Essays in Public Economics

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

vorgelegt von

Björn Uwe Saß

30. April 2010

Dekan:	Prof. Tom Krebs, Ph.D.
Referent:	Prof. Dr. Eckhard Janeba
Korreferentin:	Prof. Christina Gathmann, Ph.D.

Tag der mündlichen Prüfung: 8. Juni 2010

Acknowledgments

First and foremost, I wish to thank my advisor Eckhard Janeba for his excellent guidance throughout the whole project of writing this thesis and for the encouraging and motivating feedback he has given me. Even in a sometimes very tight schedule, he has given high priority to stimulating discussions of the content of my research as well as to the way of presenting it. It is a pleasure having him as advisor and I am grateful for his dedication to supporting my research and his confidence in me.

I also wish to thank Christina Gathmann for her immediate willingness to support the development process of the last chapter of this thesis. She has encouraged me to do empirical research and has provided me with constructive and inspiring comments.

I have enjoyed a lot the cheerful atmosphere at the chair of Public Finance and would like to thank my colleagues for the good time and the numerous constructive comments. In particular, I thank Jennifer Abel-Koch and Gonzague Vannoorenberghe for many relaxing and inspiring tea and chocolate breaks and I thank Nicole Borheier.

I further wish to thank those colleagues with whom I have cooperated a lot and those who have contributed to so many hilarious visits to the Mensa, to passionate discussions and to the joy of various tours and many beers together. Thus, I thank Galina Boeva, Sarah Borgloh, Daniel Harenberg, Tanja Hennighausen, Heiko Karle, Maryam Kazemi-Manesh, Sebastian Köhne, Michal Kowalik, Moritz Kuhn, Marie Paul, Christoph Rothe, Alexander Ruddies, Melanie Schienle, Elisabeth Schulte-Runne, Heiner Schumacher and Edgar Vogel. Besides, I thank my good friends beyond university.

I am deeply grateful to my family for everything.

Contents

Chapter 1

Introduction

This thesis examines three different topics in the area of public economics. Chapter 2 looks at distributional issues that arise from immigration into an industrialised country and possible mechanisms to compensate native losers. In Chapter 3, I examine in how far a country's tax system affects a firm's decision to evade taxes and produce in the shadow economy. Chapter 4 analyses the impact of a special childcare policy on the parents' decision to make use of publicly subsidised nurseries. Although quite different, the three topics also share some common elements.

In all three chapters, government policies play an important role, in particular tax or transfer schemes. In this field, two strains of the literature can be distinguished. While there is a wide literature on optimal tax or transfer policies, there also is a literature that focuses on policy reforms in an environment where historically grown structures are in place. In this context, reforms seek to improve the status quo recognising that the political process may not allow for the implementation of optimal, fundamentally new policies.

I contribute to the latter strain of the literature by taking specific pre-existing structures into account when analysing the three different topics. Chapter 2 looks at immigration into a welfare state that already has a redistributive policy in place. In particular, I study the consequences of immigration and the necessity to adapt the country's tax transfer scheme to it. In the third chapter, I examine adjustments of existing firm taxes that seek to reduce the extent of the shadow economy while not jeopardising government revenues. Chapter 4 analyses the consequences on the parents' childcare choice of introducing a transfer to those parents of young children who do not make use of publicly subsidised day care for their child. Here, the focus is on the evaluation of the specific policy measure taken.

A second common aspect of the different chapters is heterogeneity. The relevant agents, i.e. citizens in the second chapter, firms in the third chapter or parents of young children in the fourth chapter are heterogeneous groups. As a consequence, a particular tax-

transfer scheme does not affect every single agent in the same way. Rather there are specific agents for which the policy is highly relevant whereas it is less so for others. This aspect is linked to the third commonality of the three following chapters, namely incentives.

By taking a specific tax or transfer measure, the government changes the options of an individual or firm and consequently the behaviour of the agents concerned. The fundamental mechanism is that the government policy makes particular options more or less attractive because the policy influences the relative price of different alternatives. Specifically, this looks as follows.

In the second chapter, the government intends to redistribute from the winners to the losers of immigration. As it cannot directly observe who the winners are, it conditions its redistribution scheme on income with the aim to tax agents with higher incomes and pay the proceeds to those with lower incomes. However, if the burden levied on the earners of high incomes becomes too large, they may decide to work less and pretend being low-income agents. Consequently, the government cannot distinguish winners and losers any more and a redistributive scheme is not implementable. Taking into account the winners' incentives to reveal themselves avoids this difficulty.

The third chapter looks directly at the impact of diverse tax instruments on a firm's decision to operate in the official or the shadow economy. The combination of the tax instruments determines the attractiveness of the two options, respectively. Firms are heterogeneous in that they have different productivities. Here, the focus is on those firms that are indifferent between the official and the shadow economy as they would make the same profit in either sector. If the government can set incentives such that it is more attractive for the indifferent firms to choose the official activity, it can effectively fight the extent of the shadow economy.

In the fourth chapter, a transfer of 150€ paid to those parents who do not make use of publicly subsidised day care for their young child changes the cost of day care. After the introduction of this scheme, parents have to renounce 150€ for using day care which thus becomes more expensive. As a consequence, the policy affects the parents' childcare decision and sets incentives for home care.

Taking all these aspects into consideration, the thesis underlines the importance of accounting for heterogeneity of agents and incentives when designing a tax or transfer scheme. After having discussed the commonalities of the three following chapters, I now proceed to present these chapters in some more detail.

The second chapter examines the impact of immigration on the individuals in a country and analyses ways to compensate the losers.

Globalisation has many aspects. For instance, in recent years we have seen large increases in trade volumes as well as in factor flows with an important aspect besides

trade being immigration. While the social welfare benefits of immigration are well documented and in large parts uncontested, e.g. being it to alleviate the burden of ageing societies or overcoming the lack of particular qualifications, immigration also has important distributional consequences. As the empirical literature, see e.g., Borjas (2003) and Card (2001), documents, there are a lot of individual losers of immigration and it has increased income inequality. As a consequence, a gap between social welfare and individual well-being arises with the possibility that even the majority of citizens can be worse off due to immigration. If, however, a large part of the population loses from immigration, an elected government will have a tough time to realise the overall welfare gains.

I analyse how a sustainable immigration policy can be implemented if the majority of native voters loses from immigration individually. Sustainability in this context means that no citizen should be worse off due to (mainly unskilled) immigration so that political support for the realisation of welfare gains can be maintained. Using a model with agents of two skill types, a government that maximises the utility of the majority of unskilled natives has to exploit two dimensions of heterogeneity to achieve sustainability. It has to redistribute from skilled to unskilled natives and from immigrants to natives. Only if the government uses two different tax instruments to address both channels of redistribution, a sustainable immigration policy is feasible.

The intuitive reason for this result is that using only one tax instrument cannot assure that redistribution is focused on losers. An optimal income tax allows to redistribute from the skilled to the unskilled agents. As, however, the government budget has to be balanced, both, revenues as well as expenditures determine whether an income tax allows to compensate the losers. On the revenue side, the redistributional amounts that can be levied on the skilled agents are limited by the fact that the skilled agents cannot be forced to reveal their skill type and may reduce their effort. Thus, the tax system has to ensure that the two types remain distinguishable. On the expenditure side, the tax and transfer scheme cannot distinguish between native and immigrant unskilled agents. Thus, as the payment is conditioned on income, immigrants benefit fully of the transfers which in turn increases the group of recipients. Taken together, the upper limit to revenues and the large expenditures prevent a compensation of native losers.

On the other hand, if one simply levied a tax on the immigrants and distributed the proceeds to the native agents, the native winners and losers would benefit alike. This makes the transfer rather expensive and the potential immigrants refrain from migrating because they are better off in their country of origin.

By contrast, the combination of both instruments allows to focus the transfers on the native losers and to implement sustainable immigration. Moreover, it guarantees that immigration can be made a Pareto improvement. This is particularly important as for specific parameter combinations all natives may lose from immigration if initially a redistributive tax scheme of the Stiglitz (1982)-type is in place. A scheme that combines

redistribution between skill groups and between natives and immigrants could be implemented by linking a redistributive income tax with immigration fees or with a concept of delayed immigration. The concept of delayed immigration implies that the government excludes immigrants from transfer payments for a certain period of time.

In the third chapter, I look at the shadow economy. Usually, an important part of a country's GDP is generated in shadow activities. For instance, Bühn et al. (2007) estimate the size of the shadow economy to 15% of GDP in Germany in 2005. Taxes are an important reason for economic agents to decide to operate in the informal sector as they increase the cost of official sector activities. I analyse the interaction between different tax instruments and a firm's choice to work in the official or shadow economy. Firms are an important actor as they remit a large part of taxes and social security contributions to the government. In addition, the tax incentives set to firms affect their output choice as well as the size of firms and this has repercussions for the industry structure.

The tax instruments the government disposes of are a linear output tax, a fixed tax, a detection probability and a fine rate that is related to output. Firms are heterogeneous in their productivities and face a fixed cost of production. Using the four tax instruments, I replicate the empirical finding that more productive firms work in the official whereas less productive firms work in the shadow sector. The extent of the shadow economy as well as various industry characteristics such as productivity, overall output, and the number of producing firms are differently affected by the tax instruments.

More specifically, the firm structure in a homogeneous good sector looks as follows. Because of the fixed cost of production, the smallest, most unproductive firms do not enter the market. If firms are sufficiently productive to overcome the fixed cost, they have to decide whether to operate in the official or the shadow economy. If they are comparably small, they have difficulties to pay the fixed tax rate whereas the linear tax and the fine, both related to output, are rather irrelevant for them. In contrast, the larger a firm becomes, the more negligible the fixed tax gets. In this case, firms are more concerned by the proportional tax rate on their large output. If the expected proportional penalty for evading taxes is larger than the regular tax payment due to the linear tax, it is more attractive for large firms to produce officially, pay the fixed tax and save on the expected proportional penalty.

I show that in this framework the fixed tax is crucial for the existence of a shadow economy, whereas the linear tax and the fine rather determine its extent. As a consequence, the shadow economy can be reduced by increasing the linear tax and decreasing the fixed tax in a differential incidence. Such a reform makes use of the different reactions to the tax instruments of firms that are indifferent either between no production and shadow production or between shadow and official activities. As I examine for the particular case of the increase in the German VAT rate in a simulation, the rise of the shadow economy

due to the tax increase could have been reduced if the reform had been accompanied by a reduction in e.g. sales-independent components of the communal excise tax.

The fourth chapter is concerned with family policies with a focus on a particular reform that affects the childcare decision of parents of young children.

Many industrialised countries are confronted with the challenge of demographic change and ageing societies. As a consequence, family policy gains importance as it has the potential to alleviate the lack in qualified employees in at least two repects. Firstly, a sufficient supply of day care is expected to increase female labour force participation and to reduce gender wage inequality. Secondly, good quality childcare is a solid foundation for equal chances at school and successful educational careers (see e.g. Temple and Reynolds (2007) and Currie (2001)). Thus, it contributes to future economic growth.

In this chapter, I analyse the consequences of a particular policy concerning the parents' day-care decision in the German state (Land) of Thuringia. On July 1, 2006, the state introduced a subsidy that I refer to as a cash-for-care subsidy. It implies that parents of two-year old children are eligible for a transfer of 150€ if they do not make use of publicly subsidised day care. In effect, this cash-for-care subsidy can be considered an important increase in the cost of institutionalised day care. In addition, the state has introduced a legal claim to publicly subsidised day care for two-year olds. Given that the median household income in Thuringia is 1,470€ (in 2006), the subsidy constitutes an important transfer for many families.

The focus of my analysis is on the impact of the cash-for-care subsidy on the parents' decision to make use of publicly subsidised day care. Based on county data, I find that the introduction of the subsidy decreases the enrolment share, i.e. the share of two-year olds enrolled in day care to the cohort size, by 9 percentage points based on an initial level of 80%. Thus, the reduction is sizeable. In addition, I detect that the reduction in the enrolment share is smaller if many parents with a college or university degree live in the county whereas it is higher in counties with a large number of young children.

Thuringia is an interesting case to study as it offers a large supply of day care, so that there is no rationing. To identify the impact of the cash-for-care subsidy, I compare the treatment group of two-year olds after the introduction of the subsidy to control groups along three dimensions in a difference-in-difference-in-difference design. The three dimensions are time, age group and state. Hence, I compare the group of two-year olds before the introduction of the cash-for-care subsidy with the cohorts after its introduction. In addition, I compare the two-year olds in Thuringia with those in the neighbouring state of Saxony where no comparable policy has been introduced. Lastly, I take the group of three-year olds that is not eligible for the cash-for-care subsidy in either state as a control group. Besides the strongly negative impact of the subsidy, I identify a trend towards higher enrolment of two-year olds across states.

The German federal government discusses to introduce a similar cash-for-care subsidy for all children younger than three years. Against the background of demographic change and given the incentives that arise from the subsidy in Thuringia, I would rather recommend to invest into additional childcare supply in particular in the Western German states where rationing is severe and the enrolment share for children younger than three is in the range of 10% (against 38% for the Eastern states).

The thesis is organized in such a way that the chapters can be read independently of each other. All references are collected in the bibliography.

Chapter 2

Immigration and Redistribution

2.1 Introduction

In many cases, immigration into Western countries is welfare improving. Immigrants are important employees in the agricultural sector and for a lot of service tasks and they help alleviate the demographic change. While in various aspects the overall advantageous impact of immigration seems to be well established, not all people benefit from it individually. Often, low-skilled workers lose from low-skilled immigration because of negative wage or employment effects, in particular in the short run.[1]

To realise overall welfare gains from immigration, the immigration policy has to be sustainable and therefore needs the support of the large majority of citizens in a country in which citizens can vote on the immigration policy. As individuals base the vote on their personal well-being, differences between individual and social welfare become relevant. As a consequence, if the gains from immigration only benefit a small group of citizens or if immigration leads to an increase in inequality that is undesired by the majority, the sustainability of letting immigrants in will not be ensured. Therefore, it is necessary to consider compensation mechanisms for the losers of immigration to reconcile individual preferences and general welfare improvements and to ensure the realisation of immigration-related welfare gains in the long-run.

This chapter sheds light on ways to realise a sustainable immigration policy in particular in the case that a redistributive welfare state already exists before immigration occurs. A focus is put on analysing which tax policies are suitable to compensate the losers of unskilled immigration. I use a model with two dimensions of heterogeneity of agents, one dimension being two different skill types (skilled and unskilled) and the other being an agent's origin (native vs. immigrant). An agent's wage rate and labour supply

[1] This seems to be the consensus of a long strain of empirical literature considering the impact of immigration on the native workers, see e.g. Borjas (2003), Borjas (2006), Borjas and Katz (2005), Borjas et al. (2010), Card (2001), Card (2007), Peri and Sparber (2009)

decision are determined endogenously. The government that maximises the utility of the unskilled natives can control immigration and disposes of several tax instruments. It can levy taxes conditional on the origin of a person and conditional on income as only income and not an agent's skills are observable.

Based on this model, I find that in the absence of redistribution the skilled natives benefit whereas the unskilled natives lose from immigration. In contrast, when a redistributive tax system is in place before immigration sets in, immigration can result in all native agents being worse off. In the latter case, if the government of a welfare state wants to compensate the losers from immigration, it has to address both dimensions of heterogeneity by its choice of tax instruments. Thus, it has to combine an income-dependent tax scheme that differentiates agents based on skill with distinguishing between native agents and immigrants. Such a combination of tax instruments also allows the government to realise a Pareto improvement for all native agents. In addition, I show that immigrants can be net beneficiaries of the redistributive welfare state system even if they are discriminated against in the sense that they do not benefit from the tax transfer system in the same way as unskilled natives.

If a government does not run any sort of redistributive tax scheme initially, I find that there are specific cases where it is sufficient to only base taxation on skill or origin to compensate the native losers.[2] However, also without an initially redistributive welfare state, the government can only guarantee a compensation for the unskilled losers if it makes use of tax instruments conditioning on origin and addressing skill differences.

I contribute to the literature in that I vary the number of agents in a model with a redistributive tax scheme and endogenous wages of the Stiglitz (1982) type by allowing for immigration. It is important to note that immigration causes additional effects if wages are endogenous and if a redistributive welfare state exists in the pre-immigration situation. These effects have, to my knowledge, not been analysed before and they run through the impact of immigration on the tax system. A consequence is that immigration can result in all natives being worse off due to immigration. Thus, I go beyond Hamilton and Pestieau (2005) who look at a model with exogenous wage rates and thus exclude general equilibrium effects. As a consequence, their model depends on possible corner solutions where one group of agents does not work at all (utility is linear in labour) the result being that the welfare impact is non-linear in immigration.

Kohler and Felbermayr (2009) are also concerned with the question of whether immigration can be made a Pareto improvement. Despite the fact that they can introduce specific factor taxes on the distinguishable factors capital and homogeneous labour without informational restraints, they find that a Pareto improvement is only viable with a mechanism that does not allow immigrants to benefit from welfare benefits in the same way as natives. I confirm their result in a different environment as I cannot distinguish

[2]Only the unskilled natives lose if initially no tax scheme is in place because there is no tax effect besides the impact of unskilled immigration on wages.

factors directly and neither use a trade model nor capital mobility.

The model I use is based on Stiglitz (1982). There are two types of native agents, skilled and unskilled agents, who maximise their quasi-linear utility function over consumption and labour. The government disposes of income-dependent taxes, can levy poll taxes or subsidies and controls the extent of immigration. Its objective is to maximise the utility of the unskilled agents because they form the majority and are therefore pivotal for a sustainable policy. As an agent's skill-type is not observable, a first-best taxation cannot be realised. Instead self-selection constraints have to be used to distinguish agents based on income. Redistribution takes place via two channels. There is a direct channel via the tax amounts paid by each group of agents and an indirect general equilibrium channel. By setting marginal tax rates different from zero, the government influences the incentives to work.

When levying an extra tax burden onto the immigrants (or when exempting them from full subsidy benefits), it has to be ensured that the tax payment does not prevent every potential immigrant from migrating. Therefore, it is assumed that the minimum outside option of immigrants is zero utility. In some cases, immigrants can be net beneficiaries from the redistributional tax system but not to the same extent as the unskilled natives. The importance of the subsidy to the natives is to allow for a differentiation between the native and immigrant groups while maintaining that the government cannot distinguish skilled and unskilled agents.

The intuition for my result that both types of heterogeneity have to be addressed to guarantee a compensation to the losers from immigration is that there are leakage effects once one limits the scope of the policy instruments. If the government only taxes based on origin, the immigrants' tax payments are redistributed to the unskilled losers and to the skilled beneficiaries of immigration alike. As a consequence, the tax payments would have to be so high that immigration was inhibited. If, however, the government only relies on income-dependent taxes, the resulting transfers from the skilled to the unskilled agents also benefit the immigrants. The informational restriction that the government can condition its taxes only on income prevents that the transfer amount can be set sufficiently high to subsidise unskilled natives and immigrants.

The result that for specific parameter values immigration can make all native agents (both skilled and unskilled) worse off than they were before immigration when a redistributive tax system conditioned on income is initially in place has the following intuition. If the number of agents varies in the Stiglitz (1982) model[3], both components of redistribution, marginal tax rates and tax payments, change for both types of agents given that the tax system is adjusted optimally based on the specific level of immigration. For the skilled agents two opposing effects arise. Firstly, due to unskilled immigration the skilled wage rate increases by the general equilibrium effect. Secondly, the resulting larger wage

[3]In most of the existing literature the number of agents has been held fixed. An exception are Wilson (1980) and Wilson (1992) which have a different focus, though.

difference between the skilled and unskilled agents relaxes the incentive constraint and thus allows for levying higher tax payments on the skilled agents. If the latter effect dominates the former, skilled agents are worse off after immigration.

Discriminating between native and immigrant agents does not necessarily exclude the immigrants from being net beneficiaries in the redistributive tax-transfer system. The essential aspect is the chance to concentrate redistribution on the unskilled natives to compensate them for utility losses due to immigration. Thus, my model is able to implement the concept of delayed immigration as described in Richter (2004) or Sinn (2005) that limits the access of immigrants to benefits from a redistributional tax scheme. A different interpretation is provided by Chiswick (1982) which discusses entrance fees.

As immigration is about the movement of people, this implies for a welfare state that the number of potential beneficiaries of redistributional programs increases. Thus, trade and factor migration are not full substitutes any more as they are in the traditional Heckscher-Ohlin model (see Wellisch and Walz (1998)). Moreover, due to immigration, factor mobility and the elasticity of substitution between factors become important aspects of the welfare analysis (Michael (2006)).

The element of factor mobility also plays an important role in the traditional literature that focuses on the relation between optimal income taxation and migration, see e.g. Wilson (1980), Mirrlees (1982) and Wilson (1992). In contrast to my work, however, there the main focus has been on the impact of emigration onto a source country's tax system[4]. Wilson (1980) and Mirrlees (1982) look at the changes in an optimal tax system induced by the outside option of potential emigration of specific skill groups.

The problem that immigration harms those native groups which are close substitutes to immigrants is a well-founded empirical finding. There is a wide empirical literature on the impact of immigration on native wages[5]. A clearly negative and quantitatively important impact of immigration on the wages of competing native workers is identified in Borjas (2003) and Borjas (2006) where workers are classified into specific skill and experience groups. The effect of immigration is smaller when the focus is on local markets as in Card (2001) and Card (2007). In the general equilibrium framework used in Ottaviano and Peri (2008), the immigration effect turns positive for average native wages and for native workers without a high school degree but only in the long run. In the short run, Ottaviano and Peri (2008) find a negative effect of immigration on average native wages as well. In addition, they find that immigration has a quantitatively important negative impact on previous immigrants as these are close substitutes to new immigrants.

