Essays in Empirical Public Economics

Inauguraldissertation

zur

Erlangung des Doktorgrades

der

Wirtschafts- und Sozialwissenschaftlichen Fakultät

der

Universität zu Köln

2017

vorgelegt von

Dipl.-Volksw. Eric Sommer

aus

Freiburg

Referent: Prof. Dr. Felix Bierbrauer

Koreferent: Prof. Dr. Andreas Peichl

Tag der Promotion: 04.04.2017

Contents

List of Figures

List of Tables

Acknowledgements

Writing a dissertation is an endeavor that comes with numerous challenges. Unfortunately, these are not restricted to scientific problems. Moreover, it is an endeavor impossible to undertake alone. Throughout the past years, there is a number of people, not all mentioned here, that had an impact on this work.

First and foremost, I would like to express my thanks to **Felix Bierbrauer** for his willingness to supervise this thesis and for insightful discussions and advice. Special thanks goes to **Andreas Peichl**, not only for serving as co-supervisor. Without his encouragement, I most likely would not have started this thesis. I also like to thank **IZA** for funding and providing a vivid work environment in the past years.

I am deeply grateful to **Nico Pestel**, who has served as co-author of Chapters 2 and 5, mentor, great colleague, co-daddy of two boys and constant contact person to a novice in the academic game. I also benefited greatly, apart from co-authoring Chapter 4, from discussions and work on several projects with **Sebastian Siegloch**. Beyond, I like to thank my two long-lasting office mates **Max Löffler** and **Andreas Lichter** (also co-author of Chapter 5) for company on this joint journey trough extensive discussions, conference trips, countless coffees, not so countless beers, football evenings and much more. Further thanks to my co-doctoral students and temporary office mates **Amelie Schiprowski**, **Robert Mahlstedt** and **Florian Wozny**, as well the colleagues at IZA in general. I also like to thank my co-authors of Chapter 4, **Karina Doorley**, **Mathias Dolls** and **Alari Paulus**. Thanks to **Herwig Immervoll**, I had the opportunity to work four fantastic months at the OECD, although the outcome of that project did not make it into this thesis.

Lastly and most importantly, I have to thank my wife **Katja**, without whom this would have definitely not been possible. Although the scientific input of my sons **Felix** and **Max** has been limited, their impact on my mental well-being cannot be overestimated.

Introduction

1

Public Economics, or Public Finance, is one of the oldest strands of Economics (Musgrave, 2008). Its origins date back to the 17^{th} century, back then with an almost exclusive focus on the analysis of public budgets. Since then, the role of the government within the economy has steadily been redefined and extended. Nowadays, the public sector intervenes in every stage of economic activities. Modern democratic states not only provide a wealth of public goods, including social security, environmental protection, education and security. These expenditures are funded by a multitude of different taxes and contributions, levied on workers, firms, consumers, heirs or house-owners. Further market interventions appear in the form of labor and product regulations, tariffs, tax exemptions or direct subsidies. As of 2014, government expenditures account for 47% of GDP in the OECD on average. This figure alone underscores the importance of analyzing the role of the state in market economies. The present thesis is devoted to gain a deeper understanding of the consequences of governmental intervention by conducting four distinct empirical investigations. To start off, the field of Public Economics is briefly sketched. This is followed by an outline of the main strands of empirical methods typically applied in the field. Finally, each thesis chapter is summarized.

1.1 Spheres of government intervention

There have been several attempts to structure Public Economics as a field of research. The traditional normative strand of public finance theory aims at establishing conditions under which interventions by the government can be justified.[1] It departs from the hypothetical situation of a fully competitive market environment. This refers to a setting in which fully rational agents with complete information exchange commodities in a market with perfect competition absent of externalities or uncertainty. According

[1] Alternative strands of Public Economics analyze the interior organization of governments and its implications for government action. This entails the self-interested behavior by actors in the political system and the mechanisms by which citizens' preferences translate into policy outcomes. These *positive* theories of the state are not subject of this thesis and hence not described in detail. The same holds for *behavioral* public finance as a subfield of the general field of behavioral economics. These approaches try to explain economic phenomenons on the basis of psychological traits and thereby seek to overcome inconsistencies in mainstream theories.

to the First Welfare Theorem, such a situation is called Pareto-efficient. This implies that individual actions of self-oriented agents lead to an outcome in which there is no feasible alternative allocation that makes every individual better off. In order to make competitive equilibrium work, the minimal role of the state is to "set the rules of the economic game" (Atkinson and Stiglitz, 2015, p. 4). This includes the provision of laws on property rights and contracts, and the instruments to enforce these laws. The concept of full competition emphasizes the principle of individual sovereignty. Hence, any government intervention needs to be justified against the notion that humans should be allowed to pursue their self-interest, which is a central aspect of the enlightenment movement (Tresch, 2015, pp. 5f). While a Pareto-efficient setting is attractive as a theoretical point of reference, one or several of the assumptions made above will be violated in the real world. Each of those failures may give rise to intervention by the state. A traditional classification distinguishes three spheres for government intervention:

> "The responsibilities of the Fiscal Department [...] are derived from a multiplicity of objectives. [...] [T]hese are grouped under three headings: The use of fiscal instruments to (1) secure adjustments in the allocation of resources; (2) secure adjustments in the distribution of income and wealth; and (3) secure economic stabilization." (Musgrave, 1959, p. 5)

If market failures are believed to lead to inefficient outcomes, governments may try to alter the *allocation* of goods. There are numerous examples for this. Atkinson and Stiglitz (2015) distinguish four main types of market failures. First, the existence of economies of scale in the production of a certain commodity may imply *imperfect competition*. If the number of firms in a market becomes too low, collusion among these firms might cause the market outcome to differ from the competitive one. This typically leads to a situation where quantities (prices) are too low (high) compared to the efficient outcome. Governments may hence decide to enforce antitrust policies or, in the extreme case, provide these goods on their own. The second type of market failure occurs in the presence of *imperfect information* or uncertainty. As an example, individuals typically underestimate their probability of becoming unemployed or sick. Therefore, they do not sufficiently insure themselves against this risk, which justifies the introduction of compulsory unemployment and health insurance. A third reason for state interventions is the presence of *externalities*. Externalities appear in virtually all economic contexts. In the classic example, the industrial production of some commodity causes societal costs in the form of environmental pollution. These costs however do not enter the manufacturer's profit-maximization calculus. Implementing an environmental tax may lead to internalization of these external costs and thereby reduce the level of pollution caused by the firm. Another common form of externalities appears if individuals can't be excluded from the consumption of a certain commodity. In this case, private provision by one individual directly impacts the utility of others. This lowers the individual willingness to contribute to this commodity. As a consequence, governments have to step in providing these *public goods*. Examples include the provision of national security and basic infrastructure. A final category of market failures occurs if the societal valuation of certain goods differs from the private one. The consumption of these *meritocratic* goods (e. g. education) may thus be encouraged by the state. In the opposite case, the state may discourage the consumption of commodities that are considered of low societal value, such as drugs.

Turning to the second sphere according to Musgrave, *redistributional* objectives become relevant if market outcomes are considered not acceptable for society. The importance of the redistributional motive has been recognized by scholars very early on:

> "No society can surely be flourishing and happy, of which the far greater part of the members are poor and miserable" (Smith, 1776), cited in Sandmo (2014)

> "To determine the laws which regulate this distribution, is the principal problem in Political Economy" (Ricardo, 1821, p. 5)

In light of rising demand for welfare provision in modern societies, redistributional questions have become more evident than ever before. The need of granting support for the most disadvantaged part of the population is largely undisputed. Today's academic and political debates rather ask for the level and type of inequality that can (not) be tolerated by the society.

Normative statements on the optimal distribution of commodities are necessarily based on (explicit or implicit) value judgments. There are a few universally accepted objective criteria that may guide redistributional policies (Tresch, 2015, pp. 172ff). The redistributional objective can be summarized in the question on *how* to tax *whom*.[2] The primary criterion in tax system design is to tax individuals according to their ability to pay. Historically, this superseded the benefit principle, according to which the individual tax burden corresponds to the consumption of public goods. In addition, the principle of *horizontal equity* asks for equal treatment of equal economic circumstances. In the context of tax design, this calls for equal taxation of individuals with the same level of utility (Feldstein, 1976). Likewise, the principle of *vertical equity* requires differential taxation of individuals with unequal levels of utility. These guidelines, while widely accepted, are of little help on their own, as they do not answer a number of crucial questions. This includes the choice of tax base as a surrogate to utility or the definition of which circumstances should be considered equal. The standard methodology to analyze distributional questions is the social welfare function (SWF) approach. The SWF is supposed to capture the collective preferences regarding distributive justice

[2] Note that taxation includes the provision of public transfer payments in the form of negative taxation.

and aggregates individual utilities to some extent. The social welfare function is maximized subject to budget constraints on markets for products and factor inputs as well as the tools available to the government. The exact mechanism by which individual utilities are aggregated already entails value judgments. As an example, social welfare functions of the Rawlsian type put more value to poorer individuals than those of the utilitarian type. Further value judgments can be incorporated by, e. g., assigning different weights to different societal groups.

Compared to the first two government objectives, the issue of *stabilizing* the economy in times of downturn has gained relatively less attention in the Public Finance literature. Discretionary stabilization, i. e. policy actions that explicitly aim at smoothing the business cycle, finds little room in standard Public Finance textbooks. This is partly due to the failure of interventionist economic policy in the 1970s. By that time, mainstream economics viewed standard government tools as largely incapable of fighting recessions, mainly due to substantial time lags before such policies become effective (Gruber, 2011, p. 113). Beyond, the question of how short-term state action affects the economy is largely treated in the field of macroeconomics. A second notable strand is concerned with the role of automatic stabilizers, such as unemployment insurance or progressive taxation. In contrast to discretionary measures, automatic stabilizers work without further (time- and resource-intensive) action by the government. They can hence offset negative effects on private incomes in times of crisis. Finally, the increasing level of government debt also in developed economies raises the question of long-term sustainability of government budgets. In this context, stabilizing government budgets as a means of securing a functioning state for future generations has been added to the objective of stabilization.

The theoretical public finance literature has established that achieving these goals simultaneously is typically not feasible (Tresch, 2015, p. 38). Governments instead usually face a trade-off between conflicting objectives. The basic result of this literature states that redistribution without efficiency losses can only be obtained by means of non-distortionary (lump-sum) taxation. As this is usually not feasible, the different objectives cannot be treated separate from each other. Theoretical approaches therefore seek to characterize *second-best* allocations that employ distortionary tools to simultaneously address the goals of efficiency and equity. A seminal contribution in the realm of tax system design is Atkinson and Stiglitz (1976). They analyze the role of direct and indirect taxation, finding no justification to apply the latter.[3] The modern optimal tax literature yields expressions for optimal income tax rates that are based on few empirically observable parameters on behavioral reactions and the degree of inequality.[4] Another example for the trade-off between several government objectives is the provision of social assistance as a means of securing basic subsistence for everyone.

[3] Later contributions refined this result. See Chapter 2 for a detailed discussion.
[4] See Piketty and Saez (2013b) for a survey.

Higher social assistance usually comes at the cost of reducing individual incentives to participate in the labor market. Virtually any state intervention can be expected to affect several government objectives simultaneously. In democratic societies, this may cause political opposition from potential (or alleged) losers of a reform. It is this area of tension where empirical investigations step in. They can contribute to quantify the magnitude of these trade-offs by measuring the effect of state action. Examples include the extent to which the behavior of economic agents is altered as a consequence of certain government interventions. While a specific causal channel might be conceivable in theory, it might be inexistent or irrelevant in reality. Another example is the evaluation of redistributional instruments regarding their effectiveness (do they achieve the redistributional objective?) and their efficiency (what side-effects can be expected?). The output of such research can then be used to inform the academic and political debate and may finally contribute to better design of policies.

1.2 Empirical Methods in Public Economics

Empirical Economics is interested in measuring the causal effect of a phenomenon X on an economic outcome Y. Researchers are concerned with both the direction and the magnitude of this effect. In recent decades, a growing availability of micro data, i. e. data on individual, household or firm level as the basic economic units, boosted a new wave of empirical work (Atkinson and Stiglitz, 2015). This literature has made substantial progress on establishing causal evidence on the magnitude of economic effects, providing more credible identification than the bulk of traditional cross-sectional or time-series econometric methods (Angrist and Pischke, 2010).[5] These approaches have also found fruitful use in the field of Public Finance. Public Economics is typically interested in questions on the national (macro) scale, such as the employment effect of a tax reform or the price impact of a production subsidy. Nevertheless, there is a lot to be inferred from micro-based analyses.

If one is interested in the causal effect of X on Y, a regression of Y on X, possibly along with some control variables, does not necessarily reveal an estimate of the causal impact. As an example, the causation might actually go from Y to X. Alternatively, there might be a non-observed variable Z driving X and Y simultaneously. New research designs that rely on plausibly exogenous variation in X try to overcome these problems. These research designs aim at mimicking the experimental ideal from natural sciences, where variation in X is generated by the researcher. This is done by assigning the treatment to individuals by a random procedure. Comparing the outcomes between groups with and without treatment (or, varying assigned values of X) then delivers the causal effect. Parallel to the tremendous growth of experimental studies

[5] This section builds on Gruber (2011, chap. 3).

in economics as a whole, these designs found increasing use in public finance. As an example, Hotz et al. (2002) study the labor supply effects of a welfare program to families in California, where a share of randomly chosen recipients was assigned a lower benefit amount. Another well-known real-world experiment has been carried out by Chetty et al. (2009). They demonstrate differential consumption behavior for various degrees of tax salience, captured by whether the sales tax amount is indicated on the price tag. Experimental studies are however subject to criticism, mostly regarding their lack of external validity (Levitt and List, 2007). Generalizations beyond the sample of participants in a given experiment are hard to defend. Moreover, large-scale experiments often come with enormous costs or are constrained with ethical concerns. As a consequence, researchers increasingly turned to seek for quasi-random assignment in observational data. If a change in the economic environment creates a treatment and a control group that are identical with the exception of the treatment, the experimental ideal is fulfilled (Gruber, 2011, p. 80). The key to proper identification is that the source of variation is exogenous to the treated individuals. Opportunities for evasion, avoidance or anticipation harms identification as the membership to the treatment or the control group is affected by (non-observed) factors. Classic sources of variations can be time (unanticipated tax-benefit reforms) or geography (policies introduced only at one side of a border). In a prominent application, Saez et al. (2012) exploit a payroll tax reform that affected employees based on their first year of employment. By comparing employees just above and below the threshold, the extent to which the changed payroll tax rates transform into wage changes sheds light on the degree of payroll tax incidence. A special case in the context of taxation is to measure bunching at discontinuities in tax rates as a proxy for behavioral costs of taxation (Kleven, 2016).

Experimental and quasi-experimental studies have delivered convincing causal evidence for many economic parameters of interest. Strictly speaking however, their findings apply only to a specific parameter in a given country-context when a particular policy is applied at a specific point in time. Extrapolating these findings to other contexts is generally difficult. Beyond, quasi-experimental approaches rarely provide insights on the underlying mechanisms for a specific finding. To overcome these limitations, the broad field of structural estimations has emerged as a complementary set of approaches. Structural models are used in contexts where quasi-experiments are not feasible or where one is interested in the effect on the total population rather than a specific subgroup. Beyond, structural models allow for an ex ante analysis of counterfactual situations. After estimating the parameters of a structural model, one can investigate individual responses to changes in the economic environment, e. g. a tax-benefit reform. The advantage of structural models lies in their theoretical foundation and in their flexibility regarding the number of counterfactual scenarios. Structural models are applied in all fields where individual behavior plays a role. Examples include labor supply, transport or consumption behavior. Depending on the degree of detail in the underlying data, a wealth of economic outcomes can be investigated. It is also possible to cover heterogeneous responses across individuals. The downsides of structural approaches is their tendency to rely on strong theoretical assumptions. Examples are the specification of a utility function or the number and selection of alternatives (working hours, means of transport) an individual is assumed to choose from. Beyond, the mechanisms of structural models are harder to communicate. Validation of structural models, i. e. comparing model predictions with actual outcomes, is also difficult to achieve.

Summing up, the choice of research design depends, among other things, on the specific research question and on the availability of appropriate data. There are many economic phenomena for which natural experiments simply do not exist or where no data are available. Consequences of demographic change, as an example, take decades to materialize and are hence not feasible for a quasi-experimental setting. In such contexts, structural models may provide interesting insights despite their disadvantages discussed above. It needs to be stressed that the methodologies outlined above can also be combined. This represents a middle-ground between the two polar cases of either providing credible identification or delivering a general picture. As a notable example, Blundell et al. (1998) estimate labor supply elasticities of married women by exploiting several tax-benefit reforms over a longer period of time. Under the assumption that certain population subgroups exhibit similar behavior, differential evolution of net wages across subgroups identifies the elasticity of labor supply. A combination of ex ante and ex post methods has been carried out, among others, by Imbens et al. (2001). They estimate a structural labor supply model exploiting random variation in unearned incomes from lottery wins. In a similar vein, Bargain and Doorley (2016) use an age discontinuity in welfare eligibility to inform a structural model of labor supply. This way, the external validity of the discontinuity approach can be tested and the clearly identified parameter can be applied beyond the scope of the quasi-experiment.

1.3 Chapter summaries

In order to provide a concise overview of the thesis, Table 1.1 assigns each thesis chapter to one or several of Musgrave's spheres of government action.[6] Chapter 2 has a focus on redistributional effects of taxation, while Chapter 4 is dealing with the stability of government budgets in light of demographic change. Allocational issues play a role in every chapter. They are of particular importance in the form of welfare costs of taxation (Chapter 3) and air pollution (Chapter 5). In addition, each chapter is classified according to the applied research design — quasi-experimental or structural. Chapter 3 exploits discrete jumps in the German bequest tax schedule to obtain causal estimates

[6] This presentation is inspired by Siegloch (2013).

Table 1.1: Thesis overview

	Chap. 2	Chap. 3	Chap. 4	Chap. 5
Government Objective				
Allocation	×	×	×	×
Stabilization			×	
Distribution	×			
Type of research design				
Quasi-experimental		×		×
Structural	×		×	

of the elasticity of taxable bequests. Chapter 5 also relies on a quasi-experimental setting. Here, the causal effect of air pollution on the performance of professional football players is identified from a match schedule that is exogenous to the individual player or team. The remaining chapters feature structural models. Chapter 2 carries out estimations on labor supply and household consumption and investigates the distributional effects of a fundamental tax reform. Chapter 4, in contrast, builds on external structural estimates on the own-wage elasticities of labor supply and demand.

In the following, each chapter's key questions are briefly outlined along with a description of the methodology and the main results.

1.3.1 Chapter 2: Can a shift towards consumption taxation be justified on equity grounds?

This chapter takes up on a classic question in public finance, that is the right mix between direct and indirect taxes. Tax theory provides a number of reasons why taxing consumption might be associated with lower efficiency losses. Broadening the tax base by shifting a part of the tax burden from labor to consumption may hence raise employment while keeping government budgets constant. On the other hand, consumption taxes are often seen as falling disproportionately on poorer households. This chapter makes the point that the unfavorable distributional consequences of such a reform are substantially dampened once the positive labor supply effects are accounted for. The tax burden on labor income in Germany is particularly high, manifesting in large marginal tax rates for many employees. Lowering social security contributions hence seem to be a promising measure to increase work incentives. The chapter applies a microsimulation of a step-wise increase in the standard VAT rate, accompanied by a revenue-neutral reduction in personal income taxes or social security contributions. The microsimulation model is able to capture the full heterogeneity of the population and incorporates a structural model of household labor supply. Not surprisingly, reducing the personal income tax has a strongly regressive distributional impact. The same holds for lower social security contributions, albeit to a lower extent. This is because social security contributions are a regressive tax themselves. Accounting for increasing work incentives, some households are able to compensate their losses through higher labor earnings. As a result, the impact on aggregate inequality becomes close to zero. Pensioners remain as the main losers from the reform and would need to be compensated otherwise.

Result: If higher labor supply can be realized, the overall distributional impact of a tax shift becomes close to zero, despite the existence of some reform losers. This suggests that a heavier reliance on consumption taxation can be done in a fashion that is not only neutral to the government budget, but also to income inequality.

1.3.2 Chapter 3: How responsive are wealth transfers to taxation?

Academic and political interest in bequest taxation is likely to rise in the future as intergenerational transfers are becoming a more important source of income in developed countries. Moreover, higher bequest taxes are regularly discussed as a means of cushioning high wealth inequality. As a consequence, understanding the welfare consequences of bequest taxation will become a more pressing issue. This chapter assesses responsiveness of taxable bequests in Germany by exploiting discontinuities in marginal tax rates. This creates an exogenous variation in tax rates and allows the identification of behavioral parameters, upon which the associated welfare costs of taxation can be gauged. The institutional setting in which intergenerational wealth transfers take place allows for different tax minimization channels by the donors and recipients. The quantification of taxpayers' bunching at kink points reveals differential tax planning behavior depending on the type of wealth transfer. While there is bunching for inter-vivo gifts, there is no evidence for inheritance tax bunching. This finding is in line with the intentional character of inter-vivo gifts and ex ante tax planning by the donor. Inheritances, in contrast, occur largely accidentally. The donor has thus less control over the amount of assets each heir receives. The scope for tax planning ex ante by the recipients does not seem to be of actual relevance. Beyond, tax planning is found to be more relevant for transfers within the close family and for more valuable assets. A side results is the presence of focal points, i. e. a tendency to transfer assets amounting to round numbers.

Result: There is evidence for tax planning for wealth transfers, which is mainly driven by inter-vivo gifts. Quantitatively, the responsiveness of inherited assets is rather low, which suggests low short-run distortionary effects of bequest taxation. In the discussion of extending the taxation of bequests as a means of cushioning wealth inequality, distortionary effects of bequests taxation should hence not be a major concern.

1.3.3 Chapter 4: How will ageing societies affect public budgets in the EU?

This chapter tackles the stability of future government finances in the EU-27. In particular, it asks whether European welfare states can maintain their level of public spending in light of the ongoing demographic transition. Although demographic change has various facets, old-age dependency ratios are on the rise throughout the EU, often accompanied with a shrinking labor force. This is expected to exert substantial pressure on fiscal budgets due to higher expenditures for public health and old-age pensions. While this notion is well-established, it takes a one-sided view on the government expenditures, ignoring potential revenue channels. The inter-linkages between demographic transitions and labor market outcomes deserve special attention in this context. If, for example, a shrinking labor force is becoming better educated at the same time (as can be expected), average wages will increase. Additionally, if there is a scarcity of labor, neoclassical economic theory predicts that wages should increase in order to stimulate labor supply. Future tax revenues may therefore increase despite population shrinkage. Hence, it is crucial to account for reactions on both sides of the labor market when assessing the effects of demographic changes on future fiscal balances.

This chapter employs detailed population projections for the EU-27. An ageing and better educated labor force will lead to changing relative scarcities of the factor labor in different labor markets. Based on this, we obtain the wage effect based on external estimates on the elasticities of labor supply and demand. With an European tax-benefit calculator, the impact on household incomes and, subsequently, fiscal budgets are calculated. On top, we simulate the labor market and fiscal effects of a common policy response to demographic change by raising the statutory retirement age.

It turns out that the majority of countries is likely to face a negative outlook in fiscal terms due to demographic change. This is despite revenues from personal taxes and social security contributions are projected to rise. Accounting for labor market effects, labor scarcity leads to a strong wage growth and, in turn, small employment increases. This improves fiscal balances in most countries. Raising the statutory retirement age on top suggests to be sufficient to avoid worsening fiscal balances in nearly all countries. Although this approach is far from capturing all conceivable implications of demographic change, the results overall paint a less worrying outlook for government budgets in the future.

Result: Ageing societies are going to worsen the fiscal outlook in the EU-27, especially in those countries that are going to face major decreases in the labor force. Our analysis however suggests that labor market effects and a higher statutory retirement age might be sufficient in most countries to cover the expected expenditure increases.

1.3.4 Chapter 5: What is the short-term impact of air pollution on the performance of professional athletes?

This chapter assesses the impact of ambient air pollution on the performance of professional football players. While this is most closely related to health or labor issues, it bears interesting implications from a Public Economics perspective. Environmental protection is nowadays acknowledged as a fundamental government objective. Protection measures come in the form of pigouvian taxation or environmental regulations. These market interventions are justified by the enormous societal costs associated with environmental damage. At the same time, environmental policies usually have a redistributive component, which impedes implementation of these policies. A deeper understanding of the extent of environmental costs thus allows a better assessment of the benefits of environmental protection.

For the case of air pollution, strong negative long-term effects have been well documented in the literature. Previous empirical studies have found detrimental long-term effects of air pollution on morbidity, infants' health, human capital formation and labor supply. In contrast, evidence on the short-term impact of air pollution is relatively scarce. The existence of adverse effects on productivity of workers could substantially increase existent estimates of the societal costs of air pollution. Measuring the productivity impact of environmental factors is however connected to severe empirical challenges. Individual exposure to air pollution can typically not be considered random due to e. g. residential sorting or avoidance behavior. In this chapter, we exploit the fact that time and location of football matches in the German *Bundesliga* is exogenous to the participants. Using long-term variation within players, this quasi-experimental setting allows the causal identification the effect of pollution with particulate matter on the performance of professional football players. As measure of performance, we employ the number of passes played as the basic element of the game. Beyond, the number of passes is strongly related to physical performance and highly relevant for a team's success.

Result: We document significant negative short-term effects of air pollution on the performance of professional football players. The effects are strongly increasing with the level of pollution. As we cover a selected group of employees, generalizing these findings to a wider labor force is not straightforward. As our findings are obtained in a country with relatively low levels of air pollution, it is fair to conclude that short-term productivity effects of air pollution are substantial when evaluated on a global scale.

2 Shifting Taxes from Labor to Consumption: Distributional and Labor Supply Impact*

2.1 Introduction

The appropriate choice between direct and indirect tax instruments has been subject to an extensive debate on their respective merits and disadvantages. Although the question of the optimal mix is still open, there are reasons for a coexistence of both forms of taxation, as they address the economic policy objectives of efficiency and redistribution in different ways. Moreover, in the context of the need for fiscal consolidation, consumption constitutes an attractive and reliable source for government revenues as a stable tax base. In addition, shifting the tax burden from labor to consumption, referred to as *fiscal devaluation*, is currently considered as an alternative to nominal devaluation in order to restore competitiveness in some Euro area countries (de Mooij and Keen, 2012; Koske, 2013).

The debate on possible consequences of a tax shift from income towards consumption centers around two issues. First, according to standard economic theory, such a tax shift might be favorable with respect to employment as a consequence of lower marginal tax rates on labor income, implying higher incentives to take up work. Second, higher consumption taxes are often associated with lower tax progressivity and higher levels of inequality. However, employment increases from a tax shift may outweigh adverse distributional impacts. The degree to which there exists a trade-off between higher inequality and more employment in this context is an empirical question. We provide an analysis for Germany to gauge the extent of this trade-off and investigate whether a shift from income to consumption taxation can be justified in light of positive labor supply effects. Germany represents a particular interesting case as the tax wedge on labor income is among the highest in industrialized countries (OECD, 2014).

Despite the theoretical virtues of indirect taxes, the direct to indirect tax ratio has been on the rise over the last decades, mostly due to increasing social security contributions (Martinez-Vazquez et al., 2010). Consequently, recent years have witnessed a

* This chapter, co-authored with Nico Pestel, originated from my Diploma Thesis and is published as: N. Pestel and E. Sommer (2017). Shifting Taxes from Labor to Consumption: More Employment and more Inequality? *forthcoming in: Review of Income and Wealth*.

growing discussion on a heavier reliance on consumption taxes, such as sales taxes and the Value Added Tax (VAT) (OECD, 2007, 2010). A concrete policy implementation of such a *tax cut cum base broadening* was the 2007 VAT increase in Germany, which was compensated by simultaneously cutting unemployment insurance contributions.[1] This policy was explicitly motivated by increasing work incentives and generating revenues at the same time. In the same spirit, in 2009, Hungary financed a five percentage point reduction in the employer Social Security Contributions rate through a higher VAT. These policies followed the argumentation that the tax burden on labor in most OECD countries is too high and implies disincentives for labor market participation. Moreover, payroll taxes constitute a significant share of labor costs for employers (OECD, 2014). A shift away from income and payroll taxes towards consumption taxes could therefore release unused productive capacities by increasing labor supply *and* demand. Moreover, labor constitutes the major tax base for generating revenues in most countries, which might be questioned in light of a proper application of the Ability to Pay Principle. Broadening the tax base addresses this issue by treating all sources of income equally. The distributional consequences of a tax shift are however unclear.

In this chapter, we carry out microsimulations of several revenue-neutral policy scenarios. We simulate a step-wise increase of the standard VAT rate of currently 19% in Germany, accompanied by a reduction in personal income taxes (PIT) or social security contributions (SSC).[2] We add to the existing literature by simulating a range of revenue-neutral reforms on both PIT and SSC, accounting for labor supply responses at the same time. As the distributional analysis is differentiated along several socio-demographic dimensions, the results can help to design specifically targeted policies to compensate the potential losers from an increase in VAT rates. For example, if pensioners are found to be worse off, it might be worth considering to split the additional revenue from the higher VAT on lowering payroll taxes *and* raising old-age pensions. The analysis is carried out with the behavioral microsimulation model IZAΨMOD (Löffler et al., 2014a). Based on a representative sample of the German population from the Socio-Economic Panel Study (SOEP) and a detailed model of the German tax and transfer system, we are able to simulate changes in household budgets as well as adjustments in labor supply behavior. As the information on household consumption in SOEP is insufficient, we impute expenditures based on estimates from the German Sample Survey of Income and Expenditures (EVS). Our empirical approach is related to the studies of Decoster et al. (2009) and Bach et al. (2006), but differs in several aspects. While the former study

[1] The standard VAT rate was raised from 16% to 19%, while the total rate of unemployment insurance contributions was lowered from 6.5% to 4.2%. This specific reform has been ex ante investigated by Bach et al. (2006).

[2] There are two VAT rates in Germany. Apart from the standard rate of 19%, there is a reduced rate of 7% applied on most food commodities, public transport, books, newspapers, journals, entrance to cultural facilities and works of art. Moreover, medical, educational and financial services as well as rents are fully exempted from the VAT.

depicts only the static changes in household budgets ignoring behavioral responses, the latter does not consider a revenue-neutral reform.

We find that both scenarios of reducing the direct tax burden, either lowering PIT or SSC, imply distinct distributional impacts. Due to its strongly progressive design, a compensated reduction of personal income taxes leads to a higher level of inequality. Low-income earners, pensioners and unemployed are found to be the main losers from the policy. For payroll tax reductions, the adverse distributional effects are significantly less severe, because payroll taxes constitute a regressive tax themselves. Taking into account behavioral adjustments, we find that the distributional impacts of the tax shift are weakened. For lowering the PIT level however, a strongly regressive impact persists. Reducing payroll taxes seems particularly promising, given their potential to raise work incentives. In these scenarios, some households are able to compensate their losses through higher labor earnings. Beyond, our results suggest no systematic difference between augmenting both VAT rates or only the standard rate, which underlines the limited redistributional power that is often attributed to a differentiation of VAT rates.

The chapter is structured as follows: Section 2.2 reviews the theory on labor versus consumption taxation and the empirical evidence on tax shifts. Then, we present related empirical findings on the macro and micro level. In Section 2.3, our microsimulation approach and the underlying data base is presented. Furthermore, our method to impute expenditures in an income data set is described in detail. In the results section 2.4, the simulated labor supply reactions are presented first. Second, a detailed distributional analysis identifies winners and losers from the reform. A comparison of several aggregated measures of inequality and progressivity completes the analysis. Section 2.5 concludes.

2.2 Background and Literature

2.2.1 Theory

Taxation affects economic incentives and may therefore induce behavioral adjustments for individuals, causing efficiency costs compared to a hypothetical situation without taxes. As any feasible tax causes distortions, the theoretical question is how to characterize the second-best setting that implies minimum efficiency losses, given a fixed government revenue. Economic theory provides intuition for why a shift from income to consumption taxation might be favorable in efficiency terms, i.e., promoting growth and employment. Within a static standard utility-maximization framework, it can be shown that both taxes distort the individual decision between consumption and leisure equivalently. An income tax reduces the net wage, while a consumption tax reduces the real value of net earnings. Under non-negative wage and income elasticities of labor supply, both forms of taxation reduce work incentives (Bargain et al., 2014). While only a fraction of the population is subject to income taxation, virtually everyone pays consumption taxes. The consumption tax base is hence broader, as it includes expenditures of pensioners, benefit-recipients and capital-income earners. Hence, consumption taxes allow for obtaining the same revenue with a lower rate. If one recalls the classic insight that the excess burden of a tax rises approximately with the square of the tax rate (Auerbach, 1985), a shift towards a consumption tax induces lower aggregate efficiency costs. The intuition is that the positive effect on labor supply from the higher net wage exceeds the negative effect from a lower real income, resulting in higher aggregate labor supply.

A theoretical counter-argument is that throughout the life-cycle, income necessarily equals consumption and therefore implies an equal burden of both taxes (Caspersen and Metcalf, 1993; Metcalf, 1994). However, this argument only holds if both tax schedules are constant in the long run and if bequests are not considered. Although the only difference between (labor) income and consumption arises from consumption smoothing, this intuition is hardly relevant in the policy debate on what is understood as a regressive tax.[3] Another argument refers to the treatment of capital income. A tax levied on capital income distorts an individual's saving decision, as it implicitly taxes future consumption. If this is a normal good, an income tax discourages savings. In contrast, the savings decision is neutral to the level of consumption taxation, as the consumption tax does not alter the returns to savings. Reducing the capital income tax in favor of the consumption tax is therefore expected to increase savings and hence economic growth (Auerbach and Hines, 2002; Feldstein, 1978).

The interdependencies between both forms of taxation have regularly been addressed by the optimal taxation literature. Atkinson and Stiglitz (1976) were the first to capture the equity-efficiency trade-off of both taxes within a formal framework. Under the assumption of separable preferences and individuals that are inequality-neutral, they neglect any role for indirect taxation. Since all commodities are equally substitutable for leisure, any attempt to offset the distortion between labor and leisure is bound to cause efficiency losses.[4] Later contributions refined this argument by imposing more realistic assumptions and found commodity taxation to be a necessary component of any optimal tax structure. Among these assumptions are uncertainty about individual wages (Cremer and Gahvari, 1995), heterogeneity among agents not only in ability (Cremer et al., 2001; Saez, 2002), different underlying production technologies (Naito, 2007) or different evasion characteristics of both taxes (Boadway et al., 1994; Richter and Boadway, 2005). According to Mankiw et al. (2009), the advance of indirect taxes and VAT in particular can be attributed to findings of optimal taxation theory. Despite Atkinson and Stiglitz' wide-known result not to levy any indirect taxes, it seems worth

[3] For a treatment of lifetime inequality in a simulation context, see Creedy (1997). A recent empirical analysis of lifetime inequality among German employees can be found in Bönke et al. (2015).

[4] See also the argumentation by Sørensen (2007).

to consider whether a shift to consumption taxation might adjust the direct to indirect tax mix towards the optimum (European Commission, 2008).

A proper application of the Ability to Pay Principle might provide further justification for a heavier reliance on consumption taxation. Such arguments favor consumption (the use of income) to income (the contribution to national production) as the better measure for individual ability (Gruber, 2011, chap. 25).[5]

2.2.2 Empirical Evidence

The efficiency impact of a shift from income to consumption taxation has been investigated by a number of empirical studies, most of them based on a macrosimulation framework. They largely reveal positive, but moderate effects from a compensated SSC reduction on GDP growth rate and employment for the German case.[6] All studies suggest positive, but moderate employment effects not higher than 1% of total employment. Similar results are obtained for other countries.[7] Unions' behavior in the aftermath of the reform is found crucial for the long-run effects of the tax shift. Studies that explicitly incorporate the mode of wage bargaining draw rather pessimistic conclusions. If unions' bargaining power is assumed to be sufficiently high, wage increases as a consequence of increased living costs become likely in the medium term. Another channel that might work against the effectiveness, though not captured in these studies, are announcement effects of VAT increases that cause domestic demand to boost before and to decline in the aftermath of the policy change.

Macro approaches exhibit drawbacks when it comes to distributional questions. Any conclusions derived from macro simulations do not account for heterogeneity among individuals. As a consequence, these kinds of questions have been addressed by a number of microsimulation studies which all focus on SSC reductions. Decoster et al. (2009) provide a comprehensive study incorporating four European countries. They simulate a 25% reduction in social security contributions, compensated by a VAT increase. Their results indicate negative welfare effects for households in low income deciles, as well as for households with low-educated and unemployed heads. This is in line with O'Donoghue et al. (2004), who find a general regressive impact in twelve OECD countries, Portugal being most regressive and Belgium being nearly proportional. Similar results are obtained by Bach et al. (2006), who simulate the effect of the three percentage points VAT increase implemented by the German government in 2007. This was complemented by a cut in unemployment insurance contributions by two per-

[5] This idea dates as far back as to Thomas Hobbes: "It is fairer to tax people on what they extract from the economy, as roughly measured by their consumption, than to tax them on what they produce for the economy, as roughly measured by their income."(Gruber, 2011, p. 754)

[6] See Buscher et al. (2001), Steiner (1996), Meinhardt and Zwiener (2005), Feil and Zika (2005), Feil et al. (2006), Böhringer et al. (2005).

[7] See European Commission (2006, 2008) for a cut in income taxes in the EU as a whole, Altig et al. (2001) for a shift of the US federal income tax, and Dahlby (2003) for income tax shifting in Canada.

centage points. It should however be noted that this reform was not revenue-neutral. Thomas and Picos-Sanchez (2012) simulate a revenue-neutral shift of 5% of the SSC burden to VAT and find increasing work incentives particularly for low-income earners across several European countries. Applying the same reform, Picos-Sanchez and Thomas (2015) identify employees as particular beneficiaries. Meinhardt and Zwiener (2005) simulate a cut in SSC by two percentage points, combined with an increase in VAT by the same amount. Although the authors do not report fiscal effects, this reform is presumably not revenue-neutral as well. They identify civil servants, self-employed and unemployed as the main losers from the reform, while gains for employed persons are rather moderate. A related study is provided by Moscarola et al. (2015), who consider a shift of the tax base from labor to property, while accounting for labor market reactions.

The empirical results partly strengthen the cause for a tax shift for efficiency reasons, though the positive impact on employment and growth seems to be rather moderate. As the results for Germany indicate, the magnitude crucially depends on the institutional setting of the economy. The microsimulation studies presented here confirm a regressive impact. Low-income groups are typically worse off from a tax shift as well as unemployed and pensioners. This result is not surprising, as these groups typically face a low burden of income taxes and social security contributions.

2.3 Empirical Approach

Microsimulation models have become a standard tool in the ex-ante assessment of reforms of the tax-benefit system and therefore allow to trace changes in highly complex tax regulations. In particular, the specific institutional setting and the socio-economic structure in a given country need to be taken into account, which can hardly be accomplished by an analysis on an aggregate level.

The basic idea of microsimulation in the context of labor supply is to model the individual (or household) decision between leisure and consumption. Based on observed behavior of a representative population sample in a given institutional setting, preference parameters can be estimated. If net income (and thus consumption possibilities) changes as a consequence of a tax-benefit reform, these estimates are used to predict individual labor supply *after* the reform. The reform effect is then defined as the difference in aggregate behavior between the two institutional regimes. For this, a detailed representation of the tax-benefit system is necessary. We use the IZA Policy SImulation MODel (IZAΨMOD) of the Institute for the Study of Labor (Löffler et al., 2014a). Apart from replicating the German tax and transfer system, it comprises an econometrically estimated model of labor supply behavior. It assumes a discrete choice set of working hours, which facilitates the treatment of family labor supply. As our main database does not capture consumption expenditures, we have to extend our database. This is done by an Engel curve procedure, adopting the approach of Decoster et al. (2013).[8]

Reform Scenarios. We carry out simulations of two benchmark scenarios, in which the standard VAT rate of 19% is increased in steps of one percentage point each. For a given increase in the standard VAT rate $d\tau > 0$, we obtain the resulting additional VAT revenue from total simulated revenues. We rely on simulated, not official revenues for this, as our micro-data only capture consumption from private households living in Germany and therefore cannot depict VAT payments from public consumption, enterprises and foreigners.[9] On the basis of revenue statistics, we obtain the necessary proportional reduction on income-related taxes and apply this factor to the simulated tax liabilities.[10] This is done for personal income taxes and social security contributions separately. This procedure is repeated eleven times until in the last step, an increase in the standard VAT rate from 19% to 30% is combined with a corresponding reduction on labor-related taxes. At the same time, we provide more detailed results for a reference scenario with a standard rate of 25%. While this implies substantial tax shifts, secondary effects, such as demand for compensation by unions, are less likely to play a role than for even stronger tax shifts.

Although SSC and income tax payments flow into separate budgets, their impacts on the overall budget are highly interlinked. For many years, the German statutory pension system has been partly financed through the tax budget, since SSC revenues are not sufficient to cover public pension payments. In fact, these payments have become the largest share of federal expenses.[11] For this reason, reforms on either income taxes or SSC imply equivalent effects for the public budget as a whole. A VAT increase by six percentage points would result in additional VAT revenue of €29 bn, the corresponding relief amounting to 16.9% for the personal income tax (total status quo revenues: €174.6 bn) and 15.5% for social security contributions (total status quo revenues from employees: €190.5 bn). As a consequence, pensions and public health care would need to receive more tax funding, thus dispersing the welfare state financing away from employees.

[8] Details on the data base, the imputation procedure and the underlying labor supply model are provided in the Appendix.

[9] We however correct the simulated revenue by the under-coverage of total private consumption compared to national accounts, which amounts to 81% for the 2008 EVS.

[10] There are numerous ways for governments to reduce the burden of income-related taxes. Here, we refrain from discussing the various interdependent impacts of instruments, such as reducing marginal tax rates or raising the exemption level. Instead of providing a blueprint for a tax reform, we rather aim at gaining a rough insight on the interaction between both forms of taxation with respect to distributional questions. Therefore, we opt for the most straightforward way to reduce taxes, namely by proportional reduction. This is the standard approach in the literature.

[11] In 2009, €102 bn of tax revenues (roughly one third of total revenues) were spent on financing social security.

Income concept. For each reform step, the combined tax change alters household budget constraints which, in turn, induces adjustments in household labor supply if the expected utility of an alternative choice category is higher than the status quo. In order to account for the budget effect of an increased consumption tax, the commonly used concept of disposable income is not sufficient here, as it ignores consumption taxes. For the subsequent analysis, the quantity of interest will be *Post-VAT Income (PVI)*, which is defined as disposable income minus VAT expenses. PVI can be understood as the amount of money that *would* be left for consumption after paying the Value Added Tax. This income is of course virtual, as it is not disposable for consumption after VAT has been paid. PVI is not only the basis for the distributional analysis, but also enters the utility function and hence determines the labor supply decision. We thereby implicitly assume that households have an identical perception of their burden of direct and indirect taxes. This may be questioned in light of the experimental studies by Sausgruber and Tyran (2005) and Blumkin et al. (2012), both pointing to a lower perception of consumption taxes. If this is true, households would ignore the VAT increase to some extent, implying a higher reaction from a reduced direct tax. With positive elasticities of labor supply, our estimated labor supply reaction should hence be understood as a lower bound.

Subtracting VAT expenses from disposable income is equivalent to full and instantaneous VAT shifting from firms to consumers.[12] We therefore abstract from the fact that it may take time until firms shift the higher VAT to consumers, which is in line with the logic of static models. Our expenditure imputation is also able to depict the effect on commodity demand through income and price changes.[13] This affects savings behavior as well as adjustments in the expenditure structure across commodity groups. As the level of basic social assistance in Germany is linked to inflation rates, we address the importance of this particular channel on our results.[14]

Incidence and VAT differentiation. Subtracting the revenue-neutral deduction from household income implicitly assumes that workers bear the full burden of income taxes and social security contributions. Doubts are however justified, particularly for the case of payroll taxes, as their payment is split between employers and employees.[15] We address this issue by assuming alternative divisions of the tax incidence in a robustness

[12] Full incidence of the German VAT in the medium run has been found by the Bundesbank (2008).

[13] See Appendix for details.

[14] In practice, the level of the means-tested unemployment benefit (*Arbeitslosengeld II*) is annually adjusted by the change of an index consisting of the price change in basic goods and services (70%) and the average change in employees' net wages (30%). As 55% of all expenditures are subject to the standard VAT rate (see Table 2.4), each percentage point of higher standard VAT rate mechanically raises the price level and hence the unemployment benefit by 0.46 percentage points.

[15] The findings of Saez et al. (2012), exploiting a natural experiment in Greece, suggest that for a payroll tax increase, the long-term burden of workers is limited to the *employee* share. It is however unclear whether their findings are applicable for a different institutional setting and a payroll tax *reduction*.

check. If the incidence is low, employees benefit less from a tax reduction. We evaluate the extent to which this influences the overall distributional impact of the reform.

In a further robustness check, we alter the benchmark scenarios by increasing both VAT rates simultaneously, thereby addressing the issue of VAT rate differentiation. As in most OECD countries, expenditures for necessities are taxed with a reduced rate in Germany. The common justification for this policy are equity concerns. If the reduced rate is fulfilling its redistributional objective, a simultaneous increase of both VAT rates should imply more regressive effects than the benchmark scenarios.

