-i-Puigarnau and van Ommeren (2009), 72, 78
Haisken-DeNew and Frick (2005), 15–17, 45, 78, 106, 107
Haisken-DeNew and Hahn (2006), 15, 45, 78
Hall and Oriani (2006), 53, 56, 57
Hausman and Taylor (1981), 6, 14, 19, 27
Hausman (1978), 14
Hoffmann and Lorenz (2006), 105
Hundsdoerfer et al. (2008), 1, 32, 33
Hungerbühler (2007), 32, 65
Ichino and Winter-Ebmer (1999), 13
Ismer et al. (2008), 73, 78
Jackson et al. (2009), 45
Jacob et al. (2008), 7
Jacobs and Bovenberg (2009), 33
Jacobs and van Wijnbergen (2007), 5, 65, 125
Jappelli and Pistaferri (2003), 69, 72, 76
Judd (1998), 64
Judd (2000), 3, 5
Kalwij (2000), 15, 19
Kaplow (1996), 3, 32, 34
Kim et al. (2003), 45
King and Fullerton (1984), 89
Kloas and Kuhfeld (2003), 67
Knirsch (2007), 89
Kubik (2004), 7, 69
Kuckulenz and Maier (2006), 32
Kuckulenz and Zwick (2003), 31
Kuckulenz (2007), 32, 59
Kumar (2008), 7, 69
Kyriazidou (1997), 53
Lechner and Wunsch (2009), 32
Leuven and Oosterbeek (2004), 30, 69, 74, 76
Leuven (2005), 6
Lev and Sougiannis (1996), 53
Lintner (1965), 4
Müller (2004), 30, 64, 104
Maddala (1983), 44
Merz et al. (2006), 106
Mincer (1974), 3, 13, 31
Muehler et al. (2007), 32, 45, 47, 49, 52
Ng (2008), 72

Nielsen and Sørensen (1997), 5, 10–12, 20–23, 25–28, 32–34, 65
Niemann et al. (2003), 89
Osma and Young (2009), 32
Palacios-Huerta (2001), 5
Palacios-Huerta (2003a), 4
Palacios-Huerta (2003b), 4
Panteghini (2001), 51
Pantzalis and Park (2009), 5, 32
Parry and Bento (2001), 72
Parry (2002), 72
Richter and Söhn (2008), 73
Richter (2006), 65, 73
Robinson and Sansing (2008), 32
Rosen (1982), 30, 32
Sandmo (1990), 72
Sargan (1958), 15
Schaffer and Stillman (2006), 20
Schneider (2005), 6
Selden (2009), 72
Sharpe (1964), 4
Slovin and Sushka (1993), 5
Small et al. (2005), 75
Steenken (2002), 65, 83
Su and DeSalvo (2008), 68
Sørensen (2004), 32
Tremblay (2009), 32
Trostel (1993), 57
Vandenberghe and Debande (2008), 5, 126
van Ommeren and Fosgerau (2009), 72
van Ommeren and van der Straaten (2008), 72
Vella and Verbeek (1999), 53
Wagenhals and Buck (2007), 107
Wagner et al. (2007), 78
Wagner (2000), 3
Wagner (2008), 1
Weiss (2007), 3, 5, 6, 34, 54, 55, 60
Weiss (2009), 5
Wenger (1999), 34
Wrede (2000), 72, 73
Wrede (2001), 68, 73
Wrede (2009), 73