A different branch of the literature looks at effects of immigration on native welfare. Borjas (1999) focuses on the complementarity between immigrant and native labour in a

[4]Wilson (2006) gives an overview of this literature

[5]A survey can be found in Longhi et al. (2005) or in Freeman (2006).

two-factor, one-sector model and thereupon finds a positive welfare effect of immigration. In contrast, Davis and Weinstein (2002) concentrate on the terms-of-trade effects of immigration. Based on a Ricardian set-up they find negative welfare consequences of immigration when a country is already engaged in free trade based on technological advantages. The reason is that immigration erodes the comparative advantage of the country and therefore reduces its gains from trade. Both strands of the literature are combined by Felbermayr and Kohler (2007).

The chapter proceeds as follows. In section 2.2, I present the basic model that is used throughout the chapter. In section 2.3 the specific case of an untaxed, undistorted model is chosen to show the impact of immigration in a benchmark case. As a means to compensate the native losers of immigration, a poll-tax on immigrants is introduced. Section 2.4 extends the basic model presented in section 2.2 to incorporate a Stiglitz (1982)-type income-dependent redistributive tax. I show the specific consequences of immigration in such a model. To compensate the losers from immigration, a subsidy on natives is introduced and different characteristics of such a scheme are presented. In section 2.5, I provide numerical examples of the different model states to illustrate the consequences of the policies discussed. Section 3.5 concludes.

2.2 The model

I construct a simple model with a government and possibly immigration. The government can control immigration, i.e. it can influence the number of immigrants and it looks to the welfare of its citizens, in particular the unskilled natives as they are in the majority. Consequently, it only allows immigration in so far as the unskilled native agents are not worse off than without immigration. An immigration policy is termed sustainable once it is accompanied by measures that maintain the support of the unskilled agents.

To achieve this, the government analyses the impact (additional) immigration would have on the economy, in particular on the well-being of the unskilled natives. In case these would be worse off once immigration is allowed, the government conditions its immigration decision on the simultaneous introduction of specific tax instruments to compensate the native unskilled agents for the losses due to immigration. The instruments it has at its disposal are a poll tax on immigrants, subsidy payments to native agents and income-dependent taxes.

In using these tax instruments, separately or combined, the government is subject to informational restrictions whereby the following informational structure is assumed. The government can perfectly observe an agent's origin and thus can distinguish between natives and immigrants directly. The boarder allows to make a clear distinction between agents and as a consequence, the government can condition its tax policy on the agent's origin. In contrast, the government cannot observe a native agent's type or ability, only

her income. It would be interested to tax agents based on their skills but these are not observable and every sort of proxy for skills be it income, education or training is only a diffuse signal of underlying skills. In this aspect, immigrants are not different from natives. Hence, immigration schemes found in practice that try to screen immigrants do not facilitate the skill screening mechanism of the tax system.

This chapter discusses which of these tax instruments or combination of them are appropriate to realize a sustainable immigration process and the parameters this depends on. In the outline of the model, I start with the presentation of those features that are common to all variations and extensions of the model. Then, I go into the details of two different versions. For the first one, I assume that no taxes are raised in the economy in the beginning (see section 2.3). For the second one, I presume that the government levies a distortionary income-dependent tax in the base-line situation (see section 2.4).

2.2.1 Households

The model consists of one consumption good, two types of labour and many agents. By N^s and N^u, I denote the size of the group of skilled and unskilled native agents, respectively, whereas M denotes the number of unskilled immigrants in the economy[6]. Variables concerning immigrants will be indexed by m. I assume that $N^u > N^s > M \geq 0$ as a necessary condition for the wages of skilled agents to be higher than those of unskilled agents. It is possible that there are no immigrants in the domestic economy.

Each agent maximises her utility over consumption of a single, numeraire good c and labour l, according to the following separable utility function

$$U(c^i, l^i) = c^i - \frac{(l^i)^2}{2}, \ i = s, u, m. \tag{2.1}$$

Consumption is income adjusted by taxes and transfers. The specific expression for consumption depends on the tax instruments chosen and is presented each time when a specific tax regime is considered.

2.2.2 Firms

The one good that exists in this economy is produced by a competitive firm sector. Aggregate production takes place according to a Cobb-Douglas production function

[6]In the model, only unskilled immigration is considered. However, the essential impact of immigration in the model is that the group of unskilled agents in the domestic economy is increased relative to the group of skilled agents. Thus, it would also be possible to allow for skilled immigration as long as in the group of immigrants the ratio of skilled to unskilled agents is lower than in the native population before immigration occurs.

with Q being output.

$$\begin{aligned} Q = F((N^u + M)l^u, N^s l^s) &= (l^u(N^u + M))^\alpha (N^s l^s)^{(1-\alpha)} \\ &= l^u(N^u + M)(\tilde{n})^{(1-\alpha)}, \end{aligned} \tag{2.2}$$

where $\tilde{n} \equiv \frac{N^s l^s}{(N^u+M)l^u}$, i.e. the share of effective skilled labour to effective unskilled labour. I normalise the price of the consumption good to 1. Then, profit maximisation leads to the following first-order conditions

$$w^s = (1-\alpha)\tilde{n}^{-\alpha} \tag{2.3}$$

$$w^u = \alpha\tilde{n}^{(1-\alpha)}, \tag{2.4}$$

where agents are paid their marginal productivity.

2.2.3 The government

The objective of the government is welfare maximisation and it therefore maximises a version of the following social welfare function that weighs the utilities of the two groups of agents.

$$W = (N^u + \Delta)U^u + \mu N^s U^s. \tag{2.5}$$

To reflect the different proportions of agents in the economy, (2.5) weighs the agents' utilities by the number of agents in each group whereby the weight given to immigrants is flexible as $M \geq \Delta \geq 0$. Immigrants can have the same weight as native unskilled agents (in this case $\Delta = M$), they can have less weight or be of no importance to the government ($\Delta = 0$).

Michael (2003) and Richter (2004) choose a weighted sum of natives' utilities as a government objective function arguing that voters and thus only native agents are the relevant group to be taken into consideration. In contrast, Dolmas and Huffman (2004) consider immigrants as future voters.[7] Also Epstein and Hillman (2003) and Hamilton and Pestieau (2005) take immigrants fully into consideration in their social welfare functions. They cannot distinguish natives and immigrants as only income is observable in the tax systems.

In addition to the integration of immigrants, the weights of the two skill groups can be adapted in the social welfare function (2.5) by varying μ ($1 \geq \mu \geq 0$). For $\mu = 1$ the social welfare function becomes a utilitarian social welfare function, for $\mu = 0$ in combination with a self-selection constraint for the skilled agents it can be interpreted

[7] In particular in many European countries a lot of immigrants have voting rights for municipal elections in their destination country.

as a Rawlsian social welfare function. In the latter case the self-selection constraint prevents the skilled agents from being worse off than the unskilled agents.[8]

I mostly restrict myself to the Rawlsian social welfare function following Bucovetsky (2003) and Hamilton and Pestieau (2005) which concentrate on the majority of the unskilled natives since they are pivotal. Therefore, I use a specific form of (2.5) and that is the social welfare function

$$W = U^u.^9 \tag{2.6}$$

The number of immigrants is no argument of the government objective function but influences the utility of the unskilled natives. Choosing this government objective implies that the government is willing to refrain from the realisation of gains in overall wealth through immigration if it is detrimental to the unskilled natives' utility.

The government maximises the welfare function (2.6) by making use of different tax instruments. If it wants to focus on specific groups of agents, though, it can only condition its tax policy on observable characteristics. Each agent's origin is directly observable. There are, for example, birth certificates, official documents such as passports and in many countries people have to register. In contrast, an agent's ability or type is not observable,[10] the government can only monitor her income.

With endogenous labour supply, agents choose their income optimally given the tax policy. Therefore, they can mimic the other skill type by earning the same income once this is advantageous as compared to being treated as their particular income indicates. To avoid the emergence of pooled equilibria that prevent any attempt of redistribution, the government uses incentive compatibility constraints to distinguish agents.

The government freely decides on the immigration policy.[11] It cannot force immigrants to migrate, though.[12] Potential immigrants only migrate if they benefit from it. They have the outside option of staying where they are. Consequently, if immigrants face high tax payments when they enter the economy, those with the higher outside option, i.e. those who are relatively better off in their country of origin refrain from wanting to migrate and only the most desperate remain as potential immigrants if at all. As the minimal outside option, I assume that immigrants can reach a utility level of zero in

[8]Formally, a Rawlsian social welfare function would be described as $W = \min(U^u, U^s)$. By applying the incentive constraint on the skilled agents, I ensure that the skilled agents cannot be worse off than the unskilled agents. Hence, maximising the unskilled utility is equivalent to a Rawlsian social welfare function in my model.

[9]Thus, I set $\mu = 0$ and $\Delta = 0$, N^u is dropped because it is a fixed parameter that only scales utility.

[10]The ability cannot be deduced from other observable characteristics, either. This assumption is in line with the optimal taxation literature a la Stiglitz (1982).

[11]I refrain in the model from considering the cost of immigration control or illegal immigration as it can be found in e.g. Myers and Papageorgiou (2000).

[12]Wilson (2007) points out two aspects of immigration into a welfare state. On the one hand, the government uses immigration controls and decides whether and in how far legal immigration is allowed for. On the other hand, it determines the treatment of the immigrants, e.g. in how far they get access to the transfers that are available for otherwise identical natives.

their country of origin. Given the utility function (2.1), this level is reached when agents neither work nor consume and once they do not get any kind of transfer.

This section has introduced the basic components common to all variations and extensions of the model considered later. In the next sections, I focus on the consequences of immigration for the economy and on the impact onto the well-being of the unskilled agents in particular. Possibly negative effects of immigration provide the motivation for the analysis of compensatory tax policy measures. Throughout the following sections, I use the number of immigrants M and changes thereof as a parameter that can be controlled by the government.

I first concentrate on a version of the model where it is assumed that the economy does not apply any taxes (in section 2.3). Then, I focus on the case that a welfare state economy with redistributive income-dependent taxes exists (in section 2.4). Each time, I start by introducing the specific characteristics that are additional to the previous model outline and then do a comparative static analysis on changes in immigration. I continue by analysing tax policy measures that aim at compensating the native unskilled agents.

2.3 A simple economy without tax distortions

In this section, I specify a simple economy without any tax distortions and illustrate the effect of immigration. Then, I analyse how changes in the number of immigrants affect the endogenous variables. The tax instrument the government disposes of to compensate the losers from immigration is a poll tax on every immigrant, its proceeds are distributed as a poll subsidy to all native agents (see section 2.3.2). I examine whether the use of these tax instruments makes immigration feasible which means that at the same time two conditions are fulfilled. Firstly, the unskilled losers from immigration are at least as well off as before. Secondly, given the compensating tax instruments, immigrants are willing to migrate, i.e. they are better off than with their outside option of $U^m = 0$.

A native agent's consumption is given by

$$c^i = w^i l^i + S^p,\ i = s, u \tag{2.7}$$

where S^p is the poll subsidy paid to every native agent if it is introduced, otherwise $S^p = 0$. Consumption of an immigrant is

$$c^m = w^u l^u - T^p, \tag{2.8}$$

with T^p being a poll tax on immigrants if introduced. The price of the consumption good is normalised to 1.

The first-order condition of utility maximisation is

$$-\frac{U_l^i}{U_c^i} = l^i = w^i,\ i = s, u, m, \tag{2.9}$$

where subscripts denote partial derivatives. As all possible tax instruments are levied per head, the labour decision is unaffected by them. The tax instruments only have an income effect.

2.3.1 The equilibrium

Using the households' labour supply from (2.9) and the firms' labour demand from (2.3) and (2.4), the labour supply and demand look as follows:

$$l^s = (1-\alpha)\tilde{n}^{-\alpha} \tag{2.10}$$

$$l^u = \alpha\tilde{n}^{1-\alpha} \tag{2.11}$$

Via the labour market equilibrium the general equilibrium can be solved.

Based on the equilibrium solutions for skilled and unskilled wages, all other endogenous variables can be computed by using the respective equations for consumption ((2.7) and (2.8)) and utility (2.1). The wage rates look as follows:

$$w^s = l^s = (1-\alpha)^{\frac{2-\alpha}{2}}\alpha^{\frac{\alpha}{2}}(N^s)^{-\frac{\alpha}{2}}(N^u+M)^{\frac{\alpha}{2}} \tag{2.12}$$

$$w^u = l^u = (1-\alpha)^{\frac{1-\alpha}{2}}\alpha^{\frac{1+\alpha}{2}}(N^s)^{\frac{1-\alpha}{2}}(N^u+M)^{-\frac{1-\alpha}{2}} \tag{2.13}$$

As can be seen, the skilled wage rate is an increasing function of the number of immigrants M whereas the unskilled wage rate decreases in M. This is in line with the empirical literature on the impact of immigration onto the skill premium, see e.g. Card (2001) or Borjas (2003).

One fundamental assumption underlying the model as presented so far, is that $w^s > w^u$. The necessary assumptions concerning the group sizes $N^u > N^s > M \geq 0$ and concerning the parameter of the production function α $(1 > \alpha > 0)$ are not sufficient to guarantee this. Missing is a condition on the relation between the group sizes and α. Based on the equilibrium results, the assumption $w^s > w^u$ translates into Assumption 2.1 which establishes the missing link.

Assumption 2.1. *Together* $N^u > N^s > M$, $1 > \alpha > 0$ *and* $\frac{1-\alpha}{\alpha} > \frac{N^s}{N^u+M}$ *guarantee* $w^s > w^u$.

The group sizes determine the scarcity of the respective factor of production. The smaller is N^s relative to N^u, the scarcer is skilled labour and this increases the skilled wage rate. But α also has an important impact on wages. As α increases, unskilled labour contributes relatively more per output unit, unskilled labour becomes more productive and this increases the unskilled wage rate. Combined it holds that the higher is α the greater the difference between N^u and N^s has to be so that still $w^s > w^u$ holds.

To determine how marginal immigration affects the endogenous variables in this benchmark economy, I make a comparative static analysis of all endogenous variables with

respect to M. All signs are unambiguous as long as no taxes are levied and no subsidies paid, i.e. at $S^p = 0$ and $T^p = 0$.

The channel by which immigration changes the economy is that it influences the number of skilled to unskilled workers and via the wage rates individual labour supply. As the number of unskilled workers increases, skilled labour becomes relatively scarcer and as a consequence the wage rate for the skilled agents (w^s) increases whereas that of the unskilled agents (w^u) decreases (see (2.12) and (2.13)). As the wage rate and the labour supply can be expressed by the same function, both wage rate and labour supply undergo the same direction of change and so does income. Without tax payments, consumption is equal to income.

Concerning utility, the change in consumption always dominates the immigration impact on the labour supplied. Therefore, the utility of the skilled agents increases although they work more and the utility of the unskilled agents decreases although they work less. Hence, the unskilled native agents always lose from immigration. As they are in the majority, without a compensatory regime to make them as well off as before immigration, immigration is not a feasible policy option. Thus, I next consider compensatory tax instruments.

2.3.2 A simple poll tax and poll subsidy

A simple tax instrument to compensate the native losers from immigration and thus increase welfare by improving on the unskilled agents' utility is to levy a poll tax on every immigrant and to spend the proceeds on a poll subsidy for every native. As the government can observe an agent's origin, it is easy to distinguish agents based on origin. In contrast, the government would have to make an effort to take incentives into consideration and distinguish skilled and unskilled agents. As it is not obvious that it is necessary to discriminate between agents by skill to reach the objective of compensating the unskilled natives, I first look at the simple case. In section 2.4, I concentrate on all regimes that discriminate between skilled and unskilled agents.[13]

For a tax policy with a poll tax on every immigrant and a poll subsidy for every native, the government budget constraint becomes

$$T^p M = (N^u + N^s) S^p. \tag{2.14}$$

A system of a poll tax on immigrants and a poll subsidy for natives has to meet two requirements to allow for a sustainable immigration policy. Firstly, as the government only allows for immigration if the unskilled natives are not worse off, the poll subsidy has to be so high as to compensate this native group. The minimum level of the poll

[13]The particular case that an income-dependent tax system is introduced into an initially untaxed economy together with immigration is examined in section 2.4.7.

tax that just fulfills this requirement is denoted $\hat{T^p}$. Secondly, it has to be possible to raise the proceeds necessary to finance these subsidy payments by the poll tax. Therefore, immigrants have to be better off with paying the poll tax and immigrating to the domestic country than with their outside option in their country of origin at $U = 0$.

Proposition 2.1. *Given Assumption 2.1 and the lump-sum tax $\hat{T^p}$ that is just sufficient to make the unskilled natives as well off as before marginal immigration, the skilled natives are strictly better off. However, immigration does not always occur because the immigrants' utility is positive only if $\frac{N^u}{1+N^s} > \frac{1-\alpha}{\alpha}$.*

Proof. In the initial situation, there are no immigrants ($M = 0$). I determine the amount that has to be transferred to the unskilled natives via T^p to make them just as well off as before immigration by setting

$$U^u|_{T^p=0} - \left(U^u(T^p) + \frac{dU^u(T^p)}{dM}\right) = 0 \tag{2.15}$$

and solving this expression for T^p where $\frac{dU^u}{dM}$ is the total change of U^u due to marginal immigration. The first term of (2.15) gives the utility of the unskilled natives before immigration and at $T^p = 0$, the second bracketed term reflects the utility of the unskilled natives after immigration and including T^p. In general, utility is a function of consumption. As can be seen from the equation indicating each agent's consumption (2.7) combined with the government budget constraint (2.14), the unskilled natives' utility level as well as changes in utility are a function of T^p and immigration. Consequently, both effects, the occurrence of immigration and the introduction of T^p have to be accounted for in (2.15) and therefore, it is not sufficient to consider just $\frac{dU^u(T^p)}{dM} = 0$ to determine the compensatory level of the lump-sum tax.

The specific T^p that solves (2.15) is denoted $\hat{T^p}$. For all poll tax amounts exceeding $\hat{T^p}$ the unskilled natives are strictly better off than they were before immigration as unskilled utility is increasing in the transfer (see 2.7).

To verify whether the immigrants are better off when migrating than when staying in their country of origin at $U^m = 0$, I plug $\hat{T^p}$ into

$$U^m + \frac{dU^m}{dM} \tag{2.16}$$

The sign is not unambiguously positive, $U^m(\hat{T^p}) + \frac{dU^m(\hat{T^p})}{dM} > 0$ holds if and only if $\frac{N^u}{1+N^s} > \frac{1-\alpha}{\alpha}$ and Assumption 2.1 holds. □

The intuition for this result is as follows. Without any tax payments, the immigrants' utility is a bell-shaped function of α, it is zero for $\alpha = 0$ as well as for $\alpha = 1$ as in either case the wage rate and therefore income is zero (see (2.13)). For $0 < \alpha < 1$ the immigrant's utility is positive and an increasing function of α for low $\alpha > 0$ whereas it

is decreasing in α for large $\alpha < 1$. As $\hat{T^p}$ is a strictly positive tax amount, it reduces the immigrant's utility and may result in it becoming negative. At a specific strictly positive value of α, the immigrant's utility after the tax payment $\hat{T^p}$ is just zero.[14] This specific value of α is defined by $\frac{N^u}{1+N^s} = \frac{1-\alpha}{\alpha}$. Assumption 2.1 that guarantees $w^s > w^u$ sets an upper limit to α. At this limit the immigrant's utility is positive even after the payment of $\hat{T^p}$. Thus, no additional restriction on the parameters is needed. A numerical example for a case where immigration is not feasible is given in section 2.5.3.

There are two reasons for assuming an initial number of immigrants of zero. Firstly, it would be judicially difficult and highly controversial to introduce a discriminatory mechanism against immigrants who have already lived in a country for many years. Secondly, the critical case is whether the immigrants that migrate are able to finance the compensation for the specific losses they impose on particular native groups. If one extended the group of agents that have to pay the compensatory poll tax by also levying it on immigrants who already live in the country, the relation between cause and compensation would break down.

This section has shown that under certain parameter values it is impossible to realise a policy with positive immigration if the government only disposes of the tax instrument of a poll tax on immigrants and redistributes the proceeds to all natives via a poll subsidy. The reason is that there are parameter combinations in which the transfer that is needed to obtain the unskilled natives' support for a policy with positive immigration is so high that no immigrants are willing to migrate. At least in some respect this is due to the fact that the policy is wasteful in the sense that skilled agents also obtain a transfer even though they are already better off via the general equilibrium wage effects. However, the government cannot easily discriminate against the skilled natives because of the informational restrictions.

Before I introduce a redistributive scheme that distinguishes between skilled and unskilled agents into the framework used so far, I first present how the initial model changes when a distortionary income-dependent tax is introduced and analyse how immigration affects this economy. In addition, I address the problem of how to compensate the native losers of immigration in the context of a redistributive tax system.

[14] Clearly, the result depends on the normalisation of the outside option. The lower is the outside option, the more likely it is to make immigration feasible as immigrants would accept to pay higher tax amounts. With an outside option of $U = 0$, I choose a low outside option (assuming that no one can be forced to work for zero consumption) and even then immigration is not always feasible.

2.4 An economy with optimal income taxation

Many industrialised countries are - to varying degrees - welfare states that use taxes to redistribute income and lower the inequality of the net income distribution. In the context of welfare states, the existing literature provides contributions that look at the impact of immigration when redistributive policies are in place in the destination country and immigrants possibly become net beneficiaries of the redistributive tax-transfer system.[15] This, however, has so far not been considered in a general equilibrium model with income-dependent taxes and endogenous labour supply. Therefore, in this section, I analyse an economy with an optimal redistributive income tax system of the Stiglitz (1982) type where the government can condition taxes on income. To distinguish the skilled and unskilled agents, incentive compatibility constraints are introduced.