2.4 Results

2.4.1 Labor Supply Effects

Our microsimulation approach sheds light on whether the expectations of positive effects on work incentives can be confirmed. The labor supply effects simulated here have to be interpreted as medium-term outcomes, i. e., after households have adjusted their labor supply behavior to the new institutional environment. If one assumes a negative wage elasticity of labor demand, firms will react to higher labor supply by lowering offered wages, leading to an equilibrium outcome below the initial labor supply shift (Peichl and Siegloch, 2012).

Table 2.1: Labor Supply Effects (Standard VAT rate of 25%)

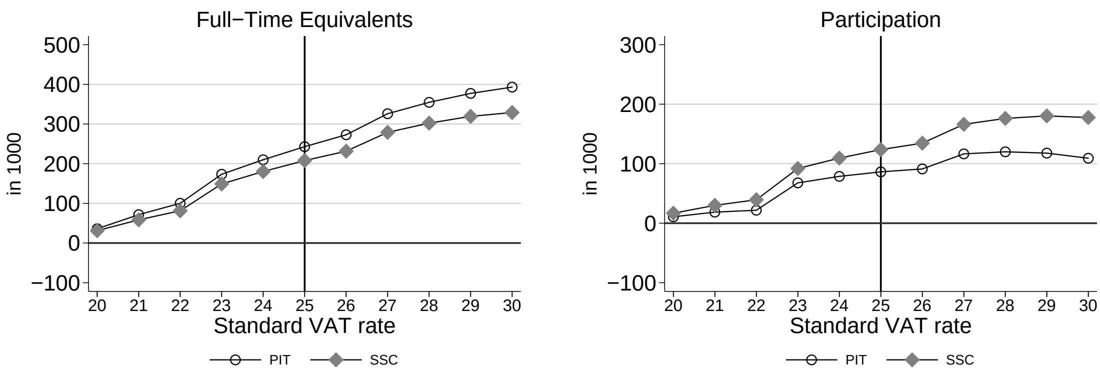
Reform Scenario	**Base**	**PIT Reduction**		**SSC Reduction**	
		with UB indexation	no UB indexation	with UB indexation	no UB indexation
		changes to Baseline			
FT Equivalents	38,039	+242.9	+286.3	+207.9	+249.9
Participation	40,344	+86.3	+125.3	+123.7	+161.6

Own calculations with IZAΨMOD v.3.0.4. Full-Time Equivalent = 40 hours per week. All figures in thousands.

The simulated labor supply responses, for an increase of the standard VAT rate from 19% to 25%, are displayed in Table 2.1. It shows the aggregate change in hours worked, measured in full-time equivalents (FTE) of 40 hours per week. The total effect is found to be positive in the order of 200,000 to 250,000 FTE for both the PIT and the SSC reduction. This corresponds to an increase in labor supply by around 0.5% of total employment. This is well in line with results obtained from CGE studies (Buscher et al., 2001, p. 466; Böhringer et al., 2005, pp. 95ff). Looking at the extensive margin of labor supply, i.e., the number of individuals entering the labor market from inactivity, we simulate an increase by 86,000 (PIT reduction) and 124,000 (SSC reduction) workers respectively, indicating substantially higher activating potential of lower social security contributions compared to lower PIT. This is not surprising, as many workers with comparably low earnings are subject to these contributions, while still exempted from the income tax. If the increase in labor supply can be mostly realized, i.e., facing limited constraints on the demand side, our simulation results confirm the theoretical expectations concerning a moderate growth in total employment.[16] In the results presented so far, unemployment benefits are indexed by the inflation rate. The additional two columns in Table 2.1 reveal that ignoring this channel would significantly overestimate potential labor supply effects. This holds particularly for the extensive margin (up to 1.5 times higher labor supply effect), as higher unemployment benefits reduce the price for leisure and hence lower work incentives.

Aggregate labor supply effects for different reform scenarios (i.e. different VAT increases) are depicted in Figure 2.1. This sheds light on the interaction between both taxes if the shift is smaller. Overall, the total hours effect increases about linearly for both scenarios, reaching 400,000 (PIT) and 330,000 (SSC) full-time equivalents respectively. For the participation margin, effects are substantial only after the fourth reform step. Moreover, the labor supply effect of the SSC reduction is stronger than the PIT reduction across the whole range of reforms for the participation margin. The inverse holds for the change in aggregate hours.

Figure 2.1: Labor Supply Effects for different VAT rate increases



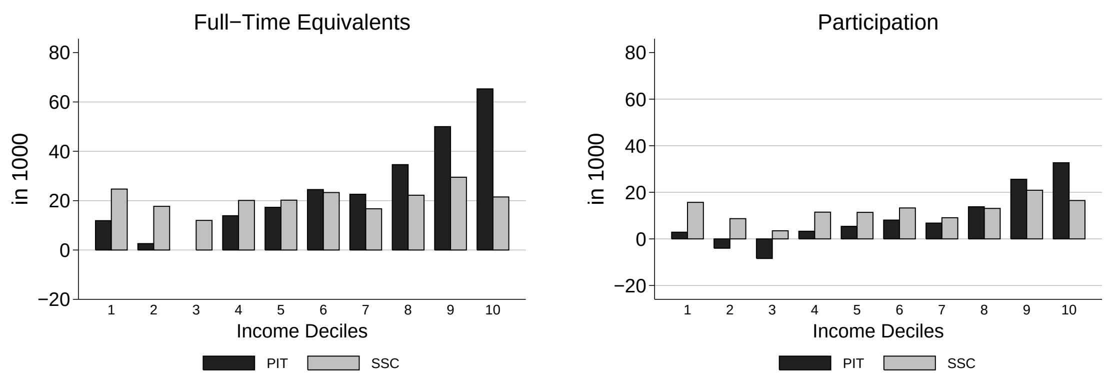
Own calculations with IZAΨMOD v.3.0.4. Full-Time Equivalent = 40 hours per week. The vertical line indicates the reference scenario that displayed in Table 2.1.

The total change in labor supply for the reference scenario is decomposed by income deciles in Figure 2.2. It can be seen that the increase in hours worked in the PIT scenario (dark gray bars) is mainly driven by higher income groups. The participation effect is even slightly negative up to the third decile, while most workers entering the labor market are in the top deciles. These are mostly secondary earners who have been previously inactive and now face a lower individual marginal tax rate. Shifting from SSC affects household budgets already at a lower income level and exceeds the hours

[16] Microsimulation approaches with demand side restrictions are provided by Creedy and Duncan (2005) and Peichl and Siegloch (2012). In both studies, at least half of the supply effect is maintained.

effect from the PIT reduction in the bottom half of the distribution, as indicated by the light gray bars. If policy-makers seek to reduce entry barriers into the labor market by reducing the tax wedge, the SSC scenario appears to be better targeted.

Figure 2.2: Labor Supply Effects by Income Deciles (VAT rate of 25%)



Own calculations with IZAΨMOD v.3.0.4. Income deciles are based on equivalized Post-VAT income. Full-Time Equivalent = 40 hours per week.

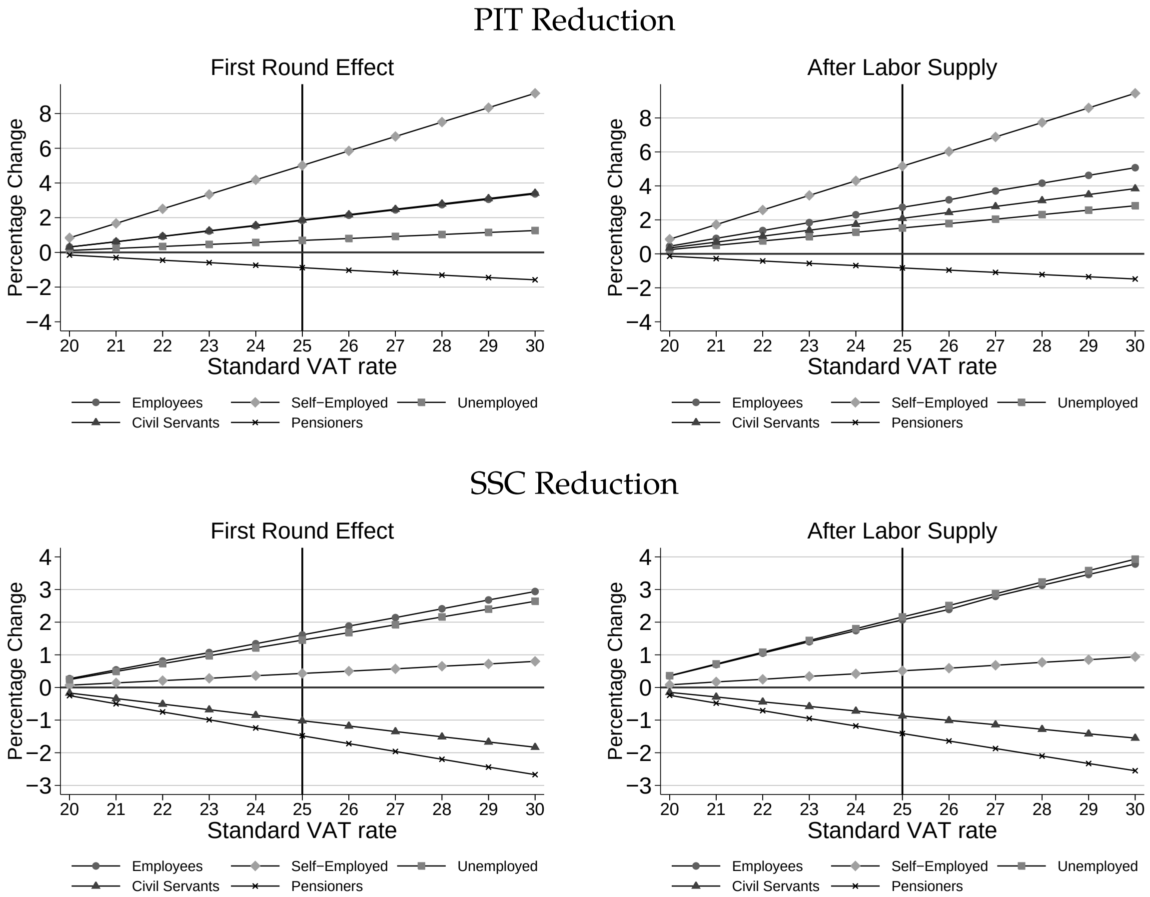
2.4.2 Distributional Impact

Employment Type. The average budget effects with respect to the employment type are illustrated in Figure 2.3.[17] Employees experience modest income gains vis-à-vis the status quo (+4% in the short run, +5% when accounting for labor supply changes). For other employment groups, the differences between the scenarios with and without behavioral adjustment are negligible.[18] Pensioners lose most from the reform, as they hardly benefit from reductions on the income side. Moreover, they are not able to cushion the adverse budget effect trough increased labor supply. For self-employed and civil servants, the picture is mixed. On the one hand, these groups significantly benefit from income tax reductions. However, they are not subject to social security payments. For this reason, civil servants turn out to be slightly worse off from the SSC reform (see also Meinhardt and Zwiener (2005)).[19]

With the exception of pensioners and civil servants, all employment groups are able to compensate a large share of their losses through increased realized labor supply. The main losers from the SSC reform are pensioners, who lose around 2% on average. In relative terms, employees and unemployed workers are the main beneficiaries. The average budget effect for unemployed (+2% and +4% respectively for the SSC reduction) is due to substantial increases for *some* unemployed. For those remaining unemployed

[17] Throughout the distributional analysis, incomes are adjusted by equivalence weights using the modified OECD scale.
[18] This is in line with Picos-Sanchez and Thomas (2015). The results for the standard VAT rate of 21% is complementary to their second reform scenario.
[19] The slightly positive budget effect for self-employed is purely due to changes in spouses' income, which cause equivalence-weighted household income to change.

Figure 2.3: Income change by employment type



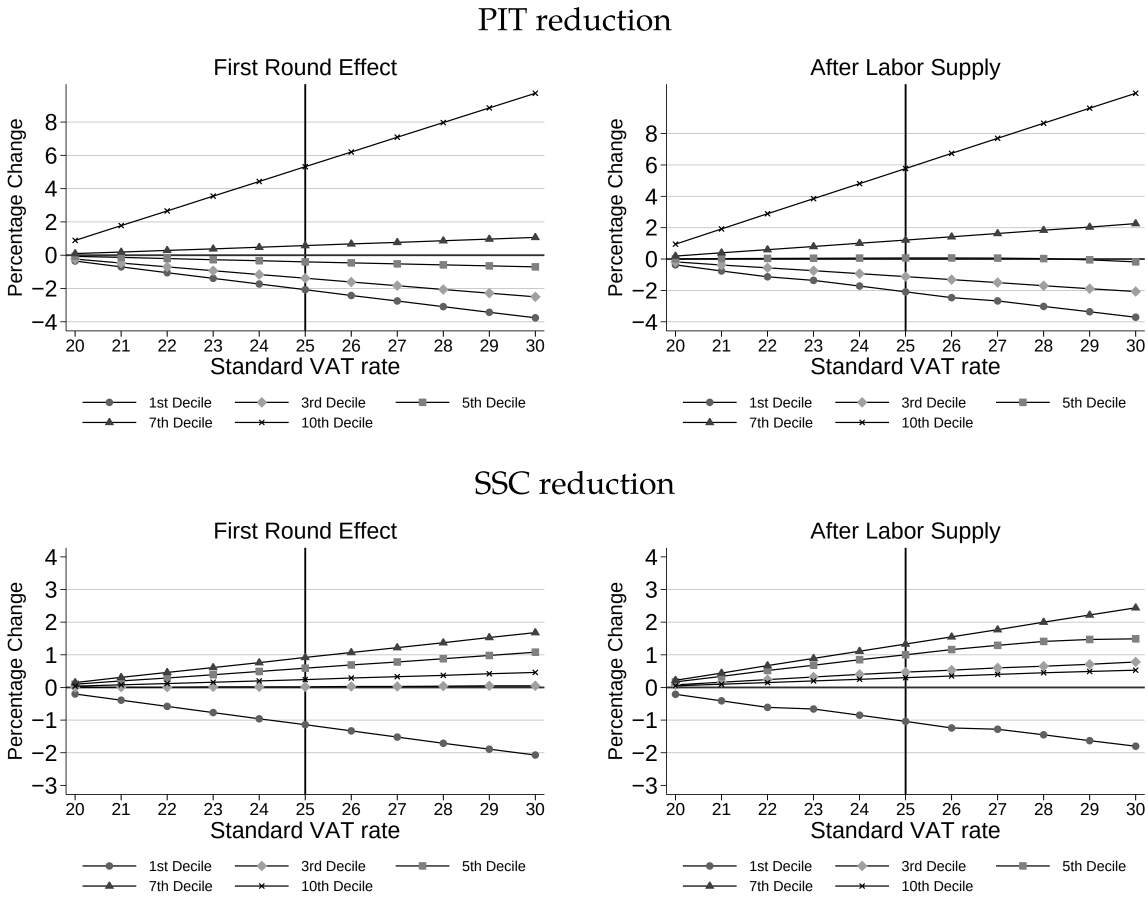
Own calculations with IZAΨMOD v.3.0.4. Income changes refer to equivalized Post-VAT income. First-Round Effects refer to the situation without labor supply reactions. The vertical line indicates the reference scenario.

the total change in PVI will be about zero, as the increase in VAT expenses is expected to correspond roughly to the increase in unemployment benefits due to indexation. In general, these results are in line with expectations as well. Those who are not affected by the tax that is reduced are, in tendency, worse off from the reform.

Income deciles. The distributional impact along the reform path, differentiated by (status quo) deciles of Post-VAT-Income, is illustrated in the upper part of Figure 2.4. It displays the relative income change due to the reform by income deciles. For a clearer exposition, we restrict the presentation to five selected deciles.

The upper panel of Figure 2.4 demonstrates the increasingly regressive impact of a shift from personal income tax to VAT. Minor VAT increases hardly affect budgets of medium-income earners, but rather let the high-income earners better off. After the final reform step (standard VAT rate of 30%), the lowest decile suffers from an income loss of around 4%, while the top decile gains more than 8%. This is in principle not surprising as one would typically expect those households to lose from a shift towards consumption taxation who bear a low burden of PIT prior to the reform. The higher

Figure 2.4: Income change by income deciles



Own calculations with IZAΨMOD v.3.0.4. Income changes refer to equivalized Post-VAT income. First-Round Effects refer to the situation without labor supply reactions. The vertical line indicates the reference scenario.

saving ratio of high-income earners exacerbates this effect. The core interest of our investigation is to analyze to what extent the regressive impact is weakened if behavioral responses are accounted for. As the right panel shows, the distributional picture however hardly changes for the PIT scenario, if we consider the budget changes after the labor supply response. The improvement to the first-round effect is one percentage point at most across income deciles, leaving the poorest decile 4% worse off compared to the baseline.

The equivalent analysis is presented in the lower panel of Figure 2.4 for the SSC reduction. While still implying a regressive impact in the short run, income gains are not larger than 1.5%. The bottom decile loses around 2% on average. Besides, the 7^{th} decile experiences larger gains than the 10^{th} income decile. This can be explained by a low marginal payroll tax burden for top-earners due to the assessment threshold and a decreasing income share of labor earnings for this group. The labor supply response causes the picture to change to some extent by raising the income effects for all deciles. Middle income groups gain relatively more than the highest income decile. The SSC reductions shift taxes from one regressive form of taxation to the other, which clearly has lower adverse distributional effects than the income tax reduction. As the burden of SSC is more dispersed over the income distribution, the budget changes from the reform are less pronounced for the second scenario.

Summing up, a shift from labor to consumption taxation indeed exhibits a regressive impact on household budgets. Lower income groups lose while receivers of high incomes benefit, in tendency, from the reform. This can be easily explained by the fact that the bottom 50% of the income distribution account only for 5% of total income tax revenues and thus hardly benefit from a reduction. The regressive impact is substantially less severe for a shift from social security contributions to VAT. Hence, reforming the personal income tax as suggested here is likely to be confronted with strong political opposition and is therefore not a realistic policy proposal. As a consequence, the in-depth analysis in Section 2.4.3 concentrates on the SSC reductions as the more attractive option for policy-makers.

Tax Progressivity and Inequality. To complete the picture on the distributional impact of the reform, Table 2.2 shows results for the degree of tax progressivity for different components of the tax-benefit system. We analyze two measures of tax progressivity. The Suits index π_{Suits} builds on the Lorenz curve for tax payments. Let $L_X(p)$ and $L_T(p)$ denote the Lorenz curves for pre-tax incomes and tax liabilities respectively. Then the *Suits* index π_{Suits} is obtained by

$$\pi_{Suits} = 2\int_0^1 [L_X(p) - L_T(p)]L'_X(p)dp \tag{2.1}$$

If π_{Suits} is calculated for some parts of the tax-benefit system (as in Table 2.2), the index for the overall progressivity is a weighted average of the partial indices, with average tax rate as weights (Suits, 1977). The index takes values in the [-1;1] interval and is an indicator for the progressivity of the tax schedule. A value of 1 would imply an extremely progressive system where only one individual would be subject to the tax. Opposed to this, the *Reynold-Smolensky* index π_{RS} captures the redistributional impact of a particular tax by the difference in pre- and post-tax income concentration.

$$\pi_{RS} = 2\int_0^1 [L_{X-T}(p) - L_X(p)]dp = \text{Gini}_{\text{PreTax}} - \text{Gini}_{\text{PostTax}} \tag{2.2}$$

The difference of both indices can be illustrated by the following example. A strongly progressive tax schedule (as measured by π_{Suits}) only exerts a redistributive impact if high marginal tax rates are paid by a significant number of taxpayers (captured by π_{RS}). Similarly to π_{Suits}, π_{RS} can be decomposed into the relative contributions of certain elements of the tax system (Lambert, 2001). For both concepts, a progressive (regressive) tax is associated with a negative (positive) value.

Table 2.2: Progressivity of Different Taxes

Reference Scenario	**Total**	**PIT**	**SSC**	**VAT**
		Base		
π_{Suits}	0.218	0.346	-0.060	-0.194
π_{RS}	0.076	0.049	-0.007	-0.012
		Reform 1: PIT Reduction		
π_{Suits}	0.185	0.345	-0.058	-0.188
π_{RS}	0.063	0.040	-0.007	-0.014
		Reform 2: SSC Reduction		
π_{Suits}	0.212	0.345	-0.059	-0.192
π_{RS}	0.074	0.049	-0.005	-0.015

Own calculations with IZAΨMOD v.3.0.4. Reform effects after Labor Supply adjustment for VAT standard rate of 25%. All reforms with indexation of basic unemployment benefit.

In the status quo, π_{Suits} for VAT amounts to -0.194, while it is 0.346 for the personal income tax. Hence, the Value-Added Tax is about half as regressive as the income tax schedule is progressive.[20] At the same time, the distributional impact of VAT as measured by π_{RS} is regressive (-0.012), but only a quarter compared to the PIT progressivity (0.049). The PIT reduction does not affect the progressivity of the tax tariff, but reduces redistribution via the income tax. It is also apparent that both reforms make the VAT schedule more regressive in distributive terms. The Reynold-Smolenksy measure for VAT is only slightly higher than for SSC. This explains why the SSC reform is close to neutral in terms of total progressivity ($\Delta\pi_{Suits} = -2.7\%$) and redistribution ($\Delta\pi_{RS} = -2.6\%$).

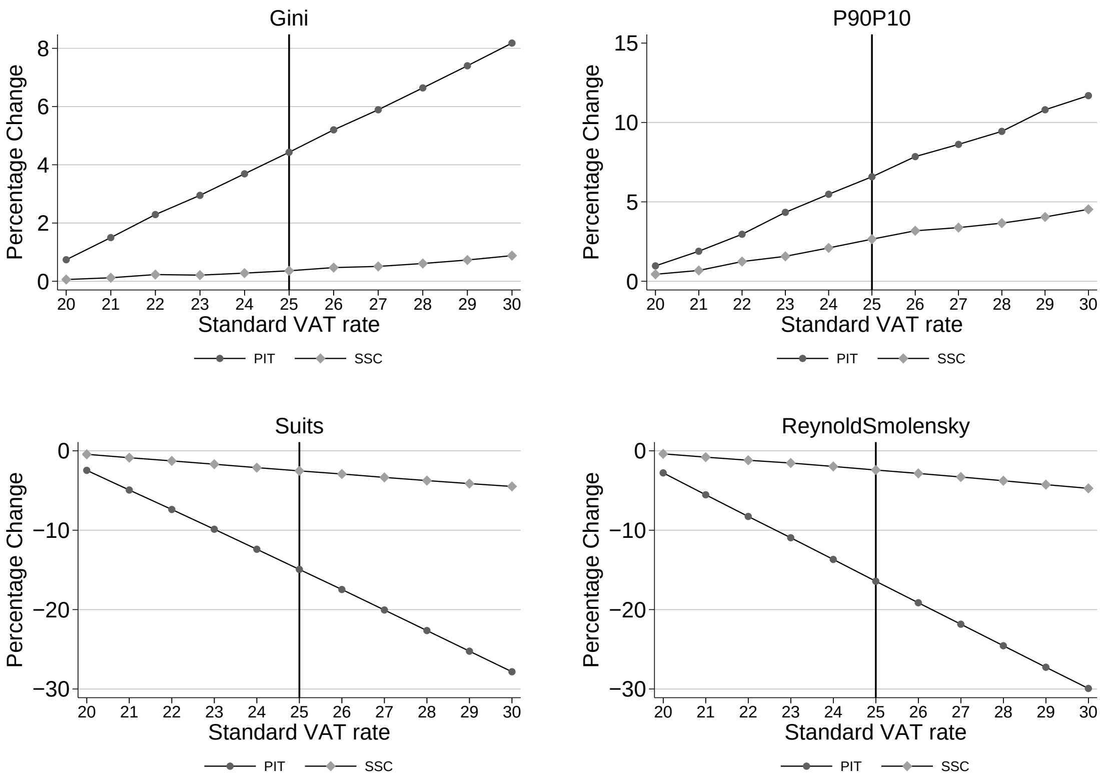
The baseline Gini for our income concepts amounts to 0.303. For the reference scenario, it increases by 0.013 for the PIT reform after Labor Supply. For the SSC reform, the Gini index increases by only 0.001, leaving inequality nearly unchanged. The percentage changes for four basic inequality and progressivity measures are depicted in Figure 2.5 for each intermediate reform step. The Gini index rises by about 8% for the full PIT reform and around 1% for the SSC reform. The $P90/P10$ ratio (upper right panel) however shows a significant increase also for the SSC reform, suggesting higher income polarization.

2.4.3 Sensitivity Analysis

Alternative payroll tax incidence. We deviate from the benchmark SSC reduction scenario by altering the assumption of full incidence of the payroll tax. This implies

[20] See Decoster et al. (2010) for other countries.

Figure 2.5: Changes in Inequality and Tax progressivity



Own calculations with IZAΨMOD v.3.0.4. The graphs show the difference in distributional indices from reducing SSC after labor supply response. Graphs without behavioral response are available upon request. The vertical line indicates the reference scenario.

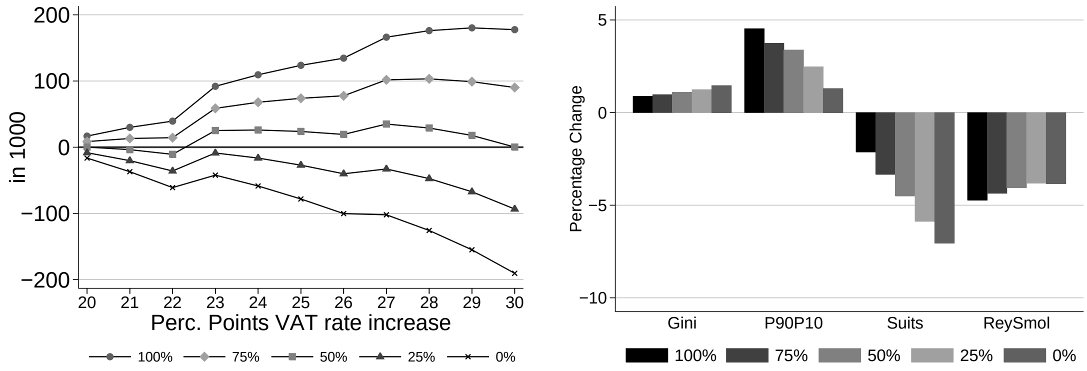
that the total payroll tax reduction falls on employees. Instead, we present changes in aggregate distributional measures in Figure 2.6 for payroll tax incidence values of 100%, 75%, 50%, 25% and 0% respectively. As labor demand is typically estimated to be more elastic than labor supply (Lichter et al., 2015a), an incidence share of more than 50% for employees seems most realistic. Incidence below 100% causes employees to gain less from a payroll tax reduction and hence weakens the positive effect on work incentives. For the extreme case of no incidence, we simply raise only the Value Added Tax. For larger SSC reductions, the labor demand channel may gain importance. As labor costs for firms decrease, hiring may become more attractive. However, this mechanism is out of our models' scope.[21]

As expected, the labor supply response is weaker if the net wage is less affected by the payroll tax change (left panel of Figure 2.6). For an incidence of 25% or less, aggregate labor supply even decreases for all scenarios. The right panel depicts the corresponding changes of various measures of inequality and tax progressivity, the left bar representing the benchmark scenario of 100%. While the Gini index is higher for lower incidence, income polarization, measured by the $P90/P10$ share, decreases. The intu-

[21] This possible positive employment impact would also need to be contrasted with lower commodity demand.

Figure 2.6: Alternative Incidence Assumptions

(a) Labor Supply Effects, Participation (b) Inequality Effects, after LS



Own calculations with IZAΨMOD v.3.0.4.

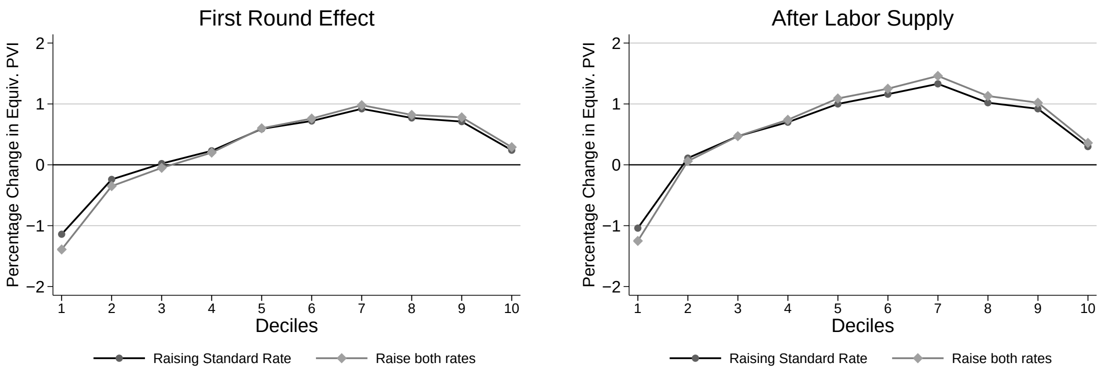
ition is that earners of higher incomes are losing disproportionately if tax incidence is lower. Tax regressivity is increasing steadily as assumed incidence decreases; the overall redistribution of the tax system does not vary much for different incidence values. For the most realistic range of 50% and above, our main conclusions with regard to the overall inequality impact however remain unaffected.

The Suits index indicates a steadily increasing overall regressivity. Interestingly, the $P90/P10$ measure slightly decreases for lower tax incidence, suggesting slightly lower income polarization.

Increasing both VAT rates. So far, our reform scenarios left the reduced VAT rate of 7% unchanged. Levying reduced VAT rates on necessities is justified, among others, by equity considerations. As a consequence, all EU countries with the exception of Denmark impose differentiated VAT rates. Nonetheless, VAT differentiation is often criticized for not achieving its social purpose (OECD/Korea Institute of Public Finance, 2014) and to distort consumers' choices. In the following, we address the question whether shifting the tax burden also on commodities that are taxed at a lower rate is particularly to the detriment of low-income earners. We alter the SSC reduction scenario such that in each reform step, we increase both rates simultaneously in steps of one percentage point. A VAT structure with rates of 23% and 11% (Status Quo: 19% and 7%) is comparable with the reference scenario with regard to the revenue effect. It is important to note that the zero-rate commodities remain exempted.

The distributional outcome of this reform is depicted in Figure 2.7, contrasted with the reference scenario for the SSC reform. As there is virtually no difference in the income changes both in the short and medium run, it is fair to conclude that raising both VAT rates instead of the standard rate does not imply a distinct distributional impact. This suggests that the reduced VAT rate in Germany hardly achieves its redistributional

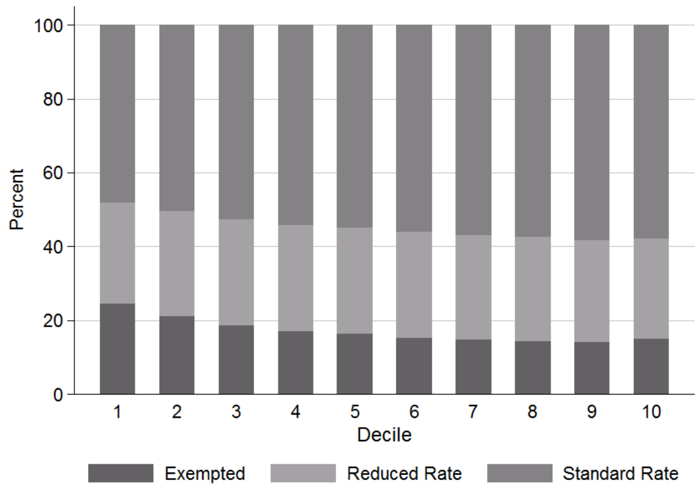
Figure 2.7: Distributional impact of raising both VAT rates for the reference scenario



Own calculations with IZAΨMOD v.3.0.4. The standard rate scenario corresponds to the baseline SSC reduction with a standard VAT rate of 25%, the second scenario applies a standard rate of 23% and a reduced rate of 11%, while reducing SSC. Income changes refer to equivalized Post-VAT income. First-Round Effects refer to the situation without labor supply reactions.

purpose. The intuition is given in Figure 2.8. Reduced-rate commodities account for about the same expenditure share across income groups. A further policy proposal often discussed is the introduction of a uniform VAT rate. If the VAT-exempted commodities are left untouched, one would expect (qualitatively) very similar distributional effects of this reform as in Figure 2.7.

Figure 2.8: VAT Tax Rates over Income Deciles



Source: EVS 2008. Income deciles for equivalized disposable income. Each bar shows mean values of expenditures shares by the respective VAT rate applied.

2.5 Conclusions

This chapter examines a partial shift of taxation from labor income to consumption in Germany. Our empirical approach combines a detailed analysis of changes in household budgets with a microsimulation of behavioral reactions on the labor market. Based on a dual data base, we carry out a microsimulation of several reform scenarios shifting a substantial share of personal income taxes or social security contributions onto the Value Added Tax. The policies are designed revenue-neutral. The expectations of positive effects on household work incentives are confirmed by the simulation. The total increase in labor supply for the reference scenario (Standard VAT rate of 25%) is expected to be rather moderate below 1% of total employment for the benchmark scenarios. This suggests a limited capacity of this policy instrument for targeting workers at the margin to enter employment.

The distributional evidence suggests that a shift from personal income tax to VAT has a regressive impact on household budgets. Negative effects are expected for low-income households, unemployed and pensioners in particular. This budget loss amounts to up to 4% of equivalized income, whereas the policy clearly favors high-income earners. The change in aggregate distributional measures supports this view by indicating higher inequality and a lower degree of the overall tax progressivity. Typically, most losers have a low burden of direct taxes and thus hardly benefit from a reduction on the income side. These static (i.e., non-behavioral) results are in line with micro-based evidence from other studies (Meinhardt and Zwiener, 2005; Picos-Sanchez and Thomas, 2015; Thomas and Picos-Sanchez, 2012).

Taking into consideration labor supply effects, the overall picture slightly improves for the SSC reduction, as income effects turn positive for the majority of people. This is for two reasons. First, SSC are a regressive form of taxation themselves. Replacing them with another regressive form of taxation hardly alters its distributional impact. Second, SSC reductions affect household budgets at a rather low income level, which bears activating potential. Reducing social security contributions overall entails lower inequality increases than a shift from personal income taxes. Besides, we demonstrate the negligible redistributive impact of the reduced VAT rate. It is worth noting that our static approach does not allow conclusions beyond the medium run. It is possible that positive employment effects vanish in the long run if unions are able to assert higher wages.

Our empirical results may serve as a point of departure for further research in several areas. First, it is worth considering possible extensions of the policy proposal in order to increase both political feasibility and effectiveness with regard to increasing work incentives for low-income groups (Thomas and Picos-Sanchez, 2012). One could think of a reform that is both *revenue-* and *inequality*-neutral. As Decoster et al. (2010) suggest, increasing the progressivity of the remaining income tax schedule is one option. Another way would be to compensate the main losers by raising old-age pensions. Our results suggest that designing such a reform is very well possible. In order to get a broader picture of the overall distribution of the consumption tax burden, incorporating excise taxes seems promising. A shift towards taxes on fuel or electricity is regularly discussed in the context of environmental tax reforms that aim at internalizing external effects.

2.A Appendix

2.A.1 Dual Database

The main database for IZAΨMOD is the German Socio-Economic Panel Study (SOEP), which is an annual panel study of households and individuals that was launched in 1984 as a representative cross-section of the population living in private households in Germany (Wagner et al., 2007).[22] Since then, the scope and size of the panel has been steadily extended. Special attention is given to the representativeness of the German population by explicitly oversampling foreigners and high-income households. As of now, it covers around 22,000 persons living in more than 12,000 households. Among others, IZAΨMOD exploits information on gross wages, household composition, working time, age and educational background of household members, as well as employment status and housing costs. These data serve as input for the tax and benefit module and for the labor supply estimation.

The 2010 SOEP wave is delivered with information on household consumption. The survey design implies non-response and heaping of values, which are dealt with by correction methods described in Marcus et al. (2013). However, the consumption categories do not cover all aspects of household expenditure, which justifies the additional effort of imputing information from an auxiliary source, namely the German Sample Survey of Income and Expenditure (*Einkommens- und Verbrauchsstichprobe*, EVS).[23] It is a cross-sectional survey conducted by the Federal Statistical Office that started in 1962/1963 and is repeated every five years. The most recent available wave was conducted in 2008.[24] It covers about 55,000 households, of which a 80% subsample is provided for scientific analyses (44,088 observations). The EVS data contain detailed information on every household member's employment, income from different sources and assets. Its main focus rests on expenditures for all types of commodities and services. All participants constantly keep record of their expenditures throughout a three-month period.

Although EVS and SOEP apply similar concepts of household and household income, both data sets are not fully comparable due to methodological differences.[25] Income is reported in more detail in the EVS. On the other hand, it shows weaknesses with respect to representing foreigners and high-income earners accurately. EVS does not sample households above a monthly gross income threshold of €18,000. In addition, middle-income groups are slightly over-represented. The measurement error

[22] In particular, we use version 29 of the Socio-Economic Panel data for years 1984–2012, doi:10.5684/soep.v29. We rely on data from the 2009 wave only.

[23] Non-covered items include vehicle purchases, home appliances and telephone costs.

[24] See Destatis (2013) for a detailed description of the methodology. EVS has been part of the European Household Budget Survey until 1998. It was then replaced by the another survey (*Laufende Wirtschaftsrechnungen, LWR*) which is carried out annually, but with a much smaller sample.

[25] See Becker et al. (2003) for further information on the comparability of EVS and SOEP.

is probably larger in SOEP than in EVS, due to the retrospective methodology of the SOEP.

2.A.2 Imputation of Expenditures

Major tax shifts, as analyzed in this chapter, are expected not only to affect household budgets, but also to assert substantial price changes. To capture the effect on household consumption, the first-best approach would be to fully characterize household consumer behavior by estimating a demand system (Banks et al., 1997). This is unfortunately not possible based on one cross-section due to lack of price variation. We hence follow a middle-path by estimating Engel curve relationships in a first step, with the aim to reproduce the expenditure patterns observed in EVS in the SOEP data. In a second step, we incorporate externally estimated price elasticities.

For the imputation of expenditures, we adopt the approach by Decoster et al. (2013), who carry out a parametric estimation of Engel curves. This approach is briefly outlined in the following.[26] A similar approach has recently been applied by Savage and Callan (2015).

As the EVS is based on quarterly, rather than annual data, significant singular purchases may bias the picture by yielding inappropriately high expenditure values. For this reason, we clear the EVS from households with negative disposable income. Furthermore, we drop observations where either the statistical difference between income and expenditures or the amount of durable expenditures exceeds twice the disposable income. In total, less than 1% of observations are excluded, leaving us with a sample of 43,632 households.

We classify expenditures into 16 categories; 15 non-durables and one category for durable consumption goods.[27] Durable expenditures require special treatment, as their purchase may not be observed in the three-month window, while the actual consumption stream persists from previous periods. For this reason, we distribute total durable expenditures equally among households with identical cells, defined by seven income

[26] The main difference to Decoster et al. (2013) is that they impute expenditures into the German EU-SILC data, which is distinct from SOEP. We do not have any reason to believe that the imputation quality should be much different when applied to SOEP, as both data sets are a comprehensive samples of the same population. Apart from that, we employ further household covariates, such as the flat size. We also do not estimate a probit model for durable commodities first, as for our choice of cells, we are left with a negligible number of zero durable expenditures. Finally, we use a tobit specification for the four expenditure groups with many zero observations.

[27] The 15 nondurable commodity classes are mostly in line with the COICOP classifications (UN, 2012). Deviations occur for household services, which encompasses items from housing (COICOP 4) and household maintenance (COICOP 5). They are defined as follows: Food and non-alcoholic beverages; alcoholic beverages; tobacco; clothes and shoes; home fuels and electricity; rents; household services; health; private transport; public transport; communication; recreation and culture; education; restaurants and hotels; other expenditures. Durable consumption entails, among others, expenditures on furniture, home appliances, means of transport and household entertainment.

groups, seven age cohorts and four household types.[28] Afterwards, two Engel curves for the total of durable and non-durable consumption are estimated.[29]

$$\ln c_{ji} = \alpha_j + \beta_j \ln y_i + \gamma_j \ln(y_i^2) + \delta_j X_i + \varepsilon_{ji} \tag{2.3}$$

where j is durable or nondurable total expenditure and y_i denotes households disposable income, X_i is a vector of households characteristics contained in both data sets, such as flat size, community size, number of children in several age groups, number of working household members and geographical region. Further covariates reflect characteristics of the household head, namely age, age squared, sex, education and employment status. We further include an interaction term between y_i and household size. Note that this framework explicitly allows negative savings, as the sum for durable and nondurable expenditures may exceed y_i. Based on these estimates, the total amount of durable and nondurable expenditure can be predicted for SOEP households. As the dependent variable in Equation (2.3) is in logs, the prediction bias needs to be corrected.

In a second step, we regress 15 nondurable expenditure *shares* on the log and log squared of total nondurable consumption. This functional form is regularly used for estimating commodity demand systems. Four nondurable categories (tobacco, rents, public transport and education) exhibit a substantial share of zero expenditures, which needs to be accounted for. We therefore fit a tobit model for these $k = (1, \ldots, 4)$ categories.

$$\omega_{ki}^* = \alpha_k + \beta_k \ln c_i^{\text{nd}} + \gamma_k \ln \left(c_i^{\text{nd}}\right)^2 + \delta_k X_i + \varepsilon_{ki} \tag{2.4}$$

$$\omega_{ki} = \max(0, \omega_{ki}^*) \tag{2.5}$$

The remaining $l = (1, \ldots, 11)$ shares are estimated by unrestricted OLS using c_i^{rest} as total expenditures. This is defined as total nondurable expenditures from these 11 categories: $c_i^{\text{rest}} = c_i^{nd} - \sum_{j=1}^{4} c_j$. This secures that predicted shares sum up to one.

$$\omega_{li} = \alpha_l + \beta_l \ln c_i^{\text{rest}} + \gamma_l \ln(c_i^{\text{rest}})^2 + \delta_l X_i + \varepsilon_{li} \tag{2.6}$$

Based on these estimates, the 15 nondurable consumption shares $\widehat{\omega_{mi}}$ with $m = (1, \ldots, 15)$ can be predicted based on imputed total amounts $\widehat{c_i^{nd}}$ and $\widehat{c_i^{rest}}$. For the tobit models, we predict the unconditional expenditure and assign a lower limit of zero. For the standard OLS estimates, negative predicted values occur in few cases. These amount however to single-digit values and are set to zero. The remaining expenditures are then corrected such that they sum up to one.

[28] This approach follows Beznoska and Ochmann (2013). Income and age groups are defined by quantiles within household types. The mean number of observations per cell is 324, with a minimum of 17.

[29] All estimations are carried out using sample weights; monetary amounts are in monthly terms.

The accuracy of the matching procedure is evaluated in Figure 2.9, comparing the mean expenditure share for each of the 16 categories, separately for four types of household composition (Single, Single Parents, Couples, Families with children). As can be seen, the major expenditure items are durable commodities, food and drinks, rents and culture and recreation. All observed expenditure shares are replicated in SOEP with minor deviations not exceeding two percentage points. Rents are the only group with larger deviations. This is however a minor issue for the accuracy of the imputed VAT burden, as the average VAT rate for this particular group is only 3.7%. One problem with our procedure may arise for top-income earners, as there are none in EVS (see Section 2.A.1). As a consequence, their imputed VAT burden might be wrong if their expenditure behavior cannot be described by extrapolation from lower income groups. The direction of this bias is however not clear.

In order to allow household expenditures to react to price increases, we have to incorporate external estimates on the price elasticities of demand. To our knowledge, the only estimation of a full demand system based on EVS is provided by (Kohn and Missong, 2003). We make use of their own-price elasticities, differentiated by six household types (Table 2.3) while using the mean value over all income levels.[30] When simulating the tax reforms, the Engel curve estimations are used to depict consumption reactions from changing disposable income. The price reaction is calculated afterwards. In the reference scenario, consumption decreases by about 1% in average. This comparable to the change from labor supply adjustments and underscores the need to account for price reactions. On the other hand, this has a minor impact on the distributional implications of the reform, as price elasticities do not show large differences across demographic groups.