After the presentation of the specific assumptions needed to adapt the general model presented above, I firstly analyse how an economy with a redistributive tax system reacts to immigration. The income-dependent tax scheme in itself does not allow to distinguish agents based on their origin because it is conditioned on income only. However, as incomes depend on M, immigration is a parameter of the tax payments and the marginal tax rates. I focus on the changes marginal immigration induces for the optimal tax structure as well as on the utilities of both types of native agents. Methodologically, the introduction of immigrants into the Stiglitz (1982) model makes the number of unskilled agents in the economy variable. Secondly, I introduce an additional tax instrument that allows to take an agent's origin into account and thus to discriminate between natives and immigrants. The instrument used here is a subsidy to all native agents that is paid out of the overall tax revenues.

2.4.1 Adjustment of the model

In order to introduce income-dependent taxation, the economy presented in section 2.2 has to be adjusted in some aspects. Also, the methodological approach changes. As the government runs a redistributive tax scheme even before it considers to permit immigration, it formally maximises the social welfare function (2.6) in the initial situation of $M = 0$.[16] Welfare maximisation is carried out via the choice of optimal income-dependent tax rates. In a second step, the government can allow immigration (it can control M as before). It does so only when immigration can be made a welfare improvement, i.e. when it benefits the unskilled natives. In addition, I analyse the feasibility of the policies considered.

[15] Examples for this are Dolmas and Huffman (2004), Epstein and Hillman (2003), Hansen (2003), Wellisch and Walz (1998), Wellisch and Wildasin (1996) and Wilson (2007).

[16] Before, there was no role for a government in the baseline economy. Only when the poll tax on immigrants was introduced, the welfare function (2.6) played a role.

As in the previous section, consumption has to be adjusted to account for the new tax and transfer scheme. Therefore, (2.7) and (2.8) change to

$$c^i = w^i l^i - T^i + S,\ i = s, u \tag{2.17}$$

$$c^m = w^u l^u - T^u = c^u - S. \tag{2.18}$$

The tax is an income-dependent tax $T^i \equiv T(y^i)$ where $y^i = w^i l^i$ is income of agent i.

In T^i no distinction is made based on origin, i.e. between unskilled natives and immigrants. As the instrument addresses the skill heterogeneity, the government only uses income to distinguish agents and in this respect immigrants are not different from native unskilled. Hence, immigrants obtain the same wage rates and have the same labour supply as unskilled natives (as (2.18) indicates). This can also be seen from the first-order condition of utility maximisation which is structurally the same for all agents:

$$-\frac{U_l^i}{U_c^i} = l^i = w^i(1 - T_{y^i}),\ i = s, u, m, \tag{2.19}$$

where subscripts denote partial derivatives. It is important to note that here labour supply depends on the net marginal wage rate.

If they are introduced, transfers $S \geq 0$ are paid lump-sum to all natives out of the overall tax revenues raised by income-dependent taxation.[17] With this instrument, the government addresses the heterogeneity based on origin which it can observe. S is only used in section 2.4.4. However, to avoid having to introduce it separately, I already present it here.

All agents are subject to the income-dependent tax structure in which tax amounts can be negative, though. Since the subsidy payments to the natives S are paid out of the overall tax revenues, the government budget constraint is

$$\begin{aligned} T^s N^s + T^u(N^u + M) &= (N^s + N^u)S \\ \Leftrightarrow N^s(y^s - c^s) + N^u(y^u - c^u) + M(y^u - c^m) &= 0 \\ \Leftrightarrow Q - N^s c^s - (N^u + M)c^u + MS &= 0. \end{aligned} \tag{2.20}$$

For the first step, I use the household budget constraints (2.17) and (2.18), for the last step the Euler theorem. As profits are zero, total sales are distributed to all input factors, i.e. $Q = N^s y^s + (N^u + M)y^u$.

2.4.2 The equilibrium

The optimisation approach here is that the government maximises the social welfare function (2.6) subject to the incentive constraint for the skilled agents and the government budget constraint (2.20). The incentive constraint for the skilled agents is

$$U^s(c^s, y^s) \geq U^s(c^u, \frac{l^u w^u}{w^s}), \tag{2.21}$$

[17] A comparable instrument can also be found in Wilson (1980), though in a different context.

i.e. skilled agent's utility has to be at least as high as if it was would they mimic the unskilled agents by earning their income. The utility of the skilled agents is higher than that of the unskilled agents who have the same income as the former are by definition more productive.

As the assumed welfare function maximises the utility of the unskilled natives, it would be welfare improving to levy very high taxes on the skilled agents and redistribute the proceeds to the unskilled agents. However, the informational structure does not allow to condition the tax system on ability, only on income. Therefore, the skilled agents can mimic the unskilled agents by earning their income and it is impossible for the government to levy a tax on them that would make them worse off than the unskilled natives. This characteristic is formalised by the introduction of the incentive constraint for the skilled agents. For the aforementioned reasons, the constraint will be binding in the optimum. In contrast, an incentive constraint on the unskilled natives will never be binding as the welfare function concentrates on maximising their utility directly. A formal proof of this structure can be found in Stiglitz (1982)[18].

The welfare maximisation approach can be solved analytically for all endogenous variables. The derivatives of the maximisation problem with respect to l^s, l^u, c^s, c^u together with the government budget constraint and the incentive constraint on the skilled agents (2.21) can be solved for l^s, l^u, c^s, c^u and the multipliers. The wage rates are determined by (2.3) and (2.4), the tax amounts by (2.17), the marginal tax rates by (2.19) and the utility levels by (2.1). All of the endogenous variables are functions of N^u, N^s, M and α. Thus, M is included as a parameter in every endogenous variable and this allows for an analysis of the consequences of immigration on an existing redistributive tax structure.

For presentational purposes I only list the wage rates and the tax scheme variables explicitely here:

$$w^s = \alpha^{\frac{\alpha}{2}}(1-\alpha)^{\frac{2-3\alpha}{2}}P^{-\frac{\alpha}{2}}(N^s)^{-\frac{\alpha}{2}}\Psi^{\frac{\alpha}{2}} \tag{2.22}$$

$$w^u = \alpha^{\frac{1+\alpha}{2}}(1-\alpha)^{\frac{3}{2}(1-\alpha)}P^{\frac{1-\alpha}{2}}(N^s)^{\frac{1-\alpha}{2}}\Psi^{-\frac{1-\alpha}{2}} \tag{2.23}$$

$$T^s = \frac{1}{2}Z(N^u+M)^2(N^s)^{-\alpha}\Omega + \frac{(N^u+N^s)S}{P} \tag{2.24}$$

$$T^u = -\frac{1}{2}Z(N^u+M)(N^s)^{1-\alpha}\Omega + \frac{(N^u+N^s)S}{P} \tag{2.25}$$

$$T_{y^u} = \frac{N^s}{P}, \quad T_{y^s} = -\frac{(N^s)^2\alpha^2}{\Psi} \tag{2.26}$$

[18] Arnott et al. (1988) show for all concave social welfare functions and separable utility functions that the so-called 'normal' case applies, i.e. that the skilled agents have to be bound from mimicking the unskilled agents.

with

$$P = N^u + N^s + M,$$

$$\Psi = (1-\alpha)^2(N^u + M)^2 - \alpha^2(N^s)^2,$$

$$\Omega = \alpha^\alpha(1-\alpha)^{3-3\alpha}P^{-\alpha-1}\Psi^{-1+\alpha},$$

$$Z = (1-\alpha)(N^u + M) - 2N^s\alpha.$$

Consequently, the wage rates depend on M via P and Ψ.

As before, an essential assumption for the model is that the skilled agents get a higher wage, i.e. $w^s > w^u$. To guarantee this in the explicit solution, I make Assumption 2.

Assumption 2.2. *Together* $N^u > N^s > M \geq 0$, $1 > \alpha > 0$ *and* $\frac{1-\alpha}{\alpha} > \frac{N^s}{N^u+M} + \frac{(N^s)^2}{(1-\alpha)(N^u+M)^2}$ *guarantee* $w^s > w^u$.

An implication of Assumption 2.2 is that $\Psi > 0$. As it was the case in the poll tax economy, Assumption 2.2 is needed to specify the relation between the group sizes (N^u, N^s, M) and the parameter of the production function α.

Given Assumption 2.2, the tax structure is of the typical Stiglitz (1982) type in a model with endogenous wage rates. The skilled agents obtain a marginal subsidy ($T_{y^s} < 0$) that induces them to work more. Hence, skilled labour becomes relatively less scarce and the unskilled wage rate is raised via the general equilibrium effect. In contrast, the unskilled agents pay a marginal tax ($T_{y^u} > 0$) which makes them work and earn less than without distortion. This results in unskilled labour becoming scarcer and thus in an increase in w^u. In addition, it becomes less attractive for the skilled agents to imitate the unskilled. As a consequence, there is redistribution from the skilled to the unskilled agents via the general equilibrium effect of the marginal tax rates.

Concerning tax payments, for a wide range of parameter values the usual structure emerges that the tax payment of the skilled agents (T^s) is strictly greater than that of the unskilled agents (T^u). Then and if there are no subsidy payments, i.e. $S = 0$, the tax system is a pure redistribution system in which $T^s > 0 > T^u$ holds so that the unskilled agents get a subsidy payment financed by the skilled agents. However, there are some exceptional cases where $T^s < 0 < T^u$ holds.

The sign of T^s and T^u is determined by Z. Whereas all other components are positive given the assumptions made, Z can be negative and in this case the skilled agents get a subsidy from the unskilled. A negative T^s can arise when the redistribution effect through the general equilibrium channel dominates the redistribution effect through the direct tax payment channel. The marginal tax rate on the unskilled agents and the marginal subsidy rate on the skilled agents are very high and lead to the overall skilled tax payment being negative.

2.4.3 The impact of immigration

After having shown how the basic tax framework looks like without immigration[19], I examine the impact of marginal immigration. I identify the winners and losers from immigration when a redistributional tax system is in place that is endogenously changed by immigration. The changes are such that the tax system is again optimal after marginal immigration has occurred. Hence, it is assumed that the tax structure is conditioned on the government's choice of immigration. The analysis is done by a comparative static analysis of the general equilibrium model with respect to immigration M. I first focus on the unskilled natives because of the government's objective. Later, I also show the effect on the skilled natives. To concentrate on the pure effects of immigration, S is set to zero throughout section 2.4.3.

In contrast to the section 2.3 model, the impact of immigrants on the well-being of the unskilled natives is not ex ante clear as there are two effects at work. In addition to the direct negative wage impact of immigration that stems from immigrants being substitutes to native unskilled agents, immigration changes the tax system. As the unskilled wage rate decreases whereas the skilled wage rate increases, there is a higher income gap between the two groups of agents after immigration, the self-selection constraint is relaxed. This result can be seen from the impact of immigration on the marginal tax rates. The marginal tax rate for the skilled (unskilled) agents increases (decreases) based on being negative (positive) in the before-immigration economy ($\frac{dT_{y^s}}{dM} > 0$, $\frac{dT_{y^u}}{dM} < 0$). Therefore, both tax rates are now closer to zero and as a consequence the redistributional tax scheme is less distortionary. Immigration causes an efficiency gain.

The relaxation of the incentive constraint potentially allows for higher redistribution from the skilled to the unskilled agents via the direct tax payment channel. In addition, it implies less distortionary marginal tax rates. If the tax effect of immigration dominate the direct wage effect, there would be no need for any sort of additional policy instruments, like e.g. the introduction of S.

As expected, the unskilled agents' gross wage rate decreases ($\frac{dw^u}{dM} < 0$) as the number of unskilled agents in the economy increases. Thus, this effect replicates the stylised facts as found e.g. in Ottaviano and Peri (2008) that natives lose in particular if immigrants are close substitutes. In contrast to the section 2.3 economy, however, the impact of immigration on unskilled labour supply is unambiguously positive, the reason being that labour supply depends on the marginal net wage rate.

If one takes the total differential of the labour supply equation (2.19) to split up the overall effect of immigration on labour supply, three effects can be identified. There is a direct wage impact of immigration and two effects related to the immigration-induced

[19] All the results presented in this section hold for the case that there are already some immigrants in the economy when marginal immigration occurs as well as that there are none ($M = 0$).

change to the marginal tax rate of the unskilled natives.

$$\frac{dl^u}{dM} = \underbrace{\frac{dw^u}{dM}}_{<0} \underbrace{- \frac{dw^u}{dM} T_{y^u}}_{>0} \underbrace{- \frac{dT_{y^u}}{dM} w^u}_{>0} > 0 \tag{2.27}$$

The sign of (2.27) as well as those of the different effects in (2.28) and (2.30) are obtained by taking the derivatives of equations (2.22) to (2.26) with respect to M.

As shown below, it is impossible that the unskilled natives are as well off as before immigration when the optimal income tax system is optimally adjusted to the marginal influx of migrants. The higher labour supply has a negative impact on the utility of the unskilled and the lower income results in an unambiguous decrease of consumption. Although the transfer payments to the unskilled agents increase for many parameters (the sign of $\frac{dT^u}{dM}$ is ambiguous), this increase is never sufficient to dominate the negative effect of marginal immigration on income. Both, higher labour supply and lower consumption clearly result in a decrease in unskilled utility ($\frac{dU^u}{dM} < 0$).

Proposition 2.2. *Given Assumption 2.2, if marginal immigration takes place in an optimal income-dependent tax system and the tax system is adjusted so as to be optimal again given the government focus on unskilled agents, unskilled native agents lose from immigration.*

For the native skilled agents the situation looks differently. The skilled agents' wage rate increases as skilled labour becomes scarcer because skilled natives are complements to immigrants ($\frac{dw^s}{dM} > 0$). The impact on labour supply is ambiguous.

$$\frac{dl^s}{dM} = \underbrace{\frac{dw^s}{dM}}_{>0} \underbrace{\underbrace{- \frac{dw^s}{dM} T_{y^s}}_{>0} \underbrace{- \frac{dT_{y^s}}{dM} w^s}_{<0}}_{<0} \tag{2.28}$$

The higher wage rate has a positive impact on labour supply. In contrast, the lower marginal subsidy makes it less attractive to work. This effect is reflected in the negative sign of the combined last two terms which holds in general given Assumption 2.2.

To determine the impact of marginal immigration on skilled utility also consumption is of relevance. Consumption is the remaining income after tax payments have been made. Whereas the income of the skilled agents increases with immigration ($\frac{dy^s}{dM} > 0$), they also have to pay higher taxes ($\frac{dT^s}{dM} > 0$). Thus, neither the impact of immigration on labour supply nor on consumption is unambiguous. Nevertheless, to structure the effects I take the total derivative of the utility function for the skilled agents (2.1).

$$\begin{aligned} \frac{dU^s}{dM} &= w^s \frac{dl^s}{dM} + l^s \frac{dw^s}{dM} - \frac{dT^s}{dM} - l^s \frac{dl^s}{dM} \\ &= \underbrace{(w^s - l^s)}_{<0 \text{ see } (2.19)} \frac{dl^s}{dM} + l^s \frac{dw^s}{dM} - \frac{dT^s}{dM} \end{aligned} \tag{2.29}$$

Using the total derivative of labour supply (2.28), the expression becomes

$$\frac{dU^s}{dM} = \underbrace{w^s \frac{dw^s}{dM}}_{>0} \underbrace{- (w^s - l^s)}_{>0} \underbrace{\left(\frac{dw^s}{dM} T_{y^s} + w^s \frac{dT_{y^s}}{dM} \right)}_{>0 \text{ see (2.28)}} \underbrace{- \frac{dT^s}{dM}}_{<0} \tag{2.30}$$

The split-up shows that immigration has a negative impact on the skilled agents' utility via taxes and a positive effect onto the combined labour supply and income component of the utility function. In the latter combined component, income is sufficiently high to compensate the increase in disutility that might arise from working more as a consequence of immigration. Thus, when the income-labour supply effect dominates, the impact of marginal immigration on skilled utility is positive and has the same sign as in the section 2.3 model. In contrast, if the negative effect of the higher tax duties dominates, skilled agents lose from immigration along with the unskilled natives and this happens in a lot of cases.

Let N^s be normalised to 10 and $M = 0$, then $\frac{dU^s}{dM}$ is always negative for $\alpha < 0.25$ with $N^u \in (10.1; 1000)$.[20] For $\alpha > 0.25$ marginal immigration has a negative impact on the skilled agents' utility for high values of N^u [21], an example is given in section 2.5.1. The existence of the tax system (without S) makes it possible that all native agents are worse off after immigration.

Proposition 2.3. *Given Assumption 2.2, if marginal immigration takes place in an optimal income-dependent tax system and the tax system is adjusted so as to be optimal again, the effect on the skilled agents is ambiguous. There exist parameter combinations where they benefit from immigration, for other parameter combinations they lose.*

Proposition 2.3 does not hinge on the specific Rawlsian social welfare function as a government objective. It holds as well if the government chooses to maximise a utilitarian social welfare function, i.e. if the utility of each type of agents is weighted with the respective group size. Formally, this means that in (2.5) $\mu = 1$ and $\Delta = M$ is used.[22] With a utilitarian welfare function, the skilled natives always lose from immigration once $\alpha < 0.34$ for $N^u > N^s$ and $M = 0$. Intuitively, the skilled agents lose in certain cases from immigration because they become more exploitable due to the impact of immigration onto the tax structure.

There are several effects at work when immigration takes place in an economy with a redistributive tax system. On the one hand, the impact of immigration onto the tax

[20] Relevant for the effect is the proportion of unskilled to skilled agents, not the number of skilled or unskilled agents.

[21] The greater is α, the higher N^u can be for the immigration impact on skilled utility to be positive.

[22] As Arnott et al. (1988) prove, in the case of a utilitarian social welfare function it is the incentive constraint on the skilled agents that always binds. Therefore, the analytical framework used so far also holds for the utilitarian social welfare function.

structure leads to less distortionary tax rates and thus to an efficiency gain. In addition, the tax payments levied on the skilled agents increase. As a consequence, the skilled agents' higher income due to immigration does not transform into higher consumption if the tax payments increase more strongly than income.

On the other hand, the higher overall tax payments by the skilled agents do not necessarily result in higher per capita transfers to the unskilled agents. As immigrants belong to the unskilled group, they are net beneficiaries in the tax-transfer scheme as well. And even if per capita transfers to the unskilled agents increase (if $dT^u/dM < 0$), the increase never compensates the unskilled agents for the negative wage effect of immigration. The integration of immigrants into the group of beneficiaries of the tax-transfer system makes it impossible that unskilled native agents are as well off after immigration than they were before.

2.4.4 The effects of a subsidy to the natives

As section 2.4.3 has shown, unskilled agents in an economy with an optimal income-dependent tax system that is optimally adjusted for immigration always lose from marginal immigration when the government only conditions its tax system on income and cannot additionally distinguish agents based on their origin. For skilled agents it depends on the parameters whether they gain or lose.[23] In this section, I make use of the S-variable as an additional tax instrument to discriminate between immigrants and native agents. Thus, I focus on compensating the native unskilled agents so that the government can consider to allow immigration. The immigrants do not fully participate in the redistribution scheme any longer. Important for the characterisation of the tax transfer system with T^u, T^s and S are the net tax payments of each group. They are $T^s(S) - S$ for the skilled, $T^u(S) - S$ for the unskilled natives and $T^u(S)$ for the immigrants.

In the adapted tax-transfer structure all agents are integrated into the Stiglitz (1982)-type tax-transfer system whereas only the natives benefit from the additional subsidy component S. If the subsidy to the natives is positive, $S > 0$, immigrants either get less out of the tax-transfer system than the native unskilled or they make net contributions whereas the unskilled natives are net beneficiaries. Which case applies depends on the optimal $T^u - S$ prevailing in equilibrium.[24]

[23] This finding is in line with Myers and Papageorgiou (2000) who find that immigration is detrimental to the native population in the destination country if immigrants have full access to redistributive public services.

[24] The above introduced scheme is equivalent to a tax-transfer scheme where S is eliminated ($S = 0$) and instead the redistribution from the immigrants to the natives takes place by a poll tax levied onto the immigrants T^m which is distributed in a lump-sum manner to all natives. In this alternative formulation, there is a separation of the redistribution between unskilled and skilled agents (independent of whether they are natives or immigrants) and the redistribution between natives and immigrants. Every

The introduction of S also changes the tax system. As can be seen from (2.24) and (2.25), T^s and T^u are both increasing in S. Therefore, it is conceivable that T^u becomes greater than zero so that immigrants are net contributors to the tax transfer system whereas native unskilled agents remain net beneficiaries as $T^u(S) - S < 0$. S allows to redistribute from the immigrants to the native agents. In contrast, the introduction of S does not make additional redistribution between the skilled and unskilled agents possible as the incentive constraint already binds at $S = 0$. Thus, all native agents benefit from the same amount of the subsidy paid. In particular, skilled natives get subsidy payments even in the case when they have already benefited from immigration at $S = 0$ ($\frac{dU^s}{dM} > 0$).

I analyse whether S allows to make the unskilled native agents as well off as before immigration. It is obvious that the native agents are always better off if an ultimately high S is introduced which increases T^u by a large amount and imposes a huge cost on the immigrants. However, as before, the immigrants' outside option has to be taken into consideration. A subsidy S that guarantees $U^m > 0$ is called feasible as immigration occurs under this regime. S is set in such a way as to make the native unskilled agents just as well off as they were before immigration.