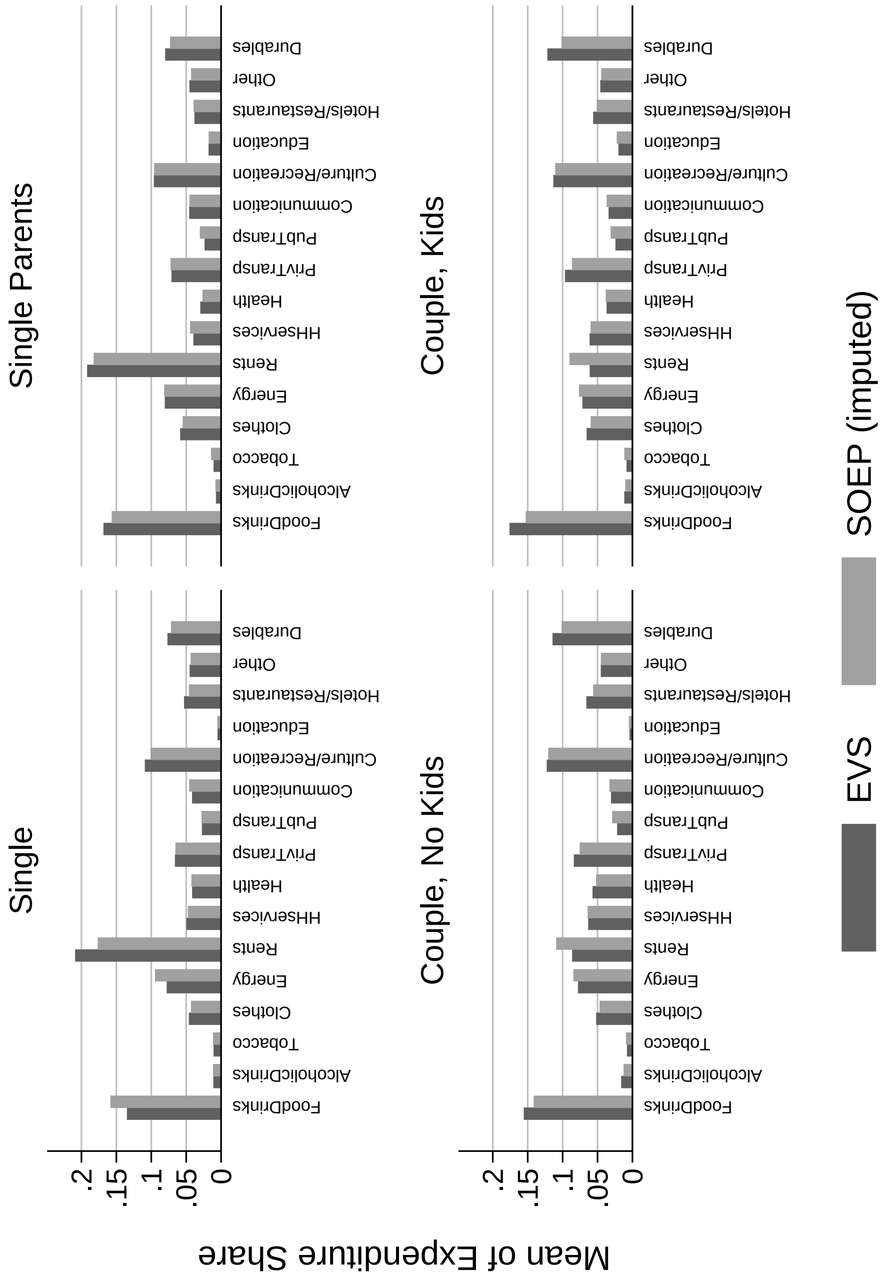
Table 2.3: Own-price Elasticities of Commodity Demand

Commodity	Household Type					
	Singles	Single Parents	Couples, no ch	Couples, 1 ch	Couples, 2 ch	Couples, 3 ch
Food, Alcohol, Tobacco	-0.391	-0.291	-0.292	-0.274	-0.262	-0.254
Clothing	-0.438	-0.264	-0.351	-0.294	-0.277	-0.272
Energy, Rents, Household Services	-0.324	-0.287	-0.325	-0.306	-0.296	-0.287
Health	-0.403	-0.273	-0.330	-0.320	-0.299	-0.296
Transport	-0.437	-0.306	-0.384	-0.325	-0.322	-0.318
Communication, Culture, Hotels, Restaurants, Other	-0.523	-0.366	-0.423	-0.402	-0.396	-0.397
Education	-0.446	-0.288	-0.400	-0.331	-0.306	-0.296

Source: Kohn and Missong (2003, Table 5), using means across income groups.

[30] The authors were not able to identify cross-price elasticities.

Figure 2.9: Mean Expenditure shares in both data sets



Calculation of VAT payments. In order to identify the consumption shares of goods with different VAT rates, we make use of the weighting scheme applied by the German Federal Statistical Office for capturing price level changes (the so-called "representative basket of products"). As an example, non-al-co-ho-lic drinks (taxed with the standard rate) are assigned a weight of 9.3% of total expenditures in expenditure category 1 (food and beverages). Therefore, 9.3% of this consumption component is allotted the standard VAT rate, while the rest is taxed with the reduced rate.

The treatment of VAT-exempted commodities for which no input tax deduction can be claimed deserves special attention. Despite their revenues being formally exempted, it is reasonable to assume that firms shift a certain share of their input tax burden to consumers through increased prices. The most notable case of exempted goods are rents. Landlords might be able to charge higher rents in order to compensate for taxes paid in connection to restorations or construction works. The extent to which this occurs is however hard to estimate, as the possibility to increase rents is restricted and depends on local market characteristics. Hence, there is no agreement in the literature on the extent to which rents are burdened with VAT.[31] Following RWI and FiFo (2007), we assume that 11% of expenditures on rents are subject to the standard VAT rate. This rather low assumption can be justified in light of the overall (rather short) time horizon of our model. Similar incidence assumptions are made for medical and financial services. Table 2.4 depicts the resulting VAT rate shares by consumption category.

The corresponding distribution of VAT payments is visualized in Figure 2.10. In line with other studies (OECD/Korea Institute of Public Finance, 2014), we clearly confirm a regressive pattern of the German VAT. As income rises, the VAT share of disposable income decreases from 12% for the first decile to 4% for decile 10, which is below OECD average. With regard to expenditures, the pattern is rather flat, the VAT burden ranging between 9% and 10%.

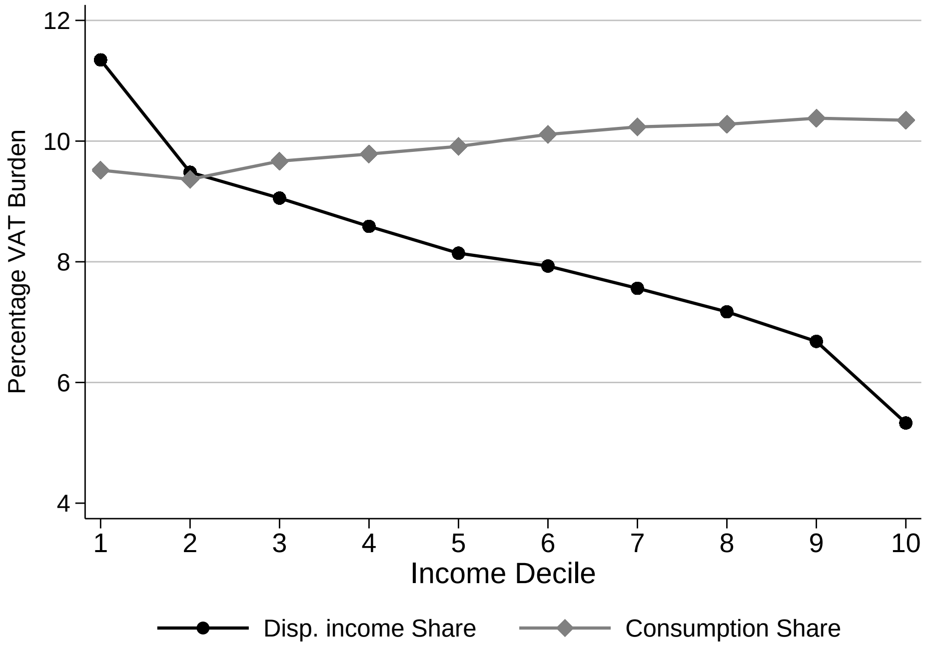
[31] See RWI and FiFo (2007), Fritzsche et al. (2003) and Bach (2005).

Table 2.4: VAT Shares by Expenditure categories

	Expenditure category	Avg. Share of Total Expenditures	Share of VAT 0% (%)	Share of VAT 7% (%)	Share of VAT 19% (%)
1	Food and beverages	16.3	—	90.7	9.3
2	Alcoholic Beverages	1.3	—	—	100.0
3	Tobacco	0.9	—	—	100.0
4	Clothing and footwear	5.6	—	—	100.0
5	Household Fuels, Energy	8.0	—	15.6	84.4
6	Rents	11.8	80.3	—	19.7
7	Household Services	2.5	—	36.2	63.8
8	Health	4.6	43.2	16.2	40.6
9	Private Transport	8.5	—	—	100.0
10	Public Transport	2.5	5.0	57.0	38.0
11	Communication	3.7	—	7.4	92.6
12	Recreation and culture	11.9	9.4	33.4	57.2
13	Education	1.0	92.4	—	7.6
14	Restaurants and hotels	6.0	—	70.2	29.8
15	Other goods and services	4.7	30.6	—	69.4
16	Durable commodities	10.8	20.2	2.4	77.4
	Total	**100.0**	**17.3**	**27.8**	**55.0**

Source: Own calculations.

Figure 2.10: VAT incidence



Own calculations with IZAΨMOD v.3.0.4. Income Deciles are computed using equivalence-weighted household incomes.

2.A.3 Tax-Benefit and Labor Supply Modules

We calculate household disposable income from gross income and household characteristics from the SOEP data by means of a tax-benefit calculator. It reproduces the regulations of the system of direct taxes, social security contributions and benefits in Germany. As the latest available EVS wave is from 2008, we apply the legal status of that year.

In the last decades, VAT has become the main source of revenue for the German government. VAT and income tax together nowadays account for about two thirds of total tax revenues. The personal income tax is designed progressively, with a rate of 14% for incomes just above the basic allowance and with a top marginal rate of 42%.[32] Social security contributions, in contrast, are calculated as a constant share of labor income until an upper threshold is reached. SSC payments are, in general, equally split between employer and employee.[33] Civil servants and self-employed do generally not contribute to the public social welfare system.

At the core of our microsimulation approach, a behavioral labor supply module estimates preference parameters for the optimal choice between leisure time and disposable income on the household level. The household decision is implemented following the discrete choice household labor supply model (see Löffler et al., 2014b, for technical details on discrete choice modelling). In case of couple households, it assumes a joint utility function for both spouses. The model is discrete in the sense that a household n can choose between a finite number of combinations in consumption and leisure, denoted with J_n. We restrict the choice set to seven time categories of weekly working hours ($0, 10, 20, 30, 40, 50, 60$ hours). For households that are flexible in their labor supply decision, this results in seven alternatives for singles and couples with one flexible spouse.[34] For couples with two flexible spouses, the choice set expands to 49 alternatives. In addition we further expand the choice set in order to account for an endogenous decision on whether or not to apply for public benefits (Hoynes, 1996). This way, we address the potential presence of welfare stigma. We specify the utility U for every choice alternative $j \in J_n$ as a function of household consumption C_{nj} (after subtract-

[32] In addition, a so-called *solidarity surcharge* is levied, amounting to 5.5% of the total income tax burden. For earners of income above €250,000 per year, a marginal tax rate of 45% is applied. In 2008, capital income was taxed the same way as earnings before introducing a dual income tax regime in 2009.

[33] In 2008, the following rates apply for employees (overall): Old Age Pension Insurance: 9.95% (19.8%); Health Insurance: 7.9% (14.9%); Unemployment Insurance: 1.65% (3.3%); Care Insurance: 0.85% (1.7%).

[34] Self-employed persons and civil servants are assumed inflexible in their choice of labor supply. Even though they might adjust their labor supply, we assume that it is based their consumption/leisure decision follows a different rationale than those of employees. Hence, we treat their labor supply as fixed.

ing consumption taxes), leisure of the spouses $L_j^s, s \in \{m, f\}$ and a dummy for welfare participation P_{nj}.[35]

$$U\left(C_{nj}, L_j^m, L_j^f, P_{nj}\right) = \boldsymbol{x_{nj}^1 \beta_1'} \ln C_{nj} + \beta_2 (\ln C_{nj})^2 + \beta_3 \ln C_{nj} \ln L_j^m + \beta_4 \ln C_{nj} \ln L_j^f + \boldsymbol{x_{nj}^2 \beta_5'} \ln L_j^m + \beta_6 (\ln L_j^m)^2 + \boldsymbol{x_{nj}^3 \beta_7'} \ln L_j^f + \beta_8 (\ln L_j^f)^2 + \delta' P_{nj} + \boldsymbol{x_{nj}^4 \gamma'} + \varepsilon_{nj} \tag{2.7}$$

Leisure follows directly from working hours, assuming a time endowment of 80 hours per week. Disposable income for counterfactual choice categories are calculated by keeping hourly wages constant. The vectors $\boldsymbol{x_{nj}^1}$ to $\boldsymbol{x_{nj}^4}$ capture individual and household characteristics, such as age, number of children, handicap status and the presence of a needy person in the household. By interacting them with leisure and consumption, we account for observed heterogeneity.[36] The parameter δ captures welfare stigma. The vector $\boldsymbol{x_{nj}^4}$ contains dummies on part-time and full-time work. The hence reflect market restrictions due to working hours regulations. Equation 2.7 simplifies for the case of single households, as leisure of the second person and interactions thereof are dropped.

Under the assumption that the error term ε_{nj} in Equation 2.7 follows an extreme value type I distribution, McFadden (1974) showed that the probability of household n to choose alternative i over all other alternatives can be expressed as follows:

$$P\left(U_{ni} > U_{nj}, \forall j \neq i\right) = \frac{\exp\left(\varphi\left[C_{ni}, L_{ni}^f, L_{ni}^m\right]\right)}{\sum_{s \in J_n} \exp\left(\varphi\left[C_{ns}, L_{ns}^f, L_{ns}^m\right]\right)} \tag{2.8}$$

The term φ captures the observed part of the utility function 2.7. The structural utility parameters can be obtained by estimating equation 2.8 via maximum likelihood. Results are given in Table 2.5. Based on these estimates, the individual reactions following a change in households' net income can be simulated. This is done by predicting the choice probabilities for each working time category under the old and the new regime. Labor supply for each household is then obtained by multiplying predicted probabilities with the respective amount of hours. The presented reform effects should thus be interpreted as an expected value of supplied hours. A similar logic applies to income amounts. The simulated labor supply elasticities are in line with extant micro-based literature and range from 0.1 for single men to 0.3 for women in couples (see Table 4 in Löffler et al. (2014a)).

[35] As Löffler et al. (2014b) demonstrate, the type of utility function is of minor importance for the implied elasticity of labor supply.

[36] We assume the coefficients β_1 to β_8 to be fixed. Assuming some of these to be random would allow for unobserved heterogeneity, but this is particularly more burdensome in terms of computation.

Table 2.5: Estimates of structural labor supply model

	(1) SingleM	(2) SingleF	(3) CoupleM	(4) CoupleF	(5) CoupleMF
C	-5.725***	-2.607	3.512*	-1.907	2.391*
	(1.360)	(1.460)	(1.451)	(1.564)	(1.037)
C^2	0.199***	0.197***	0.172***	0.118*	0.221***
	(0.0290)	(0.0303)	(0.0399)	(0.0516)	(0.0165)
$C\times$					
Age$_m$	0.0413		0.0475		0.0866*
	(0.0291)		(0.0580)		(0.0348)
Age$_m^2$	-0.000516		-0.000351		-0.00103**
	(0.000337)		(0.000744)		(0.000371)
Handc$_m$	-0.100		0.789		0.208
	(0.116)		(1.354)		(0.242)
Age$_f$		0.0958*		0.214*	-0.0336
		(0.0396)		(0.105)	(0.0386)
Age$_f^2$		-0.00112*		-0.00251*	0.000464
		(0.000480)		(0.00126)	(0.000434)
Handc$_f$		-0.322*		-0.359	0.0612
		(0.161)		(0.399)	(0.200)
Care		-0.301		-0.745	-0.0971
		(0.231)		(0.417)	(0.159)
Child $\leq$ 2y		0.441		0.0448	0.217
		(1.020)		(1.581)	(0.260)
Child 3–6y		0.132		-0.138	-0.0403
		(0.369)		(1.066)	(0.138)
Child 7–16y		0.376		0.826	-0.00863
		(0.217)		(0.570)	(0.0730)
$C \times L_1$	0.943***	-0.00331	-1.274***	-0.522*	-0.894***
	(0.246)	(0.272)	(0.233)	(0.205)	(0.125)
L_1	18.96***	23.20***	55.75***	22.64***	47.04***
	(4.312)	(4.456)	(4.862)	(3.542)	(2.641)
$(L_1)^2$	-3.095***	-2.073***	-6.586***	-1.884***	-5.452***
	(0.459)	(0.396)	(0.597)	(0.356)	(0.261)
$L_1\times$					
Age$_m$	-0.241**		-0.0361		-0.205***
	(0.0773)		(0.0868)		(0.0451)
Age$_m^2$	0.00297**		0.000845		0.00262***
	(0.000929)		(0.00100)		(0.000518)
Handc$_m$	1.844***		0.683		0.508
	(0.458)		(1.046)		(0.314)
Age$_f$		-0.424***		-0.269**	
		(0.0837)		(0.0932)	
Age$_f^2$		0.00582***		0.00396***	
		(0.00103)		(0.00108)	
Handc$_f$		-0.103		0.365	
		(0.544)		(0.555)	
Care		4.369**		-0.170	
		(1.342)		(0.623)	
Child $\leq$ 2y		2.393		1.421	
		(1.239)		(0.900)	
Child 3–6y		2.166***		2.008***	
		(0.489)		(0.568)	
Child 7–16y		1.353***		1.118***	
		(0.240)		(0.246)	
$C \times L_2$					-0.192
					(0.147)
L_2					18.16***
					(2.657)
$(L_2)^2$					-1.899***
					(0.221)
$L_2\times$					
Age$_f$					-0.221***
					(0.0518)
Age$_f^2$					0.00352***
					(0.000632)
Handc$_f$					0.434
					(0.393)

Table 2.5 — continued

Care					1.433**
					(0.487)
Child $\leq$ 2y					3.380***
					(0.394)
Child 3–6y					1.694***
					(0.188)
Child 7–16y					1.148***
					(0.0935)
$L_1 \times L_2$					0.451*
					(0.181)
Dummy Variables					
Work_m	-4.460***		-6.851***		-5.773***
	(0.380)		(0.586)		(0.239)
Parttime_m	0.135		0.834*		0.199
	(0.260)		(0.418)		(0.190)
Fulltime_m	0.928***		1.148***		1.145***
	(0.0946)		(0.116)		(0.0510)
Benefit Takeup	-1.052***	-0.992***	-1.789***	-1.448***	-1.619***
	(0.186)	(0.114)	(0.215)	(0.230)	(0.0928)
Work_f		-1.292***		-1.259***	-1.039***
		(0.186)		(0.153)	(0.0849)
Parttime_f		-0.0593		0.370***	-0.0177
		(0.117)		(0.0975)	(0.0593)
Fulltime_f		0.968***		0.704***	0.806***
		(0.0842)		(0.101)	(0.0574)
N	8237	12133	6405	8397	205897
AIC	2707.9	4249.2	2225.9	4001.8	21022.9

Own estimations with IZAΨMOD v.3.0.4, based on SOEP 2009. Standard errors in parentheses * (p<0.05), ** (p<0.01), *** (p<0.001). Columns (3) and (4) encompass men and women respectively with a non-flexible spouse (due to retirement, self-employment or disability). The choice set consists of 7 choices for weekly hours in the first 4 columns, and 49 choices in the last column. The choice set is further expanded to allow for endogeneity of benefit take-up. L_1 and L_2 indicate leisure of the first and the second adult in the household respectively. Handicap indicates a handicap degree $\geq$ 50. Care indicates the presence of a person in need of care in the household. Household utility is specified as translog (Equation 2.7.). AIC = Akaike Information Criterion.

3 Wealth Transfers and Tax Planning: Evidence for the German Bequest Tax

3.1 Introduction

For the past one hundred years, private wealth in developed economies has largely been a result of business and labor market activities. According to recent findings, this trend has reversed as inheritances are becoming an increasingly important source of income (Piketty, 2011; Piketty and Zucman, 2014). A recent rise in the capital-to-income ratio has been confirmed for a number of countries.[1] This is explained by decades of stable capital accumulation in times of peace, along with shrinking fertility rates, which causes more wealth to be distributed among fewer individuals.

Moreover, wealth is typically much more unequally distributed than income. In particular, Germany is found to display one of the highest wealth inequality levels in the EU (Carroll et al., 2014). According to Corneo et al. (2016), private wealth in Germany stems to about one third from bequests, this share being rather constant across the wealth distribution. As inherited affluence is not associated with individual labor market effort, this gives rise to equity concerns. As a consequence, the taxation of wealth and wealth transfers in order to reduce wealth inequality has re-appeared in the policy debate.[2] Taxing the stock of wealth is often seen as inefficient due to a mobile tax base and high administrative costs. In contrast, extending the taxation of wealth transfers as a source of effortless personal income is generally more accepted and seems more feasible than taxing the stock of wealth. Around three quarters of OECD countries currently levy some form of wealth transfer tax. Their overall importance is however low with revenues of 0.13 percent of GDP in average, or 0.36 percent of total tax revenues.[3]

[1] See Piketty (2011) for France, Schinke (2012) for Germany, Ohlsson et al. (2014) for Sweden and Atkinson (2013) for the UK.

[2] Whether inheritances amplify or dampen wealth inequality is an empirical question. Apart from the wealth distribution of bequest receivers and recipients, it crucially depends on the volatility of the tax base, i. e., on the extent to which taxpayers are able to circumvent the tax. Most studies find that inheritances lead to stronger wealth concentration and inequality. There are however contradictory findings. See Elinder et al. (2015) and the literature overview therein. For Germany, such a study has not been undertaken yet.

[3] Germany: 0.16% of GDP, 0.45% of tax revenue (OECD, 2013b). See Boadway et al. (2010) for an illustration of the historical development of wealth transfer taxation.

In 2016, annual bequest tax revenues increased for the second time in a row by 15%, which may reflect the rising trend in overall bequests.

The normative literature on how to solve the trade-off between equity and efficiency in the context of wealth transfer taxation is controversial. A classic result is that if the utility of bequest receivers is not regarded, there is no reason to tax bequests separately (Atkinson and Stiglitz, 1976). By allowing for altruism of parents, Farhi and Werning (2010) argue for negative optimal marginal tax rates on estates. The optimal pattern of these subsidies is progressive, i. e. higher estates are taxed at a higher (negative) rate. In their model, the amount of bequests are fully determined by individual ability. This assumption is relaxed by Piketty and Saez (2013a), by allowing for multiple sources of inequality. This reflects the fact that the distribution of bequests received is typically more skewed than the earnings distribution. They thus find a substantial role for bequest taxation, yielding optimal inheritance tax rates well above 50%.

Empirical evidence on the behavioral reactions to bequest taxation is scarce for countries outside the US. This chapter aims at closing this gap by exploiting the design of the German bequest tax schedule. I build on previous literature that identifies the elasticity of taxable income (ETI) from taxpayers' bunching at discrete jumps in the tax schedule (Saez, 2010). This chapter applies a bunching approach to the taxation of wealth transfers, which has rarely been done before.[4] Two recent studies also investigate bunching with inheritance taxes. Goupille-Lebret and Infante (2016) exploit time and age notches in the tax treatment of French life insurance schemes to estimate inter-temporal substitution in the accumulation of assets. Most closely to this chapter, Glogowsky (2016) also analyzes bunching for the German bequest tax. Relying on administrative data from 2002, he finds bunching in particular for predefined inheritances, i.e. inheritances whose exact value (as opposed to the share) is defined by a descendant's last will.[5] My findings are relevant for tax policy in at least two aspects. If transferred wealth is found to be very responsive to taxation, taxpayers might avoid bequest taxation. Apart from lowering tax revenue, this might harm the intended effects of higher bequest taxation, such as lowering wealth inequality.

Behavioral responses to taxation can broadly be classified as *real* or *shifting* responses (Kopczuk, 2013). In the context of bequest taxation, one example for a real response is the effect on investment decisions. If future transfers are anticipated, individual saving might be discouraged by bequest taxation (Kopczuk and Slemrod, 2001). Another example are labor supply responses to bequest taxation. If wealth shocks reduce individual labor supply, higher bequest taxes could prevent this.[6] Pure shifting responses might affect timing and volume of intended wealth transfers. These channels have been

[4] Most previous bunching studies focus on income taxation (see, e. g. Bastani and Selin, 2014; Chetty et al., 2011). Liu and Lockwood (2015) study value added taxes, while Best and Kleven (2016) investigate property transfer taxation.

[5] For this reason, I do not investigate bunching for predefined inheritances in this chapter.

[6] See Holtz-Eakin et al. (1993) for the US and Doorley and Pestel (2016) for Germany.

studied extensively, albeit with an almost exclusive focus on the US estate and gift tax. Previous cross-sectional studies find inter-vivo gifts to be very responsive to gift tax rates in the short run (Page, 2003), in particular for wealthy households (Bernheim et al., 2004; Joulfaian, 2005). This holds both for the decision between gifts and inheritances and for the timing of wealth transfers. Exploiting pre-announced increases in the taxation of gifts, Joulfaian (2004) reports substantial effects on gift tax revenues just before the reform. In Germany, an individual may receive tax-free inter-vivo gifts up to his exemption over the course of ten years. This makes inter-vivo gifts a promising tool for tax planning in the long run. For the US regime, which offers annual exclusion thresholds, Joulfaian and McGarry (2004) and Poterba (2001) find this tool to be underused from a tax-minimizing perspective. These results suggest that wealth-owners seek to retain some control over their assets, be it for wealth-loving or precautionary motives. Evidence for inheritance tax planning is more scarce. Most notably, Kopczuk (2007) finds evidence for tax planning if the deceased endured severe illness before death. In a similar vein, Kopczuk and Slemrod (2003) find date of deaths to be postponed to a date after an estate tax decrease, although this is likely a result from manipulating of the official time of death.

This chapter also relates to the theoretical literature on taxing inheritances versus inter-vivo gifts. Both types of transfers are usually taxed in a similar manner, but this may be questionable if both kinds of transfers are associated with distinct behavioral responses and hence different efficiency costs. Moreover, the optimal tax treatment depends on the nature of the bequest motive (Cremer and Pestieau, 2006; Kopczuk, 2010). Individuals may leave bequests for a variety of reasons. As a basic classification, bequests can be either accidental or intentional. If they are fully accidental, i. e. wealth is kept until death because of an intrinsic utility of wealth or due to precautionary reasons, efficiency costs of bequest taxation are low. In this case, the amount of bequests will not be affected by the presence or extent of bequest taxation. On the other hand, parents might plan their wealth transfers, either because they participate in their offspring's utility (altruistic motive) or because of strategic considerations. If these kind of motives are predominant, bequest taxation may have a sizable impact on the timing and amount of transferred wealth and tax-induced distortions can be sizable.

The findings can be summarized as follows. I find differential evidence for tax planning for inheritances and inter-vivo gifts. While inheritances are distributed rather smoothly around kink points, there is sharp bunching for inter-vivo gifts. This suggest that only a subset of conceivable tax planning channels are effectively used. The tax schedule hence seems to imply limited behavioral distortions, presumably due to the presence of optimization frictions. It is donors rather than receivers of bequests who seek to minimize their tax burden. Further heterogeneity analyses show that tax planning is most prominent for transfers to close relatives. Beyond, I find evidence that tax planning increases with the value of transferred wealth. Quantitatively, the overall welfare costs of bequest taxation are however estimated to be rather low; the elasticity of taxable bequests is estimated to be 0.02 at the highest.

The remainder of this chapter is structured as follows. Section 3.2 presents the institutional setting in which bequests and their taxation take place. Section 3.4 discusses the dimensions on which tax planning can be expected. Section 3.3 presents the data base and descriptive statistics. Section 3.5 delivers the empirical estimates, before section 3.6 concludes.

3.2 Institutional Background

Succession and division of estate In the event of death, an estate is in principle divided equally among the descendants of the deceased person. In case there are no descendants, the next in line of succession are parents of the deceased, including their offspring. If the deceased person was married, the spouse receives one quarter of the estate and the remaining amount is divided among the group of beneficiaries. These rules however only apply in case the deceased did not express a last will. Testators do not have full control over the distribution of their estate. The last will might determine different heirs or a different distribution of the estate among them. Alternatively, the last will might contain predefined inheritances to certain recipients (*Vermächtnis*), determining a specific asset or a specific amount of money. The minimum inheritance for descendants is half the amount they would have received in absence of a last will.[7] These restrictions obviously do not apply to inter-vivo gifts, as both the amount of the gift and the recipient can be freely determined by the donor.

Tax Treatment of wealth transfers As most EU countries, Germany imposes an *inheritance tax* that is levied on the recipients of a wealth transfer. In contrast to the *estate tax* applied in the US and the UK, the tax is levied on the transfer received and not on the total estate. This allows for granting exemptions depending on the personal characteristics of the taxpayer. A second dimension affects the treatment of inter-vivo gifts. While the US levies a federal gift tax distinct to the estate tax, both forms of transfers are treated equally in Germany. For this reason, the German system will here be referred to as a *bequest tax*, applying to wealth transfers in general.

The German Bequest Tax (*Erbschaft- und Schenkungsteuer*) is imposed on the recipient of an inheritance or a gift. All wealth types are in principle taxed, including cash, real estate, businesses and stock assets. Real estate can be transferred tax-free to spouses or children in case the property remains owner-occupied. There are substantial personal exemptions, depending on the relation between donor and recipient. Since 2009, these

[7] As an example, if there are five descendants (and no spouse), each receives 20% of the estate in absence of a last will. If the last will determines an alternative distribution or includes other beneficiaries, each descendant still receives at least the minimum share of 10%.

amount to €500k for spouses including same-sex marriages and to €400k for children and stepchildren. Lower thresholds apply for other relatives. In case of non-relatives and legal persons, the allowance is €20k. In case of inheritances, there is an additional personal deduction of €256k for spouses (*Versorgungsfreibetrag*), and an age-dependent allowance for children amounting up to €52k (see Table 3.4 for details). If someone receives a transfer from the same donor within a time span of 10 years, this transfer is additionally considered. This implies that bequests are tax-free only if the sum of transfers received from a specific person over the course of ten years is lower than the personal allowance.

Table 3.1: Tax Rates by year and taxable bequests

	Taxable bequests (1,000 €)	before 2009			Taxable bequests (1,000 €)	2009			after 2009		
		Tax Class				Tax Class					
		I	II	III		I	II	III	I	II	III
K_1	52	7	12	17	75	7	30	30	7	15	30
K_2	256	11	17	23	300	11	30	30	11	20	30
K_3	512	15	22	29	600	15	30	30	15	25	30
K_4	5,113	19	27	35	6,000	19	30	30	19	30	30
K_5	12,783	23	32	41	13,000	23	50	50	23	35	50
K_6	25,565	27	37	47	26,000	27	50	50	27	40	50
K_7	$\geq$ 25,565	30	40	50	$\geq$ 26,000	30	50	50	30	43	50

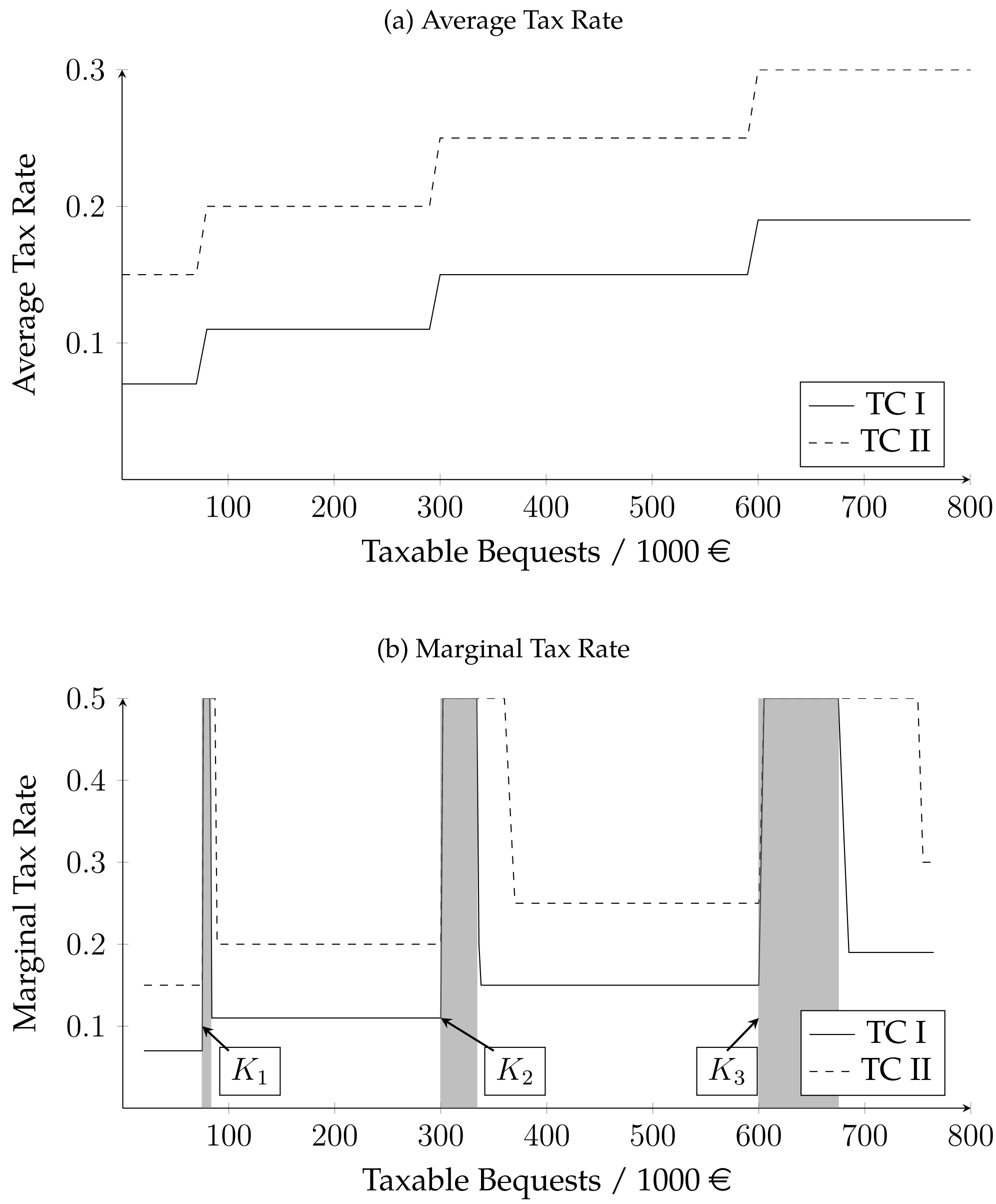
Tax Class I encompasses spouses, children, grandchildren and parents. Tax Class II encompasses siblings and their offspring, parents (for the case of inter-vivo gifts) and divorced spouses. For other recipients, including companies and foundations, Tax Class III is applied.

The tax liability is determined on the gross estate, net of liabilities and exemptions, and after adding previous transfers from the same donor. Table 3.1 shows the full tax schedule; Figure 3.1 visualizes the current schedule over the range of taxable bequests relevant for the empirical part. There are three different schedules (*tax classes (TC)*), depending on the relationship between donor and recipient. The lowest tax rates apply for Tax Class I, encompassing spouses, children, grandchildren and parents. Tax Class II affects siblings and their offspring, divorced spouses and parents (in case of inter-vivo gifts). In case of other recipients, Tax Class III is applied, which features the highest tax rates. The tariff consists of brackets in the average tax rate. For example, the tax rate that is applied to the total sum of taxable bequests flips from 7% to 11% for TC I if the taxable bequest b exceeds €75k. In absence of further rules, this would imply a discontinuity in the tax liability, i. e. a notch. The transition between two tax rates is in fact smoothened by an additional rule which effectively caps the marginal tax rate at 0.5 in the relevant range.[8] This is visualized in Figure 3.1b. The tariff induces discrete jumps in marginal tax rates (kink points). In the areas marked grey for TC1, it is particularly attractive to shift the value of total bequests towards the kink point.[9] My data base spans the years 2007 to 2011, encompassing two reforms. In 2009, tax rates were increased for tax classes II and III. Moreover, personal exemptions were substantially raised. The 2010 reform altered tax rates for tax classes II and III again.

[8] If the taxable bequest b lies between two kink points K_l and K_u (with respective tax rates τ_l and τ_u), the tax liability, depending on the tax class C, is obtained by $T^C(b) = \min(\tau_u^C b; \tau_l^C K_l + \tau_k(b - K_l))$; $\tau_k = \begin{cases} 0.5 & \text{if} \quad \tau_u^C < 0.3 \\ 0.75 & \text{if} \quad \tau_u^C \geq 0.3 \end{cases}$.

[9] As Glogowsky (2016) notes, this is essentially a setting with two kinks, a convex kink with $\Delta\tau > 0$, followed by a concave kink with a tax rate decrease ($\Delta\tau < 0$). Bunching at the second kink is however unlikely, because its value is not explicitly stated in the tax rules but has to be inferred.

Figure 3.1: The German Bequest Tax Schedule



The figure visualizes the tariff in place since 2010. Tax Class I encompasses spouses, children, grandchildren and parents. Tax Class II encompasses siblings and their offspring, parents (for the case of inter-vivo gifts) and divorced spouses. For other recipients, Tax Class III is applied with a constant ATR (MTR) of 30 % in the depicted range (not displayed). The full tariff, including higher amounts of taxable bequests, is given in Table 3.1. K_1, K_2 and K_3 indicate the kink points where bunching is going to be investigated. For Tax Class I, the grey areas indicate the regions where manipulating the sum of taxable bequests is particularly attractive due to high marginal tax rates.

3.3 Data

My empirical analysis is based on annual German administrative bequest tax return data (*Erbschaft- und Schenkungsteuerstatistik*), spanning the years 2007 to 2011. They cover the universe of bequests and gifts for which a tax claim was requested. This includes also bequests that were eventually not taxed. Due to high exemption rates, the majority of wealth transfers does not show up in the bequest tax data. For 2010, Bach et al. (2014) estimate that the tax data cover 30% of all bequests, accounting for 73% of transferred wealth in total.[10] The data years refer to the first assessment of the tax liability. They might deviate from the actual gift or death event by several years because tax authorities do not approach heirs before a couple of months after the death event. It may also take a long time until tax authorities learn about a taxable event. This holds particularly for inter-vivo gifts. Larger deviations may be caused by disputes among heirs or in case of foreign-based assets. The latter occur more often for valuable bequests entailing a multitude of assets. Lags between the taxable event and the actual tax assessment indeed increase with the total sum of bequests. Nevertheless, around 90% of tax cases get assessed within the first two years (Schinke, 2012). Independent of the time lag, the wealth transfer is always subject to the tax regime in force at the time of the taxable event. The existence of the time lag makes the sample less representative for more recent years. This sample selection does not pose a problem to the identification of the elasticity of taxable bequests if the propensity for tax planning is uncorrelated with the time lag.

Table 3.5 indicates that the majority of bequests come in the form of inheritances. As argued in section 3.4, not all types of inheritances are equally suited for tax planning. Inheritances where the heir receives exclusively predefined inheritances account for 15% of all inheritances. Adding 'standard' inheritances that accrue to one heir only yields that around one quarter of all inheritances are particularly suspect of tax planning. The Panel B of Table 3.5 reveals a three-way split of tax cases regarding the type of recipient. One third remains in the inner family, another third goes to other relatives, while the last third of bequests is received by non-family members. Bequests differ considerably in value by relationship, the value of gross bequests diminishes for non-family recipients. Gross inheritances greatly exceed gross gifts in value for most groups of recipients. After subtracting exemptions and accounting for prior transfers from the same person, taxable taxable gifts are worth around €300k on average, compared to only €160k for taxable inheritances.

[10] This implies an annual sum of inheritances and gifts of €62 bn, based on survey data (GSOEP). Estimates based on national accounts yield total annual volumes in the range of €200 bn. According to Houben and Maiterth (2013), such macro-based approaches are however less suited for estimating the bequest tax base. On the other hand, survey-based estimates might suffer from under-reporting. There is hence considerable uncertainty regarding the overall volume of annual bequests.

The bottom panel of Table 3.5 finally shows that gifts and inheritances differ not only in value, but also in asset type. Inheritances consist, in average, to 68% of financial assets and to 28% of real estate. Gifts, in contrast, consist only to 12% of financial assets. Real Estate accounts, in average, for 52% of a gifts' value.

Table 3.2: Personal characteristics of donors and recipients

		Donor		*Recipient*	
Type of Transfer		Men	Women	Men	Women
Inheritance	Shares	0.399	0.601	0.438	0.562
	Mean Age	67.7	71.6	57.0	60.4
Gifts	Shares	0.509	0.491	0.539	0.461
	Mean Age	74.6	62.7	45.3	47.5

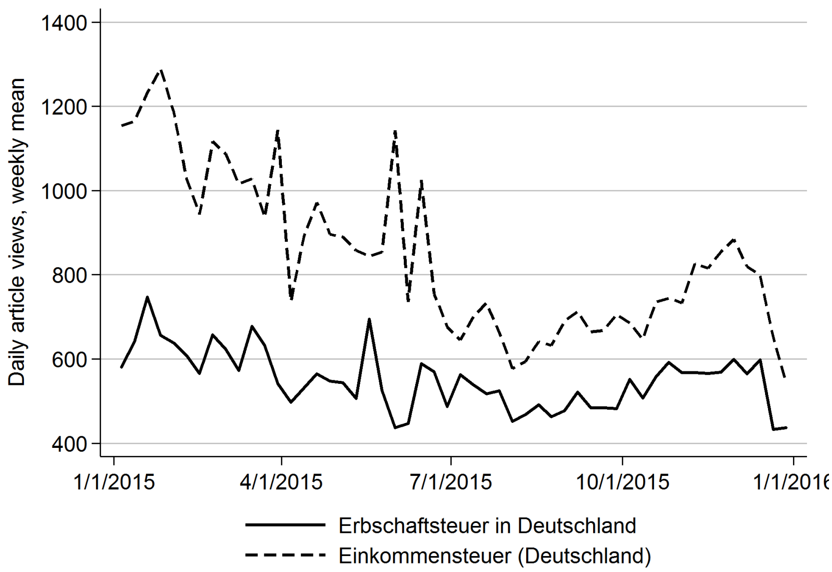
Source: Administrative bequest tax return data from 2007 to 2011.

Personal information on taxpayers is limited as usual in administrative data. In the present case, these are restricted to age and sex of both donor and recipient of the bequest (Table 3.2). Gift recipients are around 12 years younger than heirs. Nonetheless, they are already in their forties on average. As noted by Kaplow (2001), receiving gifts at an earlier stage in life, e. g. while being in education, might be more efficient as this could ease liquidity constraints for the recipients. Interestingly, women are around 12 years younger than men when leaving gifts, which could reflect a stronger altruistic motive. Finally, the majority of inheritances is left by women, which could simply be an artifact of higher female life expectancy.

3.4 Salience and Scope for Tax Planning

In order to detect bunching of taxpayers, two prerequisites need to be met. First, taxpayers need to be aware of the incentives provided by the tax tariff. External tax advisors are usually employed when large fortunes are transferred. As consultant fees can be deducted from the tax due, this information is available in the data. Tax consultant payments above the standard deduction are reported for 95% of tax cases. A further point illustrating the salience of bequest taxes is made in Figure 3.2. It compares weekly means of daily views on the German Wikipedia for the main articles on bequest and income tax in Germany.[11] Both articles provide extensive information on the respective issue and exclusively treat the German rules.[12] Information on the bequest tax is gathered around two thirds as often as on the income tax. This is in contrast to the enormous relative importance of the income tax in terms of the total number of annual taxpayers (30 million versus 200,000). Against this backdrop, it seems plausible that taxpayers (or their advisors) are sufficiently aware of the incentives to reduce their bequest tax due.

Figure 3.2: Daily Wikipedia Article Views



The graph plots daily article views at the German Wikipedia for the year 2015, comparing the main articles on bequest tax (solid line) and income tax (dashed line) in Germany. Intra-week fluctuations are eliminated by showing weekly mean values. Page view counts are obtained from https://dumps.wikimedia.org/other/pagecounts-raw.

As a second prerequisite, taxpayers need to be able to manipulate the sum of taxable bequests. In Germany, tax authorities learn about events of death directly from the reg-

[11] See Hoopes et al. (2015) for use of Wikipedia article view counts in the context of tax rule salience.

[12] This is opposed to corresponding regulations in Switzerland or Austria, for which the German Wikipedia would also be a preferred source of information.

istry. Moreover, banks, asset managers, insurance companies, notaries etc. are obliged to inform tax authorities about transfers of bank accounts, real estate or businesses. Infact, most inter-vivo gifts are indicated by notaries and hardly by taxpayers (Reis, 2005). If the tax authorities expect assets to be in a tax-relevant range, they request recipients to claim the sum of their assets. The presence of substantial third-party reporting renders full escape of bequest taxes unlikely (Kleven et al., 2011), particularly if transferred assets are home-based. There is however a considerable degree of freedom regarding the intensive margin.

Strategies to manipulate the tax base on the intensive margin are quite distinct between inter-vivo gifts and inheritances. Inter-vivo gifts clearly constitute a deliberate choice by the donor. When choosing the amount of the gift, he can target the total sum to one of the kink points in order to minimize the tax burden for the recipient. This is conceivable if the donor would have optimally chosen an amount slightly above a kink point (Nordblom and Ohlsson, 2006). Altering the taxable amount by the recipient is also possible, e. g. by misreporting the received sum to the authorities. While this chapter cannot disentangle both possible mechanisms, ex-post manipulations (misreporting, re-evaluation) however seem quite costly as some of them can be clearly classified as illegal evasion.

For inheritances, the case is somewhat different. If the deceased person did not leave a last will, there is by definition no tax planning on his behalf. Even if there is a last will, the testator would need to have a precise estimate of the total sum of assets his offspring will receive. The fact that personal exemptions might differ among recipients complicates tax planning by the testator. A third option is a predefined inheritance, where certain assets or a specific amount of money is dedicated to a specific person. This setting is highly comparable to giving an inter-vivo gift. Tax planning by the donor is hence rather unlikely for most cases of inheritances, but might be rather done by the recipient. A numerical example illustrates this: A bequest b worth €493k to a child of 18 years is, after subtracting the personal allowance of €400k and the additional age-dependent (rounded) allowance of €10k, associated with a taxable amount of €83k. Applying the tax rate of 11% (tax class I), the tax liability is €9,130. Reducing the taxable sum to €75k lowers the average tax rate to 7% and the tax due to €5,250, thus saving €3,880 in bequest taxes. The necessary reduction of €8,000 is tiny compared to the total bequest b (1.6%). It is hence sufficient to reduce the tax base by a few thousand Euros if b lies in (or slightly above) one of the grey areas in Figure 3.1. There are several ways to achieve this, e. g. by not claiming some household items or by re-evaluating real estate.[13]

[13] The value of transferred assets is subject to evaluation. Since 2009, most assets are to be valued by their market value. Before, the so-called property value was used for real estates. This concept is primarily used for the assessment of the property tax and is supposed to reflect the market value of 1964. Although the concept of market value is more appropriate in terms of the utility an individual accrues from the additional wealth, it leaves considerable scope for underreporting the sum of transferred assets.

It is important to note that the present approach measures short-term responses to bequest taxation, occurring around the time of the actual transfer. As discussed above, bunching for gifts presumably reflects tax planning by the donor, while for inheritance taxes, tax planning is conceivable by both. Concerning longer-term strategies, wealthy persons can exploit the higher personal exemptions for own children by adoption. Personal exemptions can further be used several times by giving away the wealth piecewise every 10 years.[14] When transferring property, the donor can reduce the tax liability by claiming the further right to occupy the premises for himself. In this case, the future value of this usage is subtracted from the taxed transfer. While these channels are potentially of high relevance, the time coverage of the data used in this chapter is too short to fully capture these kinds of responses. Nonetheless, the observed giving behavior might be partly motivated by exploiting the 10-year threshold.

3.5 Quantifying the amount of tax planning

3.5.1 Empirical Approach

In order to gauge the extent of tax planning, I quantify the amount of excess bunching of taxpayers. I apply the widely-used methodology of Chetty et al. (2011) who analyze bunching at kink points in the Danish income tax schedule. They assume a consumption-leisure trade-off, where only a share of households adjust their labor supply to the tax schedule. This could be due to heterogeneity in consumption-leisure preferences. Alternatively, it might be too costly to relocate on the tax schedule due to frictions, such as hours constraints or search costs. Carrying this idea to bequest taxation requires a different framework when thinking about expected behavioral responses. An inter-generational setting that features the decision between gifts and inheritances is provided by Nordblom and Ohlsson (2006). The introduction of a bequest tax induces distortions between both types of transfers and might lead to bunching at the kink point in order to minimize tax payments. Real responses to bequest taxation, e. g. on labor supply or saving behavior, are well conceivable, but they would not show up in the tax records. Shifting responses of the bequest tax base are hence the only channels that can be directly observed. This setting is comparable to the bunching evidence for house transaction taxes by Best and Kleven (2016). They find strong evidence for bunching of house prices at notches in the tax schedule. The authors explain this response with a low degree of frictions in the market due to the high number of professional real estate agents.

[14] This behavior has been documented by McGarry (2001) for the US.

According to Saez (2010), the local elasticity of taxable bequests e_B at a kink point K can be approximated by

$$e_B \simeq \frac{\widehat{b}}{K \ln\left(\frac{1-\tau_1}{1-\tau_2}\right)}, \tag{3.1}$$

where b denotes the excess mass, i. e. the extent to which tax payers cluster at the kink point relative to a hypothetical situation in absence of the kink. τ_1 and τ_2 denote tax rates below and above the kink. Intuitively, Equation 3.1 relates the amount of bunching to the size of the kink. In order to estimate $\widehat{b}$, taxable income is first grouped into equally sized bins, indicated by j. Then, a flexible polynomial function of the 7^{th} degree is fitted on the density of tax cases excluding R bins to the left and to the right of the kink point.[15]

$$C_j = \sum_{i=0}^{7} \beta_i (Z_j)^i + \sum_{i=-R}^{R} \gamma_i \cdot \mathbf{1}[j = i] + \varepsilon_j \tag{3.2}$$

As bunching is found to be very sharp under the kink point (Figure 3.3), R is set to 2, except for the first kink point ($R = 4$). Z_j denotes the number of tax cases with taxable assets falling in bin j relative to the kink point of interest. In order to provide estimates with the highest precision, small bin widths are desirable. The baseline results hence rely on a bin width of €100. This delivers the highest precision possible, as taxable amounts are in practice rounded down to the next multiplier of 100. For the more detailed analyses that build on smaller samples, the bin width is increased to €500 to secure a continuous density. The counterfactual density values $\widehat{C}_j$ are obtained from linear prediction of the coefficients. Finally, $\widehat{b}$ is calculated from the cumulated difference between both densities in the area of interest, normalized by the mean of the counterfactual values in that range.

$$\widehat{b} = \frac{\sum_{j=-R}^{R} C_j - \widehat{C}_j}{\overline{\widehat{C}_j}} \tag{3.3}$$

Plugging $\widehat{b}$ into Equation 3.1 delivers the estimate for the local elasticity. A bootstrap procedure yields standard errors for $\widehat{b}$.

3.5.2 Bunching Estimates

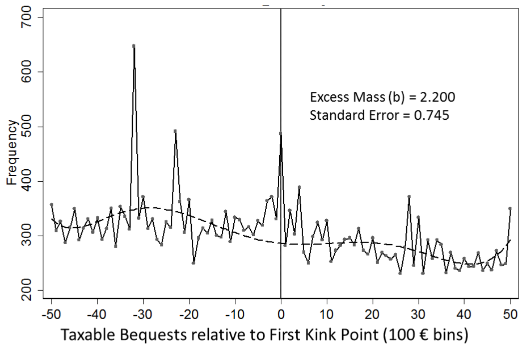
I investigate excess bunching at the first three kink points in the tax schedule.[16] Beyond the third kink, i. e., for taxable amounts of €6 millions and higher, the number of

[15] As in Best and Kleven (2016) and Chetty et al. (2011), varying the polynomial degree has a negligible effect on the estimates.

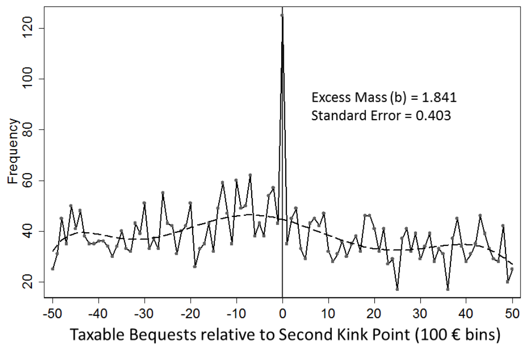
[16] As the tax return data also include non-taxed cases, one could also investigate bunching at the exemption level for gross bequests (the '$0th$ kink'). At the time of writing, these results were not made available by the tax authorities. Preliminary evidence reveals bunching at the exemption in magnitude similar to what will be shown for the first three kinks.

Figure 3.3: Bunching Graphs — Full Sample

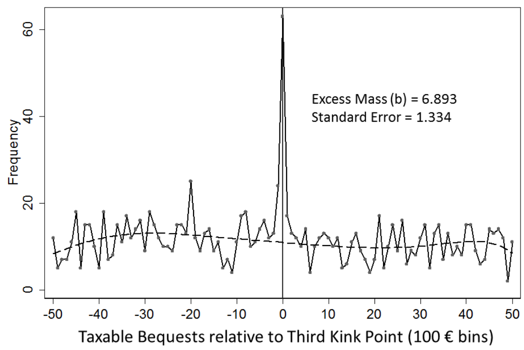
(a) First Kink Point K_1



(b) Second Kink Point K_2



(c) Third Kink Point K_3

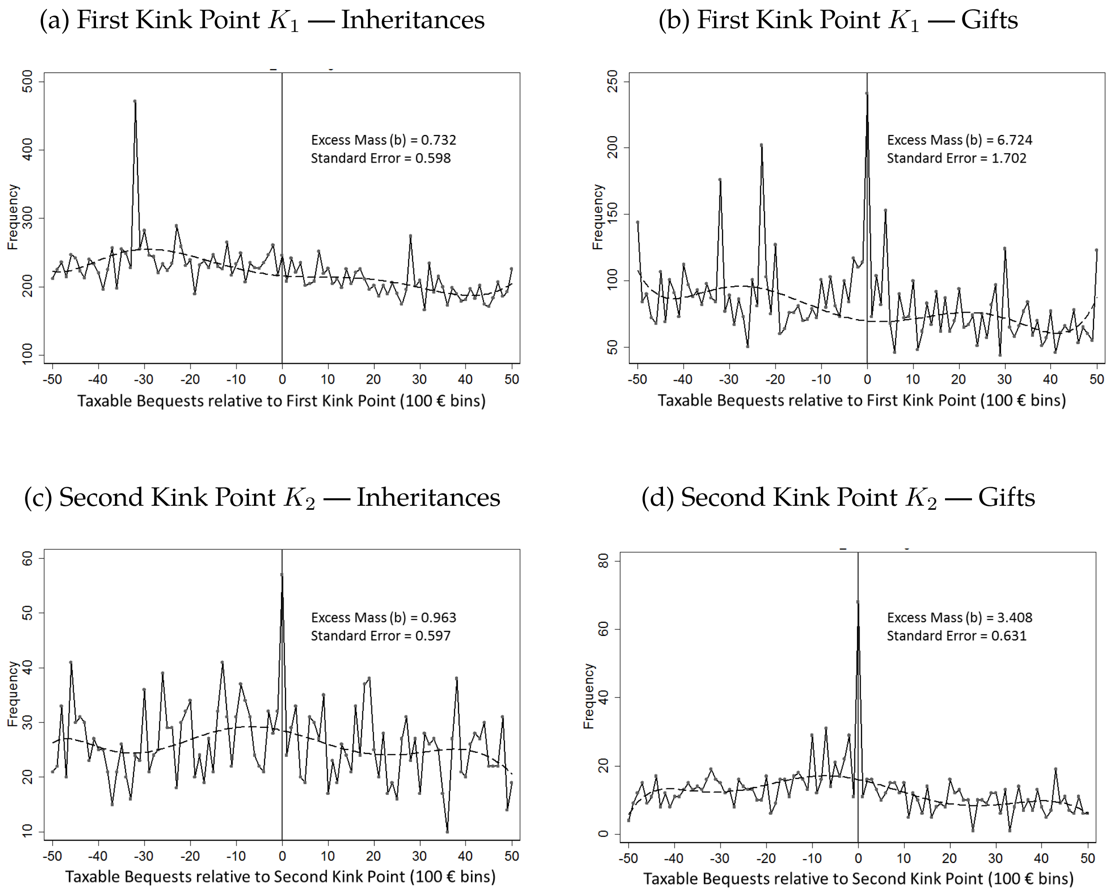


Note: Each dot represents the number of tax cases in a given €100 bin. The dashed line represents the counterfactual distribution of tax cases, based on a polynomial fit of the 7^{th} degree. b is estimated according to Equation 3.3. Graphs are created with the Stata program `bunch_count`, as used in Chetty et al. (2011). K_1 amounts to €52k before 2009 and €72k afterwards. K_2 equals €256k before 2009 and €300k afterwards. K_3 equals €512k before 2009 and €600k afterwards.

observations becomes too small. Figure 3.3 shows density plots for these kink points, underlying a pooled sample from all five years, restricted to the taxpayers who can potentially bunch. This encompasses all wealth transfers with a tax schedule featuring kinks, i. e. all transfers from tax class I, tax class II (not in 2009), and tax class III before 2009 (see Table 3.1). The horizontal axes show the difference to the respective kink point. The tax schedule, and thereby the kink points, were shifted to the right in 2009. There is visual evidence for bunching at the kink in all graphs. Each of the excess masses is estimated to be significantly different from zero. It is highest for K_3, located at €512k or €600k. This is particularly remarkable in light of less than 20 observations in each bin which is far below what related studies rely on.

The two spikes left to the kink point in Figure 3.3a stem from transfers amounting to round figures. The first spike represents transfers of exactly €100,000 to children. Subtracting the pre-2009 exemption of €51,200 results in taxable gifts of €48,800, representing the 32^{nd} bin below K_1.[17] The second spike at the 23^{rd} bin left to the kink represents transfers to grandchildren worth exactly €60,000. The tendency to transfer round-number amounts suggests the existence of reference points, induced by non-financial incentives (Kleven, 2016).

Figure 3.4: Bunching Graphs — Inheritances versus Gifts



Note: Each dot represents the number of tax cases in a given €100 bin. The dashed line represents the counterfactual distribution of tax cases, based on a polynomial fit of the 7^{th} degree. b is estimated according to Equation 3.3. Graphs are created with the Stata program `bunch_count`, as used in Chetty et al. (2011). K_1 amounts to €52k before 2009 and €72k afterwards. K_2 equals €256k before 2009 and €300k afterwards. Bunching graphs for K_3 could not be produced due to an insufficient number of observations.

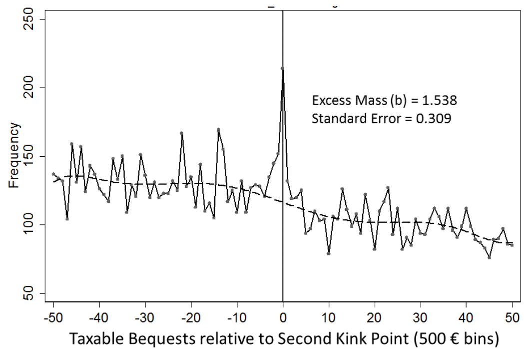
In a next step, the sample is split by considering inheritances and gifts separately. While both types of transfers are treated equally by the tax schedule, tax planning behavior can be expected to take different forms for both types of transfers. Figure 3.4 shows bunching estimates by type of transfer for the first two kink points.[18] For inheritances, no bunching can be observed at the first kink, while there is some visual, albeit not statistically significant, evidence for bunching at the second kink point. In

[17] After 2009, the exemption amounts to €400,000, rendering such transfers to children tax-free.
[18] Around the third kink point, the number of tax cases is not sufficient to form a smooth distribution.

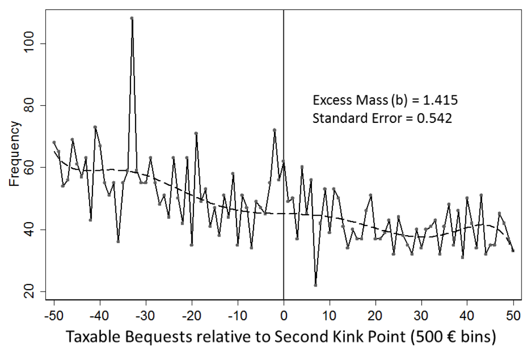
contrast, gifts display sharp and significant bunching at the first two kink points. This is supportive of the notion that tax planning for gifts is in principle easier.

Figure 3.5: Bunching graphs — by family relationship (2^{nd} kink)

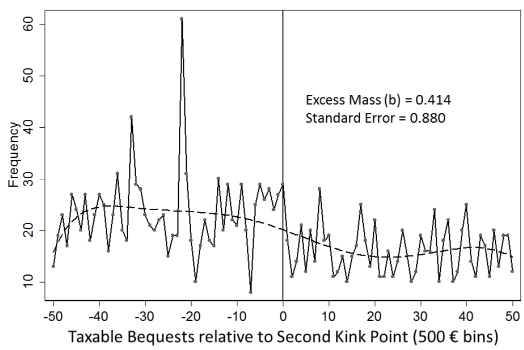
(a) Spouse & Children



(b) Other relatives



(c) Non-relatives



Notes: The graph shows bunching estimates by tax class for the second kink point K_2. Each dot represents the number of tax cases in a given €500 bin. The dashed line represents the counterfactual distribution of tax cases, based on a polynomial fit of the 7^{th} degree. b is estimated according to Equation 3.3. Graphs are created with the Stata program bunch_count, as used in Chetty et al. (2011). The second kink point K_2 equals €256k before 2009 and €300k afterwards. The density spike left to the kink in Figures 3.5b and 3.5c represents transfers worth exactly €250k.

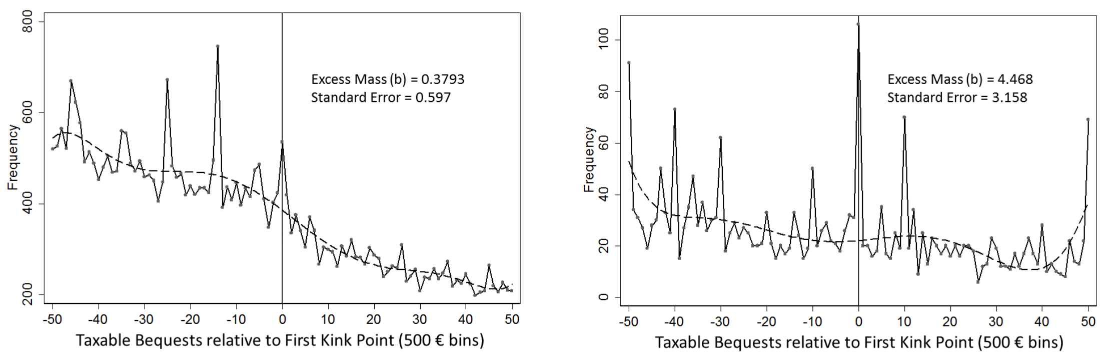
As a further decomposition, I consider differential responses by tax class, reflecting different relationships between donor and recipient (Figure 3.5).[19] For transfers to spouses, children and grandchildren, bunching is estimated to be higher than for more distant relatives. No bunching is detected for non-relatives. There are two possible explanations for this finding. On the one hand, stronger family ties could facilitate collusion between donor and recipient in order to target the transferred sum. Similarly, donors might care more about the tax due of close family members which increases the motivation to engage in tax planning. On the other hand, tax exemptions are substantially higher for family members, implying that the gross bequests to family members

[19] Figure 3.5 shows results for the second kink. Results for the first kink are shown in Figure 3.7.

are higher for comparable values of taxable bequests. The higher amount of transferred wealth then might raise the probability of tax planning.

Figure 3.6: Bunching graphs — gifts only, before and after 2009 reform (1^{st} kink)

(a) until 2009 (b) 2009 and later



Notes: Each dot represents the number of tax cases in a given €500 bin. The dashed line represents the counterfactual distribution of tax cases, based on a polynomial fit of the 7^{th} degree. b is estimated according to Equation 3.3. Graphs are created with the Stata program `bunch_count`, as used in Chetty et al. (2011). The 2009 reform implied higher tax rates for tax class II, affecting relatives other than spouses or children. Beyond, personal exemptions were raised substantially. For spouses, the personal exemption changed from €307k to €500k. For children, it increased from €51.2k to €400k. See Table 3.4 for details.

To illuminate this channel further, Figure 3.6 contrasts densities at the first kink for inter-vivo gifts before and after the 2009 reform. The reform increased tax rates for tax class II, while maintaining the tax rate difference of 5 percentage points. At the same time, personal exemptions were raised by €193k for spouses and by around €350k for children. This substantially raises the gross value of taxed transfers for the years 2009 and later. Bunching is found to be substantially higher after the reform, which is in line with a higher awareness of post-reform taxed transfers.

So far, the presentation was restricted to visual bunching evidence. In order to gauge welfare costs associated with these estimates, Table 3.3 shows the respective elasticities. They are based on Equation 3.1, assuming that taxpayers perceive a kink. As discussed above, the tax schedule however features notches at first sight. The additional rule of capping marginal tax rates receives little attention on lawyers' websites that provide advice on bequest tax avoidance. Beyond, the precise extent of the areas with high marginal tax rates is not explicitly stated in the tax code. The bunching I observe could therefore be partially a response to a perceived notch. This would imply far higher tax base elasticities, as the respective change in tax rates is lower and hence imply distinct quantitative welfare implications.

Even the sharp significant bunching estimates for inter-vivo gifts turn out to translate into rather low elasticities. The highest estimate (within the sample containing all years) is obtained at the first kink point for gifts, with a precisely estimated elasticity of

0.021. This means that the amount of bunching is small relative to the monetary gain. One could question the *economic* significance of an elasticity of 0.02 maximum. Such small effects might however become relevant for the recipient if one takes long-term wealth accumulation into account. A standard explanation for the (partial) absence of behavioral reactions is the presence of substantial optimization frictions (Chetty et al., 2011). These can take various forms. An obvious one are legal hurdles. While tax avoidance by targeting the sum of taxable bequests ex ante is obviously legal, ex post adjustments might be considered illegal.[20] It might also point to higher (pecuniary and non-pecuniary) costs of tax planning ex post versus ex ante. In light of the heavy involvement of professional tax advisors, informational frictions seem to play less of a role in the present context. The visual evidence in Figures 3.3 to 3.6 reveals the presence of bunching not only at kink points, but also at certain round numbers. The amount of bunching is in magnitude comparable to the bunching at kink points. This implies the presence of reference points other than those implied by the tax code which are of similar significance to people. The fact that a sizable number of wealth transfers amounts to round numbers suggests psychological focal points. Targeting these focal points by the tax code could be addressed by future reforms (Kleven, 2016).

The fact that gifts are more responsive than inheritances is in line with the intentional character of gifts. From a welfare perspective, a higher elasticity of gifts would suggest a preferential tax treatment of inter-vivo gifts (Piketty and Saez, 2013a). Increasing the personal allowance for inter-vivo gifts relative to inheritances might hence induce people to make gifts earlier. Recipients, in turn, might then receive gifts at a life stage with liquidity constraints (Cremer and Pestieau, 2006; Kaplow, 2001). An equivalent effect can be expected from reducing the exclusion threshold of currently 10 years in which transfers are considered for taxation.

[20] For the common bequest tax payer, the probability of being audited can be neglected. Only in case of transferred businesses, auditors might verify the prerequisites for favorable tax treatment.

Table 3.3: Local elasticities of taxable bequests

	All Years pooled			before 2009			after 2009	
	All	Close relatives	Further relatives	All	Close relatives	Further relatives	All	Close relatives
Full Sample								
K_1	0.007***	0.002	0.011	0.001	-0.002	0.013	0.009	0.048*
K_2	0.001***	0.005***	0.005*	0.003***	0.003***	0.004	0.012***	0.026***
K_3	0.003***	0.003***	0.004**					
Gifts								
K_1	0.021***	-0.006	0.039	0.006	-0.006	0.040	0.048	
K_2	0.002***			0.007***				
K_3	0.004***							
Inheritances								
K_1	0.002	0.005	0.005	-0.001	0.001	0.007	0.005	
K_2	0.001**	0.003**	0.001**	0.001	0.001	0.003	0.006***	
K_3	0.002**	0.003**	0.003					
Inheritances, Single Heirs								
K_1	0.001	-0.003	0.001	0.001	-0.009	0.002	0.004	
K_2	0.000	0.002	0.001	-0.001	0.000	0.000	0.005**	

Notes: The table shows elasticity estimates according to Equation 3.1, based on the estimated excess mass $\hat{b}$, the tax rates at the kink point (τ_1, τ_2) and the amount of taxable bequests at the kink K. The pooled estimations may correspond to different levels of tax rates and kinks due to the 2009 reform and to different tax classes. When pooling several tax classes, the lowest tax rates (for tax class I) are applied. When pooling several years, the tax rates and kinks for the pre-2009 schedule are applied. Empty cells denote cases where the taxpayer density was not sufficient to measure bunching with a bin width of €500. Significance levels correspond to those from the bootstrapped standard errors of $\widehat{b}$ in the excess mass estimation: (***) = 0.01, (**) = 0.05, (*) = 0.1.

3.6 Conclusion

Parallel to the rise of bequests as a source of income, interest in taxation of wealth transfers is likely to thrive in the future. A heavier taxation of bequests potentially cushions rising wealth inequality, but little is known about distortionary effects on the decision on whether and how much to leave. This chapter estimates shifting responses of taxable bequests in Germany. Relying on administrative data, I make use of the tax schedule featuring discrete jumps in marginal tax rates. Building on the empirical literature on bunching at kink points, I find sharp and significant bunching for taxable bequests above the basic allowance. These are confined to the subsample of inter-vivo gifts, while the distribution of inheritances is rather smooth around kink points. Gift bunching is sharp and significant even for high amounts, where the number of observations become small. These findings are in line with previous cross-sectional evidence on the responsiveness of gift behavior to taxation (Bernheim et al., 2004; Joulfaian, 2004, 2005; Page, 2003), albeit with rather low elasticities.

My findings shed light on tax planning in the context of wealth transfers. Tax planning for gifts presumably takes place on behalf of the donor. The fact that gift amounts are affected by tax incentives is in line with the intentional character of inter-vivo gifts. Inheritances, on the other hand, are rather of an accidental nature, which usually prevents tax planning on behalf of the donor. I hence do not find evidence for deathbed planning as in Kopczuk (2007). Beyond, there is hardly evidence for tax planning by the recipient. In light of high salience of bequest tax rules, this suggests the presence of optimization frictions, presumably in the form of legal hurdles and adjustments costs. The overall low level of responsiveness is surprising given that professionals are involved in most transactions. A side result is the presence of focal points not induced by the tax code for gifts and predefined inheritances. Their importance is comparable to the bunching at kink points in the tax schedule.

The main message of the chapter is that overall efficiency costs of bequest taxation are rather low. Regarding the on-going discussion on raising bequest taxes, this finding provides no reason to abstain from higher taxes for concerns of generating additional revenue. From the higher responsiveness of inter-vivo gifts, the case for a preferential tax treatment of gifts could be derived, possibly in the form of higher exemptions. This could raise the volume of deliberate giving or even advance them to an earlier point in time. Whether this is desirable from a societies' perspective however depends on the welfare weights attributed to younger and older wealth owners.

In order to investigate the issue in more detail, data covering a longer time span would be required. It is particularly promising to investigate the extent to which the ten-year exclusion threshold is exploited.

3.A Appendix

Table 3.4: Personal Exemptions by relationship between donor and recipient

	before 2009	since 2009
	in €	
Spouses[a]	307,000	500,000
Children, Stepchildren, Children of deceased (step-)children[b]	51,200	400,000
Children of living (step-)children	51,200	200,000
Parents (inheritances only)	51,200	100,000
Parents (gifts only), Siblings, Nieces, Nephews, Step-parents, divorced spouses, children and parents-in-law	10,300	20,000
Other (Non-relative, firm, trust)	5,200	20,000

[a] For inheritances, spouses are granted an additional exemption of €256,000.
[b] For inheritances, children are granted an additional age-dependent exemption. It amounts to €52,000 for children below 5 years, €41,000 for children between 5 and 10 years, €30,700 for children between 11 and 15 years, €20,500 for children between 16 and 20 years, and €10,300 for children between 21 and 27 years.

Table 3.5: Descriptive Statistics

Panel A: Observations by year and type of bequest

	2007	2008	2009	2010	2011
All observations	214,232	264,332	247,109	202,154	210,899
Inter-vivo gifts	59,830	78,681	71,337	53,058	54,755
Inheritances	154,402	185,651	175,772	149,096	156,144
thereof (shares):					
Predefined inheritance only	20.0%	20.1%	13.6%	11.4%	10.9%
Inheritances with one heir[a]	23.4%	24.0%	23.7%	25.6%	26.7%

Panel B: Value of transfers by family relationship[b]

Relationship	Observations	Inheritances: gross bequest[c]	Inheritances: taxable transfer[d]	Gifts: gross bequest	Gifts: taxable transfer
		mean value in €			
Spouses	42,631	973,324	415,830	417,007	492,806
Children	276,568	756,858	511,056	477,446	474,730
Grandchildren	32,266	446,393	302,274	274,940	285,999
Parent	25,082	158,665	91,997	38,240	24,813
Other relatives	388,144	99,505	79,739	200,795	87,510
Non-relatives	372,026	93,731	73,575	83,925	91,233
Total	*1,138,917*	*326,299*	*159,021*	*316,590*	*297,504*

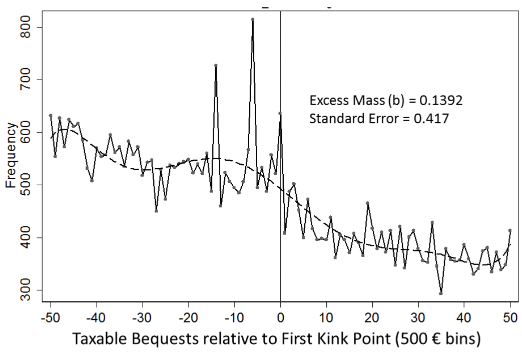
Panel C: Asset composition

	Land	Financial Assets	Real Estate	Business	Other
	Means of asset shares[e]				
Inheritances	0.7%	68.0%	28.2%	1.8%	1.3%
Gifts	3.1%	12.4%	51.8%	16.0%	16.7%

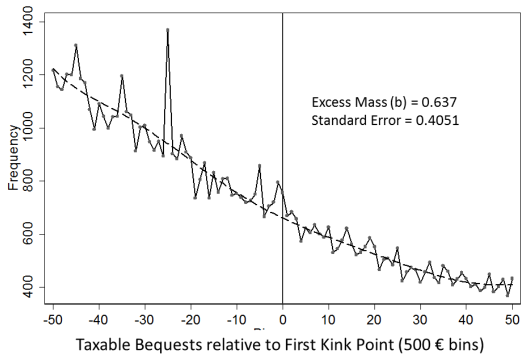
Source: Inheritance and Gift Tax Returns from 2007 to 2011.
[a] including predefined inheritances as a special case.
[b] only bequests with overall positive value (transferred wealth > transferred debts).
[c] Gross bequests equal the share of the estate accruing to the recipient, possibly after division among all heirs.
[d] Taxable transfer = gross bequest – exemptions + prior transfers within 10 years.
[e] Average share of gross bequest

Figure 3.7: Bunching graphs — by family relationship (1^{st} kink)

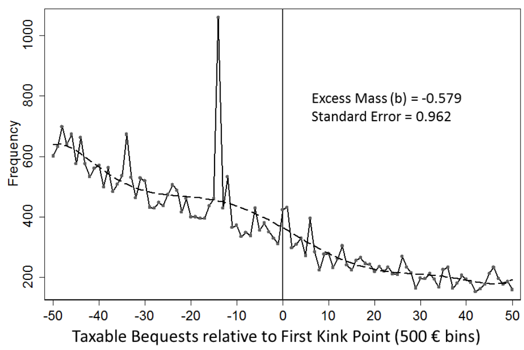
(a) Spouse & Children



(b) Other relatives



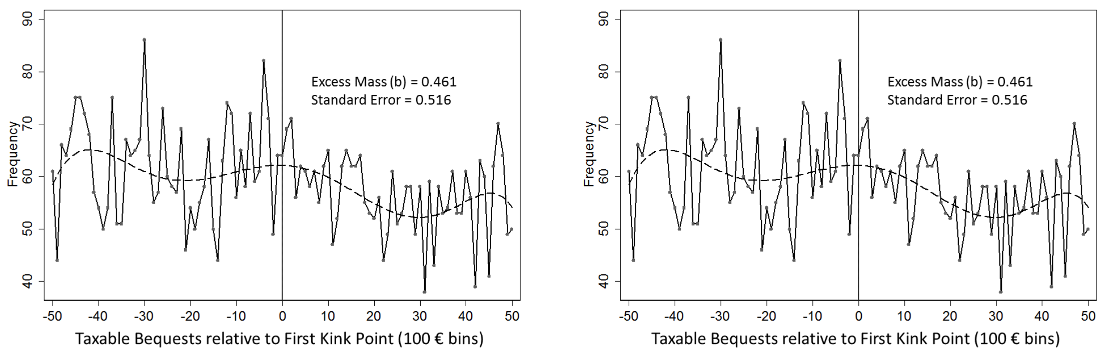
(c) Non-relatives



Notes: The Graph shows bunching estimates by tax class around the first kink point K_1, equalling €52k before 2009 and €72k afterwards. Each dot represents the number of tax cases in a given €500 bin. The dashed line represents the counterfactual distribution of tax cases, based on a polynomial fit of the 7^{th} degree. b is estimated according to Equation 3.3. Graphs are created with the Stata program `bunch_count`, as used in Chetty et al. (2011).

Figure 3.8: Bunching graphs — inheritances, single heirs only

(a) First Kink (b) Second Kink



Notes: Each dot represents the number of tax cases in a given €100 bin. The dashed line represents the counterfactual distribution of tax cases, based on a polynomial fit of the 7^{th} degree. b is estimated according to Equation 3.3. Graphs are created with the Stata program `bunch_count`, as used in Chetty et al. (2011).

Fiscal Sustainability and the Demographic Transition in the EU* 4

4.1 Introduction

Ongoing long-term demographic changes are widely considered a risk to fiscal sustainability in developed countries. A shrinking labor force, combined with a growing old-age dependency ratio, is expected to negatively affect tax revenues and raise pension expenditures. This may threaten governments' capacities to fund social welfare systems and the provision of other public goods. As a consequence, pension systems in virtually all industrialized countries have been subject to recent reforms (OECD, 2013a). While the expectations of growing pension expenditures have been supported by a number of studies, the case is less clear-cut for the evolution of fiscal revenues. The inter-linkages between demographic transitions and labor market outcomes deserve special attention in this context. If, for example, a shrinking labor force is becoming better educated at the same time (as is projected), average wages will increase. Additionally, if there is a scarcity of labor, neoclassical economic theory predicts that wages should increase in order to stimulate labor supply. Future tax revenues may therefore increase despite population shrinkage. Hence, it is crucial to account for reactions on both sides of the labor market when assessing the effects of demographic changes on future fiscal balances. Most studies however do not systematically account for labor supply and demand responses. We study fiscal sustainability in the EU, combining population projections for 2030 with micro-based elasticities of labor supply and demand, allowing us to overcome this limitation.

Specifically, this chapter outlines the extent of the challenges for public budgets from demographic changes in a four-step analysis. First, we incorporate two scenarios of projected demographic changes via a reweighting procedure into micro data sets for the EU-27 countries. In a second step, the implied wage effects are analyzed by modeling the demand and supply side of the labor market. Supply elasticities are differentiated by skill, gender and household type for each EU-27 country. On the demand side, we

* This chapter, co-authored with Mathias Dolls, Karina Doorley, Alari Paulus, Hilmar Schneider and Sebastian Siegloch, is also published as: M. Dolls et al. (2015). Fiscal Sustainability and Demographic Change: A Micro Approach for 27 EU countries. IZA Discussion Paper No. 9618.

differentiate own-wage elasticities of demand by country and skill group, drawing on a meta-analysis approach. Next, the consequences for fiscal budgets are investigated with a tax-benefit simulation. We capture personal taxes, social insurance contributions, social transfers, public pensions, and main demography-related public expenditures. Finally, we analyze the impact of an increase in the statutory retirement age, which is an obvious and widely discussed policy response to demographic change.

Our approach is micro-driven and accounts for the full heterogeneity in populations and tax-benefit rules, required to model essential interactions between demographics, labor market behavior and fiscal systems. Unlike computable general equilibrium (CGE) approaches, the only assumptions we impose concern the elasticities of labor supply and demand or stem from the demographic projections.

Our findings contribute to a broad academic debate on the consequences of demographic change. The impact of demographic ageing and decreasing population size on long-term economic growth has been treated in a number of endogenous growth models (Prettner and Prskawetz, 2010). In these models, the association between population size and economic growth is ambiguous and subject to the modeling framework. This literature regularly predicts positive growth effects from population ageing, as households seek to save more during their working life. This triggers investments and hence growth. Incorporating social security however may reverse this result, as rising payroll taxes crowd out private savings (Kotlikoff et al., 2007). Notable studies investigating the fiscal implications of population ageing in an overlapping generations setting are Fehr (2000) and Börsch-Supan et al. (2014). This literature pays particular attention to the pension system when dependency ratios rise, while treating the tax system in a rather simplistic manner. Börsch-Supan et al. (2014) argue that, while sticking to a pay-as-you-go system, living standards in Europe can be maintained in spite of population ageing if total employment can be moderately increased. A similar point is made by Ang and Madsen (2015), who show empirically, using a long-term country panel, that an ageing work force is usually more productive. This suggests that the contribution of older workers with tertiary education to national production can outweigh higher pension and health costs. Finally, Kudrna et al. (2016) explore the welfare effects from cutting pensions versus raising taxes.

Concerning the fiscal implications of demographic changes, there are a number of studies on the sustainability of pension systems. Comprehensive projections can be found in Dekkers et al. (2010), European Commission (2012) and OECD (2013a). There is however little work dealing with the impact of population ageing on public *revenues*. The complexity of existing tax-benefit system calls for micro-based approaches rather than representative agent models. Notable exceptions are Decoster et al. (2014) and de Blander et al. (2013) for Belgium and Aaberge et al. (2007) for Norway.

We aim to fill this gap by a micro-founded approach for 27 EU countries, that is able to capture heterogeneous developments between population subgroups. Our treatment of the tax and contribution systems is able to capture far more detail than macro models generally can. This comes at the cost of ignoring potential general-equilibrium effects — we return to this limitation in the next section.

Our chapter further extends the literature by exploring the scope of effective policy responses. Surprisingly, despite the relevance of the topic, there are only very few *ex-ante* studies investigating the effects of reforms to pension systems.[1] Leombruni and Richiardi (2006) set up an agent-based microsimulation model of labor supply to analyze the evolution of the Italian labor force-, taking into account demographic projections. Explicitly modeling retirement rules as well as behavior, they simulate the effects of an Italian retirement reform from the 2000s on the labor market. Mara and Narazani (2011) simulate the effects on employment and retirement behavior of a reduction in pension benefits in combination with targeted income support in Austria. They show that such a reform increases social welfare as well as the employment of middle-income males (aged 55–60). Another simulation study by Fehr et al. (2012) investigates the recent increase in the German statutory retirement age from 65 to 67 years. They show that this rise will postpone effective retirement by about one year and redistribute towards future cohorts. Yet, the reform is found to be not sufficient to offset the projected future increase in old-age poverty. None of the studies above deals with reforms of the pension system in a comparative European perspective, taking into account different country-specific fertility profiles and pension systems. Comparing the effects of pension system reforms across Europe helps to shed light on the role that systemic elements of pension policies play in shaping the fiscal budget effects.

Our results show the magnitude of fiscal strain expected from demographic change, revealing a negative outlook for the majority of countries. Taking into account labor market effects substantially improves the balance. Increasing the retirement age, as implemented in many countries, further improves fiscal outcomes, leading to mostly positive outcomes.

The chapter is structured as follows. Section 4.2 describes our approach of modeling demographic change and the labor market in more detail. Section 4.3 describes our implementation of the retirement age reform. Section 4.4 presents results on labor market and fiscal outcomes. Section 4.5 contains results on the inter-generational distribution of funding public finances. Section 4.6 concludes.

4.2 Data and methodology

Microsimulation Models (MSM) have become a standard tool for the ex-ante evaluation of tax-benefit reforms (Bourguignon and Spadaro, 2006). The basic idea of MSM is to apply different sets of tax rules to the same sample of households and compare

[1] In addition there are *ex-post* studies investigating the effects of pension reforms, see e.g. Cribb et al. (2013), Manoli and Weber (2016), Staubli and Zweimüller (2013), and Vestad (2012).

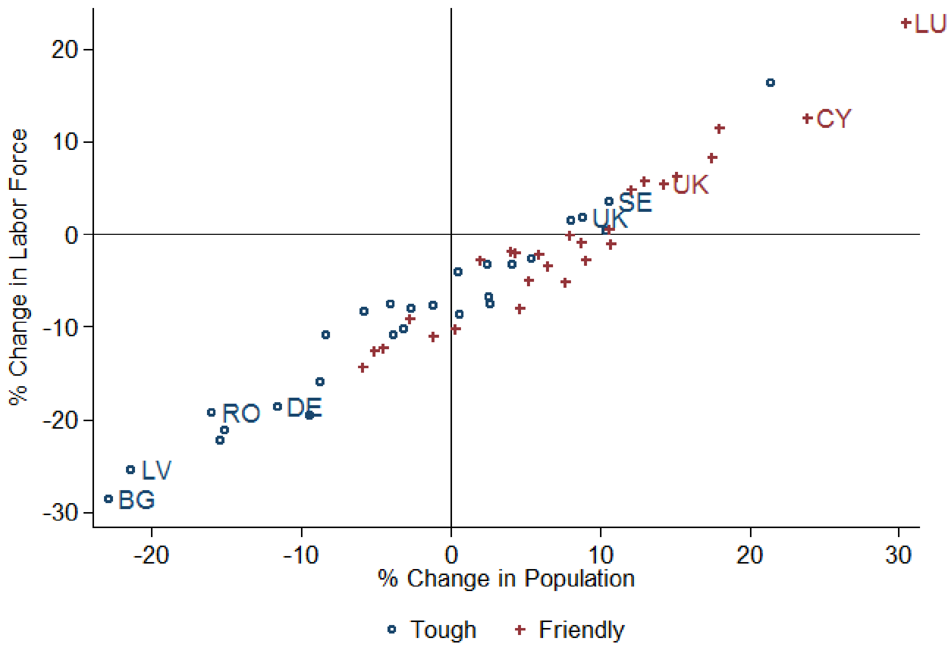
the outcomes across various dimensions such as inequality and employment. It offers a suitable framework to deal with the questions we pose due to its ability to account for the full heterogeneity within a given population. This is in contrast to approaches relying on representative agents, including CGE models. Moreover, the MSM results can be aggregated to the macro level, while this can be problematic for representative agent models due to potential biases. In the context of divergent demographic trends across EU countries, a micro-based approach is particularly useful, as we can account for the fact that the age composition, educational attainment and household composition are affected differently by demographic change across countries. In this chapter, we make two main advances in MSM that may be valuable for other research and policy analyses in the future. First, past MSM studies have been focused on modeling labor supply behavior while being relatively agnostic as far as labor demand feed-back effects were concerned. By introducing a novel labor supply and labor demand link (explained in Section 4.2.2), we overcome this shortfall and add a more realistic (partial) equilibrium notion to MSM. Second, demographic changes are accounted for by reweighting the micro data, which allows us to not only study labor market adjustments to policy reforms in current years but also in relatively distant future (see Section 4.2.1). Our chosen framework proposes, therefore, a middle ground between micro and macro approaches by making MSM outcomes more plausible when accounting for labor market effects. At the same time, the method is parsimonious, straightforward to implement and does not rest on too many assumptions, avoiding a *black box*.

The main parameters we employ, apart from assumptions underlying the demographic projections, are the elasticities of labor supply and demand. Throughout the analysis, we keep these elasticities constant, even though it is unlikely to be the case in practice. Time-persistent elasticities imply that responses of supply and demand to relative scarcities in the labor market are not changing over time. While the mechanics of the labor market might change over time, it is a priori not clear in which direction they might change and how much variation there could be. For that reason, it seems more reasonable to proceed with the assumption that there are no substantial changes to labor supply or demand elasticities in this time period.

4.2.1 Population Projections

We draw on Huisman et al. (2013) population projections for EU-27 in 2030, which are differentiated along the dimensions of age, gender, household type and education, separately for each country. The projections start from assumptions underlying the Eurostat projections, EUROPOP2010, but allow for additional variation, captured with two scenarios — the *tough* and the *friendly* scenario. The scenarios make different assumptions about international and internal migration, educational attainment, life expectancy, fertility and GDP growth. Broadly speaking, the tough scenario implies more severe challenges for European policy makers than the friendly scenario as it assumes lower fertility, lower educational attainment, less international migration and a higher life expectancy.[2] The latter scenario is assumed to cause a strong increase in the old-age dependency ratio. In contrast, the friendly scenario assumes higher net international immigration to Europe which has a positive impact on the working-age population as well as increasing the level of educational attainment.[3]

Figure 4.1: Projected change in population and labor force by 2030



Own calculations based on Huisman et al. (2013). See also Tables 4.3 and 4.4 in the Appendix.

We incorporate these projections into our micro data — European Union Statistics on Income and Living Conditions (EU-SILC) survey — by a reweighting procedure. The EU-SILC data are representative for the population in each country and contain rich information about socio-demographic characteristics and incomes of households, serving as input for the tax-benefit calculator (explained further below). Essentially, we adjust the respective sample weights for each observation proportionally to meet the target size in a given stratum.[4] By means of reweighting, we are able to analyze

[2] Huisman et al. (2013) use a cohort component model to project the age and sex distribution while education projections are based on KC et al. (2010). Comparing their population projections by skill level to those of the European Centre for the Development of Vocational Training (CEDEFOP), which provides an EU-wide population projection for 2020, shows that the two are well aligned in terms of head-counts (CEDEFOP, 2012).

[3] The recent influx of asylum seekers could not be incorporated. This is partly due to lack of reliable information on composition and size of the refugee influx. Moreover, there is huge uncertainty with regard to the length of stay in the host country. According to Hatton (2013), the rate of accepted asylum seekers dropped sharply in the course of the 1990s refugee inflow in the OECD due to tighter asylum policies. The effects on labor force composition in medium to long run is hence far from certain.