Proposition 2.4. *Given Assumption 2.2, for every parameter combination there exists a S^* such that*

1. *after marginal immigration has taken place unskilled natives are at least as well off as they were before immigration and*

2. *migrants' utility is strictly positive, i.e. S^* is feasible.*

Proof. The equilibrium utility level of the unskilled native agents is obtained by using the utility function (2.1) and plugging in unskilled labour supply (2.19) including the marginal tax rate (2.26) as well as unskilled consumption (2.17) including the unskilled wage rate (2.23) and tax payments (2.25). This leads to

$$U^u = \frac{MS}{P} + \Omega(N^u + M)^2(N^s)^{1-\alpha}. \tag{2.31}$$

I equate the utility of the unskilled natives before immigration and before the introduction of S, i.e. at $S = 0$, with the utility of the unskilled natives after immigration as a function of S,

$$U^u|_{S=0} = \left(U^u(S) + \frac{dU^u(S)}{dM}\right). \tag{2.32}$$

The equation is solved for the specific $S^* = S(N^s, N^u, M, \alpha)$ that is the minimal amount that has to be transferred to the native unskilled agents in order to make them just as

immigrant pays T^m in addition to T^u, whereas every native gets a lump-sum transfer of $\frac{M}{N^u+N^s}T^m$. In such a scheme, T^m could be interpreted as an entry fee as proposed by Chiswick (1982).

well off as they were before immigration thus ensuring their support for immigration. As $\frac{dU^u}{dM}$ is strictly increasing in S (see (2.31)), it follows that a transfer $S > S^*$ would make the unskilled natives better off than before.

For the transfer S^* to be feasible, the immigrants' utility has to be positive given S^*. There are no immigrants in the original economy ($M = 0$). Therefore, I analyse whether, given Assumption 2.2,

$$\left(U^m + \frac{dU^m}{dM}\right)\Big|_{S^*, M=0} > 0 \tag{2.33}$$

holds unambiguously. The explicit solution to U^m is obtained in the same way as the explicit solution for U^u (see above) with the only exception being that consumption for the immigrants is specified by (2.18) (instead of (2.17)) to account for the fact that the immigrants have to pay S. The post-immigration utility of the immigrants at the minimal compensatory S^* is positive given Assumption 2.2. □

The introduction of S and the feasibility of S^* solve the political economy problem that the support of the unskilled majority is needed for immigration to be a feasible policy. With S^* as an additional tax instrument, an optimal income dependent tax system and immigration become compatible from a political economy viewpoint.

2.4.5 Letting all native agents benefit from immigration

So far I have specified the minimum transfer S^* that would make an immigration policy just feasible based on the political economy argument that the unskilled agents are the majority and thus it is them who are pivotal for the feasibility of an immigration policy. However, once the instrument S that discriminates between natives and immigrants is introduced, the opportunity to extend the redistribution from the immigrants by setting a $S > S^*$ arises. Any such policy directly increases the utility of the unskilled native agents and is thus compatible with welfare maximisation.

In this section, I analyse whether $S > S^*$ can be feasibly set so that the whole native population benefits from immigration. Thus, marginal immigration would be made a Pareto improvement for all natives. In the section 2.3 model, the skilled agents always benefited from immigration and hence focusing on the well-being of the unskilled agents implied a Pareto improvement for the native population as a whole. Under the income dependent tax scheme, however, marginal immigration can lead to the skilled agents being worse off than before immigration. Compensating the skilled natives as well could be interesting for a government that seeks for unanimous support. It does not contradict the welfare maximisation objective, either.

If S^* is introduced after marginal immigration has taken place, all unskilled natives are as well off as before immigration. However, the amount S^* that makes the unskilled agents as well off as before immigration is not sufficient to compensate the skilled agents

for all parameter values for which they lose from immigration. $U^s|_{S=0} - (U^s + \frac{dU^s}{dM})\big|_{S^*}$ is not unambiguously negative, i.e. it is possible that the utility before immigration and at $S = 0$ is higher than after immigration including the subsidy payment S^*. As a consequence, the case of the skilled agents is treated separately here.

Proposition 2.5. *Given Assumption 2.2 and given that the total derivative of U^s with respect to M is negative ($\frac{dU^s}{dM} < 0$), for every parameter combination there exists a $\hat{S}$ such that*

1. *skilled natives after marginal immigration are at least as well off as they were before immigration and*

2. *migrants' utility is strictly positive, i.e. $\hat{S}$ is feasible.*

Proof. The determination of $\hat{S}$ follows the same line of argument as the determination of S^* in Proposition 2.4. $\hat{S}$ is the amount of the subsidy that makes the skilled agents just as well off as before immigration and is obtained by solving

$$U^s|_{S=0} = \left(U^s(S) + \frac{dU^s(S)}{dM}\right). \tag{2.34}$$

As it was the case for the unskilled agents, $\frac{dU^s}{dM}$ is increasing in $\hat{S}$.

For this transfer $\hat{S}$ to be feasible, the immigrants' utility has to be positive given $\hat{S}$. There are no immigrants in the original economy ($M = 0$).[25] Therefore, I analyse whether, given Assumption 2.2 and $\frac{dU^s}{dM}$,

$$\left(U^m + \frac{dU^m}{dM}\right)\Big|_{\hat{S},M=0} > 0 \tag{2.35}$$

holds unambiguously. The post-immigration utility of the immigrants at the minimal compensatory $\hat{S}$ is indeed positive given Assumption 2.2 and given that the skilled agents lose from immigration initially, i.e. $\frac{dU^s}{dM}$. □

As both S^* and $\hat{S}$ are feasible for all parameter combinations that fulfill Assumption 2.2, it is always feasible to implement the larger S as a policy. This guarantees that a Pareto improvement for all native agents is a feasible policy.

Proposition 2.6. *Given Assumption 2.2 in an economy with an income-dependent tax system used to maximise a Rawlsian social welfare function. Further assume that marginal immigration takes place and that the tax system is optimally adjusted to immigration. Using S as an additional tax instrument, a Pareto improvement for the native agents is always feasible.*

[25] The same reasoning as in section 2.3 applies for assuming $M = 0$ in the initial situation.

2.4.6 Characteristics of the tax transfer system after immigration

In this part, I characterise the tax transfer system as it prevails given a specific subsidy S^* or $\hat{S}$. The objective is to see in how far the concept of delayed immigration according to Richter (2004) or Sinn (2005) or the concept of entrance fees for immigrants apply here. The former concept implies that immigrants have to accept a delay in their access to the transfer system. They are, for example, excluded from unemployment insurance payments until they have worked and contributed to the insurance for several years. Although the concept of delayed immigration applies to a dynamic framework, it can also be considered in a static model. Then, one would either consider the model as one based on life-time income or consider the discounted tax and transfer amounts of the future in the current period. Assuming either one case, the concept of delayed immigration implies that immigrants can be net beneficiaries from the tax transfer system, i.e that the tax levied onto the unskilled agents is indeed a subsidy ($T^u(S) < 0$). A further difference between the concept of delayed immigration and my model is that in the former framework it is usually assumed that the tax system is unchanged by immigration.

The introduction of the new tax instrument S has a profound impact on the tax and transfer payments of the three groups of agents.[26] For most parameter values, unskilled agents pay positive tax amounts ($T^u(S) > 0$) whereas the tax amounts have always been negative for $S = 0$. However, there are particular parameter combinations in which $T^u(S) < 0$ is maintained although the native agents benefit from immigration. One such example is the case where $N^s = 10$, $N^u = 120$, $M = 0$ and $\alpha = 0.3$.[27] In this specific situation even immigrants benefit from the redistribution system although to a much lesser extent than the native unskilled agents. The concept of delayed immigration applies. The more radical concept of full exclusion of immigrants from the public sector (including health care and public schooling) as it is discussed in the literature[28] is not the welfare maximising policy here.

Proposition 2.7. *For specific parameter values, there exists a $S = \{S^*, \hat{S}\}$ such that one native agent group is as well off as before immigration, the other native agent group is better off than before immigration and immigrants are net beneficiaries of the tax transfer system consisting of the instruments T^s, T^u and S^* or $\hat{S}$.*

To check this result, set the parameter values indicated above and one of the two possible values of $S = \{S^*, \hat{S}\}$ into $T^u(S) + \frac{dT^u(S)}{dM}$ (from (2.25)).

A second characteristic of the tax transfer system including S concerns the net tax payments of the skilled agents. Immigration per se increases the tax payments of the

[26] In contrast, the marginal tax structure does not change with the introduction of S, it only changes with M.

[27] Section 2.5.1 shows this case in detail.

[28] The full-exclusion alternative is analysed in Myers and Papageorgiou (2000).

skilled agents as $\frac{dT^s}{dM} > 0$. However, once S is introduced, the net tax payments of the skilled agents are effectively reduced again because skilled agents only make net tax payments of $T^s(S)-S$. It might happen, therefore, that the skilled agents pay effectively less taxes after immigration and with S than before. In this case immigration would lower the burden of the tax transfer system on the skilled agents and this should increase inequality even more than just that skilled income increases whereas unskilled income decreases due to immigration.

Proposition 2.8. *After immigration has taken place and $S = \{S^*, \hat{S}\}$ is introduced as a compensatory transfer to all natives, the skilled native agents make higher net contributions to the redistribution scheme than before immigration.*

Proof. Given Assumption 2.2, analytically the following result holds unambiguously:

$$T_b^s|_{S=0} - (T_a^s(S) - S) < 0, \tag{2.36}$$

where $T_b^s|_{S=0}$ is the tax amount that is due before immigration and at $S = 0$ (see (2.24)) and $T_a^s(S) = T^s(S) + \frac{dT^s(S)}{dM}$ is the tax amount after immigration and at $S = \{S^*, \hat{S}\}$. □

2.4.7 Alternative redistribution in the undistorted model

As the mechanisms at work in the income-dependent tax scheme are characterised in detail by now, I want to go back to the model of section 3. There, immigration takes place in an initially undistorted economy and the unskilled natives are the unambiguous losers. As has been shown in section 2.3.2, the introduction of a poll tax on immigrants the proceeds of which are redistributed to all native agents is not always a feasible policy as immigrants may be worse off in such a scheme than under their outside option. An alternative is to introduce an income-dependent tax structure into the post-immigration section 2.3 model instead of the poll tax on immigrants. The benefits of an income-dependent tax are that it allows to redistribute from the skilled winners to the unskilled losers of immigration and that the group of contributors is larger as $N^s > M$. A detriment of such a policy is the efficiency loss due to the distortion in labour supply that is inherent in conditioning taxation on income.

The analysis proceeds in that I compare the undistorted before-immigration utility of the unskilled agents with the income-dependent tax utility after immigration, holding the parameters of the model (N^u, N^s and α) constant. M is increased by immigration. Unskilled utility for the undistorted case is given by (2.1) where income equals consumption and can be calculated from (2.13). For the post-immigration case with an income-dependent tax, I use $U^u|_{S=0} + \frac{dU^u}{dM}\big|_{S=0}$ where U^u can be found in (2.31).

Comparing unskilled utility in the before- and post-immigration cases, the government allows for immigration if the post-immigration utility is weakly larger than the before-

utility because of the assumed welfare function. In contrast to the previous considerations, the immigrants' outside option does not bind here as immigrants are treated as unskilled natives when they migrate.

Proposition 2.9. *Given Assumption 2.2, if marginal immigration takes place in a model without any taxation, there exist specific parameter values for which the introduction of an income-dependent tax scheme does not make the native unskilled agents as well off as before immigration.*

Proof. If M is initially equal to zero, in the case that $N^s = 0.13$, $N^u = 0.55$ and $\alpha = 0.70$ the before-immigration utility of the unskilled agents is larger than their utility after immigration and the introduction of an income-dependent tax system. □

There are many parameter values for which the introduction of an income-dependent tax scheme combined with immigration leads to a welfare improvement by an increase in the utility of the unskilled natives. However, for specific parameter values the government has to resort to the additional use of S to realise welfare-improving immigration.

2.5 Simulations

In the last two sections, I have analysed the impact of marginal immigration onto an economy that disposes in each section of a different set of tax instruments. In section 2.3, I have found that poll taxes which are paid to all natives and which are used to compensate the unskilled natives for their losses due to marginal immigration are not always feasible. In section 2.4, I have identified the influence of marginal immigration once the government runs an optimal income tax scheme. Concerning several endogenous variables, the effect of marginal immigration is ambiguous. Then, I have analysed how a government that aims at compensating the unskilled losers from immigration can make use of S. The instrument S can be used further to guarantee that all native agents are at least as well off as before immigration.

This section serves to illustrate the different phenomena found so far in a compact way. To do so, I have chosen three illustrative scenarios. All three scenarios have in common that output increases with immigration and is higher in the non-distorted model of section 2.3 than in the section 2.4 model. This reflects the efficiency loss that is due to the binding incentive constraint in the optimal income dependent tax scheme. In addition, the skilled agents' utility is always higher in the section 2.3 economy than in the section 2.4 economy and the opposite holds for the unskilled agents. This is the result of the redistribution process under the income dependent taxation.

	sec 2.3.1 bef imm.	sec 2.3.1 aft imm.	sec 2.3.2	sec 2.4.2	sec 2.4.3	sec 2.4.4	sec 2.4.5
l^s	0.8949	0.8960	0.8960	0.8852	0.8864	0.8864	0.8864
l^u	0.1691	0.1686	0.1686	0.1606	0.1602	0.1602	0.1602
Q	11.4409	11.4695	11.4695	11.1794	11.2093	11.2093	11.2093
U^s	0.4004	0.4014	0.4015	0.0554	0.0551	0.0553	0.0554
U^u	0.0143	0.0142	0.0143	0.0430	0.0428	0.0430	0.0431
U^m	0.0143	0.0142	0.0034	0.0430	0.0428	0.0150	0.0075
T^s	–	–	–	0.3354	0.3367	0.3645	0.3720
T^u	–	–	–	-0.0279	-0.0278	-0.0000	0.0074
T_{y^s}	–	–	–	-0.0013	-0.0012	-0.0012	-0.0012
T_{y^u}	–	–	–	0.0769	0.0763	0.0763	0.0763
S^p,S^*,$\hat{S}$	–	–	0.0108	–	–	0.0280	0.0355

Table 2.1: Scenario 1, parameters: $N^s = 10$, $N^u = 120$, $M = 0$, $\alpha = 0.3$

2.5.1 Scenario 1

Particular about scenario 1 (see *Table 2.1*) is the high difference between the number of skilled and unskilled agents in the domestic economy. As a consequence, there is a large gap between skilled and unskilled wage rates and their respective utility levels. This makes it unattractive for the skilled agents to mimic the unskilled natives under the optimal income tax regime which is reflected in the marginal subsidy rate on the skilled (T_{y^s}) being very close to zero. Thus, a high level of redistribution can be realised in the optimal tax section 2.4 economy and this results in utility levels there being very different from those obtained in the section 2.3 economy.

The skilled agents lose from marginal immigration as the effect of the increased tax payment dominates the combined effect of immigration on income and labour supply in equation (2.30). Because of the high number of unskilled agents and the integration of the immigrants into the group of beneficiaries the subsidy payments to the unskilled agents are reduced by marginal immigration, i.e. $\frac{dT^u}{dM} > 0$.

The introduction of S^* to compensate the native unskilled losers is not high enough to also have the skilled agents be at least as well off as before marginal immigration. Therefore, the S needed for a Pareto improvement is higher. As can be seen from U^m in *Table 2.1* which is always positive, all these policy measures are feasible.

	sec 2.3.1 bef immig.	sec 2.3.1 aft immig.	sec 2.3.2	sec 2.4.2	sec 2.4.3	sec 2.4.4
l^s	0.7022	0.7072	0.7072	0.8004	0.7998	0.7998
l^u	0.5036	0.5022	0.5022	0.4494	0.4485	0.4485
Q	17.6093	17.8629	17.8629	16.8301	17.0451	17.0451
U^s	0.2465	0.2501	0.2508	0.1565	0.1591	0.1596
U^u	0.1268	0.1261	0.1268	0.1403	0.1397	0.1403
U^m	0.1268	0.1261	0.0835	0.1403	0.1397	0.1075
T^s	–	–	–	**-0.0056**	-0.0017	0.0306
T^u	–	–	–	**0.0011**	0.0003	0.0325
T_{y^s}	–	–	–	-0.3596	-0.3400	-0.3400
T_{y^u}	–	–	–	0.1667	0.1639	0.1639
S^p,S^*,$\hat{S}$	–	–	0.0426	–	–	0.0328

Table 2.2: Scenario 2, parameters: $N^s = 10$, $N^u = 50$, $M = 0$, $\alpha = 0.72$

2.5.2 Scenario 2

Scenario 2 is characterised by a large α. Wage rates[29] and utility levels of skilled and unskilled agents in the undistorted section 2.3 economy are relatively close to each other, see *Table 2.2*. This makes it difficult to realise redistribution from skilled to unskilled agents via transfer payments in the optimal income tax scheme. Instead the general equilibrium effect becomes very important.

The strictly binding incentive compatibility constraint implies a very high marginal tax rate on the unskilled and a very high marginal subsidy rate on the skilled agents. The fact that the marginal tax rates are the essential channel of redistribution is reflected in the skilled tax payment being negative, i.e the unskilled agents make transfers to the skilled natives to induce them to work more and thus increase unskilled income via the general equilibrium effect. This phenomenon can also be seen in the huge increase in skilled labour supply in section 2.4 as opposed to the section 2.3 economy. Here clearly the high marginal subsidy has an important impact.

Marginal immigration reduces these transfers from the unskilled to the skilled by a high degree. This is due to the fact that immigration increases the number of unskilled in the economy and thus increases the wage rate gap. Therefore, the incentive compatibility constraint is relaxed and this indicates that higher transfers can be realised via the direct channel of redistribution.

The strong effect of marginal immigration onto the marginal tax rates results in a decrease of skilled labour supply despite the increase in the wage rate (the marginal tax effect in (2.28) dominates). Also skilled agents benefit from immigration. Thus, the

[29]In the section 2.3 economy $l^i = w^i$ holds.

	sec 2.3.1 bef immig.	sec 2.3.1 aft immig.	sec 2.3.2	sec 2.4.2	sec 2.4.3	sec 2.4.4
l^s	0.6878	0.6935	0.6935	0.6744	0.6782	0.6782
l^u	0.3242	0.3161	0.3161	0.2401	0.2375	0.2375
Q	6.3067	6.4118	6.4118	5.7655	5.8701	5.8701
U^s	0.2365	0.2404	0.2431	0.1329	0.1311	0.1336
U^u	0.0526	0.0499	0.0526	0.1153	0.1128	0.1153
U^m	0.0526	0.0499	**-0.0158**	0.1153	0.1128	0.0523
T^s	–	–	–	0.0721	0.0791	0.1396
T^u	–	–	–	-0.0480	-0.0496	0.0109
T_{y^s}	–	–	–	-0.0519	-0.0447	-0.0447
T_{y^u}	–	–	–	0.4000	0.3840	0.3840
S^p,S^*,$\hat{S}$	–	–	0.0657	–	–	0.0630

Table 2.3: Scenario 3, parameters: $N^s = 10$, $N^u = 15$, $M = 0$, $\alpha = 0.25$

transfer S^* to compensate the unskilled natives also guarantees a Pareto improvement for all domestic agents.

2.5.3 Scenario 3

Scenario 3 (see *Table 2.3*) stands out due to the relatively small difference in the group sizes between skilled and unskilled agents and a rather small α. The combination of both lets the immigrants' utility be negative once a general poll subsidy to natives and a poll tax on the immigrants is introduced into the non-distorted section 2.3 economy to compensate the native losers. The impact of this negative utility level is that no immigration occurs because immigrants have a better outside option in staying in their country of origin.

As in scenario 1 the skilled native agents lose from marginal immigration. However, in scenario 3 the transfer S^* to compensate the native unskilled losers is also sufficient to make the skilled agents better off than before immigration. Consequently, S^* guarantees a Pareto improvement for the natives directly.

2.6 Conclusions

Limits to immigration are often discussed controversially in Western countries, in particular in election campaigns. Politicians in favour of immigration argue for the realisation of the benefits, e.g. to fill jobs for which no natives are available. In contrast, those in favour of limits argue for the protection of native employees. Because of the widespread fear of unemployment or of being faced with strong competition, the group addressed by

those in favour of limits is important for election outcomes and their support necessary for a sustainable long-run immigration policy. One way to make an immigration policy sustainable is to compensate the losers.

This chapter analyses the impact of immigration onto an economy that runs a Stiglitz (1982)-type tax system. The effect of immigration is that not only the unskilled natives lose but also the skilled natives can be worse off. Thus, the Stiglitz (1982)-type tax system itself is not a good basis to have native voters support an immigration policy with positive numbers of immigrants and immigration has to be accompanied by additional tax measures to be sustainable. A solution to overcome the losses from immigration is to add the idea of delayed immigration as it can be found in Richter (2004) and Sinn (2005) to the model. This concept means that immigrants do not get access to the full range of welfare state benefits from the moment on when they arrive in the destination country. Rather, they first have to contribute to the tax and social security systems for a certain period of time in order to earn the privilege of full access to benefits.

As I show, the formal argument that necessitates the redistribution from migrants to native agents is that the unskilled natives lose from immigration. In an economy that initially runs a redistributive tax scheme, they can only be compensated for the losses once a tax instrument that discriminates between natives and immigrants is introduced. It has to be possible to focus redistribution on the group of the unskilled losers.

Distinguishing between natives and immigrants in the tax system is very useful as it even allows to compensate all native agents for the losses due to immigration and thus make immigration a Pareto improvement. Hence, it sustains wide support for immigration policies with positive numbers of immigrants. Concerning the immigrants feasibility is guaranteed, i.e. there are always potential immigrants willing to migrate given the discriminatory tax scheme.

If initially no tax scheme exists in the domestic economy, the introduction of a discriminatory tax based on distinguishing native agents and immigrants and aiming to compensate the unskilled losers is not always feasible. Feasibility depends on the specific parameters of the model. Due to the fact that a simple poll tax on immigrants does not enable the government to distinguish between unskilled losers and skilled winners of immigration, the burden imposed on immigrants is so high that potential immigrants refrain from migration. Combining the concept of delayed immigration that discrimination between immigrants and natives and a redistributional income tax guarantees a compensation for the unskilled natives. However, for many parameter values such a sophisticated tax structure is not necessary as simple discrimination suffices.