[4] For a similar application of sample reweighting in the context in tax-benefit microsimulation for Australia, see Cai et al. (2006).

how the European labor force will change over the course of two decades. Using the implied changes in the skill and age composition, we get a projection for the future labor force and aggregate labor supply before wage adjustments. Tables 4.3 and 4.4 detail by country how the population and the labor force can be expected to change in each European country by 2030. Figure 4.1 contrasts country-wise changes in labor force, defined as the population between age 15 and 64, and population for both scenarios. With few exceptions, the labor force is expected to shrink across countries in both the tough and friendly scenario — on average by 9.2% and 1.0%, respectively. The most drastic decreases are expected for Bulgaria, Romania, the Baltic countries and Germany. Although fertility rates are kept constant at the 2010 levels in the tough scenario, this assumption cannot be the main driver for the stark differences in headcounts between the two scenarios, as most new-born children will not be in the labor force in 2030. From all the different assumptions between both scenarios, migration has the most direct impact on the size of the labor force. As Table 4.5 shows, net migration flows are projected to be negative for the whole EU in the tough scenario. On the other hand, the friendly scenario implies a substantial overall annual inflow of 2.7 million migrants in 2030.

Figure 4.2: Structural changes in the work force composition

Projected Changes in percentage points between 2010 and 2030. Shares refer to total labor force. Older workers are defined as 50 years and older. High education is defined as completed tertiary education.

Apart from an overall decrease in size, the European labor force will undergo two major transitions, namely a shift towards older and higher-skill workers. The share of older workers is projected to rise in nearly all countries, most notably in the Southern European countries. This development is accompanied by increasing educational attainment, resulting in significant increases in the share of high-skilled workers in every country. This holds for both demographic scenarios and is particularly pronounced in the friendly scenario. In the tough (friendly) scenario, the share of high-skilled rises by only 0.9 ppt (8.0 ppt) in Germany, while other countries exhibit stronger increases, e. g. 10.7 ppt and 15.8 ppt respectively in Poland. The developments along both dimensions are visualized in Figure 4.2.

4.2.2 Labor Market Effects

In most countries, the total amount of hours worked, before accounting for wage adjustments, is projected to decrease as a result of demographic changes, ceteris paribus (Table 7, columns labeled **D**). It is unlikely that major transitions in the number of hours worked, as implied by our projections, would leave the behavior of labor market participants unaffected. In a neo-classical model of the labor market, greater scarcity of the production factor (labor) is expected to induce a wage increase which, in turn, may cause workers to supply more hours of work as potential disposable income rises. We model these wage adjustments by taking into account labor supply and demand elasticities as explained below.

Supply Side Elasticities Our estimates of labor supply elasticities stem from the analysis of Bargain et al. (2014). While the empirical literature on own-wage labor supply elasticities is vast, Bargain et al. (2014) is the first study to carry out estimations for a multitude of countries relying on a uniform methodological framework. They apply a flexible discrete choice model where couples are assumed to maximize a joint utility function over a discrete set of working hour choices. The utility function is specified to account for fixed costs of work, labor market restrictions within countries or even states, preference heterogeneity with respect to age, the presence and number of children as well as unobserved heterogeneity components. We draw on their elasticity estimates, distinguished by sex, marital status and skill level.[5] As the study covers only 17 EU countries, we use the respective country group mean (see Table 4.2) if a particular country is not covered.[6]

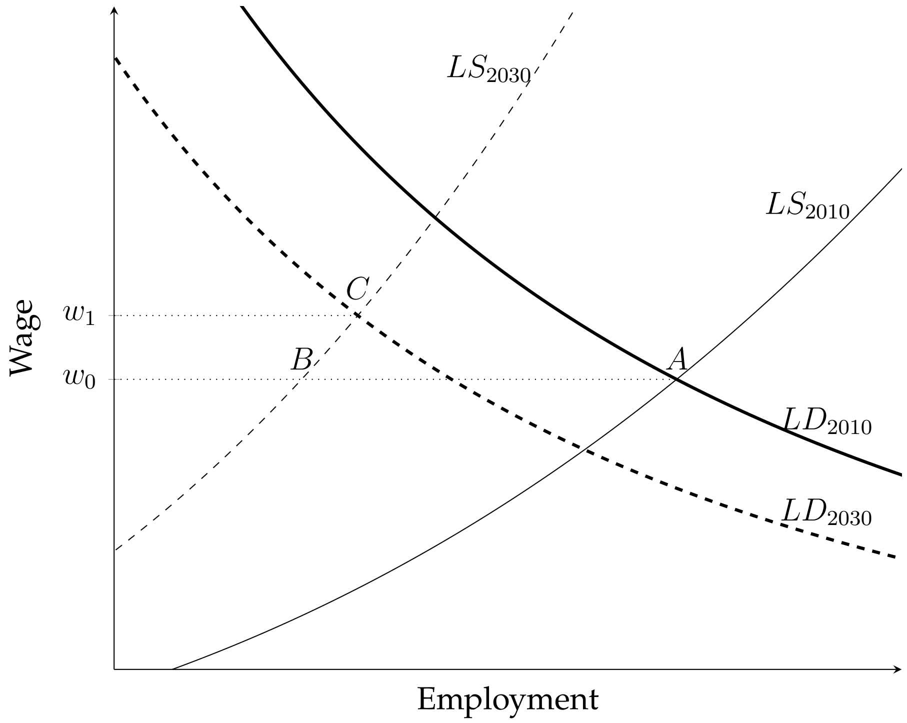
Demand Side Elasticities To capture reactions on the demand side of the labor market, we use skill-specific demand elasticities from the meta-analysis in Lichter et al. (2015a), shown at the bottom of Table 4.2. On the basis of empirical findings from 105 studies covering 30 years, the authors run a meta-regression of the estimated own-wage elasticity of labor demand. This allows them to obtain mean estimates for a given country, controlling for characteristics of the study, such as the time period or the estimation methods. We estimate a regression model on their dataset which follows their main specification (Lichter et al., 2015a, p. 101,) but adds an interaction term between skill level and country group. We then use our specification to predict conditional mean values, setting the time trend to 2030. Due to lack of available empirical studies, the demand elasticities can only be differentiated by skill level (low-skilled vs others) and country group. The latter may not be too problematic given the convergence processes

[5] See the Appendix for more details.

[6] The country groups are defined as follows. Continental: AT, BE, DE, FR, LU, NL; Nordic: DK, FI, SE; Southern: CY, EL, ES, IT, MT, PT; Eastern: BG, CZ, EE, HU, LT, LV, PL, RO, SI, SK. Anglo-Saxon: UK, IE.

among countries in the same geographic region. The meta-study reveals negative own-wage elasticities of demand which are larger than the supply side elasticities.

Figure 4.3: Linking Labor Supply and Demand



The graph illustrates the implied supply and demand shifts with overall decreasing labor supply and demand. While this is projected to happen in 15 countries in the tough scenario, the opposite may also occur (see Figure 4.1).

Labor Market Equilibrium Figure 4.3 visualizes our approach to combine both market sides to obtain the new labor market equilibrium. A formal representation is provided in the Appendix. We build on the approach of linking labor supply and demand in structural labor supply models by Peichl and Siegloch (2012).[7] In line with them, we differentiate supply-side responses by marital status, gender and skill level, leading to twelve distinct labor markets. This ensures a flexible adjustment process as it incorporates the main sources of heterogeneous labor market behavior. As we project a shrinking labor force for 18 out of 27 EU countries, even for the optimistic scenario (Figure 4.1), starting from the initial equilibrium A, the labor supply curve shifts to the left due to a shrinking labor force in the future.[8] Under constant wages, employment would

[7] The key difference to Peichl and Siegloch (2012) is the absence of a labor supply shock on the individual level. In our setting, the initial labor supply shock arises from demographic change. Our approach is more restrictive as it requires constant elasticities on both sides of the labor market. This way, the new equilibrium can be obtained analytically and does not require the iterative procedure of Peichl and Siegloch (2012).

[8] Under the assumption of constant elasticities, any supply/demand curve can be fully characterized by the elasticity and a single observation of hours. This assumption is crucial for this framework. While

change by the magnitude of the labor supply shock (Point B). This is the pure demographic effect. Negative elasticities on the demand side however imply higher wages due to greater scarcity of labor. We additionally take into account the demand shift that can be expected. As the total population is projected to decrease in the majority of countries, the aggregate demand for goods and services can be expected to decrease as well leading to a lower demand for labor. This is represented by a proportional shift of the demand curve, reflecting the relative population change in the respective country. A lower overall population hence manifests in a leftward shift of the LD curve. The new labor market equilibrium is hence defined by the intersection of LS_{2030} and LD_{2030} (Point C), featuring (in this example) higher employment and wages than in Point B.

Figure 4.4 displays the resulting average wage changes across the EU-27 for both scenarios. On average, we project wages to grow by 11.5% (12.4%) in the tough (friendly) scenario. It is crucial to note that, despite an average increase, there are many workers experiencing lower wages. With a few exceptions, average wage changes in a given country are very similar across demographic scenarios.[9] The starkest changes are projected for Germany and Austria. The smallest average wage increases are projected for Hungary, Latvia and Slovakia.

Our simulated wage changes are moderate given the time horizon of 20 years. Assuming a value of 1% for the annual productivity growth of labor over the period under consideration, one would end up with a total increase in labor productivity of 22% from 2010 to 2030.[10] Such productivity effects would add to the implied wage changes. Our labor market model does not explicitly address changing skill premiums due to technological change. The educational trends in the population projections are arguably driven to some extent by an anticipated rise in skill premiums, but they are taken exogenous in our model.

4.2.3 Tax-Benefit Calculator

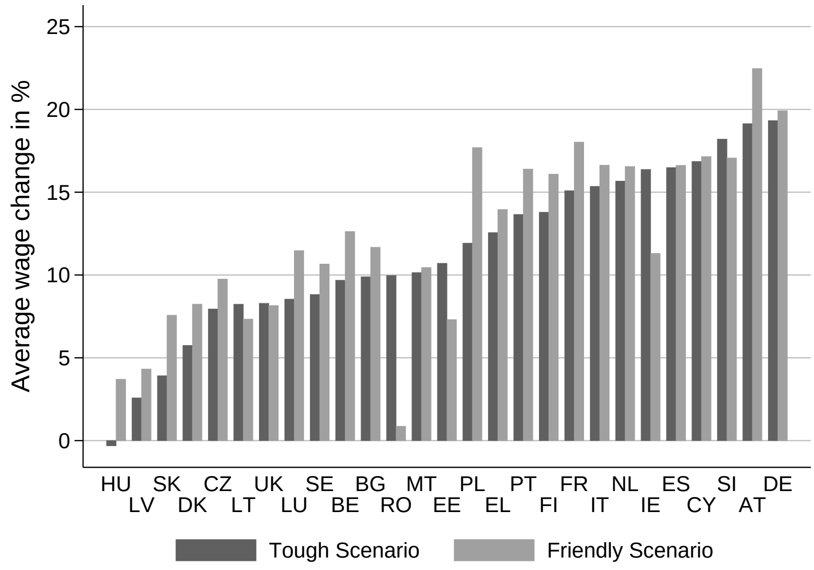
Any analysis of the fiscal effects of demographic change necessarily needs to address the full heterogeneity of the population of a country, as tax-transfer rules are highly complex and the individual burden of taxation (or eligibility for transfers) depends on personal and household circumstances. The requirements for such ex ante analysis are well met by fiscal microsimulation models (see e.g. O'Donoghue, 2014), which are commonly used in the analysis of public policies (Figari et al., 2014). Given our

behavioral responses might be quite stable over time, this may not hold under substantial wage changes. Specifying supply/demand curves with non-constant elasticities is of course possible, but the empirical foundation for this assumption would be weak.

[9] For an intuition of the wage effects, see Equation 4.6 in the Appendix. The wage change depends on the changes in total population and supplied hours, as well as on the elasticities on labor supply and demand.

[10] Comparable studies even assume an annual productivity growth rate of 1.5%, e.g. European Commission (2012, p. 75) and Börsch-Supan et al. (2014).

Figure 4.4: Average wage changes



Own calculations. Countries are sorted in ascending order by the wage change in the tough scenario.

cross-national focus and the EU-wide scope of analysis, a natural choice is to use EUROMOD, which is the only tax-benefit microsimulation model covering all EU-27 countries (Sutherland and Figari, 2013).[11] EUROMOD enables us to conduct a comparative analysis of tax and benefit systems consistently in a common framework.

EUROMOD calculates household disposable income, based on household characteristics, their market incomes and a given set of tax-benefit rules. The model covers social insurance contributions from employees, employers and self-employed, income taxes, other direct taxes as well as cash benefits. It is mainly based on nationally representative micro-data from the EU-SILC released by Eurostat, or its national counterparts where available and when they provide more detailed information. We use version F6.0 of EUROMOD with input datasets based primarily on the SILC 2008 wave.[12] The sample size for each country varies from about 10 thousand individuals for Luxembourg and Cyprus to more than 50 thousand individuals for Italy and the UK.

We define a concept of Fiscal Balance (FB) as our outcome of interest. FB encompasses the sum of all personal taxes and social insurance contributions (SIC) paid less cash benefits received, that are either simulated in EUROMOD or contained in the SILC data. We further subtract public expenditures that are closely linked to the population structure, i.e. expenditures for health care, old-age care, child care and education. As these are not provided by EUROMOD, we rely on Eurostat (2013), that provides respective per capita expenditures by age group and country. This allows us to impute

[11] As examples of recent applications, see Immervoll et al. (2007), Bargain et al. (2013) and Dolls et al. (2012).

[12] For France, the 2007 wave is used, for Malta the 2009 wave and for the UK, the Family Resources Survey 2008/09 is used.

these expenditures on the personal level.[13] This definition of fiscal balance is partial as it ignores other government expenditure items such as infrastructure or defense, and non-household or indirect taxes (corporate income tax, VAT). However, it is still an informative indicator to broadly measure *changes* in public finances collected or spent in the labor market in this context as it captures the main revenue items (income taxes, SIC) and expenditures (public pensions, health and education) affected by changes in the population structure and by retirement age policies. For the year 2010, our fiscal concept covers on average around 50% of total government revenues and 61% of total expenditures.

In order to facilitate the comparison between governments of different size, the total balance is normalised and shown as the share of total household disposable income in 2010.[14] Note that we assume an unchanged institutional environment. Our setting does not incorporate a commodity market, there are hence no price effects. All fiscal results can therefore be understood in constant prices. Equivalently, one could think of policy parameters that are uprated according to the inflation rate.

4.3 Modeling Retirement Age reform

Our policy scenario raises the gender-specific retirement age in each country by 5 years, which roughly corresponds to the average forecasted increase in the life expectancy in the friendly scenario (see Huisman et al., 2013, Table 3).[15] The statutory retirement age varied notably in 2008 (which is the reference period for our sample), from 60 in France to 68 for males in Finland — see Table 4.6 in the Appendix. Between 2008 and 2015, 21 out of 27 EU countries have implemented a raise in the statutory retirement age, mostly in the range of two to three years (European Commission, 2015, pp. 182ff). Our policy scenario hence anticipates some of the measures already undertaken.

The first complication for implementing the reform arises from the fact that average effective retirement age is usually lower than the statutory retirement age. There are substantial fractions of the population that retire before they reach the statutory retirement age, for instance due to health related concerns and/or country-specific regulations that facilitate early retirement. This is true for current retirement ages across Europe and with all likelihood also be the case after raising the legal retirement age. As

[13] Eurostat (2013) does not provide numbers for RO, BG, CY, MT, LV and LT. Similarly to the behavioral parameters, we assume in those cases the average age-related pattern of public expenditures as found in the respective country group. Although expenditure effects for these countries should be treated with caution, this facilitates cross-country comparability.

[14] A numerical example for Austria (AT) illustrates this. Here, the baseline fiscal balance amounts to €–11.2 bn and decreases to €–23.4 bn in the friendly scenario, considering demographic change only. The difference divided by the total household income in 2010 (€124.3 bn) is hence –9.82%, which is reported in Table 4.13.

[15] It also addresses the Barcelona target of raising the retirement age gradually by 5 years (European Council, 2002). We additionally ran a second reform scenario that introduces a universal retirement age of 70. The main conclusions are not fundamentally different and the results are available upon request.

a result, employment rates tend to decrease relatively smoothly around the statutory retirement age rather than exhibiting a very clear and sharp drop. This means that we need to predict employment rates under the new policy regime not only for the group of people affected by the increase of retirement age directly, i.e. those above the current age threshold and below the new one, but for a wider group of people. In the absence of a structural model determining the retirement decision (see, e.g. Manoli et al., 2015), we base the employment rate of the target group on a 5-year younger cohort (taking the three-year moving average to obtain smoother patterns).[16] We apply this approach to four separate groups of people, distinguished by gender and singles/couples to obtain new employment rates for all age groups older than 40, which is where employment rates peak in most countries, though the largest changes occur naturally for age groups around the current statutory retirement age.[17]

Figure 4.5 demonstrates our approach, taking male workers in Germany as an example: the solid lines are observed employment rates by age in the status quo (2010) under the current statutory retirement age of 65 (indicated by the first dashed vertical line). We basically assume that an increase in the statutory retirement age from 65 to 70 (under the first reform) shifts the employment curve to the right (by five years as well), shown with the dashed lines. For example, as the (smoothed) employment rate of single men at the retirement age of 65 was 0.19, we assume it will also be 0.19 at a new retirement age of 70. The area between the solid and the dashed line reflects the total increase in employment.

After deriving target employment rates, we assign a corresponding number of retirees from the affected age groups back to work. As the exit into retirement before reaching the statutory age is likely to be non-random, we need to identify individuals with the highest probability to be in employment under the new retirement rules. We estimate the probability of being in work for all individuals i between 45 and 75 years using the following probit model:[18]

$$\Pr(\text{work})_i = \Phi\left(\alpha + \beta X_i + \varepsilon_i\right) \quad \text{for } age_i \in [45; 75] \tag{4.1}$$

The probability of being employed is a function of individual characteristics X_i such as age (a cubic polynomial), the number of children, disability status, dummies for educational attainment, capital income, region, marital status as well as employment status and income of the partner.[19] Partner's status is crucial in couples, as the motivation to

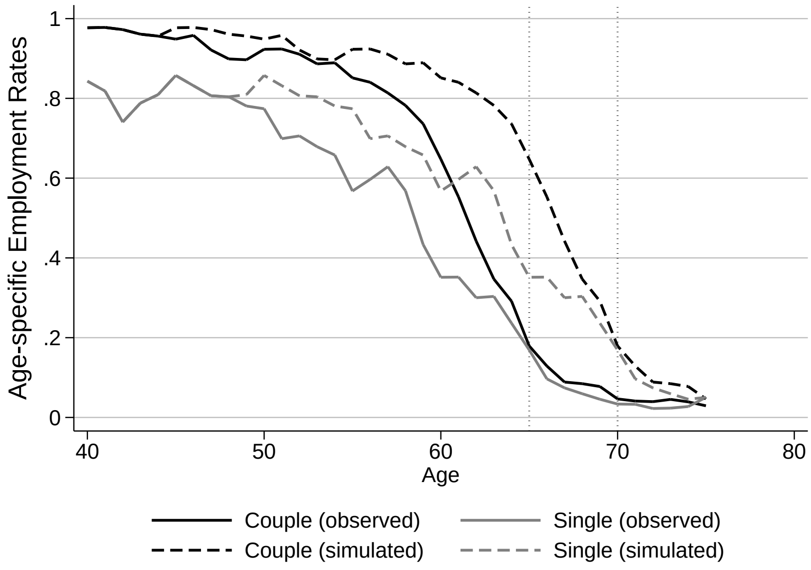
[16] We also rule out decreases in the employment rate by setting the minimum level equal to what is observed currently for a given cohort.

[17] The age variable for Malta is grouped in 5-year intervals, hence, our retirement age related adjustments are also inevitably cruder in this case.

[18] A similar approach has been used for example by Brewer et al. (2011).

[19] Some occupations or industries might bear higher health risks, implying that workers retire earlier. In order to take this into account, we would need information on pensioners' previous occupation or industry. Unfortunately, this information is not available.

Figure 4.5: Age-specific employment rates



Current observed (solid line) and predicted post-reform (dashed line) age-specific employment rates for men in Germany after a shift of the statutory retirement age from 65 to 70.

continue work might be low in the presence of a high-earning spouse. We estimate the model for each country separately for male and female workers (see Table 4.14 and 4.15 in the Appendix with the estimation results).

Having obtained the vector of coefficients β, we are able to predict the probability of being employed for those currently out of work. We then order these potential workers by the employment probability and, starting with the individuals with the highest probability, assign current retirees back to work until we meet the projected target employment rate by gender for each cohort. For those assigned into work, we assume individual labor supply to be equal to the cell-specific (defined by age, sex and education) mean value in weekly hours. The individual gross hourly wage is obtained from a regression that relates wages to observable individual characteristics and uses the standard Heckman (1979) technique to control for the unobservable factors that influence the selection into work.[20]

Once we have adjusted relevant labor market characteristics and imputed gross wage for individuals assigned back to employment, we use EUROMOD to calculate new tax liabilities and benefit entitlements. Note that we are not able to account for increased old-age pension claims from longer employment trajectories as public pensions are not simulated in the model, but taken from observed micro-data.[21] The relation between additional time in employment and the individual pension claim depends on the country-specific pension system. As an example, countries differ in the number

[20] The estimation results are available upon request.
[21] Note that pension claims might also change due to changed individual earnings (Section 4.2.2).

of years in employment on which the pension amount is calculated and how earnings from various points in time are weighted (OECD, 2013a, pp. 124f). Accounting for altered pension claims would require a dynamic modelling of individual earnings profiles, combined with the full set of institutional rules of the respective pension system, ideally also capturing interactions with private and occupational pension schemes. The European Commission (2015, p. 218) demonstrates the heterogeneity in pension claims if careers become slightly longer. For working two additional years, the change in the replacement rate (i. e. the ratio of pension entitlements to previous earnings) ranges from 0 to 20 percentage points across EU countries.

Note that the above description of deriving the market equilibrium abstracted from any policy reaction to the projected demographic transitions. Yet, the logic of our supply-demand link can be easily extended to any additional policy reform. To see how an increase in the retirement age interacts with our labor market model, return to Figure 4.3. Starting from the equilibrium with no policy reform, i.e. point C, an increase in the retirement age will increase labor supply and thus lead to an additional shift of the labor supply curve to the right. The new equilibrium point yields higher employment and lower wages compared to C.

4.4 Labor Market and Fiscal Results

In this section, we present our main simulation results. We focus on two outcomes (i) changes in hours worked and (ii) the effect of the fiscal balance. For both outcomes, we estimate effects at three different stages: (a) only taking into account demographic change (stage D), which isolates the external shock to labor supply for given wages; (b) after the demographic change and wage adjustments effects (stage DW), which captures interactions between labor demand and supply following initial supply shock; and (c) after the demographic change and the counterfactual policy reform of a 5-year increase in the retirement age (stage DRW), taking into account wage reactions. Results for the three different stages are shown estimated for both the tough and the friendly demographic scenario and for all countries. For clarity, we report the results by country group, roughly reflecting welfare regimes (Esping-Andersen, 1990; Ferrera, 1996). Detailed results by country are reported in Tables 4.7 to 4.13 in the Appendix.

The upper panel of Table 4.1 shows changes in total hours worked. The pure demographic effect (D) is -7.0% (+3.0%) in the tough (friendly) scenario for the EU-27. This represents the total labor market effect, capturing both intensive and extensive reactions. Isolating the extensive margin, i. e., the change in total employment, reveals similar effects of -7.4% and +2.5% respectively (see Table 4.8 in the Appendix). Eastern and Continental Europe are projected to face the largest declines, while total hours actually rise in both scenarios in the Nordic and Anglo-Saxon countries. Comparing changes in hours to changes in the labor force size (-9.2% and -1.0% for the tough and friendly scenarios respectively) suggests that focusing on head-count overestimates the reduction in effective labor and ignores differential labor supply behavior across socio-demographic groups. The change in hours partly compensates for the reduction in labor force. This suggests that demographic changes will increase the share of people with a stronger preference for working.

Wage reactions to initial shocks in labor supply, and accounting for demand-side adjustments at the same time, (columns labeled DW of Table 4.1) lead to additional negative effects on aggregate hours on top of what is induced by the demographic changes only. The wage adjustments to the demographic change do not, therefore, have a stabilising effect on aggregate employment.[22] The additional decrease in aggregate hours due to wage adjustment is particularly felt in southern European countries. It should be however stressed that the aggregate difference is a sum of positive and negative trends for the 12 distinct labor market simulations we employ.

As expected, the hours effects from raising the statutory retirement age by 5 years are substantial (columns labeled DRW in Table 4.1) with the change in aggregate hours going from -8.5% (1.9%) to -5.5% (16.7%). The largest improvement in hours of work is seen in Continental and Southern European countries. This suggests that undertaking this reform can counterbalance the decrease in hours worked from demographic changes even in the tough scenario. There are, however, a few countries (Bulgaria, Estonia, Latvia) where the decline in total hours still exceeds 10% (Table 4.7).

Panel B of Table 4.1 shows how the changes in total hours translate into fiscal outcomes. The figures refer to relative differences in the fiscal balance, normalized by the total disposable income in 2010:

$$\frac{\Delta \text{FB}}{\sum Y_{2010}^{\text{disp}}} = \frac{\Delta \text{SIC} + \Delta \text{Tax} - \Delta \text{Benefits} - \Delta \text{Other Exp.}}{\sum Y_{2010}^{\text{disp}}} \tag{4.2}$$

We first quantify the scale of fiscal stress which the demographic change is likely to lead to. Under constant wages (columns labeled D in Table 4.1), public fiscal balances would decrease by around 6% of household disposable income in both scenarios. The negative budgetary effect in the friendly scenario occurs despite hours increasing 3% on average. Figures 4.6 and 4.7 decompose the change in fiscal balances for the tough and friendly scenarios. The components include income tax, social security contributions, cash benefits and government expenditure (including health, old-age care, child care and educational expenditures as explained in Section 2.3). From these figures, we can see that the negative fiscal balance estimated before accounting for wage changes or introducing the retirement reform (bars labelled D in Figures 4.6 and 4.7) is driven by increased spending on (old age) cash benefits, partly counterbalanced by increased taxes and social insurance contributions though not always. The fiscal outlook is similar

[22] Considering the transition in and out of employment only gives a similar, albeit slightly more positive picture. Average changes amount to -7% and +3.8% respectively.

Table 4.1: Labor Market and Fiscal Effects by Country Groups

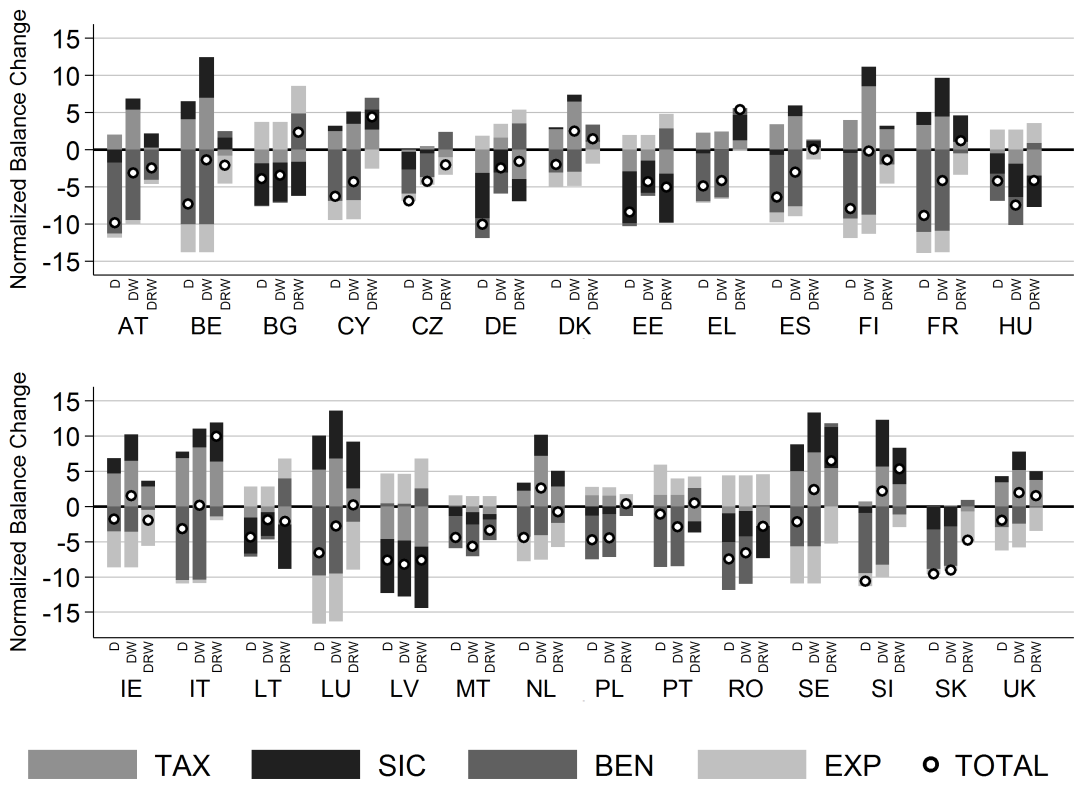
	D		**DW**		**DRW**	
	tough	friendly	tough	friendly	tough	friendly
	Panel A: Hours worked, relative change					
Continental	-10.0%	-2.4%	-10.6%	-2.5%	4.9%	14.4%
Nordic	1.8%	7.1%	0.3%	6.4%	7.7%	13.8%
Anglo-Saxon	3.1%	9.4%	2.8%	9.5%	14.1%	20.9%
Southern	-2.9%	8.1%	-6.7%	4.6%	12.6%	25.0%
Eastern	-14.5%	-2.8%	-16.0%	-4.1%	-3.0%	9.9%
EU-27 Average	*-7.0%*	*3.0%*	*-8.5%*	*1.9%*	*5.5%*	*16.7%*
EU-27 Labor Force Change	*-9.18%*	*-1.02%*				
	Panel B: Change in Fiscal Balance					
Continental	-7.9%	-8.6%	-1.9%	-0.8%	-1.1%	-0.3%
Nordic	-4.0%	-4.8%	1.5%	8.5%	2.2%	6.8%
Anglo-Saxon	-1.9%	-2.1%	1.7%	0.0%	-0.2%	-1.9%
Southern	-3.8%	-4.5%	-2.9%	-2.8%	2.4%	2.8%
Eastern	-6.8%	-5.4%	-4.8%	-3.7%	-2.1%	-0.4%
EU-27 Average	*-5.8%*	*-5.8%*	*-2.6%*	*-1.3%*	*-0.1%*	*1.1%*

D=demographic change only; DR=Retirement Age Reform. W indicates scenarios with wage effect. Panel A shows mean percentage changes in aggregate hours by country group. Results broken down by country are provided in Tables 4.4 and 4.7 in the Appendix. Panel B refer to percentage changes in the fiscal balances, normalized to household disposable income (Equation 4.2).

across countries, a few exceptions include those which are expected to face significant population growth (e.g. Sweden) or have a greater reliance on private pension schemes, such as Ireland or the UK. Another interesting finding is a positive contribution of expenditures in some Eastern European countries and Germany (tough scenario), which can be explained by large decreases in the total population.

With wage adjustments, the fiscal outlook is less bleak. The average change in fiscal balance is still negative but reduced to -3% (-1%) in the tough (friendly) scenario (columns labeled DW in Table 4.1). The net budget change in the Nordic and Anglo-Saxon countries becomes even positive, on average. The Continental countries also improve their position substantially, while improvements are less drastic for the Southern and Eastern countries. As can be seen in Figures 4.6 and 4.7, improvements in the fiscal balance from the wage change are mainly due to higher tax revenues (bars labeled DW compared to D). Revenues from contributions and spending on benefits vary relatively little. While the friendly scenario shows fiscal balances in nearly all EU countries

Figure 4.6: Decomposed balance changes by country, tough scenario

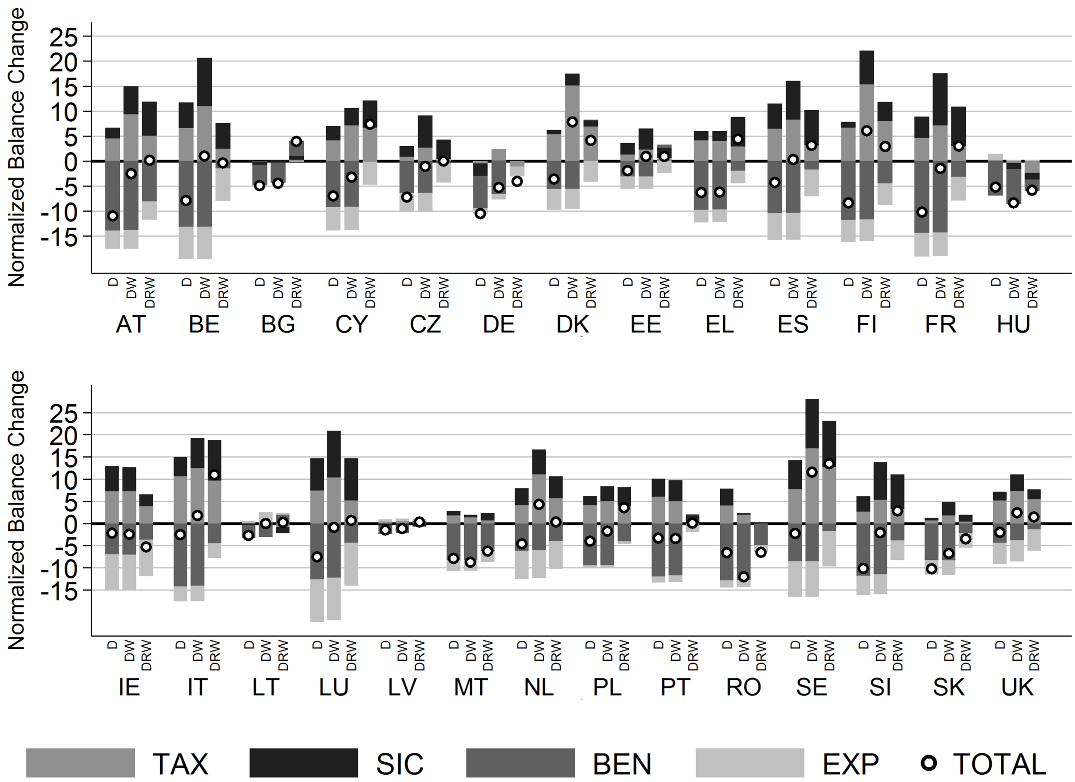


Note: The figure depicts the percentage change in the components of the normalized fiscal balance (Eq. 4.2) for each step, relative to 2010. TAX: Personal Taxes; SIC: Social Insurance Contributions; BEN: Benefit and Pension Payments; EXP: Demography-related expenditures.

close or above zero after the wage reactions, a couple of countries perform poorly in the tough demographic scenario: Hungary, Latvia, Romania and Slovakia end up with deficits above 5% of total household disposable income.

The retirement age reform brings EU average fiscal balance close to break even in both scenarios. Compared with the outcome after the demographic and wage changes (columns labeled DRW compared to DW in Table 4.1), fiscal balances improve most in the Southern and Eastern Europe, while we project stagnating or even falling balances for the other country groups. This is explained by the fact that there are two developments following the retirement age increase. Mechanically, cash benefits decrease and revenues increase with higher employment among the older cohorts. Additionally, there is a wage decrease due to higher labor supply, working against the positive revenue effect. The additional budget change from the retirement age reform is marked by a clear decrease in benefit payments (bars labeled DRW compared to DW in Figures 4.6 and 4.7). This positive effect on balances is offset by decreases in tax revenues, in some cases even dropping below the level with pure demographic changes (D).

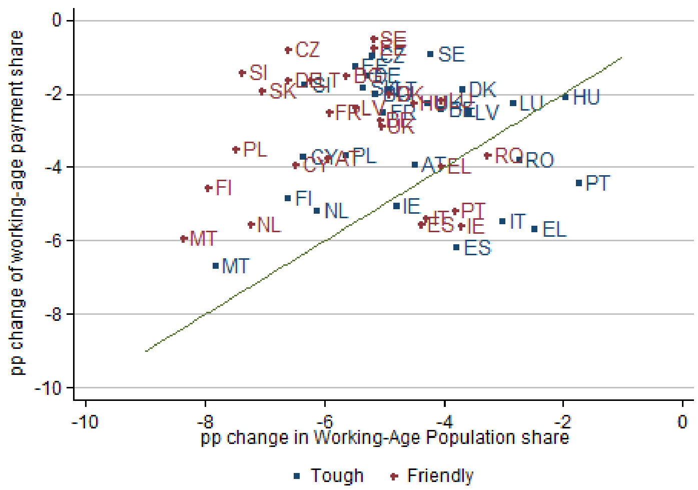
Figure 4.7: Decomposed balance changes by country, friendly scenario



Note: The figure depicts the percentage change in the components of the normalized fiscal balance (Eq. 4.2) for each step, relative to 2010. TAX: Personal Taxes; SIC: Social Insurance Contributions; BEN: Benefit and Pension Payments; EXP: Demography-related expenditures.

4.5 Intergenerational Distributional Impact

Figure 4.8: Payment burden of working-age population



The graph contrasts changes in the payment share of taxes and social security contributions by the working-age population with its change in population share. The revenue amounts take wage reactions into account (DW).

The previous section demonstrated fiscal strains for most countries from the expected demographic change. A related question is how the financing of public goods is going to be distributed across the population in the future. We therefore investigate the consequences of the demographic change on the intergenerational distribution of financial burden. Governments are financed to a large extent by the working-age population: the share of total taxes and contributions paid by people aged 15 to 64 amounts to 91% on average for the base year. Our demographic scenarios show that the share of working-age population, on average, decreases from 67.9% to 63.4% (tough) and 62.3% (friendly) respectively. Figure 4.8 plots the change in the share of the working age population between 2010 and 2030 on the horizontal axis, and the change in the share of the taxes and social security paid by the working-age population on the vertical axis (in the absence of a retirement age reform). Most of the countries (in both demographic scenarios) appear to the left of the 45-degree line. This means that, while the share of working-age people in the population is projected to decrease from 68% in average in 2010 to around 63% in 2030, the fiscal burden for this group does not decrease by the same magnitude. In other words, the working-age population pays a larger share of total tax and social security in 2030 than in 2010, relative to its share in population. The fiscal burden accrues more towards the working age population than the non-working age population. This result is intuitive on two grounds. First, it is mainly the working-age population profiting from higher average wages. Second, most income tax and contribution systems treat pension incomes preferentially (OECD, 2013a).

4.6 Conclusion

It is widely believed that ageing populations in European countries will put pressure on public finances through higher spending on old age benefits and lower tax revenues. The issue has gained even more relevance in the aftermath of the Great Recession which has weakened governments' fiscal positions ahead of demographic developments. This chapter assesses to what extent these concerns are justified and explores a raise in the statutory retirement age as one likely policy response.

Linking EU-27 demographic projections for 2030 with rich household-level data and employing microsimulation methods, we simulate the fiscal effects of demographic change, accounting for substantial population heterogeneity and the complexity of tax-benefit systems. Using the EU tax-benefit model EUROMOD, our analysis covers 27 EU countries in a consistent way in a common framework. This is complemented by a partial equilibrium model of the labor market, relying on recent micro-based empirical evidence.

We quantify the scale of fiscal stress which the demographic change is likely to incur. Assuming constant real wages, public fiscal balances would decrease by around 6% of household disposable income on average — less than the drastic fiscal adjustments carried out in European countries following the recent crisis but of a comparable magnitude.[23] This is driven by increased spending on (old age) cash benefits, in most countries partly counterbalanced by increased taxes and social insurance contributions due to the older and better educated labor force. The fiscal outlook is broadly similar across countries, a few exceptions include those which are expected to face more favorable demographic developments and have a greater reliance on private pension schemes. Overall, the results are not particularly sensitive to the underlying demographic scenarios. Under flexible wage conditions, however, labor scarcity leads to a strong wage growth and small employment increases (compared to the situation with fixed wages) which, together, notably reduce the worsening in fiscal balances though are not sufficient to withstand it entirely.

We also consider a retirement age reform which increases the current (gender-specific) statutory retirement age by 5 years — roughly corresponding to the projected increase in life expectancy. We model effective retirement ages by extrapolating current employment profiles. Our results demonstrate that such reforms could more than offset the impact of demographic processes on fiscal balances. This is due to increased taxes as there is a strong correlation between the increase in the number of people in work and improvement in the fiscal balance, though the reduction of the welfare bill also matters. These effects are, however, moderated and sometimes even reversed, by lower wages due to higher labor supply. As a result, the likely wage reaction to the demo-

[23] Replicating our fiscal balance concept with revenue statistics, EU-27 balances worsened during the Great Recession, on average, by 7.4% of disposable income.

graphic change, coupled with a retirement age reform are sufficient to avoid worsening in fiscal balances in nearly all countries. An analysis of the change of the fiscal burden reveals that under the existing tax-benefit systems, the working-age population will assume even a greater role in financing the government. Their share of payments relative to the population share is projected to rise. Overall, our results paint a less worrying outlook on the fiscal implications of the demographic change. This is line with previous findings on the country level (Aaberge et al., 2007; de Blander et al., 2013).[24]

We conclude that wage dynamics are highly relevant for the analysis as dramatic demographic shifts may engender important wage adjustments. This highlights the importance of taking interactions between the demand and supply sides of the labor market into account when evaluating retirement reforms — looking at static effects only can be highly misleading. Nevertheless, our results should be considered in light of some limitations. Extensions to our work could address broader general equilibrium effects by considering the role of technological change and associated changes in labor productivity and returns to education. A more comprehensive concept of fiscal balance, taking e.g. indirect taxes into account, could be useful. Applying our setting to a dynamic microsimulation model could improve accuracy with respect to the change in individual pension entitlements in particular. Further work can also explore alternative policy options available such as reducing public pensions and increasing the tax burden for those currently employed. Another option to counterbalance decreasing labor force is pursuing policies which encourage higher migration. Even though migrants are likely to be net fiscal contributors (see e.g. Dustmann et al., 2010), this topic remains politically highly sensitive. Lastly, this chapter examined the effect of demographic change on labor supply, wages and fiscal revenue. Other outcomes of interest include inequality and poverty levels and we leave this for future research.

[24] See also recent OECD projections on pension payments, which are projected to increase from 9.2% of GDP to 11.7% up to 2050 (OECD, 2013a, pp. 174ff).

4.A Appendix

4.A.1 Labor supply elasticities

The total supply elasticity for subgroup $g \in [1, \ldots, 12]$ in country c is defined as a percentage change in total hours in relation to the percentage change in wages: $\varepsilon^S_{gc} = \frac{\partial H_{gc}}{\partial w_{gc}} \frac{w_{gc}}{H_{gc}}$. The intensive elasticity is this ratio conditional on working at least one hour. The extensive elasticity is defined as the relative change of the employment rate E_{gc}: $\varepsilon^{S,ext}_{gc} = \frac{\partial E_{gc}}{\partial w} \frac{w}{E_{gc}}$. This corresponds to the *extensive margin (participation)* in the result tables of Bargain et al. (2014).

Looking first at single females in Table 4.2, we see that the labor supply elasticity of low skilled single females ranges from 0.1 in the Eastern European countries to just over 0.3 in the British Isles. In the medium skilled category, it is the Southern European countries which display the highest labor supply elasticity for single females (at around 0.3) while the same figure for the British Isles is almost unchanged compared to the low skilled category. The Nordic and Continental countries show a similarly low labor supply elasticity for this group of medium skilled single women. The labor supply elasticities of high skilled single women are much higher than those of low or medium skilled, ranging from 0.25 in Eastern Europe to 0.5 in the Southern European countries and in the UK and Ireland.

In general, women in couples display higher labor supply elasticities than their single counterparts (except for the high skilled category). Once again, there are discrepancies by country groups although the labor supply elasticity of women in couples displays less variability by skill group than that of single women. Eastern European women in couples have the lowest labor supply elasticity, regardless of skill type, at around 0.1. Non-single southern European women have the largest labor supply elasticities which range from 0.35 among the high skilled to 0.5 among the medium skilled. The labor supply elasticity of continental European women is fairly constant across skill groups at around 0.3 while the Nordic countries and the British Isles also have stable elasticities of around 0.2 across skill groups.

Among single men, the highest labor supply elasticities are to be found among the high and low skilled with the group of medium skilled single men displaying reasonably stable labor supply elasticities across countries of between 0.1 (in the Continental countries) and 0.2 (in the Nordic countries). Among the low-skilled single men, the British Isles have the largest labor supply elasticity of around 0.45. The smallest, of 0.15, are to be found in the Continental and Eastern European countries. Meanwhile the Nordic and Southern European low skilled single men have labor supply elasticities of around 0.25. Similar cross-country grouping patterns are found for the high-skilled with the highest elasticities found in the British Isles (0.65), followed by the Nordic (0.35) and Southern European (0.3) countries.

Finally, we observe very low labor supply elasticities for men in couples, regardless of their skill level. These range from 0.06 to 0.14 with the largest values observed for high skilled men, followed by low skilled and then medium skilled men. The Nordic countries display the largest elasticities across country groups for men in couples, regardless of the skill group.

Table 4.2: Supply and Demand Elasticities

	Skill Level		
	High	Medium	Low
(Total) Labor supply elasticities			
Single Male			
Continental	0.15	0.11	0.23
Nordic	0.27	0.21	0.34
Anglo-Saxon	0.46	0.14	0.65
Southern	0.27	0.18	0.27
Eastern	0.15	0.17	0.24
Single Female			
Continental	0.23	0.14	0.38
Nordic	0.19	0.11	0.36
Anglo-Saxon	0.32	0.20	0.51
Southern	0.26	0.29	0.48
Eastern	0.09	0.10	0.48
Married Male			
Continental	0.09	0.08	0.10
Nordic	0.11	0.09	0.14
Anglo-Saxon	0.09	0.06	0.11
Southern	0.06	0.08	0.07
Eastern	0.08	0.08	0.08
Married Female			
Continental	0.28	0.30	0.27
Nordic	0.18	0.17	0.22
Anglo-Saxon	0.20	0.23	0.19
Southern	0.40	0.49	0.36
Eastern	0.11	0.12	0.11
Labor demand elasticities			
Continental	-0.53		-0.62
Nordic	-0.48		-0.54
Anglo-Saxon	-0.66		-0.91
Southern		-0.58	
Eastern		-0.66	

Note: Supply elasticities based on estimations from Bargain et al. (2014). The values refer to mean value by country group. Where possible, elasticities are country-specific. If a specific country was not covered in the initial study, it was assigned the mean value within the country group. Demand elasticities are from Lichter et al. (2015a), by adding an interaction between skill and country group to the main specification and setting the time trend to 2030. Due to insufficient empirical estimates, we had to partly aggregate skill levels for the demand side.

4.A.2 Analytical derivation of new labor market equilibrium

Denoting total hours worked with H and the average wage w, the labor demand elasticity η with respect to wage is defined by $\eta = \frac{\partial H}{\partial w}\frac{w}{H} = H'(w)\frac{w}{H}$. We assume an isoelastic demand curve of the form $H_D(w) = cw^\eta$, where c is derived from the observed combination of hours and (average) wages.

$$H_D(w) = c_0^{LD} w^\eta = \frac{H_0}{w_0^\eta} w^\eta. \tag{4.3}$$

Assuming an equilibrium state initially, both the supply and the demand curve go through this point. Defining the wage elasticity of labor supply ε analogously[25], the analytical labor supply curve looks as

$$H_S(w) = c_0^{LS} w^\varepsilon = \frac{H_0}{w_0^\varepsilon} w^\varepsilon \tag{4.4}$$

Now suppose a labor supply shock due to demographic change, i. e. $H_1^S = \lambda H_0$. This shifts the labor supply curve (4.4) by manipulating c_0, i. e. $c_1 = \lambda \frac{H_0}{w_0^\varepsilon}$.

At the same time, we mimic general equilibrium effects from demographic change on the labor demand side by scaling c_0^{LD} in Eq. 4.3 in proportion to the population change π. The new labor market equilibrium is found at the intersection of both equations

$$\underbrace{\frac{\pi H_0}{w_0^\eta} w^\eta}_{\text{new LD curve}} \overset{!}{=} \underbrace{\frac{\lambda H_0}{w_0^\varepsilon} w^\varepsilon}_{\text{new LS curve}} \tag{4.5}$$

This yields the new equilibrium wage

$$w^* = \left(\frac{\lambda}{\pi}\right)^{\frac{1}{\eta-\varepsilon}} w_0 \tag{4.6}$$

The relative wage effect $\frac{w^*}{w_0} = \left(\frac{\lambda}{\pi}\right)^{\frac{1}{\eta-\varepsilon}}$ for the respective population subgroup can then be fed into the tax-benefit calculator to compute labor market reactions on the individual level, and, finally, fiscal effects. Note that measurement error in the individual wage does not constitute a problem here, as $\frac{w^*}{w_0}$ is independent of w_0. We distinguish individual reactions by extensive and intensive labor supply elasticities. First, people in work adjust their number of hours according to the intensive elasticity. In a next step, the number of people in work is adjusted such that the employment rate changes according to the extensive elasticity.

[25] At this stage, the total labor supply elasticities are used.

Table 4.3: Projected Total Population in 2010 and 2030

	Million People			% Change	
	Base	tough	friendly	tough	friendly
AT	8.4	8.3	9.1	-1.2	8.7
BE	10.8	11.7	12.5	8.1	15.1
BG	7.6	5.8	7.2	-22.9	-4.5
CY	0.8	0.9	1.0	10.4	23.9
CZ	10.5	10.1	11.2	-3.8	6.5
DE	81.8	72.3	80.8	-11.6	-1.2
DK	5.5	5.7	6.0	2.5	7.9
EE	1.3	1.1	1.4	-15.4	5.9
EL	11.3	10.9	11.8	-4.0	4.4
ES	46.0	44.8	52.0	-2.6	13.0
FI	5.4	5.5	5.8	2.6	7.6
FR	62.8	66.2	69.5	5.4	10.6
HU	10.0	9.2	9.7	-8.3	-2.7
IE	4.5	4.7	5.3	4.2	18.0
IT	60.3	60.6	67.6	0.5	12.1
LT	3.3	2.8	3.1	-15.2	-5.9
LU	0.5	0.6	0.7	21.4	30.4
LV	2.2	1.8	2.1	-21.4	-5.1
MT	0.4	0.4	0.4	-9.5	4.6
NL	16.6	17.0	18.1	2.6	9.0
PL	38.2	34.8	38.3	-8.8	0.3
PT	10.6	10.0	11.1	-5.8	4.0
RO	21.5	18.0	21.9	-16.0	2.0
SE	9.3	10.3	11.0	10.6	17.5
SI	2.0	2.1	2.3	0.6	10.8
SK	5.4	5.3	5.7	-3.2	5.2
UK	62.0	67.5	70.8	8.8	14.2
Mean				-2.7	7.9
Population-weighted mean				-2.3	7.4

Note: Own calculations based on projections in Huisman et al. (2013) applied to EU-SILC data for the EU-27.

Table 4.4: Projected Total Labor Force in 2010 and 2030

	Million Workers			% Change	
	Base	tough	friendly	tough	friendly
AT	5.7	5.2	5.6	-7.7	-0.9
BE	7.1	7.3	7.6	1.5	6.3
BG	5.2	3.7	4.6	-28.6	-12.3
CY	0.6	0.6	0.6	0.4	12.4
CZ	7.4	6.6	7.2	-10.9	-3.5
DE	53.9	43.8	47.9	-18.7	-11.1
DK	3.6	3.5	3.6	-3.3	-0.1
EE	0.9	0.7	0.9	-22.3	-2.2
EL	7.5	7.0	7.4	-7.6	-2.0
ES	31.4	28.9	33.2	-8.0	5.8
FI	3.6	3.3	3.4	-7.6	-5.2
FR	40.7	39.6	40.9	-2.7	0.5
HU	6.9	6.1	6.2	-10.9	-9.1
IE	3.0	2.9	3.4	-3.2	11.5
IT	39.7	38.0	41.5	-4.1	4.8
LT	2.3	1.8	2.0	-21.2	-14.4
LU	0.3	0.4	0.4	16.3	22.7
LV	1.5	1.2	1.4	-25.5	-12.6
MT	0.3	0.2	0.3	-19.6	-8.0
NL	11.1	10.4	10.8	-6.8	-2.8
PL	27.2	22.9	24.5	-16.0	-10.2
PT	7.1	6.5	7.0	-8.3	-1.9
RO	15.0	12.1	14.6	-19.3	-2.8
SE	6.1	6.3	6.6	3.4	8.2
SI	1.4	1.3	1.4	-8.6	-1.0
SK	3.9	3.5	3.7	-10.3	-5.1
UK	41.0	41.7	43.2	1.8	5.5
Mean				-9.2	-1.0
Population-weighted mean				-8.7	-1.4

Note: Own calculations based on Huisman et al. (2013) applied to EU-SILC data for the EU-27. Labor force is defined by the population aged between 15 and 64.

Table 4.5: Assumed annual net migration flows in 2030

Country	last observed net flow	projected net flows	
		tough	friendly
	in 1000		
AT	21.1	12.9	58.4
BE	64.1	23.3	62.0
BG	-15.7	-57.4	50.8
CY	1.8	1.8	9.2
CZ	28.3	-5.6	56.7
DE	-10.7	-100.3	366.2
DK	15.3	5.9	18.0
EE	0.0	-11.8	11.2
EL	35.1	11.2	60.3
ES	50.3	-5.8	513.8
FI	14.6	5.6	13.9
FR	70.0	5.0	169.0
IE	-27.6	-3.9	45.4
IT	311.6	128.2	549.1
HU	17.3	18.4	25.9
LT	-15.5	-9.5	7.4
LV	-4.7	-12.6	13.5
LU	6.6	2.1	4.7
MT	-0.2	-1.7	2.4
NL	38.5	-13.4	37.0
PL	-1.2	-85.3	91.7
PT	15.4	10.1	64.4
RO	-1.6	-144.2	150.6
SE	62.6	9.0	43.0
SI	11.5	-0.6	11.9
SK	4.4	-2.6	18.9
UK	201.3	100.8	255.5
EU-27	*892.6*	*-120.4*	*2710.9*

Source: Huisman et al. (2013), Table 4.

Table 4.6: Overview: Pension Systems in the EU

Country	Public Pension System	Occup. Pension System	Private Pension System	Statutory Retirement Age (2010)		recent RA reforms	Share workers/pensioners	Fertility Rate
				Men	Women		2012	2012
AT	ER	M		65	60	RA+5(w)	3.77	1.44
BE	ER	V		65	65		3.76	1.79
BG	ER	V	M (*after 1959&SP)	63	60	RA+2(m);RA+3(w)	3.60	1.50
CY	ER	M (Publ); V (Priv)		65	65		5.22	1.39
CZ	ER			62y2m	58y8m	RA+:2m/birth coh.(m/w)	4.17	1.45
DE	ER	V		65	65	RA+2(m/w)	3.14	1.38
DK	FR and MTS	QM		65	65	RA+2(m/w)	3.68	1.73
EE	ER since 1999	M (*after 1982)		63	61	RA+2(m); RA+4(w)	3.74	1.56
EL	ER			65	60	RA+2(m);RA+7(w)	3.29	1.34
ES	ER (Priv); FRw (Publ)	M (Publ);V (Priv)		65	65	RA+2(m/w)	3.83	1.32
FI	ER	V		63	63		3.53	1.80
FR	ER	V		65	65	RA+2(m/w)	3.65	2.01
HU	ER		V	62	62	RA+3(m/w)	4.02	1.34
IE	FR	M (Publ); V (Priv)		66	66	RA+2(m/w)	5.54	2.01
IT	ER	V		65y4m	60y40	RA+1y8m(m);RA+6y8m(w)	3.09	1.43
LV	ER			62	62		3.62	1.44
LT	ER			62y6m	60	RA+1.5(m);RA+3(w)	3.72	1.60
LU	ER	V		65	65		4.94	1.57
NL	FR	M		65	65	RA+2(m/w)	4.01	1.72
PL	ER	V	M(*after 1969);V	65	60	RA+2(m);RA+7(w)	5.06	1.30
PT	ER	mostly V		65	65		3.43	1.28
RO	ER		M	64	59	RA+1(m);RA+2(w)	4.22	1.53
SE	ER	QM	M	61	61		3.39	1.91
SI	ER	mostly V		63	61		4.05	1.58
SK	ER		M; V for NLME	62	57y11m	RA+4.1(w)	5.53	1.34
UK	ER (V)	V		65	60	RA+3(m);RA+8(w)	3.84	1.92

Source: European Commission (2012), OECD (2013a); With regard to the private pension, a voluntary scheme exists in all countries. RA+ = Increase Retirement age by number of given years for women (w) and/or men (m); no entry: does not exist; ER = earnings-related; FR = flat rate; FRw = flat rate by wage categories; MTS = means tested supplement; S = Supplement; Priv = Private Sector Employees; Publ = Public Sector Employees; V = Voluntary Participation; M = Mandatory Participation; QM = Quasi-mandatory; SP = selected professions; NLME = new labor market entrants.

Table 4.7: Hours worked in 2010 and 2030

	Base	D		DW		DRW	
		to	fr	to	fr	to	fr
	Mill. Hours	% change					
AT	148.0	-11.1	-3.1	-9.4	0.1	2.4	14.1
BE	177.9	4.7	11.8	3.5	11.3	24.0	32.4
BG	179.0	-27.4	-8.6	-30.7	-12.1	-17.4	2.1
CY	17.0	2.5	16.7	2.5	16.6	18.7	34.0
CZ	221.4	-8.6	1.0	-8.9	1.8	3.2	14.3
DE	1442.5	-20.3	-10.1	-19.8	-9.4	-7.4	4.8
DK	109.4	-0.3	3.9	0.1	5.2	5.4	10.8
EE	29.8	-20.9	2.3	-23.1	-0.6	-12.4	11.7
EL	218.1	-9.0	-2.9	-10.0	-3.2	4.8	11.6
ES	910.5	-6.2	9.5	-8.0	7.7	10.5	27.7
FI	110.2	-4.4	-0.3	-5.9	-1.0	3.9	8.8
FR	1080.2	-0.2	4.5	-2.6	2.6	18.5	25.0
HU	188.5	-8.2	-5.0	-14.0	-9.9	-1.6	2.9
IE	82.1	2.0	19.0	-0.4	15.2	11.9	28.3
IT	1169.6	1.1	10.3	-5.0	4.6	16.6	27.1
LT	69.6	-17.6	-7.5	-19.6	-9.1	-1.3	9.8
LU	9.5	17.3	24.7	17.4	26.1	48.9	56.9
LV	51.6	-22.6	-6.5	-25.6	-9.9	-17.4	-0.7
MT	6.7	-12.5	2.1	-13.7	1.6	-7.7	8.0
NL	270.6	-4.6	0.9	-4.1	2.2	4.0	10.9
PL	745.3	-11.1	-3.2	-12.3	-3.9	-1.3	7.8
PT	206.5	-4.4	0.5	-7.1	-1.7	8.2	15.1
RO	376.3	-22.3	-0.5	-22.2	-1.3	-5.9	17.8
SE	177.9	7.0	13.7	4.3	11.8	11.6	18.8
SI	40.3	-8.8	-0.0	-10.4	-1.2	6.8	17.0
SK	119.5	-7.3	-0.8	-8.1	-0.8	7.5	14.7
UK	1033.9	3.1	8.6	3.1	9.1	14.3	20.3
Mean		*-7.0*	*3.0*	*-8.5*	*1.9*	*5.5*	*16.7*

Notes: Own calculations based on Euromod input datasets reweighted for 2010 and 2030 and labor demand and labor supply elasticities. Hours refer to total hours worked per week. D=demographic change only; DR=Retirement Age Reform. W indicates scenarios with wage effect.

Table 4.8: Employment in 2010 and 2030

	Base	D		DW		DRW	
		to	fr	to	fr	to	fr
	Millions	% change					
AT	3.6	-11.3	-3.7	-9.9	-0.5	4.8	17.7
BE	4.5	3.8	10.1	4.8	12.8	29.2	37.7
BG	3.9	-27.5	-8.6	-30.1	-10.9	-14.6	5.0
CY	0.4	1.7	15.7	5.0	19.6	27.8	44.5
CZ	4.7	-8.5	0.9	-3.8	7.4	12.7	24.7
DE	35.2	-20.2	-10.9	-18.0	-7.4	-3.0	13.3
DK	3.0	0.3	4.7	0.9	5.8	11.3	17.1
EE	0.7	-20.9	1.7	-21.7	1.1	-8.1	16.8
EL	4.7	-8.9	-2.5	-6.8	0.8	11.5	18.1
ES	21.5	-6.3	9.5	-5.4	11.7	18.2	36.6
FI	2.7	-4.5	-0.6	-4.7	0.0	9.2	14.3
FR	27.5	-1.3	2.9	-1.8	3.1	22.5	28.8
HU	4.5	-9.1	-6.3	-12.9	-9.0	2.1	6.6
IE	2.1	1.8	18.8	0.8	17.7	18.1	35.9
IT	26.8	0.2	9.4	-2.5	7.7	20.5	30.7
LT	1.7	-17.7	-7.9	-19.0	-8.4	-0.6	11.0
LU	0.2	15.2	21.6	20.3	28.7	48.7	57.6
LV	1.2	-22.1	-6.9	-23.5	-7.7	-12.6	4.4
MT	0.2	-13.8	0.1	-13.8	1.9	-5.9	9.4
NL	7.8	-5.0	0.2	-4.5	1.9	8.4	16.1
PL	16.2	-10.6	-2.4	-8.9	-0.1	5.1	14.9
PT	4.9	-6.0	-0.3	-5.6	-0.6	12.9	19.7
RO	8.5	-22.1	0.3	-19.7	3.0	-0.4	25.8
SE	5.3	6.8	13.2	3.7	10.9	13.0	20.1
SI	0.9	-9.7	-1.1	-9.5	-0.3	7.6	18.2
SK	2.7	-7.5	-0.7	-8.3	1.7	8.7	16.2
UK	28.7	3.5	8.9	4.5	10.6	22.3	28.8
Mean		*-7.4*	*2.5*	*-7.0*	*3.8*	*10.0*	*21.9*

Notes: Employment is defined by working a positive amount of hours. Own calculations based on Euromod input datasets reweighted for 2010 and 2030 and labor demand and labor supply elasticities. Notes: Hours refer to total hours worked per week. D=demographic change only; DR=Retirement Age Reform. W indicates scenarios with wage effect.

Table 4.9: Tax revenues in 2010 and 2030

	Base	D		DW		DRW	
		to	fr	to	fr	to	fr
	bn. €per year	% change					
AT	21.6	11.6	26.1	30.7	53.6	-3.0	23.8
BE	37.0	15.3	24.6	26.0	40.9	-3.0	9.3
BG	1.2	-24.1	-0.9	-23.0	0.7	-21.8	2.7
CY	1.0	28.3	47.0	39.9	81.8	30.8	76.1
CZ	4.4	-3.7	11.1	5.8	34.8	-17.7	4.3
DE	221.0	-16.3	-2.4	8.4	12.4	-20.6	-15.4
DK	45.7	6.8	13.4	16.0	37.5	2.6	17.2
EE	1.0	-18.9	8.4	-9.7	14.7	-20.8	3.6
EL	11.8	18.0	32.6	19.2	31.4	9.8	22.6
ES	65.0	23.5	44.8	31.2	57.6	1.7	21.9
FI	29.6	11.9	20.0	25.7	46.2	8.2	24.0
FR	129.0	21.5	29.8	28.9	46.3	6.9	19.6
HU	3.7	-3.8	8.2	-13.7	-2.9	-25.2	-17.0
IE	10.9	29.2	45.7	40.6	45.2	17.8	24.1
IT	218.2	25.0	38.8	30.7	45.9	23.4	35.3
LT	1.5	-12.3	1.2	-6.4	7.6	-20.1	-6.6
LU	2.2	32.1	45.6	41.9	64.1	15.8	32.1
LV	2.1	-19.4	1.7	-20.2	2.4	-23.8	-0.2
MT	0.2	2.0	25.4	-11.9	20.3	-16.4	9.4
NL	58.9	9.9	18.1	31.7	48.5	12.4	25.0
PL	28.4	8.4	22.4	8.1	26.9	2.7	20.6
PT	10.1	33.5	45.8	21.3	38.0	-16.5	-4.4
RO	6.2	-5.8	23.5	-4.0	11.2	-16.2	-0.1
SE	47.0	15.9	24.7	24.3	53.8	17.3	40.5
SI	2.3	5.2	19.6	42.6	39.9	23.5	24.4
SK	1.2	-1.0	12.0	2.4	33.7	-21.2	7.0
UK	243.4	13.5	20.7	20.5	29.0	14.9	21.8

Notes: Own calculations based on Euromod input datasets reweighted for 2010 and 2030 and labor demand and labor supply elasticities. Taxes include taxes on personal income, capital income and property. D=demographic change only; DR=Retirement Age Reform. W indicates scenarios with wage effect.

Table 4.10: SSC revenues in 2010 and 2030

	Base	D		DW		DRW	
		to	fr	to	fr	to	fr
	bn. €per year	% change					
AT	42.9	-5.1	6.2	4.4	16.2	4.3	18.6
BE	47.2	7.1	15.0	16.0	28.2	4.8	15.3
BG	3.6	-26.3	-3.2	-24.2	-1.3	-21.4	3.2
CY	1.5	5.4	21.4	12.1	26.1	19.9	41.3
CZ	20.6	-6.6	5.9	-1.4	17.9	-5.7	11.1
DE	357.7	-19.6	-8.1	-10.0	-0.1	-9.7	-0.5
DK	11.3	2.6	7.9	9.1	23.9	1.4	12.2
EE	2.3	-20.0	6.5	-12.4	12.0	-18.9	5.9
EL	23.9	-2.1	7.0	-0.2	7.7	13.3	23.2
ES	140.0	-2.4	16.4	4.7	25.1	2.9	22.7
FI	25.5	-1.6	4.2	9.1	23.9	1.6	13.4
FR	357.2	4.2	10.2	12.2	24.6	8.3	18.4
HU	9.2	-8.0	0.8	-13.1	-3.7	-12.4	-4.0
IE	14.1	10.6	27.6	18.2	26.9	3.9	13.0
IT	239.0	3.2	14.9	8.9	22.4	18.5	30.7
LT	4.4	-13.6	-1.0	-9.0	4.2	-16.8	-3.6
LU	3.1	20.6	31.3	29.1	45.1	28.3	40.5
LV	3.1	-22.1	-1.4	-23.0	-0.8	-25.1	-1.9
MT	0.4	-9.2	7.3	-11.6	3.6	-5.0	11.4
NL	99.1	2.9	10.0	7.8	14.8	5.8	12.9
PL	41.6	-4.9	7.6	-4.0	12.3	-0.5	16.1
PT	21.5	5.6	14.9	4.3	17.2	-5.7	6.5
RO	9.6	-15.1	14.1	-13.3	1.4	-17.1	-0.0
SE	54.8	10.4	17.7	15.6	30.5	15.8	28.6
SI	6.2	-2.6	9.7	18.2	23.4	14.2	21.5
SK	8.9	-8.2	1.5	-7.2	7.9	-10.0	4.2
UK	151.3	5.7	12.6	16.6	23.9	8.0	13.9

Notes: Own calculations based on Euromod input datasets reweighted for 2010 and 2030 and labor demand and labor supply elasticities. Social security contributions from employees, employers, self-employed and pensioners are captured. D=demographic change only; DR=Retirement Age Reform. W indicates scenarios with wage effect.

Table 4.11: Benefit payments in 2010 and 2030

	Base	D		DW		DRW	
		to	fr	to	fr	to	fr
	bn. €per year	% change					
AT	40.1	29.5	43.1	29.4	42.9	12.3	25.0
BE	33.1	41.9	54.7	41.8	54.7	-3.3	7.6
BG	3.3	0.5	20.1	0.6	20.3	-29.3	-14.6
CY	1.8	42.7	57.1	42.3	56.7	-10.0	0.6
CZ	13.9	13.5	26.5	13.4	26.5	-9.8	2.0
DE	370.1	8.3	20.2	8.7	20.5	-10.9	-0.3
DK	30.5	11.5	21.0	11.1	20.5	-8.0	-0.4
EE	1.0	2.7	20.8	2.8	21.0	-19.2	-4.6
EL	20.3	29.5	44.6	29.2	44.3	-4.3	8.9
ES	70.2	49.8	67.3	49.3	66.7	-1.3	10.8
FI	23.7	33.2	44.5	33.0	44.0	7.5	17.0
FR	263.0	35.3	45.8	34.9	45.7	1.7	10.1
HU	9.1	10.8	20.5	11.2	20.7	-2.6	6.8
IE	17.7	13.8	27.0	13.9	27.2	2.0	14.3
IT	200.6	41.7	56.6	41.4	56.1	5.7	17.6
LT	2.3	2.4	14.8	2.7	15.4	-20.2	-10.1
LU	3.0	43.7	55.9	42.4	54.2	9.7	19.4
LV	1.2	-3.2	14.7	-2.8	14.7	-19.6	-4.0
MT	0.6	19.6	35.8	19.4	35.7	12.5	27.0
NL	47.1	23.9	34.2	22.7	33.0	13.0	21.6
PL	38.1	24.8	38.2	24.6	37.8	5.0	16.1
PT	18.2	36.8	51.2	36.3	50.2	-11.1	-1.0
RO	11.6	21.1	39.6	20.8	39.2	-0.5	14.9
SE	35.4	24.0	36.0	23.9	36.0	-2.3	7.2
SI	4.1	35.6	49.1	34.4	47.8	4.7	15.8
SK	5.0	25.3	37.2	25.5	37.8	-4.2	10.0
UK	175.7	16.2	23.9	13.6	20.8	0.9	7.4

Notes: Own calculations based on Euromod input datasets reweighted for 2010 and 2030 and labor demand and labor supply elasticities. Benefits includes pensions (old-age, widower, orphan, disability pensions) and other social transfers including benefits for children, housing, education, unemployment, maternity as well as tax credits. D=demographic change only; DR=Retirement Age Reform. W indicates scenarios with wage effect.

Table 4.12: Other public expenditures in 2010 and 2030

	Base	2030	
		tough	friendly
	bn. €per year	% change	
AT	35.6	2.0	13.0
BE	43.7	11.9	20.6
BG	3.3	-18.2	0.7
CY	1.9	15.1	27.5
CZ	14.3	4.0	15.3
DE	290.2	-7.3	4.2
DK	35.7	6.1	13.2
EE	1.4	-9.0	11.1
EL	21.0	0.9	11.2
ES	114.8	5.2	21.2
FI	22.4	10.4	17.4
FR	267.7	9.0	14.8
HU	8.2	-8.9	-0.7
IE	19.4	18.0	28.8
IT	175.0	2.3	15.5
LT	2.8	-11.5	-1.8
LU	3.8	23.8	33.7
LV	1.9	-20.1	-2.4
MT	0.7	-5.6	9.1
NL	90.6	9.9	18.3
PL	27.9	-6.7	3.7
PT	19.3	-6.5	5.7
RO	12.2	-13.0	4.8
SE	46.1	17.2	26.1
SI	3.9	8.0	19.4
SK	5.9	2.7	12.4
UK	221.5	14.6	21.1

Notes: Imputed values based on per capita expenditures by country and age group from Eurostat (2013). Expenditures include health care, old-age care, child care and education. Expenditures are assumed irresponsive to changes in wages or employment and are hence fix across all scenarios. In lack of country values for RO, BG, CY, MT, LV and LT, the age-specific amount is imputed by the age-specific mean from the other countries in the respective country group.

Table 4.13: Changes in normalized budgets between 2010 and 2030

	D		DW		DRW	
	to	fr	to	fr	to	fr
AT	-9.82%	-10.96%	-3.17%	-2.65%	-3.57%	-1.24%
BE	-7.29%	-7.94%	-1.37%	1.01%	-2.15%	-0.63%
BG	-3.92%	-4.92%	-3.43%	-4.44%	3.25%	3.67%
CY	-6.30%	-6.98%	-4.32%	-3.25%	4.38%	7.35%
CZ	-6.93%	-7.25%	-4.29%	-1.09%	-2.05%	0.01%
DE	-10.06%	-10.55%	-2.46%	-5.32%	-1.61%	-4.09%
DK	-2.03%	-3.61%	2.47%	7.84%	1.42%	4.11%
EE	-8.38%	-1.94%	-4.32%	0.94%	-5.05%	0.90%
EL	-4.89%	-6.32%	-4.20%	-6.23%	5.39%	4.38%
ES	-6.40%	-4.31%	-3.04%	0.33%	0.03%	3.13%
FI	-7.93%	-8.36%	-0.23%	6.07%	-1.41%	2.91%
FR	-8.87%	-10.18%	-4.18%	-1.45%	1.20%	2.96%
HU	-4.21%	-5.26%	-7.47%	-8.41%	-4.19%	-5.81%
IE	-1.83%	-2.19%	1.54%	-2.48%	-1.97%	-5.34%
IT	-3.17%	-2.55%	0.14%	1.76%	9.95%	10.97%
LT	-4.34%	-2.69%	-1.90%	-0.06%	-2.14%	0.23%
LU	-6.61%	-7.53%	-2.77%	-0.91%	0.18%	0.64%
LV	-7.65%	-1.47%	-8.20%	-1.13%	-7.62%	0.33%
MT	-4.40%	-7.89%	-5.66%	-8.76%	-3.36%	-6.30%
NL	-4.45%	-4.64%	2.58%	4.30%	-0.78%	0.31%
PL	-4.74%	-3.98%	-4.49%	-1.75%	0.35%	3.49%
PT	-1.10%	-3.29%	-2.92%	-3.43%	0.48%	0.05%
RO	-7.49%	-6.60%	-6.58%	-12.02%	-2.82%	-6.48%
SE	-2.18%	-2.32%	2.38%	11.54%	6.48%	13.43%
SI	-10.63%	-10.11%	2.17%	*-2.12%*	5.31%	2.82%
SK	-9.57%	-10.27%	-9.04%	-6.77%	-4.79%	-3.48%
UK	-1.99%	-1.99%	1.94%	2.43%	1.50%	1.48%
Mean	-5.82%	-5.78%	-2.62%	-1.34%	-0.13%	1.10%

Notes: Balance change to 2010 according to Equation 4.2. Own calculations based on Euromod input datasets reweighted for 2010 and 2030 and labor demand and labor supply elasticities. D=demographic change only; DR=Retirement Age Reform; 2. W indicates scenarios with wage effect.

Table 4.14: Probit estimations for entering employment: Men

Country	AT	BE	BG	CY	CZ	DE	DK	EE	EL	ES	FI	FR	HU
Age	-1.218*	0.684*	0.342*	-3.310*	-5.801*	-1.664*	-2.013*	-3.780*	2.888*	-1.187*	-1.033*	-3.489*	-5.354*
Age^2	0.0285*	-0.00968*	-0.00395*	0.0651*	0.119*	0.0356*	0.0452*	0.0779*	-0.0531*	0.0261*	0.0244*	0.0783*	0.109*
Age^3	-0.000*	0.000*	0.000	-0.000*	-0.000*	-0.000*	-0.000*	-0.000*	0.000*	-0.000*	-0.000*	-0.000*	-0.000*
Disabled	0	-1.603*	-1.943*	-2.986*	-2.691*	-1.983*	-2.009*	-1.966*	-2.584*	-2.468*	-1.420*	-3.046*	-1.992*
MedSkilled	0.369*	0.428*	0.342*	-0.317*	0.875*	0.233*	0.223*	0.357*	-0.224*	0.121*	0.174*	0.300*	0.335*
HighSkilled	0.826*	0.825*	0.838*	-0.009	1.136*	0.591*	0.334*	0.791*	0.206*	0.375*	0.278*	0.419*	0.564*
# of Children	0.115*	0.0655*	-0.0221*	0.0871*	0.0945*	0.276*	0.0667*	0.0574*	0.170*	0.112*	0.0548*	0.0178*	0.00596*
CapIncome	0.769*	0.478*	()	-0.724*	-0.0748*	1.196*	0.0000226*	0.899*	0.568*	0.297*	0.749*	0.359*	-0.599*
Spouse Empl.	0.548*	0.704*	0.180*	0.412*	0.542*	0.904*	0.715*	0.436*	0.441*	0.587*	0.598*	0.409*	0.508*
Spouse Not Empl.	-0.211*	-0.331*	0.0164*	-0.176*	-0.505*	-0.519*	-0.540*	-0.212*	-0.313*	-0.265*	-0.238*	-0.199*	-0.468*
Constant	18.36*	-13.18*	-6.194*	58.20*	94.12*	26.40*	30.77*	61.68*	-49.35*	19.04*	15.00*	52.64*	88.29*
Pseudo R^2	0.375	0.365	0.242	0.342	0.425	0.398	0.42	0.37	0.277	0.333	0.304	0.306	0.435

Country	IE	IT	LT	LU	LV	MT	NL	PL	PT	RO	SE	SI	SK	UK
Age	-1.149*	2.515*	-3.845*	3.168*	-4.665*	0.679	-3.471*	-0.353*	0.386*	-1.564*	-2.380*	-2.826*	-3.220*	-0.140*
Age^2	0.025*	-0.045*	0.077*	-0.057*	0.091*	-0.013	0.072*	0.010*	-0.006*	0.032*	0.049*	0.057*	0.068*	0.005*
Age^3	-0.000*	0.000*	-0.001*	0.000*	-0.001*	0	-0.000*	-0.000*	0.000*	-0.000*	-0.000*	-0.000*	-0.000*	-0.000*
Disabled	-2.166*	-1.856*	-2.347*	-3.146*	-1.560*	-3.572*	-2.256*	-1.201*	-2.104*	-2.167*	-2.159*	-1.985*	-2.708*	0
MedSkilled	0.390*	0.318*	0.256*	0.160*	0.079*	0.580*	-0.015*	0.147*	0.151*	0.116*	0.266*	0.037*	0.356*	-0.180*
HighSkilled	0.447*	0.699*	0.571*	0.643*	0.296*	0.996*	0.239*	0.678*	0.474*	0.745*	0.404*	0.659*	0.525*	-0.039*
# of Children	0.033*	0.139*	0.043*	0.218*	0.108*	0.050*	0.280*	0.070*	0.035*	0.016*	0.085*	0.094*	-0.047*	-0.002*
CapIncome	0.276*	0.383*	0.167*	0.735*		-0.250*	0.443*	0.119*	0.239*	-0.443*	0.919*	0.390*	0.535*	0.412*
Spouse Empl.	0.725*	0.375*	0.709*	0.131*	0.386*	0.473*	0.773*	0.723*	0.819*	0.330*	0.276*	0.423*	0.809*	0.766*
Spouse Not Empl.	-0.316*	-0.181*	-0.137*	0.113*	-0.336*	0.109*	-0.408*	-0.585*	-0.204*	-0.317*	-0.432*	-0.226*	-0.508*	-0.582*
Constant	18.373*	-44.021*	64.738*	-55.058*	80.334*	-9.767	56.943*	5.067*	-6.522*	27.620*	39.639*	48.005*	51.518*	1.944*
Pseudo R^2	0.35	0.328	0.411	0.546	0.172	0.439	0.466	0.303	0.238	0.284	0.335	0.283	0.312	0.176

Dependent Variable: Employment Dummy. Sample is restricted to workers between 40 and 70 years. Significance levels: (*) = 0.01 level, (#) = 0.05 level, (+) = 0.1 level.

Table 4.15: Probit estimations for entering employment: Women

Country	AT	BE	BG	CY	CZ	DE	DK	EE	EL	ES	FI	FR	HU
Age	-1.149*	2.515*	-3.845*	3.168*	-4.665*	0.679	-3.471*	-0.353*	0.386*	-1.564*	-2.380*	-2.826*	-3.220*
Age^2	0.025*	-0.045*	0.077*	-0.057*	0.091*	-0.013	0.072*	0.010*	-0.006*	0.032*	0.049*	0.057*	0.068*
Age^3	-0.000*	0.000*	-0.001*	0.000*	-0.001*	0	-0.000*	-0.000*	0.000*	-0.000*	-0.000*	-0.000*	-0.000*
Disabled	-2.166*	-1.856*	-2.347*	-3.146*	-1.560*	-3.572*	-2.256*	-1.201*	-2.104*	-2.167*	-2.159*	-1.985*	-2.708*
MedSkilled	0.390*	0.318*	0.256*	0.160*	0.079*	0.580*	-0.015*	0.147*	0.151*	0.116*	0.266*	0.037*	0.356*
HighSkilled	0.447*	0.699*	0.571*	0.643*	0.296*	0.996*	0.239*	0.678*	0.474*	0.745*	0.404*	0.659*	0.525*
# of Children	0.033*	0.139*	0.043*	0.218*	0.108*	0.050*	0.280*	0.070*	0.035*	0.016*	0.085*	0.094*	-0.047*
CapIncome	0.276*	0.383*	0.167*	0.735*		-0.250*	0.443*	0.119*	0.239*	-0.443*	0.919*	0.390*	0.535*
Spouse Empl.	0.725*	0.375*	0.709*	0.131*	0.386*	0.473*	0.773*	0.723*	0.819*	0.330*	0.276*	0.423*	0.809*
Spouse Not Empl.	-0.316*	-0.181*	-0.137*	0.113*	-0.336*	0.109*	-0.408*	-0.585*	-0.204*	-0.317*	-0.432*	-0.226*	-0.508*
Constant	18.373*	-44.021*	64.738*	-55.058*	80.334*	-9.767	56.943*	5.067*	-6.522*	27.620*	39.639*	48.005*	51.518*
Pseudo R^2	0.35	0.328	0.411	0.546	0.172	0.439	0.466	0.303	0.238	0.284	0.335	0.283	0.312

Country	IE	IT	LT	LU	LV	MT	NL	PL	PT	RO	SE	SI	SK	UK
Age	-0.140*	-1.149*	2.515*	-3.845*	3.168*	-4.665*	0.679	-3.471*	-0.353*	0.386*	-1.564*	-2.380*	-2.826*	-3.220*
Age^2	0.005*	0.025*	-0.045*	0.077*	-0.057*	0.091*	-0.013	0.072*	0.010*	-0.006*	0.032*	0.049*	0.057*	0.068*
Age^3	-0.000*	-0.000*	0.000*	-0.001*	0.000*	-0.001*	0	-0.000*	-0.000*	0.000*	-0.000*	-0.000*	-0.000*	-0.000*
Disabled	0	-2.166*	-1.856*	-2.347*	-3.146*	-1.560*	-3.572*	-2.256*	-1.201*	-2.104*	-2.167*	-2.159*	-1.985*	-2.708*
MedSkilled	-0.180*	0.390*	0.318*	0.256*	0.160*	0.079*	0.580*	-0.015*	0.147*	0.151*	0.116*	0.266*	0.037*	0.356*
HighSkilled	-0.039*	0.447*	0.699*	0.571*	0.643*	0.296*	0.996*	0.239*	0.678*	0.474*	0.745*	0.404*	0.659*	0.525*
# of Children	-0.002+	0.033*	0.139*	0.043*	0.218*	0.108*	0.050*	0.280*	0.070*	0.035*	0.016*	0.085*	0.094*	-0.047*
CapIncome	0.412*	0.276*	0.383*	0.167*	0.735*		-0.250*	0.443*	0.119*	0.239*	-0.443*	0.919*	0.390*	0.535*
Spouse Empl.	0.766*	0.725*	0.375*	0.709*	0.131*	0.386*	0.473*	0.773*	0.723*	0.819*	0.330*	0.276*	0.423*	0.809*
Spouse Not Empl.	-0.582*	-0.316*	-0.181*	-0.137*	0.113*	-0.336*	0.109*	-0.408*	-0.585*	-0.204*	-0.317*	-0.432*	-0.226*	-0.508*
Constant	1.944*	18.373*	-44.021*	64.738*	-55.058*	80.334*	-9.767	56.943*	5.067*	-6.522*	27.620*	39.639*	48.005*	51.518*
Pseudo R^2	0.176	0.35	0.328	0.411	0.546	0.172	0.439	0.466	0.303	0.238	0.284	0.335	0.283	0.312

Dependent Variable: Employment Dummy. Sample is restricted to workers between 40 and 70 years. Significance levels: (*) = 0.01 level, (#) = 0.05 level, (+) = 0.1 level.

5 Productivity Effects of Air Pollution: Evidence from Professional Football*

5.1 Introduction

Air pollution is considered an important environmental risk factor for human health. The European Environment Agency reports that exposure to air pollution causes more than 550,000 premature deaths per year in Europe alone (EEA, 2016). Against this backdrop, a vast body of empirical literature has been devoted to quantitatively assessing the health impacts of exposure to various air pollutants to comprehensively trade off social benefits from limiting hazardous emissions against negative impacts such as reduced industrial activity and employment.[1] While existing studies have typically focused on infant health and mortality[2], the costs of ambient air pollution are not limited to specific parts of the population: high concentrations of air pollution have been found to increase premature death among the elderly as well as hospitalization rates among the working-age population (Deschênes et al., 2012; Schlenker and W. R. Walker, 2016).

In addition to causing damage to individual health, exposure to ambient air pollution may also trigger further costs for societies by lowering individual labor market performance and consequently hindering economic growth. Recent studies show that poor air quality reduces workers' labor supply (Hanna and Oliva, 2015) as well as the short-run performance of low-skilled agricultural and factory workers (Chang et al., 2016; Graff Zivin and Neidell, 2012; He et al., 2016).

This chapter contributes to this strand of the literature by analyzing the effect of ambient air pollution on the physical performance of young adults who are positively selected with respect to their general physiological condition, namely professional ath-

* A previous version of this chapter, co-authored with Andreas Lichter and Nico Pestel, is also published as: A. Lichter et al. (2015b). Productivity Effects of Air Pollution: Evidence from Professional Soccer. IZA Discussion Paper No. 8964.

[1] See Graff Zivin and Neidell (2013) for an overview. In addition, there is a large epidemiological literature on various other aspects of air pollution (see, e.g., Pope III, 2000; Pope III and Dockery, 2006).

[2] For studies on infant mortality, see Chay and Greenstone (2003), Luechinger (2014) and Tanaka (2015). Effects of air pollution on infants' general health are studied by Currie and Neidell (2005), Currie et al. (2009), Lleras-Muney (2010), Currie and R. Walker (2011), Coneus and Spiess (2012) and Janke (2014). Ebenstein et al. (2016) and Isen et al. (2016) show that exposure to air pollution in early life may also affect long-term educational attainment and earnings.

letes.[3] For this purpose, we make use of panel data on the universe of players and teams in Germany's top professional football league (*Bundesliga*) over the period from 1999 to 2011. The setting of our analysis offers a useful "laboratory" to study the causal relationship between air pollution and individual performance (see Kahn, 2000, for an assessment of the sports business for labor market research). First, we are able to exploit rich data on individual performance, which is measured in a consistent and comparable way in 2,956 matches at 32 different stadiums throughout the country on a weekly basis over a twelve-year period. Such a comprehensive coverage is unavailable for most other occupations. For each match, we combine information on individual performance with hourly data on the concentration of particulate matter (PM10) and ozone (O_3) in spatial proximity to the stadium at the time of kick-off. Both air pollutants negatively affect the respiratory system of the human body. We measure players' performance by the total number of passes per match, which is strongly related to physical performance and holds relevance for a team's success. Second, due to match scheduling rules that are beyond the control of teams and players, individual exposure to ambient air pollution can be considered as exogenous. Hence, our approach overcomes endogeneity concerns arising from residential sorting and avoidance behavior.

Overall, our analysis shows significant negative effects of ambient air pollution on players' performance. Using within-player variation and controlling for weather conditions on the matchday as well as a variety of player, team and match variables we find that a one percent increase in the concentration of particulate matter leads to a 0.02% reduction in the number of total passes played. While this linear effect is small in magnitude, allowing for a non-linear dose-response relationship reveals substantial negative effects: performance significantly decreases if the concentration of particulate matter exceeds the European Union (EU) regulatory threshold. For particulate pollution levels above 50 micrograms per cubic meter, the elasticity is around -0.15. However, negative effects of pollution are found well below the current limit set by the EU, starting at around 15 micrograms per cubic meter. By contrast, we do not find a statistically significant effect of ozone concentration on players' performance once controlling for weather characteristics and the time of kick-off. We show that this is most likely due to limited variation in ozone pollution in our estimation sample, given that *Bundesliga* matches are generally not held during the summer months when ozone concentrations peak due to intense solar radiation and high temperatures.

[3] Professional sports data have been frequently used to analyze economic questions. For example, Parsons et al. (2011) study behavioral implications of racial/ethnic discrimination using umpire decisions in professional baseball. Brown (2011) shows how workers' effort in competitions depends on the relative ability among competitors by exploiting the presence of a superstar in professional golf tournaments. Kleven et al. (2013) exploit cross-country differences in labor market regulation and income taxation to analyze the effect of top tax rates on international migration decisions of football players in Europe. Using data on professional golf players, Rosenqvist and Skans (2015) investigate how past success can boost future performance.

Our analysis further points to considerable heterogeneity across individuals. We find that negative effects of particulate matter on short-run performance increase with age and are largest for players aged above 30. Moreover, the performance of midfielders and defenders is particularly affected by pollution because these players are more attached to the game by exerting a larger number of passes. Our analysis further suggests that particulate pollution causes marginal reductions in individual pass accuracy and style of play, with athletes substituting short with long passes at high levels of pollution. Finally, we show that aggregate team- and match-level regressions offer similar results, suggesting that interaction effects between pollution, individual performance and the player's team-mates' or opponents' physical capacity are either small or cancel out.

The remainder of this chapter is organized as follows. Section 5.2 describes the institutional background and the data. Section 5.3 introduces the empirical strategy. The results are presented in Section 5.4 before Section 5.5 concludes.

5.2 Background and Data

Professional football in Germany. In our analysis, we exploit rich data on athletes' performance in matches of the *Bundesliga*, Germany's top professional league of men's football. Every season, eighteen teams face each other at home and away (see Figure 5.1 for the geographic spread of teams across Germany). Therefore, a season comprises 34 matchdays and 306 matches, which are typically held on weekends between late-August and May.[4] At the beginning of every season, the German Football League determines the match schedule for the entire season and specifies the weekend on which a specific matchday takes place as well as which teams face each other at which stadium. The exact day and time of each match is determined several weeks in advance and is subject to a set of factors, such as international football competitions, television agreements or security considerations.[5] Importantly, match schedules are thus beyond any control of teams and players.