Chapter 3

Taxation and Shadow Economic Activities of Firms

3.1 Introduction

Empirically, shadow economic activities have a considerable extent. One important cause for the creation of shadow economic activities are taxes.[1] The intuitive reason for this is that taxes form a wedge between gross and net incomes, revenues or prices. The higher is this wedge, the more attractive it is to try to avoid the tax payments. Firms are an important actor in the process of tax collection. Not only do they have to pay firm-specific taxes, but also and more importantly via withholding they remit a huge part of overall taxes such as pay-roll taxes and social security contributions to the government. As Shaw et al. (2008) find for the United States, 80% of taxes are remitted by firms.

Empirical evidence suggests that evading firms are different from complying or official firms. De Paula and Scheinkman (2008) find, for example, that informal firms have a higher cost of capital and are less efficient and more labour-intensive than official firms. Alm et al. (2003), Almeida and Carneiro (2009) and Dabla-Norris et al. (2008) detect that informal firms are smaller. In summary, it seems to be a stylised fact that shadow firms have a lower (labour) productivity.

A lot of the existing literature concentrates on the tax evasion decision of individuals. If the focus is on firms and shadow economic activities, the literature mostly assumes homogeneous firms or different forms of strategic interaction. However, the existing literature has little to say on the influence of a specific tax policy on a single firm's

[1] This is e.g. emphasized in Schneider (2000), Bühn et al. (2007), Giles and Tedds (2002) and Dell'Anno (2003). Hill and Kabir (1996) find that marginal tax rates are more important than average tax rates for a firm's decision to evade taxes, whereas Johnson et al. (1998) focus on the ineffective and discretionary application of the tax system as the main determinant.

decision to operate in the official or in the shadow economy. In addition, it does not analyse how the tax policy influences the extent of the shadow economy and by which channels the stylised facts of shadow economic firms that are found empirically can arise.

The heterogeneous firm literature provides a good framework to address these shortcomings. By allowing for heterogeneity of firms in some aspect, as for example differing levels of productivity, my model is well capable of replicating the empirically observed distribution of sales (Axtell (2001)) and the distribution and sizes of firms. In addition, the framework allows to examine the impact of different tax policies on the marginal firms that are indifferent between the official and shadow economy or between market entry and no production. As linear and fixed tax instruments affect firms that differ in productivity in diverse ways, additional insights can be won. Hence, I investigate the consequences of the tax instruments on the single firm's output and shadow decision, for the productivity of the firms in the shadow and the official economy as well as for the extent of the shadow economy. Thus, the model also incorporates effects on the tax base. In contrast, so far the existing heterogeneous firms literature has focused on official sector firms only.

This chapter builds a model that sheds light on the missing aspects. In a heterogeneous firm model in which firms can decide to operate in the shadow economy, I show that taxes induce the small and less productive firms to operate in the shadow economy. Thus, I replicate the empirically found stylised facts concerning shadow firms. Furthermore, I identify the impact of different tax instruments whereby I use a fixed tax and proportional taxes such as a linear tax and a fine rate. The fixed tax is not related to a firm's sales or revenues but applies to every firm operating in the official economy. Such a tax can be interpreted as the revenue-independent parts of the communal excise tax in Germany or as specific registration taxes in other countries.

For the effects of these instruments on the formation and the extent of the shadow economy, I find that the fixed tax is crucial for the creation of a shadow economy, whereas tax instruments such as a linear tax rate or the scale of fines rather determine its size. I assume that the economy is not in an optimum initially and thus I rather focus on reform options. A government that disposes of all of the tax instruments described before and makes use of them initially can increase its revenues to some extent while holding the size of the shadow economy constant. The mechanism to achieve this is to reduce those tax instruments as the fixed per firm fee, which have a strong impact on the choice to opt for the shadow economy, and increase the more innocuous proportional tax rate. Nevertheless, such a policy comes at the cost of reducing overall productivity in the economy because it provides incentives for many rather unproductive firms to enter the production process. In an alternative scenario where the two tax instruments are varied at the same time, a government can reduce the shadow economy without dispensing with any revenue by varying the mix of tax instruments.

Usually, the analysis of tax evading firms focuses on one tax instrument and on the

extent of evasion only, whereas other effects of a government's tax policy are neglected. A government's tax policy, however, does not only have an impact on a firm's decision to evade taxes. It also affects other decisions of the firm. I investigate the impact of the tax instruments on a firm's productivity, output and the overall productivity in the economy, first by analytical comparative statics, then in a calibrated model. Firms adapt to the tax policy by changing their level of output, declaring more or less output and by entering or leaving the market. Thus, overall the number of firms and the productivity within a firm sector are affected by the governments tax policy and there are additional effects of tax increases other than just a smaller tax base that the government should take into consideration. I address these in the chapter.

An additional focus of my work is to quantitatively analyse the effect of tax increases on different firms characteristics. As an example I choose the increase in the German value-added tax in 2007 and find a severe increase in shadow economic activities whereas average productivity remains rather stable. An important insight of the simulation is that the increase in the shadow economy could have been significantly reduced if the increase in the value-added tax rates had been accompanied by a reduction in fixed tax components.

The model I use is a modification of Dharmapala et al. (2008)[2] in which I incorporate a shadow economy. Within a single industry, there are many firms which are heterogeneous with respect to their cost functions due to different productivities. Each firm can decide to produce or not. If it produces, it can opt for either the official or the shadow economy. This decision and the output choice will depend on the government's tax policy. The government has several tax instruments at its disposal. It can levy a linear tax on output, a fixed-firm fee, a detection probability as well as a fine rate as it is found e.g. in Allingham and Sandmo (1972).

The literature on heterogeneous firms and shadow economic activities is still sparse. Two papers stand out, however, the works by Tuzova (2008) and by Prado (2008). Both use a monopolistic competition set-up with heterogeneous firms and differentiated products whereas I look at a homogeneous good sector with many firms producing one good. In contrast to the model of Tuzova (2008) in which the firm that is just indifferent between the official and the shadow economy is given exogenously[3], I endogenise the firm's choice of operating in the official or the shadow economy. Tuzova (2008) finds that the shadow economy can be reduced by lowering taxes or monitoring activities which reflects the view of the classical tax evasion literature. Going beyond, I find that it is possible to reduce the extent of the shadow economy holding revenues constant. Prado (2008) puts

[2]Thus, my model is inspired by Hopenhayn (1992a) and Hopenhayn (1992b) in using heterogeneity in a homogeneous good sector. Whereas Dharmapala et al. (2008) focus on optimal taxation and firms structures taking in particular administrative cost into account, I introduce the opportunity of tax evasion.

[3]Tuzova (2008) assumes a specific output level above which firms are controlled and below which they are not.

a special focus on red tape regulation to get an endogenous sector choice of firms. In his model, reforms of regulation are found to be more important than changes in taxes or monitoring. The tax system is not considered in detail.

The wider firm and tax evasion literature concentrates on strategic interaction and homogeneous firms. Thus, it considers e.g. the tax shifting and tax evasion decision of monopolistic firms (see Marrelli (1984)). Another field of interest has been the effect of tax evasion on the strategic interaction of firms in oligopoly markets as in Marrelli and Martina (1988) and Goerke and Runkel (2006). If there are many homogeneous firms, heterogeneous cost for evading taxes have to be assumed to get official as well as shadow firms (see e.g. Alexeev et al. (2004)).

The importance of shadow economic activities and the relevance of tax evasion are confirmed by the empirical literature. Based on different indirect methods of measurement[4], the size of the shadow economy is estimated to be around 8.7% of GDP in the USA in 2000 (Alm et al. (2003)) or 15% of GDP in Germany in 2005 (Bühn et al. (2007)). Quantitatively, tax evasion is an important phenomenon as well. Although the extent of evaded taxes is difficult to measure, estimates, as for instance Slemrod and Bakija (2004), suggest that in 2002 18% of personal and corporate income tax is not paid to the U.S. tax authorities due to evasion.

The chapter proceeds as follows. In section 3.2, I first present the heterogeneous firm model putting an emphasis on the characteristics of the official and shadow firms. Then, I study analytically the impact of the various tax instruments on the firms and their choice to operate in the official or shadow economy. In section 3.3, I calibrate the model to allow for numerical comparative statics and look at two scenarios in which two tax instruments are varied at once. In addition, I simulate the increase in the German VAT. Section 3.4 provides two extensions to the model, whereas section 3.5 concludes.

3.2 A model with taxation and heterogeneous firms

Within an industry, firms usually differ in size. There are, for instance, small and huge grocery markets, single consultants and large consultancy firms, etc. In their specific industry, these firms produce the same or a similar good but have quite different productivities and size. I want to look at such an industry consisting of many heterogeneous firms and analyse the impact of different tax instruments onto the firms, their productivity, average size and output when firms have the opportunity to avoid taxation by not declaring their revenues.

My model starts with the households which consume the single good of the economy. Then, I characterise the basic elements of the firm sector, defining in particular the firms'

[4]For an overview of the different approaches used to measure the extent of the shadow economy see Schneider and Klinglmair (2004).

heterogeneous productivity. After the tax instruments are introduced, I concentrate on the choices of the firms, deriving in detail how profits in the official economy and expected profits in the shadow economy (the shadow firms face a detection risk) look like conditional on the government tax policy. Thus, I complete the information firms need for *i*) deciding for one of the three options, no production, production in the shadow or in the official economy and *ii*) the consequential choice of optimal output in the two production states. The firm sorting across these options is investigated and I examine the comparative statics of changes in the tax instruments on the firm's sector and output choices.

3.2.1 Households

There is a representative consumer who maximises utility $U(X)$ defined over a homogeneous good X. The consumer supplies labour inelastically to the firms in the X-producing industry and receives all profits as she is also the owner of the firms. Labour serves as the numeraire, so the firms' costs are measured in units of labour. In order to have demand functions that only depend on the good's own consumer price q, the utility function is separable and quasi-linear in labour.

3.2.2 Firms

As in Chaney (2008), there is a fixed mass of firms M in the X good industry each of which maximises profits over output given its firm-specific cost as a function of output. Each firm in the industry faces two types of cost for production, per-firm fixed cost and convex variable cost. The fixed cost have to be paid for establishing the firm and include cost for basic equipment and informational cost about customers. Whereas the fixed cost are industry-specific and thus common to all firms, the variable cost of production differ across firms. The cost of a type-φ firm that produces output y are denoted $c(y, \varphi)$. The productivity parameter φ takes on values over an interval $\left[\varphi^l, \varphi^h\right]$ where $\varphi^l > 0$. It is defined as in Dharmapala et al. (2008) where a higher value of φ decreases the marginal cost of production:

$$\frac{\partial c(y, \varphi')}{\partial y} < \frac{\partial c(y, \varphi'')}{\partial y} \text{ for } \varphi' > \varphi'' \text{ and } y \tag{3.1}$$

An assumption on the cost function that is used further on is the following:

Assumption 3.1. *The variable costs are increasing and convex, i.e.* $\frac{\partial c(y,\varphi)}{\partial y} > 0$ *and* $\frac{\partial^2 c(y,\varphi)}{\partial y^2} > 0$. *The cost function* $c(y, \varphi)$ *is differentiable with respect to* φ *and the cross-partials with respect to* y *and* φ *exist and are continuous.*

Combining the definition of φ as given by (3.1) and Assumption 3.1, it follows that

$$\frac{\partial^2 c(y,\varphi)}{\partial y \partial \varphi} < 0, \tag{3.2}$$

i.e. marginal cost decrease in productivity. This is in fact an extension of the definition of productivity φ as given by (3.1).

3.2.3 Government

The government disposes of four tax instruments which it uses to collect revenues. These are a linear tax on output $t \geq 0$, a detection rate $0 \leq \chi \leq 1$, a fine rate $f \geq 0$ which is proportional to undeclared output and a fixed per-firm fee b that is mostly assumed to be positive[5]. As a firm's productivity is unobservable, taxation can only be based on output, it cannot be conditioned on productivity. Thus, a firm's output decision is important to determine the government revenues arising from a specific tax policy.

The linear tax t relates to output.[6] The detection rate χ represents the government's effort to discover firms that do not declare their taxes.[7] As output is the tax base and there is a linear tax on output, nothing would change qualitatively if f was applied to the taxes evaded instead of undeclared output.

The fixed fee b is a fixed per-firm tax that is independent of sales. Tax structures in various countries comprise such a tax component. In Germany, for example, a part of the communal excise tax is raised independently of the revenues of a firm. In other countries, specific fees exist that firms have to pay for their registration, etc. and that are also independent of sales or revenue. Thus, I interpret b as such a fixed, sales-independent per-firm tax payment.

It is important to note that my model only looks at one specific sector with all firms producing in this sector being subject to the same fixed cost. Hence, the fixed per-firm tax b is equivalent to a proportional tax on a firm's fixed cost and effectively taxes fixed capital of all sorts including its financing cost. As a consequence of applying to the sector-specific fixed cost, b need not be the same across sectors.

An alternative interpretation is to think of b as the fixed cost of starting a business that are related to the public administration. Empirically, there are relevant cost of obtaining legal status to operate a firm as Djankov et al. (2002) calculate. They set up a measure for official expenses that include fees, legal and notary charges, fiscal stamps and the like

[5] In section 3.2.5, I analyse the consequences of assuming b to be negative.

[6] In section 3.4.1, I also take a look at a linear tax on profits. This does not affect the basic results of the model.

[7] In principal it is costly to control firms. I refrain from cost considerations in my model as they do not contribute additional insights.

taking a standard firm[8] as a basis. Taking all the cost related to the public administration into account, they estimate these cost of setting up a business, for example in Germany, to 16% of GDP per capita in 1999. Dharmapala et al. (2008) interpret b as the tax payment that covers the fixed administrative cost of the government when collecting its tax revenue.

3.2.4 Profit maximisation

Knowing its productivity, each firm can decide not to produce at all and thus avoid the fixed cost of production. It certainly does so if it expects to make negative profits when producing.[9] If a firm produces, it has to choose whether to operate in the official economy (the official sector), variables are then indexed by o, or in the shadow economy (sector), here variables are indexed by s.[10] Working in the official economy implies that a firm fully complies to all tax obligations.

The profit maximisation of an official firm looks as follows:

$$\pi = (q - t)y - c(y, \varphi) - b \tag{3.3}$$

The official firm gets the consumer price q for the output y of the good it sells. Besides having to cover the cost of production, it pays the proportional output tax t that distorts the optimal output choice as compared to a world without taxes. In addition, it pays the (non-distortionary) fixed fee b. Because of the fixed cost and the variable convex cost, profit maximisation results in a unique optimal output choice that is denoted $y^o(q-t, \varphi)$. Thus, the profit π^o of an official firm is

$$\begin{aligned} \pi^o(q - t, b, \varphi) &= (q - t)y^o(q - t, \varphi) - c(y^o(q - t, \varphi), \varphi) - b \\ &= \hat{\pi^o} - b. \end{aligned} \tag{3.4}$$

$\hat{\pi^o}$ is defined as the gross profit before the payment of b.

Alternatively, a firm can choose not to declare any revenues. Such a firm is further on called a shadow firm. As the existence of this firm is known to the government, each shadow firm is confronted with the risk that its non-declaration of revenues to the tax authorities is detected whereby the risk of detection is given by the detection rate χ. In case a firm is not detected, it escapes all tax obligations. Consequently, a shadow firm maximises the following expected profit function:

$$\pi = (1 - \chi)\left[qy - c(y, \varphi)\right] + \chi\left[(q - t - f)y - c(y, \varphi) - b\right] \tag{3.5}$$

[8]The standard firm is a domestically-owned firm with limited liability that is situated in the largest city and does not engage in trade.

[9]If a firm decides to start production by paying the fixed cost, its existence is known to the government.

[10]I use the terms official (shadow) economy and sector synonymously.

The first additive term shows the profit of the firm if it is not detected. Thus, it does not contain either a proportional tax or a fixed fee payment. The second additive term indicates the profit of the shadow firm in case of detection. In addition to the normal taxes (as in (3.3)), the firm has to pay the proportional fine at rate f.

Given this structure, the expected profit of a shadow firm is

$$\begin{aligned} \pi^s(q,t,f,b,\chi,\varphi) &= (q-\chi(t+f))y^s(q,t,f,\chi,\varphi) - c(y^s(q,t,f,\chi,\varphi),\varphi) - \chi b \\ &= \hat{\pi^s} - \chi b, \end{aligned} \tag{3.6}$$

with $y^s(q,t,f,\chi,\varphi)$ being the optimal choice of output from maximising the expected profit of shadow economic activities (3.5) and $\hat{\pi^s}$ being shadow gross profits.

The optimal output level for both, an official and a shadow firm, is an increasing function of the good's price q and a firm's specific productivity φ. The higher the productivity, the larger is the optimal output because of the strictly convex cost functions. This implies a positive correlation between productivity and firms size measured in output units.

3.2.5 The firm's official - shadow sector choice

Granted that it is profitable to produce at all, each firm decides whether to produce in the official or the shadow sector. A firm chooses the official economy and pays all its tax obligations if its official profit is higher than the expected profit from a shadow activity, i.e. if $\pi^o > \pi^s$. This occurs if

$$(1-\chi)b \;<\; \hat{\pi^o} - \hat{\pi^s} = (q-t)y^o - c(y^o,\varphi) - \left[(q-\chi(t+f))y^s - c(y^s,\varphi)\right]. \tag{3.7}$$

For $b \geq 0$, the left hand side of equation (3.7) is a weakly positive constant (it can be equal to zero, if either the fixed fee b is set to zero or the government exerts full control on all firms, i.e. $\chi = 1$). The size of the (expected) profit difference on the right hand side of (3.7) depends crucially on the productivity φ of the specific firm considered. The following lemma establishes that its sign is either positive for all φ or negative for all φ depending on whether the tax payment in the official economy t or the expected payment in the shadow economy $\chi(t+f)$ is higher.

Lemma 3.1. *$\hat{\pi^o} - \hat{\pi^s} > 0$ iff $\chi(t+f) > t$ for given q.*

Proof. Given its specific type φ, a firm chooses its optimal output so that marginal revenue equals marginal cost. In the shadow sector, a firm's marginal revenue is $q - \chi(t+f)$, in the official sector it is $q-t$. Assuming that $\chi(t+f) > t$ holds, the marginal revenue is higher in the official than in the shadow sector. Given the strictly convex cost function, for a specific φ-type firm this implies that $y^o > y^s$, i.e. a firm's output is higher in the official sector. As for each firm-type profit is a concave function of output

and thus increases in y as long as $y < y^*$ where y^* is the undistorted profit-maximising output, it results that the gross profit $\hat{\pi^o}$ at the official output y^o is higher than the gross profit $\hat{\pi^s}$ at the lower shadow output y^s, i.e. $\hat{\pi^o} > \hat{\pi^s}$. For $\chi(t+f) < t$, the line of arguments is accordingly. □

As $\chi(t+f) > t$ guarantees that the right-hand side of (3.7) is positive, it is a necessary condition for an equilibrium in which at least some firms produce in the official economy (given that $b \geq 0$). To guarantee the existence of some official firms, the right hand side of equation (3.7) has to be greater than $(1-\chi)b$ for at least some φ. Hereby, the following lemma is of help.

Lemma 3.2. *Given Assumption 3.1 and assuming $\chi(t+f) > t$, the right hand side of equation (3.7) is an increasing function of φ, i.e.*

$$\frac{\partial(\hat{\pi^o} - \hat{\pi^s})}{\partial\varphi} >^{!} \; 0 \tag{3.8}$$

Proof. Using the optimality of a firm's output choice with the definitions of gross profits $\hat{\pi^o}$ and $\hat{\pi^s}$ in (3.4) and (3.6) it follows that

$$\frac{\partial(\hat{\pi^o} - \hat{\pi^s})}{\partial\varphi} = -\frac{\partial c(y^o,\varphi)}{\partial\varphi} + \frac{\partial c(y^s,\varphi)}{\partial\varphi}. \tag{3.9}$$

Then

$$\frac{\partial(\hat{\pi^o} - \hat{\pi^s})}{\partial\varphi} > 0 \Leftrightarrow \frac{\partial c(y^o,\varphi)}{\partial\varphi} < \frac{\partial c(y^s,\varphi)}{\partial\varphi}. \tag{3.10}$$

Given Assumption 3.1, Young's Theorem applies to equation (3.2). It follows that

$$\frac{\partial^2 c(y,\varphi)}{\partial y \partial\varphi} = \frac{\partial^2 c(y,\varphi)}{\partial\varphi \partial y} < 0 \tag{3.11}$$

and the latter inequality guarantees that (3.10) holds.[11] This proves that, given Assumption 3.1, (3.8) holds and thus the right hand side of (3.7) is an increasing function of productivity φ. □

The result that the right hand side of equation (3.7) is an increasing function of productivity allows for a productivity φ such that (3.7) holds with equality. However, at the limits of the support for φ as well as for $b(1-\chi) = 0$ corner solutions may arise. The following assumption ensures an interior solution.

Assumption 3.2. *Assume $(1-\chi)b > \hat{\pi^o}(\varphi^l) - \hat{\pi^s}(\varphi^l)$ and $\hat{\pi^o}(\varphi^h) - \hat{\pi^s}(\varphi^h) > (1-\chi)b$.*

[11] Remember that $y^o > y^s$ holds because of $\chi(t+f) > t$ (see Lemma 3.1) and that $\frac{\partial c(y,\varphi)}{\partial\varphi} < 0$.