Performance of football players. Information on players' performance is provided by *deltatre*, a commercial enterprise collecting data on professional sports and serving as an external service provider to the media and sports clubs. The dataset comprises information on all *Bundesliga* matches for every season from 1999/2000 to 2010/2011 and contains detailed information for each match (location, date and kick-off time, home

[4] The season pauses for a winter break, generally lasting from late-December until late-January. After each season, the worst three teams are relegated, while three teams are promoted from the second division (*2. Bundesliga*).

[5] For example, rivaling teams from the same city or neighboring areas do not play matches at home on the same day.

Figure 5.1: Geographic distribution of stadiums across Germany

Note: Location of Bundesliga stadiums across Germany. The number of stadiums in each city is indicated in brackets. The Rhine-Ruhr Area comprises (from west to east) the stadiums in Mönchengladbach (2), Düsseldorf, Duisburg, Gelsenkirchen, Bochum and Dortmund.

and away teams) and each player who was on the pitch at any point during the match.[6] For every player, we observe the number of minutes played (up to the full duration of 90 minutes), the team played for, the position played (defender, midfielder or striker) as well as various measures of performance. We exclude goalkeepers from our analysis as they constitute a very different style of play.

We use players' total number of ball passes during a match as our main measure of interest, a widely used statistic for assessing football players' individual performance. Although teams might generally be successful by pursuing a rather defensive style of play and thus passing the ball less often than their opponent, research shows that the number of passes represents an essential element of team success. For example, Redwood-Brown (2008) shows that a team's number of completed passes significantly increases in the five minutes before scoring. Moreover, focusing on the major European football leagues (including the *Bundesliga*), Collet (2013) documents a strong positive relationship between the number of passes and various measures of team success, such as the number of goals or points per season. Players' running distance might serve as another interesting measure of performance. Unfortunately, information on running distance was not collected prior to the 2010/2011 season, which precludes us from analyzing this measure. However, using publicly available data for the seasons from 2013/2014 to 2015/2016, we show in Appendix Figure 5.5 that the relationship

[6] Even if matches end in a draw after 90 minutes, there is no single elimination element or overtime/penalty shootout in the *Bundesliga*.

between players' running distance and the number of passes per match appears to be strongly positive.[7]

Air Pollution and Weather. We combine the data on football players' performance with detailed information from the air pollution monitor system of the German Federal Environment Agency (*Umweltbundesamt*). For each match, we extract all available hourly monitor readings for the concentration of particulates with a diameter smaller than ten micrometers (PM10) and ozone (O_3) in ambient air within a radius of ten kilometers (about 6.2 miles) around the stadium at the hour of kick-off and compute inverse-distance weighted means for both pollutants. Matches without pollution readings within this radius are dropped from the sample (716 out of 3,672).[8]

We expect that both pollutants — particulate matter and ozone — negatively affect the physical performance of football players. A high concentration of particulate matter in ambient air is particularly harmful to humans' health as it enters deep into the lungs and affects the pulmonary and cardiovascular functioning.[9] While particulate pollution may have natural sources (e.g. wildfires, sandstorms or volcano eruptions), there are various man-made emission sources such as automobile exhaust, electricity generation or any other industrial activity involving combustion processes. By contrast, ozone is not directly emitted to ambient air but emerges from complex chemical interactions between specific precursors (nitrogen oxides or volatile organic chemicals) under conditions of substantial solar radiation and high temperatures. Ozone has been shown to impede the respiratory system of the human body by irritating lung airways.[10]

Since weather conditions are important environmental confounders of air pollution, we further supplement our dataset with a rich set of weather controls. The data are provided by the German Meteorological Service (*Deutscher Wetterdienst*) and contain daily information on temperature, precipitation, humidity, air pressure and wind speed. As before, we derive inverse-distance weighted means from monitor readings in proximity to each stadium (40 kilometers or 24.9 miles) on each matchday.

[7] The underlying data were obtained from the German football magazine *Kicker*, see www.kicker.de.

[8] The results of our analysis are robust to relying on monitor readings from the closest station within a ten kilometer radius only.

[9] The sports medicine literature provides evidence of a negative relationship between ambient air pollution and performance through particulate matter inhalation (see Rundell, 2012, for an overview).

[10] Alternatively, one could also study the effect of other air pollutants such as sulfur dioxide (SO_2) or nitrogen dioxide (NO_2). However, particulate matter acts as a suitable proxy for air pollution in general given that it is positively correlated with all other main air pollutants except for ozone (Ziebarth et al., 2013). Rather than studying the effect of PM10, one could also analyze effects of fine particulate matter (PM2.5), given that smaller particulates can penetrate deeper into the human body and thus are even more harmful than PM10. Unfortunately, the concentration of PM2.5 has been only monitored since 2008 in Germany, which substantially limits both the sample size (by around 83%) and variation in air pollution. Nevertheless, we find a negative — albeit insignificant — effect of PM2.5 on player performance. Regression results are available upon request.

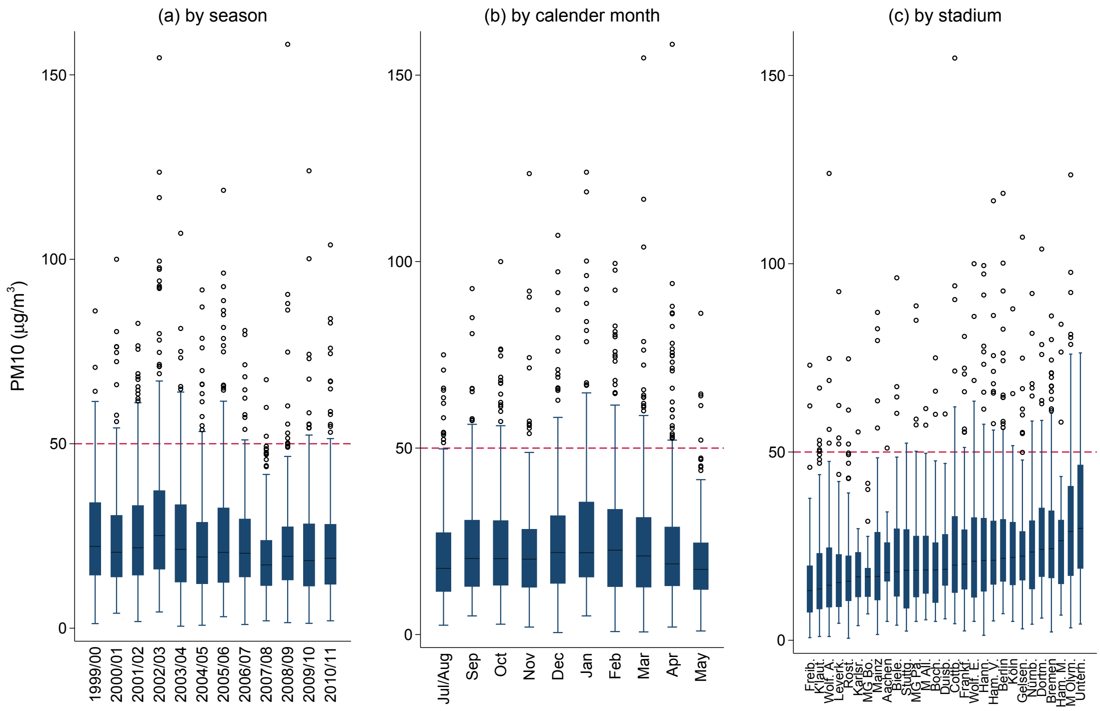
Descriptive Statistics. Our final dataset covers twelve seasons (1999/2000 — 2010/2011) and comprises 1,771 professional athletes playing for 29 different teams in 2,956 matches in 32 stadiums across Germany, totaling to 75,163 player-match observations. On average, we observe 42 matches per player. Table 5.1 provides descriptive statistics at the player and match level. Panel A summarizes players' characteristics and measures of performance. On average, a player passes the ball about 26 times per match (34 times conditional on playing for the full match) at an average accuracy of 77%, with more than 90% of these passes being short ones (less than 30 meters). On average, professional athletes in the *Bundesliga* are aged 27, play 71 minutes per match and cover the full duration (90 minutes) in 58% of games played.

Table 5.1: Summary statistics on player- and match-level variables

Variable	Mean	Std. Deviation	Median	Minimum	Maximum	Obs.
	Panel A: Player-level data					
Player performance						
Total passes	26.3	16.48	25	0	138	75,163
Total passes (full-time players)	34.2	14.47	32	1	138	43,346
Log of total passes	2.92	1.11	3.22	-2.3	4.93	75,163
Short passes	23.95	15.06	23	0	123	75163
Pass accuracy (completed over total)	.77	.15	.79	0	1	73,832
Pass ratio (long over short)	.1	.15	.06	0	6	73,718
Player characteristics						
Age (years)	26.81	3.93	26.71	16.92	39.55	75,163
Tenure (matches for team)	51.22	52.08	34	1	390	75,163
Played full-time (0/1)	.58	.49	1	0	1	75,163
Minutes played	70.6	28.97	90	1	90	75,163
Home match (0/1)	.5	.5	1	0	1	75,163
	Panel B: Match-level data					
Pollution variables						
PM10 ($\mu g/m^3$)	23.76	16.22	20.05	.53	158.28	2,956
Ozone ($\mu g/m^3$)	55.06	34.27	53.21	.13	248	2,956
Ln(PM10)	2.96	.68	3	-.64	5.06	2,956
Ln(Ozone)	3.69	1.01	3.97	-2	5.51	2,956
Weather variables						
Maximum temperature (°C)	12.59	7.6	12.35	-12.15	36.06	2,956
Precipitation (mm/m^2)	1.97	3.81	.2	0	34.03	2,956
Dewpoint (°C)	4.47	5.69	4.56	-17.06	18.79	2,956
Air pressure (hpa)	992.99	22.74	998.73	924.84	1044.7	2,956
Wind speed (m/sec)	3.61	1.84	3.2	.8	16.6	2,956
Match characteristics						
Stadium attendance (in 1,000s)	38.34	17.18	36.09	6	83	2,956
Matchday: Tuesday (0/1)	.02	.15	0	0	1	2,956
Matchday: Wednesday (0/1)	.03	.17	0	0	1	2,956
Matchday: Thursday (0/1)	0.01	.04	0	0	1	2,956
Matchday: Friday (0/1)	.05	.23	0	0	1	2,956
Matchday: Saturday (0/1)	.69	.46	1	0	1	2,956
Matchday: Sunday (0/1)	.2	.4	0	0	1	2,956
Kick-off: 3pm (0/1)	.69	.46	1	0	1	2,956
Kick-off: 5pm (0/1)	.18	.38	0	0	1	2,956
Kick-off: 6pm (0/1)	.02	.15	0	0	1	2,956
Kick-off: 7pm (0/1)	0	.03	0	0	1	2,956
Kick-off: 8pm (0/1)	.11	.31	0	0	1	2,956

Note: This table provides descriptive statistics on our estimation sample. Information on players' performance and match characteristics was provided by deltratre. Data on pollution levels and weather conditions was obtained from the German Federal Environment Agency (*Umweltbundesamt*) and the German Meteorological Service (*Deutscher Wetterdienst*), respectively.

Figure 5.2: Variation of particulate matter across matches

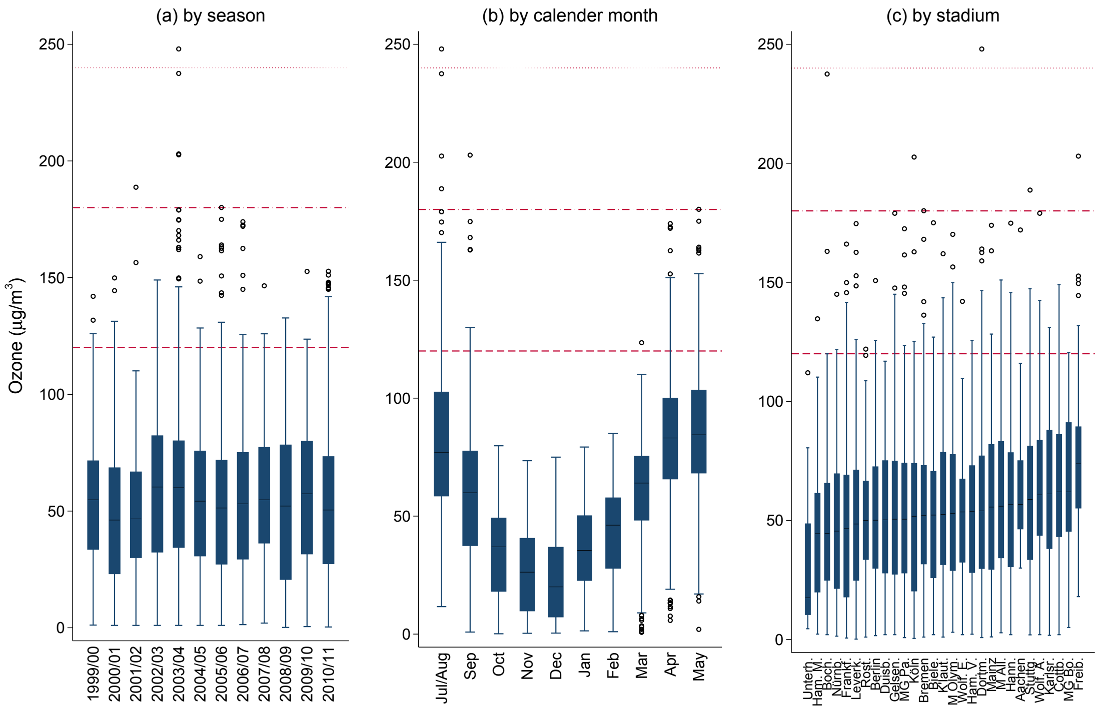


Note: Boxplots of the concentration of particulate matter (PM10) at the hour of match kick-off by seasons (panel (a)), calendar months (panel (b)) and across different stadiums (panel (c)). Monitor readings from July (N=4) are merged with August. There are no matches scheduled in June. Stadiums with less than ten match observations are excluded from this graph. The horizontal line indicates the EU regulatory threshold of 50 $\mu g/m^3$. Dots indicate matches where very high levels of PM10 concentration have been observed.

Panel B of Table 5.1 shows summary statistics for our two pollution variables of interest as well as the set of weather and match controls. The mean concentration of particulate matter at the time of kick-off is 23.8 micrograms per cubic meter ($\mu g/m^3$) but substantially varies from 0.5 to 158 $\mu g/m^3$. The mean concentration of ozone is 55 $\mu g/m^3$ in our sample, varying from 0.1 to 248 $\mu g/m^3$ at the hour of kick-off. The table further reveals that matches usually take place at moderate weather conditions on a weekend afternoon.

In order to better understand the observed variation in pollution levels, we present additional information on differences in the emergence of the two pollutants by season, calendar month and location. In Figure 5.2, panel (a) reveals that there is no remarkable time trend of PM10 concentration across seasons: at most, there is a slight reduction in average PM10 levels after 2005 when the EU regulation became binding. Panel (b) further shows a weak seasonal pattern in PM10, with pollution levels being slightly higher during winter months. Finally, panel (c) ranks stadiums with respect to the median level of PM10 concentration, reflecting patterns of population size and density as well as the degree of industrialization in proximity to the stadiums to some extent. The EU regulation threshold of 50 $\mu g/m^3$ – indicated by the horizontal line – is exceeded

Figure 5.3: Variation of ozone across matches



Note: Boxplots of the concentration of ozone (O_3) at the hour of match kick-off by seasons (panel (a)), calendar months (panel (b)) and across different stadiums (panel (c)). Monitor readings from July (N=4) are merged with August. There are no matches scheduled in June. Stadiums with less than ten match observations are excluded from this graph. The horizontal lines indicate the EU target, information and alert values at 120, 180 and 240 $\mu g/m^3$. Dots indicate matches where very high levels of O_3 concentration have been observed.

in around 7% of the observed matches.[11] Note that the EU threshold value is exceeded in all seasons, calendar months and stadium locations at least once.

Analogously, we plot differences in the level of ozone in Figure 5.3. As for PM10, panel (a) provides no evidence of a strong time trend across seasons. By contrast, panel (b) displays a very strong seasonal pattern. High levels of ozone are only observed for those matches scheduled in summer months, which reflects the fact that ozone emerges under heat and sunlight. The EU target value of 120 $\mu g/m^3$ – the bottom horizontal line – is exceeded in 4% of the matches covered, virtually all of them taking place between April and September. The EU information and alert thresholds at 180 and 240 $\mu g/m^3$ – indicated by the upper two horizontal lines – are only exceeded in a tiny number of matches (0.02% and 0.003%, respectively).

[11] In EU member states, air pollution is regulated by the "*Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe*". Since 1 January 2005, the daily mean concentration of PM10 may not exceed the limit value of 50 $\mu g/m^3$ on more than 35 days per calendar year. The annual mean may not exceed 40 $\mu g/m^3$ (Annex XI).

5.3 Empirical Strategy

Researchers face several empirical challenges when estimating the causal effect of air pollution on short-run performance (see Graff Zivin and Neidell, 2013). Residential sorting of healthier and more productive individuals into areas with lower levels of pollution may bias cross-sectional estimates of the performance-pollution relationship. Moreover, workers may opt to reduce their labor supply at short notice when being subject to high concentrations of air pollutants on a given day. The present setting provides a suitable framework to overcome these challenges as match scheduling rules are beyond the control of both players and teams, exposing these athletes to varying levels of pollution for exogenous reasons.[12] Moreover, avoidance behavior is virtually impossible in our context, given that there is no option to reschedule matches and players cannot evade air pollution on the field even in case of awareness.

Our empirical strategy exploits this exogenous variation in players' exposure to pollution. The underlying empirical model reads as follows:

$$\text{Passes}_{imt} = \beta_1 \text{PM10}_m + \beta_2 \text{Ozone}_m + X'_{imt}\gamma + W'_m\delta + M'_m\mu + C_{tm} + T_{ts} + \alpha_i + \varepsilon_{imt}, \quad (5.1)$$

where the number of passes of player i in match m for team t (Passes_{imt}) is regressed on the concentration of particulate matter and ozone at the hour of kick-off (PM10_m and Ozone_m). In our main specification, both the dependent variable as well as the two pollution variables enter our model in logarithmic form. This allows us to interpret the coefficients of interest (β_1 and β_2) as elasticities.[13]

We control for individual player characteristics (X'_{imt}) such as age (age squared), overall tenure (tenure squared), minutes played (minutes squared) and indicators for the position played in a particular match (defender, midfielder or striker).[14] Moreover, we include a dummy variable indicating a team's home stadium advantage. Given that weather conditions are important confounders for air pollution and may also have a direct effect on individual performance (Adhvaryu et al., 2014; Graff Zivin and Neidell, 2013), we further include controls for the maximum temperature, total precipitation, humidity, air pressure and wind speed at a daily level in spatial proximity to the sta-

[12] Note that even in the very unlikely case of athletes self-selecting into teams in low-pollution locations, only half of a season's matches are held at the home stadium, while away matches take place at stadiums across Germany (see Figure 5.1).

[13] We assign zero passes a log value of zero and account for the difference between one and zero passes by means of a corresponding dummy variable. When dropping all observations with zero passes from our sample (N=1,331; 1.8% of the sample), the results remain unaffected. In columns (1)–(4) of Appendix Table 5.6, we show that our results are robust to using log-level and level-level specifications of the pollution-performance relationship instead.

[14] Minutes played may be affected by the level of pollution in cases where coaches selectively substitute at an earlier stage of the game. If this were true, the variable would be a "bad control". However, the estimates in columns (1) to (4) of Table 5.7 show that there is no effect of pollution on the minutes played or the probability of playing the full duration in a given match. Note that our results are robust when including higher order polynomials of the minutes played in the regressions.

dium for a given match (W'_m). In our baseline specifications, all weather controls are standardized and enter the model in linear and quadratic form.[15] Moreover, match controls M'_m account for features of the particular match. We further include indicators for the day of the week and the time at kick-off and also control for stadium attendance (in 1,000s).[16] In addition, we add fixed effects for the coach of team t in match m (C_{tm}) as well as team-by-season fixed effects (T_{ts}) to capture all factors that are specific to a team within a season s, such as the squad's composition, style of play or the club's budget. Finally, we add player fixed effects to control for unobserved time-invariant differences across players (α_i). Identification of our model thus relies on exogenous variation in players' exposure to pollution at different locations over time. Accordingly, standard errors are clustered at the match level.

5.4 Results

We begin our analysis by examining the baseline linear effect of exposure to air pollution on player performance. Subsequently, we address the question concerning the extent to which the relationship between pollution and passes is non-linear in nature, before showing results on effect heterogeneity by player characteristics as well as for additional outcomes.

Baseline estimates. In Table 5.2, we present the results for different specifications of the model laid out in equation (5.1), subsequently including the different sets of control variables.[17] We present the results from our leanest specification in column (1), controlling for player characteristics only. We find statistically significant negative – albeit economically small – effects of both PM10 and ozone on players' performance. In column (2), we add player fixed effects to our model and thus identify the effect of air pollution on performance by using within-player variation only. Both coefficients of interest remain statistically significant and of similar magnitudes, suggesting that there is no systematic selection of players with respect to the degree of air pollution on matchdays. Next, we add the defined set of weather controls in column (3). While the effect of PM10 becomes even more negative, the coefficient for ozone and its statistical significance slightly decline. This is not particularly surprising because high temperatures act as a precondition for the emergence of ozone in ambient air. Hence, the inclusion of temperature as a control variable substantially reduces the remaining variation of ozone

[15] Appendix Table 5.8 shows that results for the baseline regression model are robust to different specifications of the set of weather controls.

[16] We do not find any evidence that stadium attendance is significantly affected by PM10 concentration. Hence, we can rule out the notion that our results are driven by pollution-induced reductions in support from the spectators.

[17] For the sake of clarity, we abstain from reporting the coefficients of the control variables. Detailed regression outputs are available upon request.

Table 5.2: The effect of air pollution on performance: Baseline effects

	(1)	(2)	(3)	(4)	(5)	(6)
Ln(PM10)	-0.021***	-0.019***	-0.023***	-0.019***	-0.020***	-0.021***
	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Ln(Ozone)	-0.008***	-0.008***	-0.007**	-0.004	-0.004	-0.005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Player controls	Yes	Yes	Yes	Yes	Yes	Yes
Player FE	No	Yes	Yes	Yes	Yes	Yes
Weather controls	No	No	Yes	Yes	Yes	Yes
Match controls	No	No	No	Yes	Yes	Yes
Coach FE	No	No	No	No	Yes	Yes
Team × Season FE	No	No	No	No	No	Yes
Observations	75,163	75,163	75,163	75,163	75,163	75,163
Adjusted R-Squared	0.770	0.746	0.746	0.749	0.752	0.756

Note: Dependent variable: Log number of passes. Player controls: age (squared), tenure (squared), position (defender, midfielder, striker), minutes played (squared), home match indicator. Weather controls on daily basis: maximum temperature (squared), precipitation (squared), dew point (squared), wind speed (squared), air pressure (squared). Match controls: day of week, kick-off time, stadium attendance and home match indicator. Standard errors are clustered at the match level (N=2,956). The usual significance levels apply: 0.1 (*), 0.05 (**), and 0.01 (***).

conditional on weather conditions. However, when further controlling for match-level characteristics (see column (4)), and in particular the time of kick-off, the effect of ozone declines even more and becomes statistically insignificant. By contrast, the estimated coefficient for PM10 remains almost unchanged with the inclusion of this additional set of controls. In columns (5) and (6), we subsequently add coach and team×season fixed effects to our model. Point estimates are not affected by either set of controls. Overall, the results from our preferred specification presented in column (6) indicate a statistically significant elasticity between the concentration of PM10 and the number of passes of -0.021, implying that an increase in PM10 by 1% reduces the number of passes by about 0.02%.[18] Put differently, this estimate thus implies that a one-standard-deviation increase in PM10 concentration (by roughly 16 $\mu g/m^3$) reduces passing by around 1.6% on average (see column (2) of Appendix Table 5.6). In turn, we find a small and negative — albeit statistically insignificant — effect of ozone on player performance.

As previously indicated, the results suggest that the observed differences in the impact of weather and match-level controls on the coefficients of PM10 and ozone reflect the fact that variation in ozone concentration in our sample is rather limited, especially when conditioning on temperature and the time of kick-off. This is mainly because almost no matches take place during June or July, when levels of ozone peak in Germany due to long sunshine duration and high temperatures (see panel (b) of Figure 5.3). By

[18] In Appendix Table 5.6, we present baseline estimates for the subsample of full-time players. The results are virtually identical to the full sample. We also use passes per minute as an alternative outcome measure. We find a statistically significant negative effect on passes per minute. In another, unreported specification we additionally control for stadium × home field fixed effects. The results are hardly affected by the inclusion of these additional control variables, with the point estimate being -0.017 (0.005). However, as this specification limits the identifying variation in PM10 to home matches within one season, we abstain from controlling for these fixed effects in the remainder of this chapter.

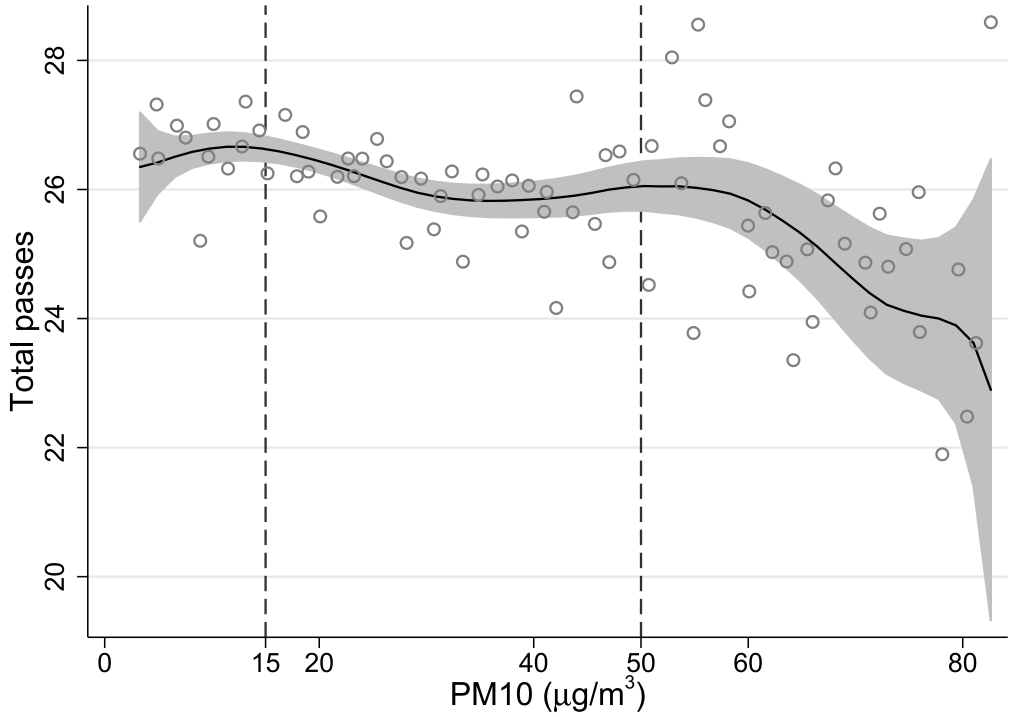
contrast, elevated levels of particulate pollution are observed throughout the year (see panel (b) of Figure 5.2). For this reason, we will focus on the effect of PM10 in the remainder of the chapter, while controlling for log ozone concentration at kickoff time in all specifications.

Non-linear dose-response relationship. We next test for non-linear effects in the relationship of particulate matter and player performance. Moderate levels of PM10 might not affect individual performance, while substantial negative effects may arise once pollution levels exceed a given threshold. In order to investigate this relationship more explicitly, we first examine potential non-linear effects in a non-parametric way, running a kernel-weighted local fourth-order polynomial regression of match-level PM10 concentration on players' total passes without any further controls. The results of this regression are presented in Figure 5.4, which plots the predicted number of passes against the level of pollution. At low levels of pollution (up to approximately 15 $\mu g/m^3$, equivalent to around the bottom third of the observed distribution of PM10 concentration), player performance does not seem to be impeded. Beyond this value, we observe a moderate negative effect of PM10 on the number of total passes. This negative effect becomes significantly larger for PM10 levels above the EU's daily limit value, which is set at 50 $\mu g/m^3$.[19]

Based on this non-parametric evidence, we further test whether the observed non-linear effect is confirmed in our regression framework. We allow for different marginal effects below and above the threshold values that have been identified by the non-parametric model visualized in Figure 5.4. As displayed in columns (1) and (2) of Table 5.3, the results of our analysis first confirm that player performance is not impeded at low levels of air pollution (PM10 $\leq$ 15 $\mu g/m^3$), with the effects being small and not significantly different from zero. By contrast, performance significantly decreases at higher levels of PM10 concentration. In columns (3) and (4), we next allow for different marginal effects below and above the EU threshold of 50 $\mu g/m^3$. As indicated in Figure 5.4, we find effects of air pollution to be particularly pronounced at very high levels of pollution. At PM10 levels above 50 $\mu g/m^3$, our estimates imply that a 1% increase in particulate pollution reduces the number of passes by around 0.15%. However, negative effects can be already observed at lower levels of PM10 concentration, implying that the linear effect displayed in Table 5.2 is not entirely due to very high concentrations of particulate matter. Finally, we allow for different marginal effects for the three identified segments. The regression results displayed in columns (5) and (6) confirm the overall pattern displayed in Figure 5.4. Negative effects of PM10 are already found

[19] For the sake of clarity, we exclude information from the bottom and top 1% of the PM10 distribution in the polynomial regression underlying Figure 5.4. These observations are outliers with respect to the overall distribution and would inflate the confidence intervals at the tails to a considerable extent. However, note that we include these observations in all other regressions.

Figure 5.4: The effect of particulate matter on performance: Non-linear relationship



Note: Non-linear relationship between PM10 concentration and players' number of total ball passes. The line shows the result from a kernel-weighted local fourth-order polynomial regression of players' passes on the concentration of particulate matter in ambient air absent any further controls. For the sake of clarity, the bottom and top percent of the PM10 distribution are discarded from this particular analysis. Dots indicate the mean number of passes at a given level of PM10 concentration rounded to the next integer value.

at moderate levels well below the current EU threshold of 50 $\mu g/m^3$ but become much stronger at very high levels of PM10 concentration.

Heterogeneity by player characteristics. We further investigate whether the observed effect of air pollution varies for different types of players. While professional football players constitute a rather homogeneous group that is positively selected with respect to their overall physical condition, we proxy general physiological status by age and test whether (slightly) older athletes may be more sensitive to pollution than their younger counterparts. When interacting the log of PM10 with the players' age in columns (1) and (2) of Table 5.4, we first find that the negative effect of PM10 increases with age. The results displayed in columns (3) and (4) further show that the performance of the youngest group of players (athletes aged below 21) is not affected by PM10, whereas all other players are negatively affected. The estimated effect is strongest for players older than 30 years, although differences are not statistically significant among the three older age groups. Nevertheless, we take these results as suggestive evidence that older individuals appear more vulnerable to detrimental environmental conditions, even among this group of professional athletes.

Table 5.3: The effect of air pollution on performance: Non-linear effects

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(\text{PM10}) \times (\text{PM10} \leq 15)$	0.004	0.012			0.004	0.012
	(0.011)	(0.011)			(0.011)	(0.011)
$\ln(\text{PM10}) \times (\text{PM10} > 15)$	-0.019**	-0.033***				
	(0.008)	(0.008)				
$\ln(\text{PM10}) \times (\text{PM10} \leq 50)$			-0.017***	-0.021***		
			(0.005)	(0.006)		
$\ln(\text{PM10}) \times (15 < \text{PM10} \leq 50)$					-0.028**	-0.040***
					(0.011)	(0.012)
$\ln(\text{PM10}) \times (\text{PM10} > 50)$			-0.127***	-0.153***	-0.127***	-0.155***
			(0.041)	(0.041)	(0.041)	(0.041)
Player controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather/Ozone controls	No	Yes	No	Yes	No	Yes
Match controls	No	Yes	No	Yes	No	Yes
Coach FE	No	Yes	No	Yes	No	Yes
Team × Season FE	No	Yes	No	Yes	No	Yes
Observations	75,163	75,163	75,163	75,163	75,163	75,163
Adjusted R-Squared	0.746	0.743	0.746	0.743	0.746	0.743

Note: Dependent variable: Log number of passes. All regressions include player fixed effects. Player controls: age (squared), tenure (squared), position (defender, midfielder, striker), minutes played (squared), home match indicator. Weather controls on daily basis: maximum temperature (squared), precipitation (squared), dew point (squared), wind speed (squared), air pressure (squared). Match controls: day of week, kick-off time, stadium attendance. Standard errors are clustered at the match level (N=2,956). The usual significance levels apply: 0.1 (*), 0.05 (**), and 0.01 (***).

We further account for differential effects of PM10 concentration by players' position, given that strikers, midfielders and defenders generally fulfill varying tasks during a game and may thus exert different levels of physical effort. This is partly reflected in the number of passes played, ranging from an average of 15.7 passes per match for strikers to 32.7 for defenders. However, our results in columns (5) and (6) of Table 5.4 show that PM10 affects the performance of all players irrespective of their position. At most, we find that defenders and midfielders are slightly stronger affected than strikers. While differences are not statistically significant, this finding is consistent with the notion that defenders and midfielders exhibit more physically demanding tasks and play a more active role in the game than strikers, and that pollution generally affects the performance of workers with more strenuous tasks to a larger extent.

Effects on pass accuracy and the style of play. We next investigate whether high levels of pollution affect players' pass accuracy, i.e., the number of completed over total passes, or style of play. Pass accuracy has been shown to serve as an important indicator of performance as it increases a team's overall ball possession and thereby raises the chances of scoring (Collet, 2013; Oberstone, 2009; Redwood-Brown, 2008). As displayed in columns (5) and (6) of Appendix Table 5.7, we find a statistically significant — albeit very small — negative effect of PM10 concentration on pass accuracy once the full set of control variables is included. We further test whether players adjust their style of play — measured by the ratio of long to short passes — when exposed to high

Table 5.4: The effect of air pollution on performance: Heterogeneous effects by player characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Ln(PM10)	0.016	0.014				
	(0.016)	(0.017)				
Ln(PM10) × Age	-0.001**	-0.001**				
	(0.001)	(0.001)				
Ln(PM10) × (Age < 21)			-0.007	-0.004		
			(0.010)	(0.010)		
Ln(PM10) ×(Age 21–25)			-0.011**	-0.018***		
			(0.005)	(0.006)		
Ln(PM10) ×(Age 26–30)			-0.018***	-0.023***		
			(0.005)	(0.006)		
Ln(PM10) × (Age > 30)			-0.024***	-0.032***		
			(0.007)	(0.007)		
Ln(PM10) × Defender					-0.020***	-0.026***
					(0.005)	(0.006)
Ln(PM10) × Midfielder					-0.015***	-0.022***
					(0.004)	(0.005)
Ln(PM10) × Striker					-0.009	-0.015**
					(0.006)	(0.006)
Player controls	Yes	Yes	Yes	Yes	Yes	Yes
Match controls	No	Yes	No	Yes	No	Yes
Coach FE	No	Yes	No	Yes	No	Yes
Team × Season FE	No	Yes	No	Yes	No	Yes
Observations	75,163	75,163	75,163	75,163	75,163	75,163
Adjusted R-Squared	0.746	0.743	0.746	0.743	0.746	0.743

Note: Dependent variable: Log number of passes. All regressions include player fixed effects. Player controls: age (squared), tenure (squared), position (defender, midfielder, striker), minutes played (squared), home match indicator. Weather controls on daily basis: maximum temperature (squared), precipitation (squared), dew point (squared), wind speed (squared), air pressure (squared). Match controls: day of week, kick-off time, stadium attendance. Standard errors are clustered at the match level (N=2,956). The usual significance levels apply: 0.1 (*), 0.05 (**), and 0.01 (***).

levels of pollution. While passing the ball over longer distances is riskier in terms of retaining ball possession, it may reduce the overall physical burden. As displayed in column (7) and (8) of Appendix Table 5.7, we find that pollution slightly increases this ratio.

Team- and match-level effects. So far, our analysis has been conducted at the individual athlete level. However, a given football player's performance should also be dependent upon his team-mates' and opponents' performance levels, which are also affected by pollution. Hence, we might miss important interactions within a player's team and/or with the opponent when using variation at the player level. While negative effects of pollution on opponents' performance should bias our estimates towards zero (as it becomes easier to pass), the opposite may be true for the effect of pollution on team-mates' performance. To account for this potential concern, we estimate equation (5.1) at the aggregate team and match level, respectively. Interaction effects between

Table 5.5: The effect of air pollution on performance: Team- and match-level effects

	Team-level regressions			Match-level regressions		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(PM10)	-0.029***	-0.020***	-0.015***	-0.025***	-0.030***	-0.017***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Player controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather/Ozone controls	Yes	Yes	Yes	Yes	Yes	Yes
Match controls	No	Yes	Yes	No	Yes	Yes
Coach FE	No	Yes	Yes	No	No	No
Team × Season FE	No	Yes	Yes	No	No	No
Opponent × Home FE	No	No	Yes	No	No	No
Season FE	No	No	No	No	No	Yes
Observations	5,912	5,912	5,912	2,956	2,956	2,956
Adjusted R-Squared	0.091	0.307	0.426	0.075	0.098	0.193

Note: Dependent variable: Log number of passes. Player characteristics: mean age (squared) and tenure (squared) at team or match level, respectively. Weather controls on daily basis: maximum temperature (squared), precipitation (squared), dew point (squared), wind speed (squared), air pressure (squared). Match controls: day of week, kick-off time, stadium attendance. Standard errors are clustered at the match level. The usual significance levels apply: 0.1 (*), 0.05 (**), and 0.01 (***).

levels of PM10 concentration, the individual's and his team-mates' and/or opponents' performance are explicitly accounted for at these aggregate levels.

Displayed in Table 5.5, we find that team- and match-level estimates are of similar magnitude and hence in line with our baseline results. Controlling for player and weather characteristics, column (1) suggests that a 1% increase in the level of pollution reduces the total number of team passes by around 0.03%. This effect declines to -0.015% when subsequently adding match-level, team$\times$season or opponent$\times$home fixed effects to the model (see columns (2) and (3)), but remains highly statistically significant. A similar conclusion arises when turning to the match-level regressions displayed in columns (4)–(6). The number of match-level passes statistically significantly decreases with higher concentrations of particulate matter in ambient air, even in our most comprehensive specification including season fixed effects. Overall, these findings suggest that interaction effects between pollution, the individual's and his team-mates' and/or opponents' performance either cancel out or are of minor importance.

In a last exercise, we test whether air pollution impairs the overall attractiveness of the game, focusing on the number of scored goals per match and a potential asymmetric effect on the odds of winning the game, i.e., whether high levels of pollution affect the probability of an "underdog victory".[20] The corresponding results are displayed in Appendix Table 5.9. While the linear effect on goals per match is close to zero and not statistically significant, pollution values above the EU threshold of 50 $\mu g/m^3$ substantially reduce the number of scored goals per match. Ceteris paribus, the coefficient indicates that an increase in the level of pollution from 50 to 60 $\mu g/m^3$ would reduce the number of goals per match by about 0.2 goals, which is sizable compared to the

[20] We define the match result as an underdog victory in case a team finishing among the top three in a given season loses against a lower-ranked team. Out of 839 matches, 18% are won by the underdog.

overall mean of 2.9 goals per match. At the same time, we do not find evidence of underdogs being more likely to win. This is consistent with the notion that pollution has a symmetrical effect on players of both teams and does not affect their relative strength.

5.5 Conclusion

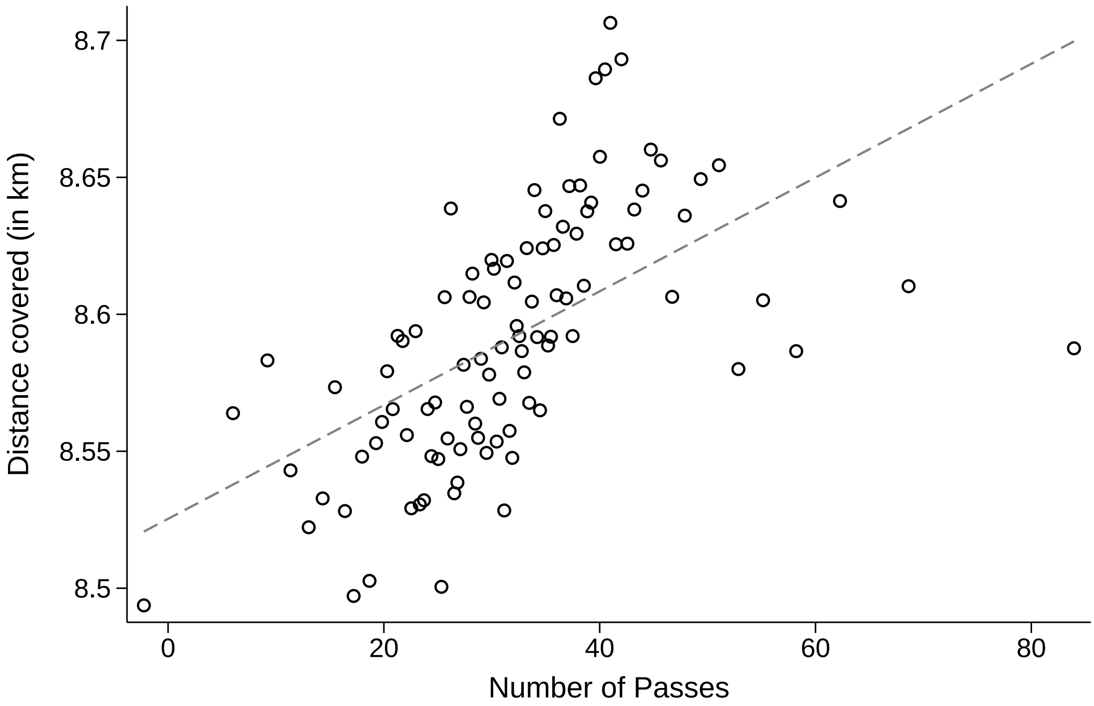
In this chapter, we estimate the causal effect of ambient air pollution on individual performance. Using panel data on the universe of football players in the German *Bundesliga* and hourly information on the concentration of particulate matter and ozone in spatial proximity to the stadium at the hour of kick-off, we exploit exogenous variation in players' exposure to air pollution due to match scheduling rules that are beyond the control of the players and teams.

Our results show that the concentration of particulate matter in ambient air has a negative effect on players' performance, measured by the total number of passes. While the linear effect is statistically significant and robust, the magnitude is quite small. When allowing for a non-linear dose-response relationship, we find substantial negative effects: performance significantly decreases when the concentration of particulate matter exceeds the EU regulatory threshold of 50 micrograms per cubic meter. Nevertheless, we find significant negative effects at levels well below this threshold, emerging at around 15 micrograms per cubic meter. When accounting for heterogeneous effects across players, we further find that the overall effect is mainly driven by players of relatively older age and those playing in positions that require more physical exertion.

Overall, our analysis complements previous empirical evidence on the negative effects of air pollution on the performance of low-skilled agricultural and factory workers from single-plant case studies (Chang et al., 2016; Graff Zivin and Neidell, 2012; He et al., 2016). Even moderate concentrations of particulate matter commonly experienced in developed countries impede the performance of professional athletes to a considerable extent. While our data allows us to consistently measure individual performance over a long period, generalizing our findings beyond the realm of sports is not straightforward. As young professional athletes are in better shape than the average individual, we might expect to observe even stronger effects of pollution on other workers' performance. However, professional sports is also associated with substantially higher physical strain compared to more common occupations, suggesting that athletes are particularly vulnerable to ambient air pollution. Hence, future research on the effects of air pollution for different types of workers may broaden our knowledge on the benefits of environmental regulation.

5.A Appendix

Figure 5.5: Relationship between passes and running distance



Note: Relationship between players' running distance and number of passes, controlling for player fixed effects and the number of minutes played. The data stem from the seasons 2013/2014 to 2015/2016 and were obtained from the German football magazine "kicker", see www.kicker.de. The graph is based on Stata's "binscatter" command, which groups the number of passes into 100 equally sized bins. Data bins are depicted by circles, the dashed line indicates a linear fit.

Table 5.6: The effect of air pollution on performance: Alternative specifications

Sample	All players				Full-time players		All players	
Outcome	Ln(passes)		Passes		Ln(passes)		Passes per minute	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PM10 ($\mu g/m^3$)	-0.001***	-0.001***	-0.016***	-0.025***				
	(0.000)	(0.000)	(0.004)	(0.005)				
Ln(PM10)					-0.015***	-0.021***	-0.005***	-0.007***
					(0.004)	(0.005)	(0.001)	(0.002)
Player characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	No	Yes	No	Yes	No	Yes	No	Yes
Match controls	No	Yes	No	Yes	No	Yes	No	Yes
Coach FE	No	Yes	No	Yes	No	Yes	No	Yes
Team × Season FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	75,163	75,163	75,163	75,163	43,346	43,346	75,163	75,163
Adjusted R-Squared	0.746	0.743	0.436	0.465	0.017	0.083	0.142	0.060

Note: All regressions include player fixed effects. Player characteristics: age (squared), tenure (squared), position dummies (defender, midfielder, striker), minutes played (squared), home match indicator. Weather controls on daily basis: maximum temperature (squared), precipitation (squared), dew point (squared), wind speed (squared), air pressure (squared). Match controls: day of week, kick-off time, stadium attendance. Standard errors are clustered at the match level. The usual significance levels apply: 0.1 (*), 0.05 (**), and 0.01 (***).

Table 5.7: The effect of air pollution on alternative outcomes

Outcome	Minutes Played		Playing full-time		Pass accuracy		Pass ratio	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(PM10)	-0.100	-0.038	-0.000	-0.000	-0.002	-0.003*	0.001	0.003***
	(0.089)	(0.094)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Player characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather/Ozone controls	No	Yes	No	Yes	No	Yes	No	Yes
Match controls	No	Yes	No	Yes	No	Yes	No	Yes
Coach FE	No	Yes	No	Yes	No	Yes	No	Yes
Team × Season FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	75,163	75,163	75,163	75,163	73,832	73,832	73,718	73,718
Adjusted R-Squared	0.038	0.064	0.027	0.045	0.002	0.026	0.019	0.036

Note: All regressions include player fixed effects. Player characteristics: age (squared), tenure (squared), position dummies (defender, midfielder, striker), minutes played (squared), home match indicator. Weather controls on daily basis: maximum temperature (squared), precipitation (squared), dew point (squared), wind speed (squared), air pressure (squared). Match controls: day of week, kick-off time and stadium attendance. Standard errors are clustered at the match level. The usual significance levels apply: 0.1 (*), 0.05 (**), and 0.01 (***).

Table 5.8: The effect of air pollution on performance: Alternative specifications of weather controls

	(1)	(2)	(3)	(4)	(5)	(6)
Ln(PM10)	-0.022***	-0.020***	-0.023***	-0.021***	-0.022***	-0.020***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Ln(Ozone)	-0.006**	-0.004	-0.007**	-0.005	-0.008**	-0.006*
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)
Maximum temperature (°C)	-0.005	-0.007	-0.002	-0.005		
	(0.006)	(0.006)	(0.012)	(0.012)		
Precipitation (mm)	-0.002	-0.000	-0.002	-0.005		
	(0.003)	(0.003)	(0.007)	(0.007)		
Dew point (°C)	0.003	0.003	-0.002	0.000		
	(0.005)	(0.006)	(0.007)	(0.007)		
Wind speed (m/sec)	-0.004	-0.007**	-0.002	-0.007		
	(0.003)	(0.003)	(0.008)	(0.008)		
Air pressure(hpa)	0.006**	0.000	-0.156	-0.112		
	(0.003)	(0.003)	(0.238)	(0.239)		
Maximum temperature^2			-0.003	-0.003		
			(0.011)	(0.011)		
Precipitation^2			-0.000	0.005		
			(0.008)	(0.007)		
Dew point^2			0.008	0.007		
			(0.005)	(0.005)		
Wind speed^2			-0.001	0.001		
			(0.008)	(0.008)		
Air pressure^2			0.162	0.112		
			(0.237)	(0.238)		
Air pressure dummies (omitted: Air pressure < 975)						
$975 \leq$ Air pressure < 1000					0.014*	0.003
					(0.008)	(0.009)
$1000 \leq$ Air pressure < 1010					0.024***	0.008
					(0.009)	(0.009)
Air pressure ≥ 1010					0.014	-0.003
					(0.009)	(0.009)
Wind speed dummies (omitted: Wind speed < 2.5)						
$2.5 \leq$ Wind speed > 3.5					0.007	0.003
					(0.007)	(0.007)
$3.5 \leq$ Wind speed > 4.5					0.003	-0.001
					(0.008)	(0.008)
Wind speed ≥ 4.5					-0.003	-0.010
					(0.008)	(0.008)
Precipitation dummies (omitted: Precipitation = 0)						
$0 <$ Precipitation < 2.5					0.001	0.005
					(0.007)	(0.007)

Table 5.8 — continued

$2.5 \leq$ Precipitation < 10					-0.014*	-0.014*
					(0.009)	(0.008)
Precipitation ≥ 10					-0.002	0.012
					(0.015)	(0.014)
Max. temperature dummies (omitted: Max. temperature < 7.5)						
$7.5 \leq$ Max. temp. < 15					-0.012	-0.014
					(0.009)	(0.009)
$15 \leq$ Max. temp < 22.5					-0.013	-0.010
					(0.012)	(0.012)
Max. temp. ≥ 22.5					-0.018	-0.019
					(0.016)	(0.016)
Player controls	Yes	Yes	Yes	Yes	Yes	Yes
Match controls	Yes	Yes	Yes	Yes	Yes	Yes
Team $\times$ Season FE	No	Yes	No	Yes	No	Yes
Coach FE	No	Yes	No	Yes	No	Yes
Observations	75,163	75,163	75,163	75,163	75,163	75,163
Adjusted R-Squared	0.746	0.756	0.746	0.756	0.746	0.756

Note: Dependent variable: Log number of passes. All regressions include player fixed effects. Player characteristics: age (squared), tenure (squared), position (defender, midfielder, striker), minutes played (squared), home match indicator. Match controls: day of week, kick-off time, stadium attendance. Standard errors are clustered at the match level (N=2,956). The usual significance levels apply: 0.1 (*), 0.05 (**), and 0.01 (***).

Table 5.9: The effect of air pollution on alternative outcomes at the match level

	Goals per match		Underdog victory	
	(1)	(2)	(3)	(4)
Ln(PM10)	0.005	0.012	-0.025	-0.022
	(0.057)	(0.064)	(0.025)	(0.027)
Ln(PM10) $\times$ (PM10 > 50)		-0.964**		-0.087
		(0.454)		(0.167)
Player Controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Match controls	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	2,956	2,956	839	839
Adjusted R-Squared	0.009	0.010	0.089	0.087

Note: Player controls: mean age (squared) and tenure (squared) at the match level. Weather controls on daily basis: maximum temperature (squared), precipitation (squared), dew point (squared), wind speed (squared), air pressure (squared). Match controls: day of week, kick-off time, stadium attendance. Standard errors are clustered at the match level. The usual significance levels apply: 0.1 (*), 0.05 (**), and 0.01 (***).

Bibliography

Aaberge, R., U. Colombino, E. Holmøy, B. Strøm, and T. Wennemo (2007). "Population Ageing and Fiscal Sustainability: Integrating Detailed Labour Supply Models with CGE Models". In: *Modelling Our Future: Social Security and Taxation*. Ed. by A. Harding and A. Gupta. Vol. I. Amsterdam: Elsevier, pp. 259–290.

Adhvaryu, A., N. Kala, and A. Nyshadham (2014). The Light and the Heat: Productivity Co-benefits of Energy-saving Technology. mimeo.

Altig, D., A. J. Auerbach, L. J. Kotlikoff, K. A. Smetters, and J. Walliser (2001). Simulating Fundamental Tax Reform in the United States. *American Economic Review* **91** (3), 574–595.

Ang, J. B. and J. B. Madsen (2015). Imitation versus innovation in an aging society: international evidence since 1870. *Journal of Population Economics* **28** (2), 299–327.

Angrist, J. D. and J.-S. Pischke (2010). The Credibility Revoultion in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics. *Journal of Economic Perspectives* **24** (2), 3–30.

Atkinson, A. (2013). Wealth and inheritance in Britain from 1896 to the present. CASEpaper 178, London School of Economics.

Atkinson, A. and J. E. Stiglitz (1976). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics* **6** (1-2), 55–75.

— (2015). Lectures on Public Economics. Princeton University Press.

Auerbach, A. J. (1985). The Theory of Excess Burden and Optimal Taxation. In: *Handbook of Public Economics*. Ed. by A. J. Auerbach and M. Feldstein. Vol. 1. Amsterdam: North-Holland. Chap. 2, 61–127.

Auerbach, A. J. and J. R. Hines (2002). Taxation and Economic Efficiency. In: *Handbook of Public Economics*. Ed. by A. J. Auerbach and M. Feldstein. Vol. 3. North-Holland. Chap. 21, 1347–1421.

Bach, S. (2005). Mehrwertsteuerbelastung der privaten Haushalte. Deutsches Institut für Wirtschaftsforschung (DIW), Data Documentation 10.

Bach, S., P. Haan, O. Hofmeister, and V. Steiner (2006). Increasing the Value-Added Tax to Re-Finance a Reduction of Social Security Contributions? A Behavioral Microsimulation Analysis for Germany. DIW Berlin.

Bach, S., H. Houben, R. Maiterth, and R. Ochmann (2014). Aufkommens- und Verteilungswirkungen von Reformalternativen für die Erbschaft- und Schenkungsteuer. DIW Berlin: Politikberatung kompakt 83.

Banks, J., R. Blundell, and A. Lewbel (1997). Quadratic Engel Curves and Consumer Demand. *Review of Economics and Statistics* **79** (4), 527–539.

Bargain, O., M. Dolls, C. Fuest, D. Neumann, A. Peichl, N. Pestel, and S. Siegloch (2013). Fiscal union in Europe? Redistributive and stabilizing effects of a European tax-benefit system and fiscal equalization mechanism. *Economic Policy* **28** (75), 375–422.

Bargain, O. and K. Doorley (2016). The Effect of Social Benefits on Youth Employment: Combining RD and a Behavioral Model. *Journal of Human Resources, forthcoming*.

Bargain, O., K. Orsini, and A. Peichl (2014). Comparing Labor Supply Elasticities in Europe and the US: New Results. *Journal of Human Resources* **49** (3), 723–838.

Bastani, S. and H. Selin (2014). Bunching and non-bunching at kink points of the Swedish tax schedule. *Journal of Public Economics* **109**, 36–49.

Becker, I., J. R. Frick, M. M. Grabka, R. Hauser, P. Krause, and G. G. Wagner (2003). A Comparison of the Main Household Income Surveys for Germany: EVS and SOEP. In: *Reporting on Income Distribution and Poverty*. Ed. by R. Hauser and I. Becker. Heidelberg: Springer, 55–90.

Bernheim, B. D., R. J. Lemke, and J. K. Scholz (2004). Do estate and gift taxes affect the timing of private transfers? *Journal of Public Economics* **88** (12), 2617–2634.

Best, M. C. and H. J. Kleven (2016). Housing Market Responses to Transaction Taxes: Evidence From Notches and Stimulus in the UK. *Review of Economic Studies, forthcoming*.

Beznoska, M. and R. Ochmann (2013). The interest elasticity of household savings: a structural approach with German micro data. *Empirical Economics* **45**, 371–399.

Blumkin, T., B. J. Ruffle, and Y. Ganun (2012). Are income and consumption taxes ever really equivalent? Evidence from a real-effort experiment with real goods. *European Economic Review* **56** (6), 1200–1219.

Blundell, R., A. Duncan, and C. Meghir (1998). Estimating labor supply responses using tax reforms. *Econometrica*, 827–861.

Boadway, R., E. Chamberlain, and C. Emmerson (2010). Taxation of Wealth and Wealth Transfers. In: *Dimensions of Tax Deisgn — The Mirrlees Review*. Ed. by S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles, and J. Poterba. Oxford University Press, 737–814.

Boadway, R., M. Marchand, and P. Pestieau (1994). Towards a theory of the direct-indirect tax mix. *Journal of Public Economics* **55** (1), 71–88.

Böhringer, C., S. Boeters, and M. Feil (2005). Taxation and unemployment: an applied general equilibrium approach. *Economic Modelling* **22** (1), 81–108.

Bönke, T., G. Corneo, and H. Lüthen (2015). Lifetime Earnings Inequality in Germany. *Journal of Labor Economics* **33** (1), 171–208.

Börsch-Supan, A., K. Härtel, and A. Ludwig (2014). Aging in Europe: Reforms, international diversicification and behavioral reactions. *American Economic Review: Papers and Proceedings* **104** (5), 224–29.

Bourguignon, F. and A. Spadaro (2006). Microsimulation as a tool for evaluating redistribution policies. *Journal of Economic Inequality* **4** (1), 77–106.

Brewer, M., J. Browne, and R. Joyce (2011). Child and Working-Age Poverty from 2010 to 2020. IFS Commentary C121.

Brown, J. (2011). Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars. *Journal of Political Economy* **119** (5), 982–1013.

Bundesbank (2008). Preis- und Mengenwirkungen der Mehrwertsteueranhebung zum 1. Januar 2007. Bundesbank Monatsbericht April.

Buscher, H. S., H. Buslei, K. Göggelmann, H. Koschel, T. F. Schmidt, V. Steiner, and P. Winker (2001). Empirical macro models under test. A comparative simulation study of the employment effects of a revenue neutral cut in social security contributions. *Economic Modelling* **18** (3), 455–474.

Cai, L., J. Creedy, and G. Kalb (2006). Accounting for Population Ageing in Tax Microsimulation Modelling by Survey Reweighting. *Australian Economic Papers* **45** (1), 18–37.

Carroll, C. D., J. Slacalek, and K. Tokuoka (2014). The Distribution of Wealth and the MPC - Implications of New European Data. ECB Working Paper No. 1648.

Caspersen, E. and G. Metcalf (1993). Is a Value Added Tax Progressive? Annual versus Lifetime Incidence Measures. NBER Working Paper No. 4387.

CEDEFOP (2012). Future skills supply and demand in Europe: Forecast 2012. European Centre for the Development of Vocational Training Research Paper No. 26. URL: www.cedefop.europa.eu/files/5526_en.pdf.

Chang, T., J. Graff Zivin, T. Gross, and M. Neidell (2016). Particulate Pollution and the Productivity of Pear Packers. *American Economic Journal: Economic Policy* **8** (3), 141–169.

Chay, K. Y. and M. Greenstone (2003). The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession. *Quarterly Journal of Economics* **118** (3), 1121–1167.

Chetty, R., J. Friedman, T.Olsen, and L.Pistaferri (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: evidence from Danish tax records. *Quarterly Journal of Economics* **126** (2), 749–804.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and Taxation: Theory and Evidence. *American Economic Review* **99** (4), 1145–1177.

Collet, C. (2013). The possession game? A comparative analysis of ball retention and team success in European and international football, 2007–2010. *Journal of Sports Sciences* **31** (2), 123–136.

Coneus, K. and C. K. Spiess (2012). Pollution exposure and child health: Evidence for infants and toddlers in Germany. *Journal of Health Economics* **31**, 180–196.

Corneo, G., T. Bönke, and C. Westermeier (2016). Erbschaft und Eigenleistung im Vermögen der Deutschen: Eine Verteilungsanalyse. *Perspektiven der Wirtschaftspolitik* **17** (1), 35–53.

Creedy, J. (1997). Lifetime Inequality and Tax Progressivity With Alternative Income Concepts. *Review of Income and Wealth* **43** (3), 283–295.

Creedy, J. and A. Duncan (2005). Aggregating Labour Supply and Feedback Effects in Microsimulation. *Australian Journal of Labour Economics* **8** (3), 277–290.

Cremer, H. and F. Gahvari (1995). Uncertainty, Optimal Taxation and the Direct versus Indirect Tax Controversy. *The Economic Journal* **105** (432), 1165–1179.

Cremer, H. and P. Pestieau (2006). Wealth transfer taxation: a survey of the theoretical literature. In: *Handbook of the Economics of Giving, Altruism and Reciprocity*. Ed. by S.-C. Kolm and J. M. Ythier. Vol. 2. North Holland, 1107–1134.

Cremer, H., P. Pestieau, and J.-C. Rochet (2001). Direct versus Indirect Taxation: The Design of the Tax Structure Revisited. English. *International Economic Review* **42** (3), 781–799.

Cribb, J., C. Emmerson, and G. Tetlow (2013). Incentives, shocks or signals: labour supply effects of increasing the female state pension age in the UK. IFS Working Paper.

Currie, J. and M. Neidell (2005). Air Pollution and Infant Health: What Can We Learn from California's Recent Experience? *Quarterly Journal of Economics* **120** (3), 1003–1030.

Currie, J., M. Neidell, and J. F. Schmieder (2009). Air pollution and infant health: Lessons from New Jersey. *Journal of Health Economics* **28**, 688–703.

Currie, J. and R. Walker (2011). Traffic Congestion and Infant Health: Evidence from E-ZPass. *American Economic Journal: Applied Economics* **3**, 65–90.

Dahlby, B. (2003). Restructuring the Canadian tax system by changing the mix of direct and indirect taxes. In: *Tax Reform in Canada: Our Path to Greater Prosperity*. Ed. by H. Grubel. The Fraser Institute, 77–108.

De Blander, R., I. Schockaert, A. Decoster, and P. Deboosre (2013). The Impact of Demographic Change on Policy Indicators and Reforms. FLEMOSI Discussion Paper No. 25.

De Mooij, R. and M. Keen (2012). Fiscal Devaluation and Fiscal Consolidation: The VAT in troubled times. NBER Working Paper No. 17913.

Decoster, A., X. Flawinne, and P. Vanleenhove (2014). Generational accounts for Belgium: fiscal sustainability at a glance. *Empirica* **14**, 663–686.

Decoster, A., J. Loughrey, C. O'Donoghue, and D. Verwerft (2010). How regressive are indirect taxes? A microsimulation analysis for five European Countries. *Journal of Policy Analysis and Management* **29** (2), 326–350.

Decoster, A., R. Ochmann, and K. Spiritus (2013). Integrating Indirect Taxation into EUROMOD - Documentation and Results for Germany. EUROMOD Working Paper No. EM 20/13.

Decoster, A., K. D. Swerdt, J. Loughrey, C. O'Donoghue, and D. Verwerft (2009). Comparative analysis of the distributional impact of combined changes in direct and indirect taxes and benefits for five European countries. EUROMOD AIM-AP Deliverable 3.6, University of Leuven.

Dekkers, G., H. Buslei, M. Cozzolino, R. Desmet, J. Geyer, D. Hofmann, M. Raitano, V. Steiner, and F. Verschueren (2010). The Flip Side of the Coin: The Consequences of the European Budgetary Projections on the Adequacy of Social Security Pensions. *European Journal of Social Security* **12** (1), 94–120.

Deschênes, O., M. Greenstone, and J. S. Shapiro (2012). Defensive Investments and the Demand For Air Quality: Evidence From the NOx Budget Program and Ozone Reductions. NBER Working Paper No. 18267.

Destatis (2013). Einkommens- und Verbrauchsstichprobe – Aufgabe, Methode und Durchführung der EVS. Fachserie 15 Heft 7.

Dolls, M., K. Doorley, A. Paulus, H. Schneider, S. Siegloch, and E. Sommer (2015). Fiscal Sustainability and Demographic Change: A Micro Approach for 27 EU countries. IZA Discussion Paper No. 9618.

Dolls, M., C. Fuest, and A. Peichl (2012). Automatic stabilizers and economic crisis: US vs. Europe. *Journal of Public Economics* **96** (3-4), 279–294.

Doorley, K. and N. Pestel (2016). Labour supply after Inheritances and the Role of Expectations. IZA Discussion Paper No. 9822.

Dustmann, C., T. Frattini, and C. Halls (2010). Assessing the Fiscal Costs and Benefits of A8 Migration to the UK. *Fiscal Studies* **31** (1), 1–41.

Ebenstein, A., V. Lavy, and S. Roth (2016). The Long Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution. *American Economic Journal: Applied Economics* **8** (4), 36–65.

EEA (2016). European Environment Agency: Air Quality in Europe – 2016 Report. EEA Report No. 28/2016.

Elinder, M., O. Erixson, and D. Waldenström (2015). Inheritance and Wealth Inequality: Evidence from population registers. Uppsala Center for Fiscal Studies Working Paper No. 2015:3.

Esping-Andersen, G. (1990). The Three Worlds of Welfare Capitalism. Princeton University Press.

European Commission (2006). "Macroeconomic Effects of a Shift from Direct to Indirect Taxation: A simulation for 15 EU Member States". In: *72nd meeting of the OECD Working Party No. 2 on Tax Policy Analysis and Tax Statistics*.

— (2008). The Efficiency of Tax Systems: shifting taxation from labour to consumption. Note for the EPC Working Group on Quality and Public Finances.

European Commission (2012). The 2012 Ageing Report: Economic and budgetary projections for the 27 EU Member States (2010-2060). European Union.

— (2015). The 2015 Pension Adequacy Report: current and future income Adequacy in old age in the EU. Directorate-General for Employment, Social Affairs and Inclusion.

European Council (2002). Barcelona European Council: Presidency Conclusions. 15 and 16 March.

Eurostat (2013). The distributional impact of public services in European countries. Eurostat Methodologies and Working Papers.

Farhi, E. and I. Werning (2010). Progressive estate taxation. *The Quarterly Journal of Economics* **125** (2), 635–673.

Fehr, H. (2000). Pension Reform during the Demographic Transition. *Scandinavian Journal of Economics* **102** (3), 419–443.

Fehr, H., M. Kallweit, and F. Kindermann (2012). Pension reform with variable retirement age: a simulation analysis for Germany. *Journal of Pension Economics and Finance* **11**, 389–417.

Feil, M., S. Klinger, and G. Zika (2006). Sozialabgaben und Beschäftigung. Simulation mit drei makroökonomischen Modellen. IAB Discussion Paper No. 22/2006.

Feil, M. and G. Zika (2005). Less contribution, more employment? What will cutting social-security contributions accomplish? In: *Determinants of employment. The macroeconomic view*. Ed. by C. Dreger, H. P. Galler, and U. Walwei. Schriften des Instituts für Wirtschaftsforschung Halle No. 22. Nomos, 123–156.

Feldstein, M. (1976). On the theory of tax reform. *Journal of Public Economics* **6** (1–2), 77–104.

— (1978). The Welfare Cost of Capital Income Taxation. English. *Journal of Political Economy* **86** (2), 29–51.

Ferrera, M. (1996). The 'Southern Model' of Welfare in Social Europe. *Journal of European Social Policy* **6 (1)**, 17–37.

Figari, F., A. Paulus, and H. Sutherland (2014). Microsimulation and Policy Analysis. In: *Handbook of Income Distribution*. Ed. by A. B. Atkinson and F. Bourguignon. Vol. 2. Elsevier-North Holland. Chap. 25.

Fritzsche, B., R. Kambeck, and H. D. von Loeffelholz (2003). Empirische Analyse der effektiven Inzidenz des deutschen Steuersystems. Untersuchungen des Rheinisch-Westfälischen Instituts für Wirtschaftsforschung, Heft 42. Essen.

Glogowsky, U. (2016). Behavioral Responses to Wealth Transfer Taxation: Bunching Evidence for Germany. URL: http://hdl.handle.net/10419/145922.

Goupille-Lebret, J. and J. Infante (2016). Behavioral responses to inheritance tax: Evidence from notches in France. mimeo. URL: http://www.parisschoolofeconomics.eu/docs/goupille-jonathan/goupille_infante(2015).pdf.

Graff Zivin, J. and M. Neidell (2012). The Impact of Pollution on Worker Productivity. *American Economic Review* **102** (7), 3652–3673.

— (2013). Environment, Health, and Human Capital. *Journal of Economic Literature* **51** (3), 689–730.

Gruber, J. (2011). Public Finance and Public Policy. 3rd ed. Worth Publishers.

Hanna, R. and P. Oliva (2015). The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City. *Journal of Public Economics* **122**, 68–79.

Hatton, T. J. (2013). Refugee and Asylum Migration: A Short Overview. In: *International Handbook on the Economics of Migration*. Ed. by A. Constant and K. F. Zimmermann. Edward Elgar.

He, J., H. Liu, and A. Salvo (2016). Severe Air Pollution and Labor Productivity: Evidence from Industrial Towns in China. IZA Discussion Paper No. 8916.

Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica* **47** (1), 153–161.

Holtz-Eakin, D., D. Joulfaian, and H. S. Rosen (1993). The Carnegie Conjecture: Some Empirical Evidence. *The Quarterly Journal of Economics* **108** (2), 413–435.

Hoopes, J. L., D. H. Reck, and J. Slemrod (2015). Taxpayer Search for Information: Implications for Rational Attention. *American Economic Journal: Economic Policy* **7** (3), 177–208.

Hotz, V. J., C. H. Mullin, and J. K. Scholz (2002). Welfare, Employment, and Income: Evidence on the Effect of Benefit Reductions from California. *American Economic Review* **92** (2), 380–384.

Houben, H. and R. Maiterth (2013). Erbschaftsteuer als "Reichenbesteuerung" mit Aufkommenspotential? *Vierteljahreshefte zur Wirtschaftsforschung* **82** (1), 147–175.

Hoynes, H. W. (1996). Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP. English. *Econometrica* **64** (2), 295–332.

Huisman, C., J. de Beer, R. van der Erf, N. van der Gaag, and D. Kupiszewska (2013). Demographic scenarios 2010 - 2030. NEUJOBS Working Paper D 10.1. http://www.neujobs.eu/publications/working-papers/demographic-scenarios-2010-2030.

Imbens, G. W., D. B. Rubin, and B. I. Sacerdote (2001). Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players. *American Economic Review* **91** (4), 778–794.

Immervoll, H., H. J. Kleven, C. T. Kreiner, and E. Saez (2007). Welfare Reform in European Countries: A Micro-Simulation Analysis. *The Economic Journal* **117** (1), 1–44.

Isen, A., M. Rossin-Slater, and R. Walker (2016). Every Breath You Take – Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970. *Journal of Political Economy, forthcoming*.

Janke, K. (2014). Air pollution, avoidance behaviour and children's respiratory health: Evidence from England. *Journal of Health Economics* **38**, 23–42.

Joulfaian, D. (2004). Gift taxes and lifetime transfers: time series evidence. *Journal of Public Economics* **88** (9), 1917–1929.

Joulfaian, D. (2005). Choosing between gifts and bequests: How taxes affect the timing of wealth transfers. *Journal of Public Economics* **89**, 2069–2091.

Joulfaian, D. and K. McGarry (2004). Estate and Gift Tax Incentives and Inter Vivos Giving. *National Tax Journal* **57** (2), 429–444.

Kahn, L. (2000). The Sports Business as a Labor Market Laboratory. *Journal of Economic Perspectives* **14** (3), 75–94.

Kaplow, L. (2001). A Framework for Assessing Estate and Gift Taxation. In: *Rethinking Estate and Gift Taxation*. Ed. by W. Gale, J. Hines, and J. Slemrod. Brookings Institution Press.

KC, S., B. Barakat, A. Goujon, V. Skirbekk, W. Sanderson, and W. Lutz (2010). Projection of populations by level of educational attainment, age, and sex for 120 countries for 2005-2050. *Demographic Research* **22**, 383–472.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark. *Econometrica* **79** (3), 651–692.

Kleven, H. J. (2016). Bunching. *Annual Review of Economics* **8** (1).

Kleven, H. J., C. Landais, and E. Saez (2013). Taxation and International Migration of Superstars: Evidence from the European Football Market. *American Economic Review* **103** (5), 1892–1924.

Kohn, K. and M. Missong (2003). Estimation of Quadratic Expenditure Systems Using German Household Budget Data. *Jahrbücher für Nationalökonomie und Statistik* **223** (4), 422–448.

Kopczuk, W. (2007). Bequest and Tax Planning: Evidence from Estate Tax Returns. *Quarterly Journal of Economics* **122** (4), 1801–1854.

— (2010). Economics of Estate Taxation: A Brief Review of Theory and Evidence. NBER Working Paper No. 15741.

— (2013). Taxation of Intergenerational Transfers and Wealth. In: *Handbook of Public Economics*. Ed. by A. Auerbach, R. Chetty, M. Feldstein, and E. Saez. Vol. 5. North Holland, 329–390.

Kopczuk, W. and J. Slemrod (2001). The Impact of the Estate Tax on the Wealth Accumulation and Avoidance Behavior of Donors. In: *Rethinking Estate and Gift Taxation*. Ed. by W. G. Gale, J. R. Hines, and J. Slemrod. Brookings Institution, 299–343.

— (2003). Dying to Save Taxes: Evidence from Estate-Tax Returns on the Death Elasticity. *Review of Economics and Statistics* **85** (2), 256–265.

Koske, I. (2013). Fiscal Devaluation - Can it help to boost competetiveness? OECD Economics Department Working Paper No. 1089.

Kotlikoff, L. J., K. Smetters, and J. Walliser (2007). Mitigating America's demographic dilemma by pre-funding social security. *Journal of Monetary Economics* **54** (2), 247–266.

Kudrna, G., C. Tran, and A. Woodland (2016). Facing Demographic Challenges: Pension Cuts or Tax Hikes. Cesifo Working Paper No. 5644.

Lambert, P. J. (2001). The Distribution and Redistribution of Income. 3rd ed. Manchester University Press.

Leombruni, R. and M. Richiardi (2006). LABORsim: An Agent-Based Microsimulation of Labour Supply. An application to Italy. *Computational Economics* **27**, 63–88.

Levitt, S. D. and J. A. List (2007). What do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? *Journal of Economic Perspectives* **21** (2), 153–174.

Lichter, A., A. Peichl, and S. Siegloch (2015a). The Own-Wage Elasticity of Labor Demand: A Meta-Regression Analysis. *European Economic Review* **80**, 94–119.

Lichter, A., N. Pestel, and E. Sommer (2015b). Productivity Effects of Air Pollution: Evidence from Professional Soccer. IZA Discussion Paper No. 8964.

Liu, L. and B. Lockwood (2015). VAT Notches. CESifo Working Paper Series No. 5371.

Lleras-Muney, A. (2010). The Needs of the Army – Using Compulsory Relocation in the Military to Estimate the Effect of Air Pollutants on Children's Health. *Journal of Human Resources* **45** (3), 549–590.

Löffler, M., A. Peichl, N. Pestel, S. Siegloch, and E. Sommer (2014a). Documentation IZAΨMOD v3.0: The IZA Policy Simulation Model. IZA Discussion Paper No. 8553.

Löffler, M., A. Peichl, and S. Siegloch (2014b). Discrete Choice Labor Supply Models and Wage Exogeneity. IZA Discussion Paper No. 8281.

Luechinger, S. (2014). Air pollution and infant mortality: A natural experiment from power plant desulfurization. *Journal of Health Economics* **37**, 219–231.

Mankiw, N. G., M. Weinzierl, and D. Yagan (2009). Optimal Taxation in Theory and Practice. *Journal of Economic Perspectives* **23** (4), 147–174.

Manoli, D., K. J. Mullen, and M. Wagner (2015). Policy Variation, Labor Supply Elasticities, and a Structural Model of Retirement. *Economic Inquiry* **53** (4), 1702–1717.

Manoli, D. and A. Weber (2016). The Effects of the Early Retirement Age on Retirement Decisions. IZA Discussion Paper No. 10154.

Mara, I. and E. Narazani (2011). Labour-incentive reforms at preretirement age in Austria. *Empirica* **38**, 481–510.

Marcus, J., R. Siegers, and M. M. Grabka (2013). Preparation of Data from the New SOEP Consumption Module. DIW Data Documentation No. 70.

Martinez-Vazquez, J., V. Vulovic, and Y. Liu (2010). Direct versus Indirect Taxation: Trends, Theory and Economic Significance. International Studies Program, Andrew Young School of Policy Studies, Georgia State University. International Studies Program Working Paper Series at AYSPS.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In: *Frontiers in Econometrics*. Ed. by P. Zarembka. New York: Academic Press, 105–142.

McGarry, K. (2001). The cost of equality: unequal bequests and tax avoidance. *Journal of Public Economics* **79**, 179–204.

Meinhardt, V. and R. Zwiener (2005). Gesamtwirtschaftliche Wirkungen einer Steuerfinanzierung versicherungsfremder Leistungen in der Sozialversicherung. DIW Politikberatung Kompakt 7.

Metcalf, G. (1994). Life Cycle versus Annual Perspectives on the Incidence of a Value Added Tax. In: *Tax Policy and The Economy*. Ed. by J. M. Poterba. Vol. 8. MIT Press, 45–64.

Moscarola, F. C., U. Colombino, F. Figari, and M. Locatelli (2015). Shifting Taxes from Labour to Property: A Simulation under Labour Market Equilibrium. IZA Discussion Paper No. 8832.

Musgrave, R. A. (1959). The Theory of Public Finance. McGraw Hill.

— (2008). Public Finance. In: *The New Palgrace Dictionary of Economics*. Ed. by S. N. Durlauf and L. E. Blume. 2nd ed. Chap. Public Finance.

Naito, H. (2007). Atkinson-Stiglitz Theorem with Endogenous Human Capital Accumulation. *The B.E. Journal of Economic Analysis & Policy* **7** (1).

Nordblom, K. and H. Ohlsson (2006). Tax avoidance and intra-family transfers. *Journal of Public Economics* **90** (8–9), 1669–1680.

Oberstone, J. (2009). Differentiating the Top English Premier League Football Clubs from the Rest of the Pack: Identifying the Keys to Success. *Journal of Quantitative Analysis in Sports* **5** (3), 10.

O'Donoghue, C., ed. (2014). Handbook of Microsimulation Modelling. Emerald.

O'Donoghue, C., M. Baldani, and D. Mantovani (2004). Modelling the Redistributive Impact of Indirect Taxes in Europe: An Application of Euromod. EUROMOD Working Papers No. EM7/01.

OECD (2007). Consumption Taxes: the Way to the Future? OECD Policy Brief.

— (2010). Choosing a Broad Base - Low Rate Approach to Taxation. OECD Tax Policy Studies No. 19.

— (2013a). Pensions at a Glance 2013. OECD.

— (2013b). Revenue Statistics.

— (2014). Taxing Wages 2014. OECD Publishing.

OECD/Korea Institute of Public Finance (2014). The Distributional Effects of Consumption Taxes in OECD Countries. OECD Tax Policy Studies No. 22, OECD Publishing.

Ohlsson, H., J. Roine, and D. Waldenström (2014). Inherited Wealth over the Path of Development: Sweden, 1810 – 2010. IFN Working Paper No. 1033.

Page, B. R. (2003). Bequest Taxes, inter vivo gifts and the bequest motive. *Journal of Public Economics* (87), 1219–1229.

Parsons, C. A., J. Sulaeman, M. C. Yates, and D. S. Hamermesh (2011). Strike Three: Discrimination, Incentives, and Evaluation. *American Economic Review* **101** (4), 1410–1435.

Peichl, A. and S. Siegloch (2012). Accounting for Labor Demand Effects in Structural Labor Supply Models. *Labour Economics* **19** (1), 129–138.

Pestel, N. and E. Sommer (2017). Shifting Taxes from Labor to Consumption: More Employment and more Inequality? *forthcoming in: Review of Income and Wealth*.

Picos-Sanchez, F. and A. Thomas (2015). A Revenue-neutral Shift from SSC to VAT: Analysis of the Distributional Impact for 12 EU-OECD Countries. *FinanzArchiv: Public Finance Analysis* **71** (2), 278–298.

Piketty, T. (2011). On the Long-Run Evolution of Inheritance: France 1820–2050. *Quarterly Journal of Economics* **126** (3), 1071–1131.

Piketty, T. and E. Saez (2013a). A Theory of Optimal Inheritance Taxation. *Econometrica* **81** (5), 1851–1886.

— (2013b). Optimal Labor Income Taxation. In: *Handbook of Public Economics*. Ed. by A. J. Auerbach, R. Chetty, M. Feldstein, and E. Saez. Vol. 5. Elsevier, 391–474.

Piketty, T. and G. Zucman (2014). Capital is Back: Wealth-Income Ratios in Rich Countries 1700–2010. *The Quarterly Journal of Economics* **129** (3), 1255–1310.

Pope III, C. A. (2000). Epidemiology of Fine Particulate Air Pollution and Human Health: Biologic Mechanisms and who is at Risk? *Environmental Health Perspectives* **108** (4), 713–723.

Pope III, C. A. and D. W. Dockery (2006). Health Effects of Fine Particulate Air Pollution: Lines that Connect. *Journal of the Air and Waste Management Association* **56** (6), 709–742.

Poterba, J. (2001). Estate and gift taxes and incentives for inter vivos giving in the US. *Journal of Public Economics* **79**, 237–264.

Prettner, K. and A. Prskawetz (2010). Demographic change in models of endogenous economic growth. A survey. *Central European Journal of Operations Research* **18** (4), 593–608.

Redwood-Brown, A. (2008). Passing Patterns Before and After Goal Scoring in FA Premier League Soccer. *International Journal of Performance Analysis in Sport* **8** (3), 172–182.

Reis, D. (2005). Von der Sterbefallanzeige zur Steuerfestsetzung. In: *Erbschaft- und Schenkungsteuerstatistik 2002 — Möglichkeiten und Grenzen*. Ed. by N. Zifonum and J. Schupp. Statistisches Bundesamt, 8–24.

Ricardo, D. (1821). On the Principles of Political Economy and Taxation. 3rd ed. reprinted by: Dover Publications, 2004.

Richter, W. and R. Boadway (2005). Trading Off Tax Distortion and Tax Evasion. *Journal of Public Economic Theory* **7** (3), 361–381.

Rosenqvist, O. and O. N. Skans (2015). Confidence Enhanced Performance? The Causal Effects of Success on Future Performance in Professional Golf Tournaments. *Journal of Economic Behavior and Organization* **117**, 281–295.

Rundell, K. W. (2012). Effect of Air Pollution on Athlete Health and Performance. *British Journal of Sports Medicine* **46**, 407–412.

RWI and FiFo (2007). Der Zusammenhang zwischen Steuerlast- und Einkommensverteilung. RWI Projektberichte, Forschungsprojekt für das Bundesministerium für Arbeit und Soziales. URL: http://www.rwi-essen.de/media/content/pages/publikationen/rwi-projektberichte/PB_Steuerlast-und-Einkommensverteilung.pdf.

Saez, E. (2002). The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics* **83** (2), 217–230.

— (2010). Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy* **2** (3), 180–212.

Saez, E., M. Matsaganis, and P. Takloglou (2012). Earnings Determination and Taxes: Evidence from a cohort-based payroll tax reform in Greece. *Quarterly of Journal of Economics* **127**, 493–533.

Sandmo, A. (2014). The Principal Problem in Political Economy: Income Distribution in the History of Economic Thought. In: *Handbook of Income Distribution*. Ed. by A. B. Atkinson and F. Bourguignon. Vol. 2. Elsevier, 3–66.

Sausgruber, R. and J.-R. Tyran (2005). Testing the Mill hypothesis of fiscal illusion. *Public Choice* **122** (1), 39–68.

Savage, M. and T. Callan (2015). Modelling the Impact of Direct and Indirect Taxes Using Complementary Datasets. IZA Discussion Paper No. 8897.

Schinke, C. (2012). Inheritance in Germany 1911 to 2009: A Mortality Multiplier Approach. DIW SOEPpapers No. 462.

Schlenker, W. and W. R. Walker (2016). Aiports, Air Pollution, and Contemporaneous Health. *Review of Economic Studies* **83** (2), 768–809.

Siegloch, S. (2013). Labor Market Effects of Public Policy. PhD thesis. Universität zu Köln.

Smith, A. (1776). The Wealth of Nations. New York: Modern Library Edition.

Sørensen, P. B. (2007). The theory of optimal taxation: what is the policy relevance? *International Tax and Public Finance* **14** (4), 383–406.

Staubli, S. and J. Zweimüller (2013). Does raising the early retirement age increase employment of older workers? *Journal of Public Economics* **108**, 17–32.

Steiner, V. (1996). Employment and Wage Effects of Social Security Financing. ZEW Discussion Paper No. 96-14.

Suits, D. (1977). Measurement of Tax Progressivity. *American Economic Review* **67** (4), 747–752.

Sutherland, H. and F. Figari (2013). EUROMOD: the European Union tax-benefit microsimulation model. *International Journal of Microsimulation* **6** (1), 4–26.

Tanaka, S. (2015). Environmental regulations on air pollution in China and their impact on infant mortality. *Journal of Health Economics* **42**, 90–103.

Thomas, A. and F. Picos-Sanchez (2012). Shifting from Social Security Contributions to Consumption Taxes: The impact on Low-Income Earner Work Incentives. OECD Taxation Working Papers, No. 11, OECD Publishing.

Tresch, R. W. (2015). Public Finance — A Normative Theory. 3rd ed. Elsevier.

UN (2012). COICOP - Classification of Individual Consumption According to Purpose. http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=5&Lg=1.

Vestad, O. L. (2012). Labour supply effects of early retirement provision.

Wagner, G. G., J. R. Frick, and J. Schupp (2007). The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements. DIW SOEPpapers, No. 1.

Ziebarth, N. R., M. Schmitt, and M. Karlsson (2013). The Short-Term Population Health Effects of Weather and Pollution: Implications of Climate Change. IZA Discussion Paper No. 7875.

Curriculum Vitae

Personal Details

Name	Eric Sommer
Date of Birth	14 March 1987
Citizenship	German

Professional career

Since 08/2016	**Research Associate** IZA Institute of Labor Economics, Bonn
11/2013 – 02/2014	**Secondee** OECD, Directorate for Employment, Labor and Social Affairs, Paris
08/2012 – 07/2016	**Resident Research Affiliate, Doctoral Student** IZA Institute of Labor Economics, Bonn

Education

Since 10/2012	**Ph.D. candidate in Economics** University of Cologne, Supervisor: Felix Bierbrauer
11/2008 – 05/2012	**Diplom-Volkswirt** (equiv. M.A. Economics and Political Science) University of Cologne
01/2010 – 07/2010	**Graduate Studies in Economics and Political Science** University of Copenhagen
04/2007 – 10/2008	**Vordiplom** University of Cologne
06/2006	**Abitur** Rotteck-Gymnasium, Freiburg

Publications

Refereed Journals

2017 Shifting Taxes from Labor to Consumption: More Employment and more Inequality?, Review of Income and Wealth, forthcoming (with N. Pestel).

Refereed German Journals

2016 Ist eine Glättung des Mittelstandsbauchs finanzierbar? Eine Mikrosimulationsstudie, Perspektiven der Wirtschaftspolitik, 17(3), 264–275 (with N. Pestel, R. Schnabel, S. Siegloch and A. Spermann).

2014	Bundestagswahlkampf 2013: Klientelpolitik durch Steuerreform? Perspektiven der Wirtschaftspolitik, 15(2), 182–194. (with A. Peichl, N. Pestel and S. Siegloch).

Non-refereed Publications

2017	Verteilungswirkungen ökonomischer Anreize energetischer Sanierungen, in: Großmann, K. A. Schaffrin and C. Smigiel (eds): Energie und soziale Ungleichheit, Springer, 425–453 (with A.-L. Guske, K. Jacob, N. Pestel and C. Range).
2016	Analyse der Verteilung von Einkommen und Vermögen, German Ministry for Labor and Social Affairs.
	Verteilungswirkungen umweltpolitischer Maßnahmen und Instrumente, UBA Texte 73/2016, Umweltbundesamt (with K. Jacob, A.-L. Guske, S. Weiland, C. Range, N. Pestel and J. Pohlmann).
2015	Productivity Effects of Air Pollution: Evidence from Professional Soccer, IZA Discussion Paper No. 8964 (with A. Lichter and N. Pestel). Fiscal Sustainability and the Demographic Transition: A micro approach for 27 EU countries, IZA Discussion Paper No. 9618 (with K. Doorley, M. Dolls, H. Schneider, S. Siegloch and A. Paulus).
2014	IZAΨMOD 3.0: The IZA Policy Simulation Model, IZA Discussion Paper No. 8553 (with M. Löffler, A. Peichl, N. Pestel and S. Siegloch).

Presentations

International Conferences

2016	IIPF Congress, Lake Tahoe. MaTax Conference, Mannheim.
2015	IMA World Congress, Esch-sur-Alzette.
2014	APPAM International Conference, Segovia; IMA European Congress, Maastricht.
2013	IIPF Congress, Toarmina; NTA Conference, Tampa; SMYE, Aarhus.

Workshops

2016	IZA/CMR Workshop, Bonn, IZA Brownbag Seminar, Bonn.
2015	IZA/OECD Employment Seminar, Paris; IZA Conference on Fiscal Policy Tools and Labor Markets during the Great Recession, Bratislava.
2014	Social Policy Division (SPD) Research Workshop, OECD, Paris.
2013	Center for Macroeconomic Research (CMR) Seminar, U of Cologne; Tax Policy Workshop at Statistics Norway, Oslo; Young Scholar SOEP Symposium, Delmenhorst.
2012	IZA Brown Bag Seminar, Bonn.

Other Professional Activities

Research Projects: Member of the Project team

2015	People to jobs, or jobs to People? (on behalf of Randstad)
2014 – 2015	Analyse der Verteilung von Einkommen und Vermögen in Deutschland, 5. Armuts- und Reichtumsbericht der Bundesregierung (on behalf of the German Ministry for Labour and Social Affairs)
2013 – 2015	Verteilungswirkungen umweltpolitischer Maßnahmen und Instrumente (on behalf of the German Federal Environment Agency)
2012 – 2014	Employment 2025: How will multiple transitions affect the European Labor Market (NEUJOBS, on behalf of the European Commission)
2011	Energiepreisentwicklung und Landnutzung (E-LAN, on behalf of the German Ministry for Education and Research)

Refereeing

European Journal of Political Economy, Public Finance Archive

Ph.D. Courses and Summer Schools

2013	Research Methods in Public Economics — Andreas Peichl, U of Cologne Identification Strategies in Econometrics — David A. Jaeger, U of Cologne Microsimulation Training Course — Alan Duncan, University College London Advanced Microeconomics I — Oliver Gürtler, U of Cologne Advanced Applied Econometrics — Oleg Badunenko, U of Cologne
2012	EUROMOD Training Course, U of Essex 8^{th} Ruhr Graduate Summer School — Trade and Climate Policy Analysis with GAMS and MPSGE — Christoph Böhringer, U Duisburg-Essen