Proposition 3.1. *Under Assumptions 3.1 and 3.2 and furthermore $\chi(t+f) > t$, there exists a unique threshold value $\varphi^* \in [\varphi^l, \varphi^h]$ where $\pi^s = \pi^o$, i.e. there exists a firm-type which is indifferent between producing in the official or in the shadow sector. Producing firms with $\varphi > (<)\varphi^*$ work in the official (shadow) economy.*

Proof. Assuming $\chi(t+f) > t$ guarantees that the right hand side of (3.7) is positive (by Lemma 3.1). Whereas the left hand side is a constant, the right hand side is an increasing function of φ (Lemma 3.2 with Assumption 3.1). Thus, if the least productive firm chooses to work in the shadow economy, there will be a firm indifferent between the shadow and the official sector as φ increases the right hand side of (3.7) by Lemma 3.2. Assumption 3.2 guarantees that the indifferent firm-type lies within the support of φ.[12] □

Proposition 3.1 implies that the more productive and larger firms work in the official economy whereas the smallest and least productive firms work in the shadow economy (if they operate at all). For the small firms the advantage of saving on the fixed tax (by Assumption 3.2, $b > 0$ holds) is very important whereas for the large firms this saving is negligible. They benefit from the higher net price for their large amounts of output in the official economy.

So far, I have looked at the particular case that expected tax payments of shadow firms (including fine) are higher than official tax payments ($\chi(t+f) > t$) and that the fixed tax levied on firms (b) is positive. Hence, I have posed two specific restrictions on the tax instruments the government disposes of. Beyond the case studied so far, three other combinations of restrictions on the relation between $\chi(t+f)$ and t as well as on the sign of b are imaginable. Each of the other combinations results in different effects on the firms' sector choice and is investigated below. *Figure 3.1* summarises the results. I start by maintaining the assumption $\chi(t+f) > t$ but assuming $b \leq 0$ in addition.[13]

Proposition 3.2. *Take Assumption 3.1 and furthermore assume $\chi(t+f) > t$ and $b \leq 0$. Then all producing firms work in the official economy.*

Proof. The proof is straightforward given the proof of Proposition 3.1. By Lemma 3.1, $\chi(t+f) > t$ guarantees that the right hand side of (3.7) is strictly positive. The left hand side is zero (for $b = 0$) or negative (for $b < 0$), thus the official profit is always higher than the expected shadow profit. □

Abandoning the fixed fee b or paying a fixed per firm subsidy ($b < 0$) eliminates the shadow economy if $\chi(t+f) > t$. The reason is that the obstacle that lets small firms opt for the shadow sector, i.e. the additional tax cost of official firms, disappears. A

[12] If φ^h is infinite, only the first part of Assumption 3.2 is needed.

[13] Because of Assumption 3.2, $b = 0$ is not viable under Proposition 3.1.

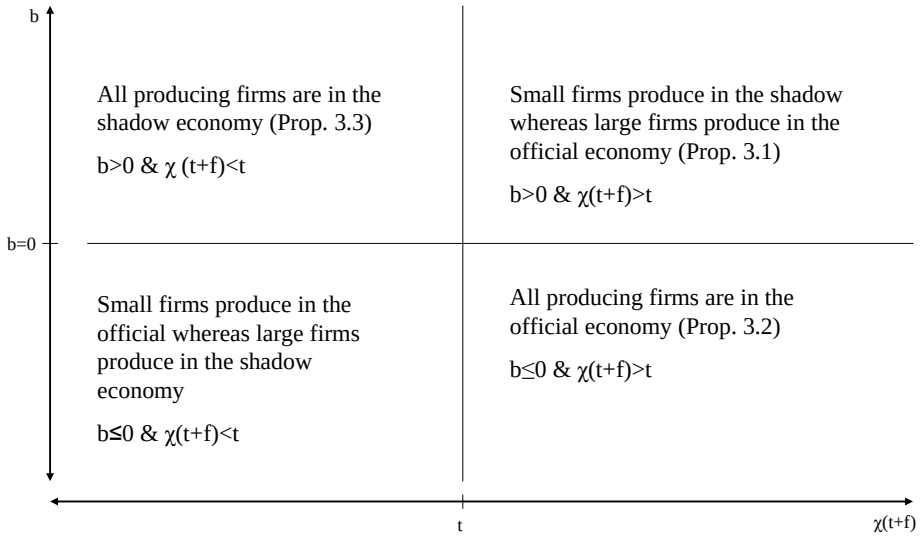


Figure 3.1: The firm's official - shadow sector choice

further deviation from the previous assumptions would be to return to $b \geq 0$ but assume $\chi(t+f) < t$ in addition.

Proposition 3.3. *Take Assumption 3.1 and furthermore assume* $\chi(t+f) < t$ *and* $b \geq 0$. *Then all producing firms work in the shadow economy.*

Proof. By Lemma 3.1, $\chi(t+f) < t$ results in the right hand side of (3.7) being strictly negative. In contrast, the left hand side of (3.7) is weakly positive for $b \geq 0$. Thus, the official profit is always lower than the expected shadow profit. □

The remaining case is that official firms get a subsidy $b < 0$ and the expected revenue from producing in the shadow economy is larger than from official production ($\chi(t+f) < t$). Thus, both sides of (3.7) are negative. Large firms benefit more from saving on the linear tax component than they would from getting the fixed subsidy payment b in the official economy. Thus, they produce in the shadow sector. In contrast, small firms produce few output and therefore, they cannot save a lot on avoiding linear tax payments in the shadow sector. However, for them the subsidy paid in the official economy is important and so they opt for official production. As a result, in this case small firms produce in the official whereas large firms produce in the shadow economy. The intuition is just the inverse that given to Proposition 3.1.

3.2.6 Firm sorting

For the further analysis of this chapter, I return to the case as specified in Proposition 3.1 and assume that Assumptions 3.1 and 3.2 as well as $\chi(t+f) > t$ hold. Firstly, it is

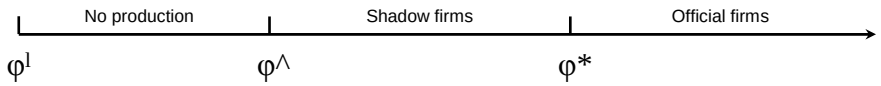


Figure 3.2: Firm sorting

one of the two potentially interesting cases because it allows to have at the same time an official and a shadow economy with their extent depending on the rate of the tax, fine and fee set by the government. Secondly, this is the empirically relevant case. As studies such as those by Schneider and Enste (2000), Cebula (1997) and Loayza (1996) document, the tax burden in a country is an important factor for explaining the extent of the shadow economy. The model presented here finds that the tax system causes a firms structure such that large firms work in the official economy whereas small firms work in the shadow economy thereby replicating the stylised firm distribution as found e.g. in Fortin et al. (1997), Dabla-Norris et al. (2008) and De Paula and Scheinkman (2008).

In Proposition 3.1 I have specified the assumptions which guarantee that amongst the producing firms the less productive firms operate in the shadow economy whereas the more productive firms prefer the official economy. To make comparative statics of variations in the different tax instruments and to analyse the interaction between taxes and the firms' decisions further, some more structure is needed. Therefore, in this section I complete the description of the firm behaviour over the full range of productivities φ and analyse the firm characteristics that go along with it.

A firm only produces if it makes a weakly positive profit, given its productivity parameter φ, the government tax policy and the fixed cost of production. As profit is an increasing function of productivity, some low-productivity firms exist that do not enter the market because the fixed cost cannot be covered. Thus, there is a firm that is just indifferent between not-entering the market and producing and the zero profit condition

$$(q - \chi(t + f))y^s(q, t, f, \chi, \hat{\varphi}) - c(y^s(q, t, f, \chi, \hat{\varphi}), \hat{\varphi}) - \chi b = 0 \tag{3.12}$$

defines its productivity $\hat{\varphi}(q, t, f, b, \chi)$. If a firm's productivity is higher than $\hat{\varphi}$, it is profitable for the firm to produce whereas if productivity is lower, it is not.

Comparably, at a specific φ, firms are indifferent between producing in the shadow and in the official economy. This φ is formally given by

$$\pi^s(q, t, f, b, \chi, \varphi^*) = \pi^o(q - t, b, \varphi^*) \tag{3.13}$$

(see equations (3.4) and (3.6)) and defines $\varphi^*(q, t, f, b, \chi)$. Thus, the sorting of firms based on their productivity can be represented by *Figure 3.2* where $\hat{\varphi} < \varphi^*$ holds for the two cut-offs.

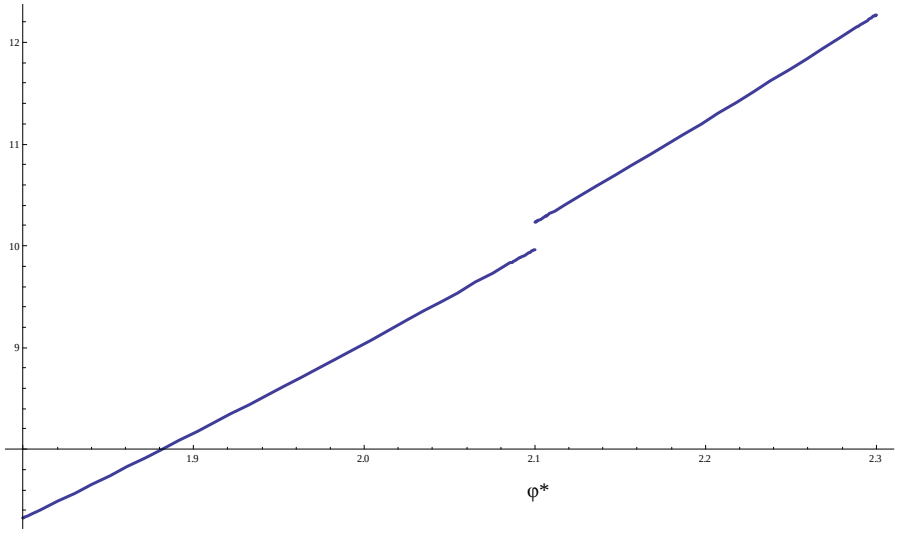


Figure 3.3: A firm's output as a function of φ

A characteristic of the firms is their specific output choice which changes with productivity in the following way: For productivities below $\hat{\varphi}$ firms do not enter production and output is zero. Because of the fixed cost of production, there are no very small firms and the output of a producing firm with productivity $\hat{\varphi}$ is strictly greater than 0. In the range of the shadow economic firms, i.e. for $\varphi \in [\hat{\varphi}, \varphi^*]$ as well as in the range of the official firms, i.e. for $\varphi \in [\varphi^*, \varphi^h]$ output is strictly increasing in productivity. At the cut-off itself, however, i.e. for a φ^*-type firm, optimal output is strictly greater if the firm produces in the official than in the shadow economy. As a consequence, there is a jump in output which is due to the reduction in the marginal cost of production from $\chi(t+f)$ to t[14] and can be seen in *Figure 3.3*.

It is also interesting to look at the relation between a firm's productivity and its tax payments. Within the shadow or official sector, respectively, expected tax payments increase in firm size because the higher output increases the proportional tax payments whereas the fixed b component is unchanged. At φ^* the following holds:

Proposition 3.4. *Firms at φ^* make higher expected payments to the government if they produce in the official sector than if they choose the shadow economy, i.e.*

$$ty^o(\varphi^*) + b > \chi(t+f)y^s(\varphi^*) + \chi b \tag{3.14}$$

Expected tax payments are increasing in φ for $\varphi > \hat{\varphi}$.

Proof. Because of (3.13) and with profits in the shadow and official economy as indicated by (3.4) and (3.6), equation (3.14) is equivalent to

$$qy^s - c(y^s, \varphi^*) < qy^o - c(y^o, \varphi^*) \tag{3.15}$$

[14]Intuitively, this jump is necessary because the indifferent firm at φ^* has to pay more on b (going from $(1-\chi)b$ to b) and saves these additional costs by making higher revenues on its output.

Given φ, a non-taxed firm's profit $\pi^{nt} = qy - c(y)$ is a function of y only. Because of convex cost, profit maximisation results in $q = \frac{\partial c}{\partial y}$. As the marginal revenues for y^o and y^s are smaller than at q (without taxes), it is further known that $y^s < y^o < y^{nt}$. Because of optimality and the strict concavity of π (because of the increasing, convex cost), for the profits it holds that $qy^s - c(y^s, \varphi^*) < qy^o - c(y^o, \varphi^*) < \pi^{nt}$. □

3.2.7 The equilibrium

The model is closed via the goods market equilibrium where the consumer price is determined. Aggregate demand is given by $X(q)$, the representative consumer's demand as a function of the consumer price. Aggregate supply is the aggregate output that all firms produce either in the shadow (first additive term) or in the official (second additive term) economy whereby M accounts for the total mass of firms. Thus, market clearing is given by:

$$M\left[\int_{\hat{\varphi}(q,t,f,b,\chi)}^{\varphi^*(q,t,f,b,\chi)} y^s(q,t,f,\chi,\varphi)f(\varphi)d\varphi + \int_{\varphi^*(q,t,f,b,\chi)}^{\varphi^h} y^o(q-t,\varphi)f(\varphi)d\varphi\right] - X(q) = 0. \tag{3.16}$$

where $f(\varphi)$ is the density function of φ. (3.16) determines the consumer price $q(t, f, b, \chi)$.

3.2.8 Comparative statics

So far I have concentrated on the firms and their characteristics given a specific government policy, i.e. a particular mix of the tax instruments. Now, I focus on changes in the tax instruments and therefore onto the interaction between the tax policy and the industry considered since the different instruments have distinct effects on each firm's output and sector choice. For the analysis, I first look at how changes in different tax instruments influence the optimal choice of output. Then, I examine how the tax instruments impact the sector choice (official-shadow) of a firm and how the number of firms in the different categories changes. Finally, I try to put all these elements together to present the influence changes of tax instruments have on overall government revenues, when firms adapt optimally to the change in the instruments. As there are certain limits to the analytical analysis, section 3.3 provides numerical solutions and discusses further questions. Throughout the analysis I assume that $b(1-\chi) > 0$ and $\chi(t+f) > t$.

Implicit differentiation of the first order conditions of the firm's profit maximisation (from (3.3) and (3.5)) leads to the following comparative statics concerning output. The optimal output of each firm increases in the consumer price q because of the convex cost whereas the linear tax rate t has a negative impact on output as it decreases marginal revenue. Concerning the output choice of shadow economic firms, a rise in the detection rate χ and the fine f decreases output as both reduce marginal revenues.

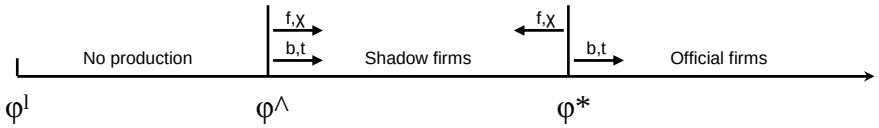


Figure 3.4: Comparative statics of the cut-offs with respect to the tax instruments

Very important is the impact of the government policy on the firms' sector and production choice. By influencing the decision of the marginal firms to opt for production or no-production at $\hat{\varphi}$ or for the official of shadow economy at φ^*, the tax policy changes the firm structure. At $\hat{\varphi}$ as well as at φ^* there are direct effects of the tax instruments and indirect effects via the impact of the tax policy on the consumer price (see (3.16)).

I first concentrate on the direct effects of the tax instruments neglecting general equilibrium price effects because for the direct effects unambiguous results can be obtained. Increases in the detection rate χ or the fine f make shadow production more costly. As a consequence, some firms at or close to $\hat{\varphi}$ leave the market and some of the firms at φ^* change into the official economy (φ^* decreases). Hence, the mass of firms in the shadow sector decreases whereas the mass of firms in the official economy increases. *Figure 3.4* gives a graphical overview of the direct effects of every tax instrument on the cut-offs.

Proposition 3.5. *Holding the consumer price constant, the mass of shadow firms decreases in χ and f, as $\frac{\partial\hat{\varphi}}{\partial\chi} > 0$, $\frac{\partial\hat{\varphi}}{\partial f} > 0$, $\frac{\partial\varphi^*}{\partial\chi} < 0$, $\frac{\partial\varphi^*}{\partial f} < 0$.*

Proof. The above Proposition as well as Proposition 3.6 are derived by applying the implicit function theorem to the two equations determining $\hat{\varphi}$ and φ^*, i.e. to (3.12) and (3.13). □

The linear tax rate t and the fixed rate b make production in general less attractive because they decrease revenues overall. For the official firms the cost increase is higher than for the shadow firms. As a consequence, marginal firms, i.e. those at φ^*, leave the official sector for the shadow economy and the firms indifferent between production and no production stop producing in the shadow economy and leave the market. In effect, these two tax instruments reduce the overall mass of producing firms and in particular the mass of official firms.[15] As φ^* increases, the average productivity φ of firms in the official sector rises. The impact of b and t on the range of shadow firms $(\varphi^* - \hat{\varphi})$ is analytically ambiguous.

Proposition 3.6. *Holding the consumer price constant, the mass of producing firms decreases in t and b as $\frac{\partial\hat{\varphi}}{\partial t} > 0$ and $\frac{\partial\hat{\varphi}}{\partial b} > 0$. In addition, increases in t and b make the average productivity of official firms larger as $\frac{\partial\varphi^*}{\partial t} > 0$ and $\frac{\partial\varphi^*}{\partial b} > 0$.*

[15] The opposite effect is achieved by an increase in the consumer price q which makes production in general more attractive.

Taking these firm level effects into account, I would like to examine the impact of changes in a single tax instrument onto overall government revenues when firms adapt to the taxes and make optimal choices conditional on the government policy. Government revenues are defined as

$$T = M\left[\int_{\hat{\varphi}}^{\varphi^*} (\chi(t+f)y^s + \chi b)\, f(\varphi)d\varphi + \int_{\varphi^*}^{\varphi^h} (ty^o + b) f(\varphi)d\varphi\right]. \tag{3.17}$$

The first additive term indicates the tax payments of the shadow firms, the second additive term reflects those of the official firms. However, as can be seen in (3.17) the total derivatives of government revenues with respect to each of the tax instruments are not very tractable. Thus, the effects of the tax instruments on overall government revenues are analytically ambiguous. In particular, they depend crucially on whether the direct effect of the tax instruments on the cut-offs or the indirect effect via the market clearing consumer price q dominate and how the distribution function looks like. Also a strong interdependence between the various tax instruments exists.

In this section, I have shown the impact of the tax instruments on a firm's choice of output, the sector sizes and average productivity. In order to obtain additional results and in particular to look at revenue effects of changing the mix of instruments, I calibrate the model in the next section.

3.3 Numerical solutions

Given the many channels of influence and the high dimensionality of the problem, the difficulty to obtain analytical results for the impact of changes in the tax instruments on government revenues is not surprising. The impact of the tax policy on the industry when general equilibrium effects are taken into account is even less tractable analytically. Therefore, in this section, I simulate the model using a calibration based on German data. The purpose of this section is threefold. Firstly, I want to give additional comparative static results. Besides, I want to extend the analysis and look at the impact of changes in the tax instruments on various parameters that can be observed and/or are of interest for economic policy considerations. Secondly, I want to analyse in how far different policy objectives can be combined or are mutually exclusive. To do so, I hold overall government revenues constant and investigate whether the extent of the shadow economy can be reduced or the productivity in the economy increased by a differential incidence using b and t. Thirdly, I assess the quantitative impact of the increase in the value-added tax rate in Germany in 2007 by three percentage points on the variables of interest. For the calibration I do not assume that the economy is initially in its optimum. Rather I start from an initial set-up that is as close as possible to the data.

3.3.1 Calibration and parameterisation

To simulate the model, a variety of parameters have to be calibrated and several functional forms have to be parameterised in order to give the model a sufficient structure. In the choice of parameter values, I use German data as much as possible.

A cost function that satisfies the assumptions made is

$$c(y, \varphi) = \xi + \left(\frac{y}{\varphi}\right)^2, \tag{3.18}$$

where ξ is the fixed cost component. With this cost function, outputs are given as

$$y^s = 0.5\varphi^2 \left[q - \chi(f + t(q))\right] \text{ and } y^o = 0.5(q - t(q))\varphi^2 \tag{3.19}$$

for the firms producing in the shadow and in the official economy, respectively. The cut-off values for φ are

$$\hat{\varphi} = \frac{2\sqrt{\xi + b\chi}}{q - \chi(f + t(q))} \text{ and } \varphi^* = \frac{\sqrt{b(1-\chi)}}{0.5\sqrt{(q - t(q))^2 - (q - \chi(f + t(q)))^2}}. \tag{3.20}$$

For the distribution of productivities, I follow the heterogeneous firm literature[16] and base my calibration on the distribution of sales as found in Axtell (2001). Transforming the Pareto distribution of sales into a distribution of productivities, I get the following density function

$$g(\varphi) = 2\alpha \left(\frac{s_0}{s(\varphi)}\right)^{\alpha} \varphi^{-1}, \tag{3.21}$$

where s_0 and $s(\varphi)$ are the shape and scale parameter, respectively, and $\alpha = 1.06$ is the parameter for the sales distribution in Axtell (2001). For s_0 a small parameter value of 1 is chosen.[17] $s(\varphi)$ reflects sales, therefore it is defined as $s(\varphi) = qy^s(\varphi) = 0.5q(q - \chi(t + f))\varphi^2$ (see (3.19)). I choose the sales of the shadow firms and not those of the official firms because the shadow firms determine the lower end of the distribution. I set the upper limit of the support of φ to $\varphi^h = 1,000,000$. Higher values for the upper limit of the support do not change the results.

To determine the goods market equilibrium, it is necessary to assume a specific demand function. For simplicity, I take a linear demand function of the form

$$d(q) = D - q \tag{3.22}$$

where D is a constant that I set equal to $15,000$.[18]

[16] See e.g. Chaney (2008), Ghironi and Melitz (2005), Vannoorenberghe (2008) and Abel-Koch (2010).
[17] Doubling s_0 would result in the same consumer price and fine rate but a smaller mass of firms.
[18] The qualitative results of the calibrated model do not hinge on this specific choice.

The model assumes output to be the tax base for the linear tax rate t. An existing sales-based linear tax in Germany is the value-added tax (VAT).[19] It also fits the model in that the firms are responsible for transferring it to the government. Since 2007 the standard VAT rate in Germany is 19%. To fit it into the model as presented in section 3.2, which assumes a unit tax instead of an ad valorem tax, it is transformed to $t(q) = \frac{0.19q}{1.19}$.

The specification of a detection risk χ poses some difficulties as the extent of the shadow economy is unknown. Therefore, I take the audit probability as detection probability instead.[20] Bernasconi (1998) finds an audit probability of at most 9% for private US households, for large firms it is almost 100% (Slemrod and Bakija (2004)). As there are no data available for Germany, I am geared to the US data and take 7% as a start value.[21]

The fixed fee b may be interpreted as the revenue-independent components of the German communal excise tax or as fees that have to be paid to government agencies when registering a new firm. Hence, it is difficult to find a decent calibration value for it. I choose a small value of 1.4 for b. This value is smaller than the fixed cost ξ of getting access to the heterogeneous technology (see below).[22]

The remaining parameters, i.e. the fine rate f, the mass of firms in the industry M and the fixed cost component ξ are calibrated so as to replicate certain stylised facts of the German economy. These are the size of the shadow economy and the relative fixed cost of setting up a business. In addition, a degree of freedom is needed to assure market clearing. Hence, I determine q as a function of the parameter values.

The size of the shadow economy is mostly defined as the estimated value of shadow economic activities in a country's official GDP. A variety of methods has been developed to measure the extent of shadow economic activities (for a survey of the methods see e.g. Schneider and Klinglmair (2004)). Most of the macro estimation methods find a size of the shadow economy in Germany in the range of 14% to 17%.[23] For my calibration, I choose 15% of GDP as the size of the shadow economy that is to be replicated.

[19] In practice, the input VAT is deductible, wages, interest payments and the like are not, though. A way to incorporate this would be to assume that a specific share of sales is due to input values and reduce the tax base accordingly.

[20] In the US an audit takes place when inconsistencies arise either within a declaration of taxes or between the data that have been transmitted to the IRS directly and the tax declaration (Slemrod and Bakija (2004)). Thus, if an audit actually takes place, some evasion is likely.

[21] In this model a detection probability of $\chi = 1$ would fully eliminate the shadow economy. However, in real life there are important cost of auditing so that it is not optimal for a government to try to achieve full control.

[22] I hereby assume that the cost to start a production process are higher than those for completing the official registration process.

[23] Bühn et al. (2007) estimate the shadow economy to be 17% in 2005 using a MIMIC method, Schneider and Enste (2000) estimate 14.7% based on a currency demand approach for the year 2000, by a DYMIMIC model Schneider (2005) estimates the shadow economy to be 16.8% of GDP in 2003 whereas the most recent outlook for 2009 published by IAW (2009) forecasts a value of 14.6% for 2009.

It is difficult to get data on firms' fixed cost. Djankov et al. (2002) provide estimates of the cost that have to be incurred to set up a business. Hereby, the focus is on cost for all procedures somehow related to the public administration. Therefore, they take the number of administrative procedures as well as the monetary and time cost into account and give a value for the fixed cost as a percentage of GDP per capita. For Germany, their value is 15.69% of nominal GDP per capita in 1999. I calibrate the model to this value because more recent data are not available. As the model presented here only consists of one industry, the sales of this industry are the total creation of value in the economy. Therefore, it is easy to determine GDP.

I solve the model numerically with the objective to guarantee market clearing, a size of the shadow economy of 15% and a fixed cost share in GDP of 0.1569%[24] and get a price of $q = 5.53$, a fine rate of $f = 13.58$ and a mass of firms of $M = 6,233$ with an assumed fixed cost of $\xi = 2.3$.

The fine rate that emerges from this calibration is higher than anything that can be directly observed in reality. For risk-neutral actors as in my model this problem is known in the literature. Lower fine rates in practice may be due to tax morale, for instance (see Traxler (2010) for a survey of the relevant literature on this), or to fines with a high leverage. For example, firms which engage in shadow economic activities and are convicted of doing so can be excluded from public contracts in specific sectors in Germany.

3.3.2 Numerical comparative statics

The model is now fully calibrated to fit the size of the shadow economy and the (public administration related) fixed cost of starting a business in Germany. The associated parameter values for the four tax instruments considered are $b = 1.4$, $t = 0.72$ (this corresponds to a value added tax rate of 19%), $f = 13.58$ and $\chi = 0.07$. The calibration allows me to study detailed comparative statics for changes in one tax instrument when the other three are held constant.

Besides the effect on overall government revenues, I focus on the effect on several characteristics of the firms and the economy. Thus, I look at the size of each sector in terms of the number of firms and output, the productivity in each sector and that of the entire economy as well as on firm profits. The number of producing firms in the economy is given by

$$N^w = N^s + N^o = M \int_{\hat{\varphi}}^{\varphi^h} g(\varphi) d\varphi \tag{3.23}$$

[24]The result of Djankov et al. (2002) relates the fixed cost to GDP per person. As the baseline calibration has around 130 producing firms, I divide overall GDP by 100 and thus assume there to be approximately as many consumers as firms. In an alternative calibration, I have used a factor of 1,000. The qualitative results and relative quantities are robust to this change.

where index w refrains to the entire economy. Total output is

$$y^w = y^s + y^o = M \int_{\hat{\varphi}}^{\varphi^*} y^s g(\varphi) d\varphi + M \int_{\varphi^*}^{\varphi^h} y^o g(\varphi) d\varphi. \tag{3.24}$$

As an observable measure of productivity I choose average firm size which is defined as output relative to the number of firms producing it, i.e. $\frac{y^i}{N^i}$ with $i = w, s, o$. The profit of the shadow and official firms is given by (3.4) and (3.6), government revenues by (3.17).

To compute the comparative statics, I hold three of the four tax instruments constant and let the fourth vary in small steps. Then, I first calculate the market clearing price for each parameter step of the instrument and subsequently the values of the variable of interest. The basic parameters that characterise the economy such as the mass of firms, the parameters of the cost function, the demand function and the density function of productivity φ are held constant throughout. The ranges within which I let the tax instruments vary are the following: $b \in [0.3, 8]$, $VAT \in [0.12, 0.22]$, $f \in [11.5, 18]$, $\chi \in [0.063, 0.1]$.[25]

Comparative statics for changes in χ and f

Table 3.1 gives an idea of the effects of an increase of each tax instrument by around 7% and summarises the effects explained here in detail. The fine rate f and the detection rate χ are substitutes in that they have the same qualitative impact on all of the variables considered. Their effects go into the same direction at both cut-offs. This can be seen most easily for f because the fine rate f is always connected to χ in a multiplicative way. Of course, χ is also relevant in the components χb and $(1-\chi)b$ that determine the cut-off productivities $\hat{\varphi}$ and φ^*, respectively. These additional effects, though, do not have a dominant effect on the result.

Increases in the detection rate χ and the fine rate f decrease the number of firms in the shadow economy as the indifferent firms at $\hat{\varphi}$ and φ^* switch into non-production or into the official economy, respectively. This shows that the direct effect of these tax instruments as identified in section 3.2.8 dominates the indirect effect (via q). As the number of shadow firms decreases[26], so do the total profits of shadow firms and their total output whereas the opposite holds for the official firms. Consequently, the number of official firms increases and also official output and overall profits rise. In contrast, profits per firm decrease as rather small firms enter the official sector.

Within each sector the average firm size and therefore productivity decreases. This is very plausible for the official economy because the firms switching from the shadow to the official economy are unproductive relative to the existing firms. For the shadow

[25] Going beyond the lower boundary for f, b or χ as well as going beyond the upper boundary for t corner solutions would arise. The remaining boundaries of the intervals are chosen so as not to leave the equilibrium values too far.

[26] Large increases in either one tax instrument eventually eliminate the shadow sector.

	χ	**f**	**b**	**t**
	0.07 ⇒ 0.075	13.5 ⇒ 14.5	1.4 ⇒ 1.5	0.19 ⇒ 0.204
	(+7.14%)	(+7.41%)	(+7.14%)	(+7.37%)
Δ number of firms	-2.90%	-2.55%	-0.24%	2.13%
Δ number of official firms	58.26%	57.99%	-7.19%	-46.73%
Δ number of shadow firms	-9.19%	-8.46%	0.48%	7.16%
Δ productivity sha. firms	-34.51%	-34.42%	7.21%	78.92%
Δ productivity off. firms	-12.61%	-12.61%	2.37%	18.13%
Δ overall productivity	2.99%	2.61%	0.24%	-2.09%
Δ government revenues	1.27%	1.21%	0.13%	5.79%

Table 3.1: The effects of a 7% increase in each tax instrument

economy this is not so obvious. Here, the most and the least productive firms leave the sector. It seems to be the case that the output loss at the top end is relatively large compared to the overall number of firms leaving the shadow economy. In contrast to the productivity decrease in each sector, over all sectors productivity increases because the firms with the lowest productivity leave the market. At first glance, this result might look contradictorily, it is not, though. As the shadow sector becomes smaller, its weight in the determination of overall productivity diminishes. Thus, the high level of productivity in the official sector gets more important for the average and this increases overall productivity.

Government revenues are an increasing function of χ and f because the tax payments of the shadow firms increase in both tax instruments. Every firm that switches from the shadow to the official sector increases its payments to the government (see Proposition 3.4). These two effects dominate the loss of tax revenues due to firms leaving the market. As an example, *Table 3.1* presents the effects of a 7% increase in each tax instrument.

Comparative statics for changes in b and t

The general effect of both the linear tax rate t and the fixed per-firm rate b is to let some official firms change into the shadow economy and have some shadow firms give up production. Thus, the direct effect of these tax instruments onto the cut-off productivities dominates. This is reflected in the number of firms in each sector. Whereas the number of official firms is clearly decreasing in t and b, the effect on the number of firms in the shadow economy varies (see also the graphical illustration in *Figure 3.5*). For low values of b the number of shadow firms increases whereas it decreases for high values of b. This reflects that for increases based on originally low values of b more firms enter the shadow economy from the official economy than quit production. Concerning t, the number of shadow firms is unambiguously increasing. The reason is that t is the more important for the sector choice the larger a firm's output. Thus, the effect at the upper cut-off dominates the one at the lower cut-off, although many more firms exist at $\hat{\varphi}$ than

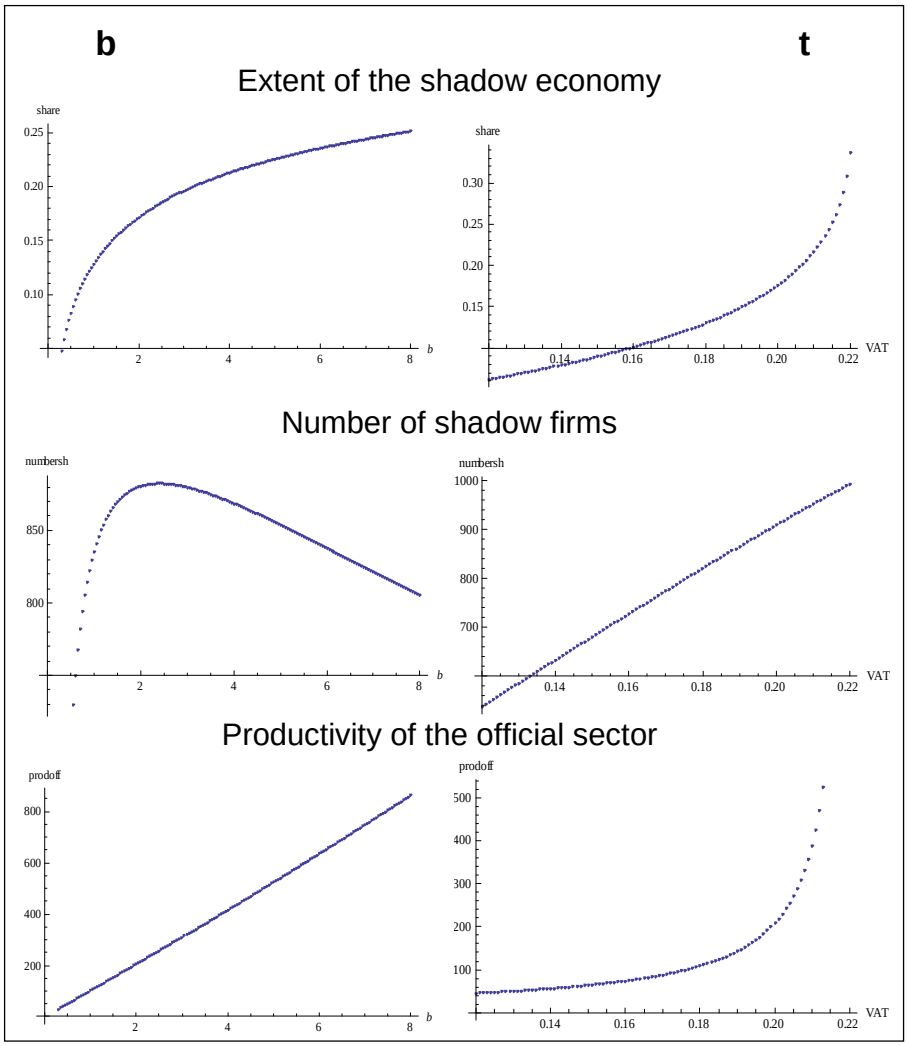


Figure 3.5: Some comparative statics: b and t

at φ^*.

Another consequence of the fact that firms switch from the official into the shadow sector and from the shadow sector into no-production due to increases in b or t is that productivity increases within each sector and across the entire industry. The least productive official firms become the most productive shadow firms and the least productive shadow firms quit. The increases in productivity are convex in b and concave in t. The reason for this is that increases in b get less and less important to the marginal firms because the marginal firms become bigger. In contrast, t increases output proportionally. Therefore, an increase in t always drives at about the same mass of firms out of the official economy into the shadow economy. Overall productivity decreases in t.

The output of the shadow economy increases in t or b. In contrast, output of the official economy decreases because the output of the large number of firms leaving the official economy cannot be balanced by price induced additional output of the remaining firms. Taken together the extent of the shadow economy defined as the share of sales of shadow firms to GDP (overall sales) increases. As profits and outputs are directly linked (profits are an increasing function of output), profits behave just like output. Profits per firm increase in the official economy because only the most productive firms remain. Government revenues first increase in t and b but start to decrease when the effect of the loss in the number of firms dominates the additional revenues from the remaining firms.

3.3.3 Differential incidence of the linear and fixed tax

The previous section has shown the effect of a change in a single tax instrument on interesting characteristics of the economy. This section addresses possible advantages of using one tax instrument instead of another one. Hence, I look at the interaction between the two tax rates, the linear output tax t and the fixed tax b.[27] This is of interest because these two tax instruments allow to steer different parameters of the economy in various ways. A government that wants to achieve specific goals might want to weigh the instruments accordingly. In a first part, I assume that overall government revenues stay constant on the level of the basic calibration case and look at the differential incidence of increasing one tax rate while lowering the other one. In a second part, I look at whether it is possible to increase overall government revenues when the objective is to hold the extent of the shadow economy constant.

Holding government revenues constant

I start with the analysis of the differential incidence of b and t. When the linear output tax t is reduced from a level of 19.2% to a level of 15%, the fixed tax b has to be increased from 0.82 to 44 to hold overall government revenues constant. The marginal rate of

[27] f and χ are held constant over all changes in b and t.

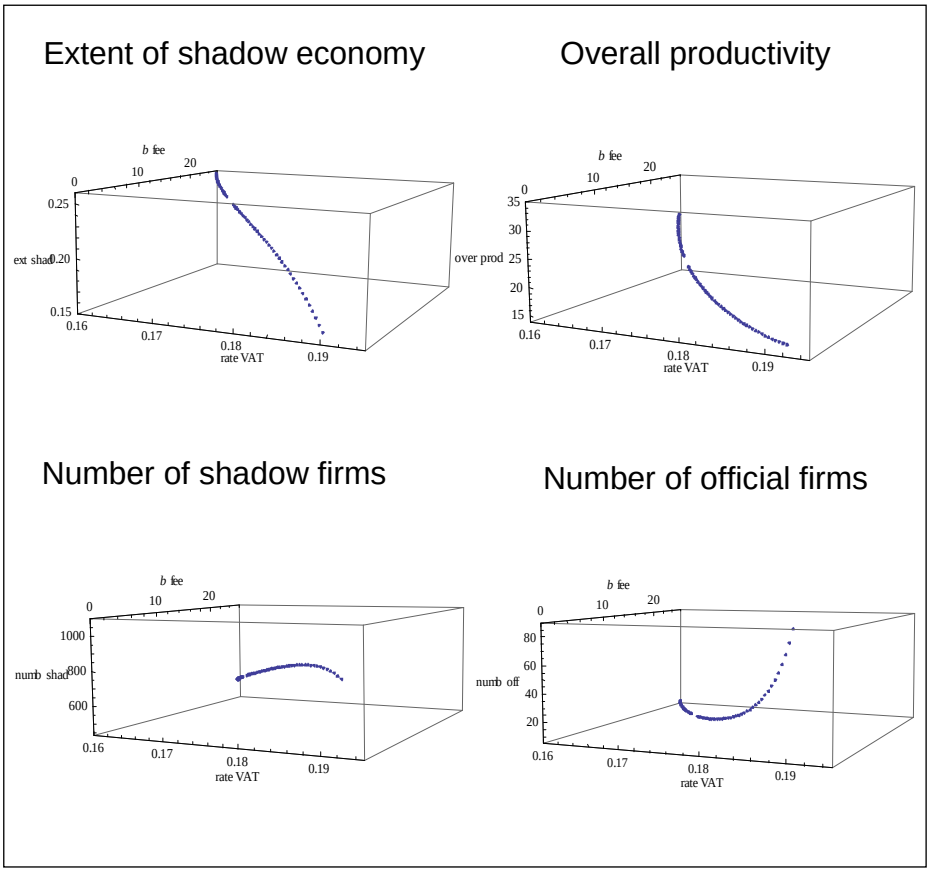


Figure 3.6: Differential incidence of b and t holding overall government revenues fixed

substitution between the two tax rates is decreasing such that for the same decrease in t, increases in b become bigger and bigger the lower the level of t is. This result is explained by effects of b and t onto both cut-offs. Principally increases in b or t go into the same direction, letting some firms switch from official to shadow production and others leave the market. However, in the differential incidence case, both tax rates are changed in opposite directions. Thus, the effects are opposing. Which effect dominates is essential for the outcome and depends on the Pareto distribution of productivities.

When t is lowered from 19.2% and b is raised so as to hold overall government revenues constant, GDP stays (almost) constant. The consumer price in equilibrium adapts to this incidence (increases slightly) in such a way that overall output is left unchanged before and after the tax changes. This does not imply, however, that other variables stay constant as well.

The two lower graphs in *Figure 3.6* show the impact of a differential incidence where t is varied between 15% and 19.2% on the number of firms in each sector. The number of shadow firms first increases when t is decreased from 19.2% on and then falls whereas the number of official firms always becomes lower though at a decreasing rate. When b is low, there are many comparably small firms close or at the cut-off φ^* in the official economy. For these small official firms an increase in b is relatively more important than for the large firms and increases in b are more relevant than decreases in t because output is comparably low. This explains the trend in the number of official firms. For the shadow economy, the influx of firms from the official economy first dominates the outflow of firms to no production at $\hat{\varphi}$ where χb is decisive. The higher b gets, though, the lower is the influx of firms at φ^*. Thus, eventually the number of firms switching to no-production at $\hat{\varphi}$ dominates. For these firms b is the essential factor as their output is negligeable and the lower becomes t the higher has to be the increase of b to compensate the revenue loss. Correspondingly, the effect at $\hat{\varphi}$ becomes more and more important.

The differential incidence of increasing b on the average output per firm and thus on productivity is unambiguous. Decreases in t and consequential increases in b increase the productivity of both shadow and official firms and also overall productivity of the X industry (see *Figure 3.6*). For the official firms the effect is directly linked to the number of firms. As only the most productive firms remain in this sector, this clearly increases productivity. An analogous argument can be found for the shadow economy. Here, relative productive firms enter the sector - they are the firms that leave the official economy at φ^* - whereas relative unproductive firms stop producing. Both effects go into the same direction and increase productivity in the shadow economy.[28]

The extent of the shadow economy as measured by the ratio of the sales of the shadow firms to GDP can be reduced by decreasing the fixed fee b and increasing t (see *Figure 3.6*). For this to happen, the sales of the official sector have to grow relatively to those

[28] In contrast, increases in t and consequential decreases in b decrease overall productivity of the X industry.

of the shadow sector. In fact, it can be shown graphically that shadow output decreases whereas the output in the official sector increases. There are two effects at work here. Firstly, whereas the biggest and most productive firms leave the shadow economy to produce officially, many additional firms enter the shadow sector from non-production at $\hat{\varphi}$. Thus, output in the shadow economy decreases. Secondly, there are more official firms and the effect of the number of firms seems to be stronger than the reduction in each official firm's output due to a higher distortion because of the increased t.

Taking all theses effects together, a policy that aims at increasing the productivity in the economy by increasing b and decreasing t in a differential incidence has to accept that the extent of the shadow economy increases. This does not have to imply that also the number of shadow firms increases though. A policy oriented towards fighting the shadow economy should refrain from fixed-fee like tax components and concentrate on proportional tax instruments once the sector consists of heterogeneous firms. A lot of additional official firms emerge but they are on average less productive.

Result 3.1. *A government can reduce the extent of the shadow economy while holding overall government revenues constant by increasing the linear output tax t and decreasing the fixed tax b. Such a policy decreases overall productivity of the industry. See Figure 3.6 for a graphical illustration.*

Holding the extent of the shadow economy constant

Assuming it is the government's objective to hold the extent of the shadow economy constant, it can still improve on other policy objectives by varying b and t. Methodically, I vary the VAT between 12% and 22% and solve the model for the market clearing price q and the value of the fixed fee b[29] which ensures that the extent of the shadow economy stays constant. I then look at the changes in the characteristics of the economy these tax changes induce.

Increases in the linear tax t and counterbalancing decreases in the fixed tax b allow to increase overall government revenues by a large extent (see *figure 3.7*). The number of firms in the shadow as well as in the official economy increases. Thus, the effect of b on the cut-offs seems to outweigh the effect of t on them. Productivity decreases in each sector and overall. This is the direct result of the additional number of firms in the economy because the new firms can only come from the cohort of the previously non-producing firms and these are the least productive ones.

Result 3.2. *A government can increase overall revenues while holding the extent of the shadow economy as measured by shadow sales in total GDP constant by increasing the linear output tax t and decreasing the fixed tax b accordingly. Decreasing t and increasing b while holding the extent of the shadow economy constant would result in lower government revenues. See Figure 3.7 for a graphical illustration.*

[29] Over this range of VAT rates, b decreases from 6.3 to 0.002.

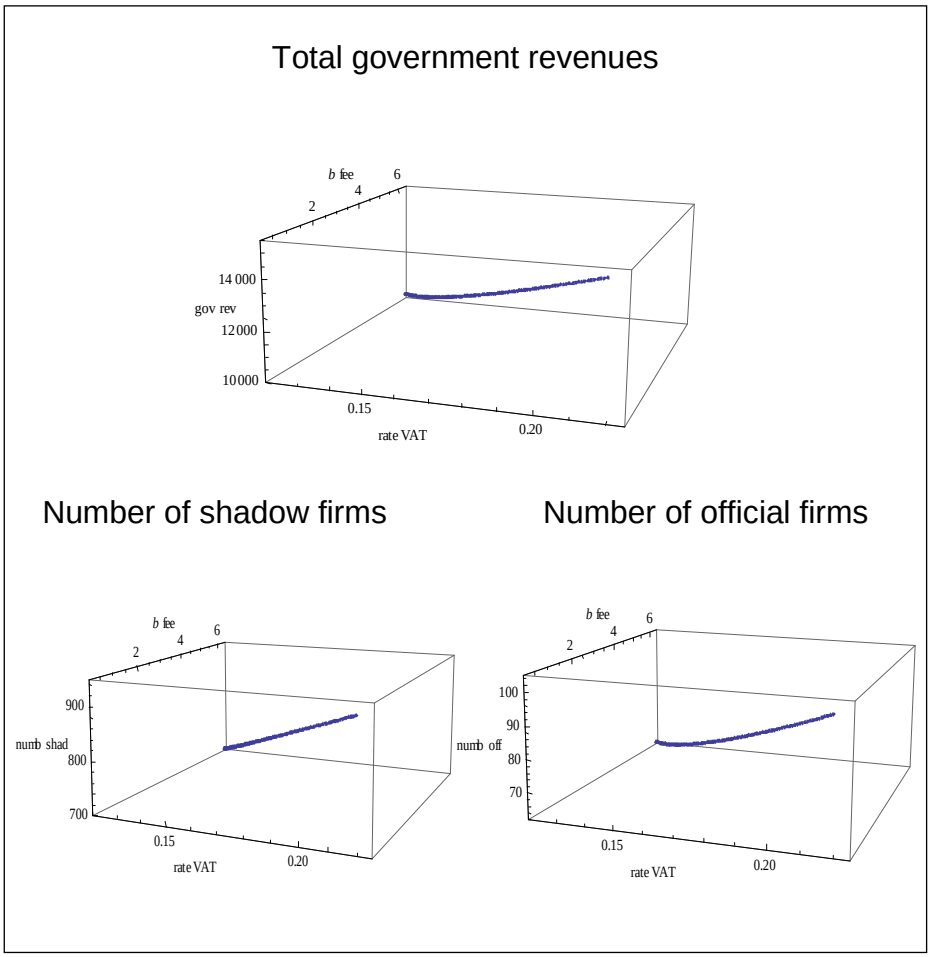


Figure 3.7: Differential incidence of b and t holding the extent of the shadow economy fixed

Thus, the simulations allow to analyse the impact of the tax instruments on several structural variables that characterize the industry. It can be of interest for economic policy setting to take such considerations into account.

3.3.4 Simulation of the 2007 VAT increase in Germany

So far, I have considered qualitative effects of changes in the tax instruments. Now I want to present an idea of the relevance of these effects. As an example, I take the rise of the value-added tax rate in Germany from 16% to 19% on January 1, 2007 . The scenario is calculated in the following way: I calibrate the economy with a VAT rate of 16% to match the size of the shadow economy in 2006 of 15% of GDP and the fixed cost share in GDP just as I have done in section 3.3.1. The fine rate that matches this shadow size and fixed cost is $f = 11.56$. Then, I hold all parameters and tax instruments constant and only increase the VAT rate in the formula that specifies t ($t(q) = \frac{VATq}{1+VAT}$) by three percentage points. The values presented (see also *Table 3.2*) result from this change and the induced rise in the consumer price to guarantee market clearing.

The increase in the value-added tax rate from 16% to 19% (i.e., an increase by 18.75%) results in an increase of government revenues in the model by around 14%[30]. The aggregate variables remain comparably stable. The total number of firms is increased by 4.4%, overall productivity is decreased by 4.32% and total output stays constant.[31] Thus, looking at these three aggregated characteristics of the economy, the increase in the value-added tax does not result in significant effects. What changes decisively, however, is the size of the shadow economy measured as the share of sales of shadow firms to GDP. It more than doubles.

The increase of the shadow economy can be explained by analysing the changes at the cut-offs. At the lower cut-off, $\hat{\varphi}$ the firms are very small. As they produce very little, any change in the proportional tax rate barely has an impact on them. Thus, only around 40 (out of a total of 939 producing) firms quit production.

In contrast, very important changes take place at the upper cut-off. For firms at φ^*, the linear output tax t is much more relevant than the fixed fee b. In addition, because of the Pareto distribution the mode of the distribution of official firms is at φ^*. Thus, any change to these firms has a large quantitative impact. For the details of the quantitative effects of this scenario see *Table 3.2*.

In a nutshell, the impact of the VAT increase on the size of the shadow economy is enormous. Almost all official firms change into the shadow economy. The switching firms reduce output a bit because the marginal cost of production at φ^* are higher for

[30] As an important part of government revenues comes from the other tax instruments (b, f), this is not very surprising (and, of course, much more than in reality).

[31] Government revenues T are defined by (3.17), the total number of firms N^t is given by (3.23), total output y^t is given by (3.24) and overall productivity is $\frac{y^t}{N^t}$.

	VAT = 16%	VAT = 19%	change
Total government revenues	11,463.90	13,071.40	14.02%
Size of the shadow economy	0.15	0.33	119.51%
Average profit of a shadow firm	3.44	8.99	161.19%
Average profit of an official firm	324.12	5,759.14	1,676.87%
Number of shadow firms	849.65	976.00	14.87%
Number of official firms	88.98	3.98	-95.53%
Total number of firms	938.63	979.97	4.40%
Productivity in the shadow sector	2.63	5.03	91.10%
Productivity in the official sector	143.40	2,538.19	1,669.96%
Overall productivity	15.97	15.30	-4.21%
Output of shadow sector	2,234.21	4,904.56	119.52%
Output of official sector	12,759.60	10,089.90	-20.92%
Total output	14,993.81	14,994.46	0.00%
Consumer price	5.30	5.40	1.89%

Table 3.2: Effects of the VAT increase in Germany

the shadow firms. However, as this effect only holds for a minority of firms and as the total number of firms stays almost constant, the overall productivity is rather stable. The gain in the productivities in both sectors hinges, of course, on the fact that the least productive official firms become the most productive shadow firms thus increasing productivity in each sector.

As the example indicates, an increase in the linear output tax affects the extent of the shadow economy severely. In contrast, the overall characteristics of the economy like the number of producing firms, their output and productivity remain rather unaffected. At first glance the outcome of the simulation seems to contradict the Results 3.1 and 3.2. The simulation shows a severe increase in the extent of the shadow economy due to a higher VAT rate whereas an increased t is used in Result 3.1 to reduce the size of the shadow economy. The essential difference between both outcomes is that in the Results 3.1 and 3.2 I make use of both tax instruments, b and t are varied at the same time. In contrast, in the VAT simulation the countervailing effect of b is missing and thus only the direct negative impact of an increase in t remains. As a consequence, when determining the consequences of a policy change it is essential to take additional effects into consideration.

3.4 Extensions

In this section, I present two extensions of the model. Both of them are based on a weakening of one or more of the previously made assumptions, the objective being to

show the robustness of the model to changes in specific assumptions. Since corporate tax systems often use a firm's profit as the tax base, in a first extension I replace the linear output tax by a linear profit tax. In a second extension, I introduce a third sector of activity for firms. In addition to the official or the shadow sector, firms can choose unregistered production which implies that they cannot be detected by the government.

3.4.1 Profit tax

So far the linear tax t has been set proportionally to output. Now, this tax instrument is replaced by a linear tax θ that is proportional to a firm's profits. With such a profit tax θ, the profits of an official economy firm look as follows:

$$\pi^{op} = (1-\theta)(qy^{op} - c(y^{op}, \varphi) - b) \tag{3.25}$$

where p is added to the index to indicate the profit tax case. The expected profit of a shadow firm then is

$$\pi^{sp} = (1-\chi(\theta+\gamma))(qy^{sp} - c(y^{sp}, \varphi)) - \chi(1-\theta-\gamma)b \tag{3.26}$$

(instead of (3.6)). Here, γ is the fine rate that is proportional to undeclared profits. Again, firms have two choices. They can decide on entering the market or not producing at all and if they enter they can opt for either the shadow or the official sector. A firm decides for the alternative which offers the highest profit. Therefore, it works in the official economy if $\pi^{op} > \pi^{sp}$, i.e.

$$(1-\chi)b < (\chi(\theta+\gamma) - \theta)(qy - c(y(q,\varphi),\varphi)). \tag{3.27}$$

Analogously to Proposition 3.1, the larger and more productive firms work in the official sector whereas the less productive but still producing firms work in the shadow economy. This is summarized in the following Proposition:

Proposition 3.7. *Take Assumptions 3.1 and 3.2 with profits as given in equations (3.25), (3.26) and furthermore assume* $(\chi(\theta+f) > \theta)$*, for* $b(1-\chi) > 0$*. There exists a unique threshold value* $\varphi^I \in \left[\varphi^l, \varphi^h\right]$ $\pi^{sp} = \pi^{op}$*, i.e. there exists a firm type which is indifferent between producing in the official or in the shadow sector. Producing firms with* $\varphi > (<)$ φ^I *work in the official (shadow) economy.*

The proof follows along the same lines as the one of Proposition 3.1. The assumption $(\chi(\theta+f) > \theta)$ guarantees that the right hand side of equation (3.27) is positive and is a necessary condition for an equilibrium with official and shadow firms. The left hand side is a positive constant whereas the right hand side is an increasing function of φ because profits are an increasing function of productivity.

With a profit tax, firms which are indifferent between producing in the shadow or in the official economy choose the same optimal output level because the marginal output decision is undistorted (in contrast to the linear output tax case and thus there is no jump in output here). The comparative static results concerning the impact of the parameters on the cut-off values of φ are parallel to those found in the linear output tax case.

3.4.2 Non-registered firms

Another extension of the model concerns the firms' choice options. I add an additional option for firms which, in effect, replaces the no production option. Whereas the existence of firms in the two previous firm categories (official and shadow firms) was known by the government, there can also be firms that are not. These firms are called unregistered (indexed by ur). The essential characteristic of these unregistered firms is that they do not pay the fixed cost of production and thus do not have access to the heterogeneous technology. As argued before, the fixed cost for a legal establishment pave the way towards various financing means, patents, etc. thus give access to the heterogeneous production technology. The unregistered firms may produce with the lowest productivity φ^l and can thus enter the market. This assumption on the unregistered firms replicates the stylised fact that informal firms face higher cost of production and are less competitive (see e.g. Fortin et al. (1997))

The profit of such an unregistered firm is:

$$\pi^{ur}(q,\varphi^l) = qy^{ur}(q,\varphi^l) - c(y(q,\varphi^l),\varphi^l) + \xi = \hat{\pi^{ur}} + \xi \tag{3.28}$$

where ξ is the fixed cost component that has so far been integrated in the general cost function $c(y,\varphi)$. A firm decides to pay the fixed cost and works in the shadow economy if its shadow profit is higher than either the profit as an unregistered firm or the profit as an official firm,

$$\hat{\pi^s} - \chi b > \max\{\hat{\pi^o} - b, \hat{\pi^{ur}} + \xi\}. \tag{3.29}$$

From the first inequality it can be seen that the least productive firms always work as unregistered firms. If theses firms produced in the shadow economy, profits would be negative or only slightly positive for them as they would have to pay the fixed cost. Without the fixed cost, however, the profit is always positive. Since the profit of the unregistered firm π^{ur} is a constant, whereas $\hat{\pi^s}$ is an increasing function of productivity φ, there is a cut-off value $\tilde{\varphi}$ such that all firms with a productivity greater than $\tilde{\varphi}$ work in the shadow economy (if the comparison is limited to that of unregistered vs. shadow economy). The second inequality is the same as before so that the largest firms produce in the official economy for sure.

If one allows for the three sectors to coexist, it is not automatically guaranteed that there are always firms in each sector. Depending on the parameters, it might happen

that no firm chooses to work in the shadow economy because at $\tilde{\varphi}$, the official economy might be more attractive than the shadow economy. All parameter combinations which guarantee an outcome according to (3.29) imply the existence of three sectors. Then the smallest firms are unregistered, the medium productivity firms are observable and dispose of a heterogeneous technology but evade taxes and the most productive firms always choose to work in the official economy.

3.5 Conclusions

In contrast to the literature that looks at optimal tax policies, I consider reforms that take place in an environment without existing optimal taxes. I thus concentrate on policy options intended to improve the status quo when it is not realistic to introduce a fully-new tax structure from scratch and where the tax system has developed to its actual structure that is not optimal for many historical or political reasons. In this context, it is important to take a broad view on the many consequences of a tax reform. I have shown the effect of different tax instruments on the firm's decision to produce in the shadow sector and various firm characteristics as well as the subsequent revenue effects for the government.

I have introduced a firm's choice to operate either in the official or the shadow economy into the homogeneous good heterogeneous firms model of Dharmapala et al. (2008). A firm's decision to produce in the shadow economy responds to the government's choice of tax instruments which are a linear output tax, a fixed tax, a fine rate and a detection probability. Given these instruments, I can replicate the stylised fact that the most productive firms work in the official economy while the less productive firms opt for operating in the shadow economy. The least productive firms do not take up production at all.

The analysis shows that the fixed tax is essential for the existence of a shadow sector. In contrast, the linear tax rate and the penalty fee rather determine the extent of the shadow sector. Thus, I have specified the impact of the different tax instruments. A higher detection risk and a higher fine rate reduce shadow economic activities because they increase the expected penalty for tax evasion. A higher linear tax and a higher fixed tax rate make production per se less attractive because they increase the cost of any kind of production. Nevertheless, each of these two instruments increases the overall productivity of the economy as some of the small shadow firms quit production.

Although the effect of the linear and of the fixed tax rates go into the same direction, the two instruments are not perfect substitutes. The reason is that firms which are indifferent between shadow production and no production are relatively small. Thus, the fixed tax is highly relevant for them. In contrast, the linear tax has a bigger impact on the decision of the comparably large firms that are indifferent between operating

in the official or the shadow economy. Because of this differential impact of the two tax instruments, it is possible for the government to reduce the extent of the shadow economy while holding overall revenues constant by increasing the linear tax and by decreasing the fixed tax at the same time. Alternatively, the government can increase revenues without extending the shadow economy through adapting the composition of its tax instruments. Thus, the analysis emphasises that the tax mix has important implications for the firm sector that go beyond its impact on the extent of the shadow economy.

In the numerical simulations of the increase in the German VAT rate, I show that the increase of 2007 could have had a much smaller impact onto the extent of the shadow economy if at the same time fixed tax components like e.g. the revenue independent part of the communal excise tax[32] had been reduced.

[32]This tax is not fixed over all firm sizes but independent of output at least in the short term.

Chapter 4

Incentives in Early Childhood Care

4.1 Introduction

In Germany, family policy and in particular policies concerning child care are a topic of controversial debate. Two traditions exist in the different parts of Germany. In East Germany, there is traditionally a large supply of full-time, publicly subsidised child care and high enrolment shares of children in day-care centres go along with a relatively large female labour supply participation. In contrast, in West Germany, relatively few children below the age of three attend publicly subsidised day-care centres and availability is limited. In addition, the majority of day-care slots is for part-time care only (see e.g. Wrohlich (2006)). To improve the situation and to give mothers of young children the choice of whether or not they want to work, the federal parliament has enacted a law that gives every parent of a less than three year old child a legal claim to a child care slot from the year 2013 on.[1] The new conservative-liberal government plans to complement the legal claim with the introduction of a 'cash-for-care subsidy' to which all those parents are eligible who do not make use of their claim to publicly subsidised childcare for their under three year old.

The envisaged reform has already caused a lively debate about potential consequences and policy objectives. The conservative liberal coalition forming the federal government emphasises its willingness to give parents the freedom of choice between educating children at home or giving them into early childhood care. They also want to honour the efforts made by parents for raising their children. In contrast, the opposition parties stress that the lump-sum subsidy sets incentives in particular for low-income households to refrain from publicly subsidised nurseries[2], thus depriving specific groups of children from quality early childhood education.[3]

[1] For children older than three such a claim already exists.

[2] Anecdotal evidence indeed seems to confirm this argument.

[3] There is an extensive literature that analyses the impact of early childhood care on skill formation,

In this chapter, I evaluate a particular policy in the German state (Land) of Thuringia that is close to the proposed federal policy to draw conclusions on the expected outcome of the introduction of the federal policy by 2013. On July 1st, 2006, Thuringia has introduced a legal claim to a child care slot for every two-year old child. In addition, parents of two-year old children who do not make use of full-time publicly subsidised child care become eligible for a cash-for-care subsidy of 150€ per month. The subsidy is paid for the 12 months between the child's second and third birthday. Although the policy focuses on two-year old children only, the combination of a legal claim and a cash-for-care subsidy as well as the amount paid are close to the federal proposal.

I evaluate the impact of the cash-for-care subsidy on the parents' childcare choice. Given a median net household income in Thuringia of 1,470€ (for the year 2006), the amount of the subsidy of 150€ per two-year old and month is quantitatively important.

A parent of a young child has to decide whether he wants to work and make use of public or private childcare provision or whether he cares for the child instead of working. An important factor of influence on the childcare choice are the cost of childcare facilities as these are opportunity cost of working. The cash-for-care subsidy increases the opportunity cost by a large amount of 150€ as the subsidy is paid conditional on not using the standard childcare in a publicly subsidised nursery.

Prior to the introduction of the cash-for-care subsidy, 80% of mothers of two-year olds made use of such nurseries. As the treatment group is the particular group of parents of two-year olds, the law is only introduced in one state and the introduction takes place in a sharp design applying to every two-year old from July 1, 2006 on, I use the law as a quasi-natural experiment.

The effect of the cash-for-care reform on the enrolment of two-year olds in publicly subsidised day care is studied by using the (potentially) exogenous eligibility of the particular age-group of two-year olds in Thuringia. The idea is to compare the treatment group, i.e. the cohort of two-year old children, to control groups along three dimensions: time, age groups and state. The two-year olds that are eligible for the policy are compared to the same cohort before the introduction of the policy in Thuringia in 2006. As the new policy is introduced at a fixed date, the identification of both groups is not difficult. I use a difference-in-difference (DD) design comparing the treatment group to the cohort of two-year olds in other states where no cash-for-care reform has taken place to capture general shocks on the age group. Then, I extend the DD estimation to a difference-in-difference-in-difference (DDD) design by adding the three-year olds as a comparison group and account for possible common shocks to young children in general.

From a theoretical perspective, one would expect parents to consider the cost of day care relative to their household income and therefore expect a decrease in the enrolment

social behaviour and success in school. This literature finds positive effects of early childhood care in particular if certain quality standards are respected, see e.g. Carneiro and Heckman (2003), Love et al. (2005) and Temple and Reynolds (2007).

of children from families with lower incomes. Moreover, there are negative effects on female labour supply and thus potentially on life-time earnings inequality. Standard labour supply theory predicts that the introduction of a cash-for-care subsidy reduces female labour supply as both, the income as well as the substitution effect, go into the same direction. The size of these effects depends on parents' preferences over child care time as well as on the availability, cost and quality of child care.

This chapter seeks to enrich the political debate with empirical evidence exploiting the case of Thuringia. Based on county-level data, I find that the introduction of the cash-for-care subsidy has significantly reduced the enrolment of two-year old children in day care. The share of two-year olds enrolled in nurseries to the total cohort has decreased by 9 percentage points from an initial level of 80%. Thus, the subsidy has a quantitatively important effect on the parents' child care choice. In particular I show that the cash-for-care subsidy mostly affects the extensive child care decision whereas no clear effect can be found on the intensive choice. Further findings are: *i*) the causal effect is larger than the effect one would expect from looking at descriptive statistics. *ii*) I identify a positive trend in the enrolment of two-year olds across states. *iii*) I detect that the reduction in the enrolment share of two-year olds due to the introduction of the cash-for-care subsidy is the higher, the larger is the number of young children per woman and the higher are the voting shares of the conservative CDU party in the respective county. In contrast, the reduction in the enrolment share is the lower, the better educated the workforce is.

The cash-for-care subsidy reform is a rather rare policy reform. Comparable policies can only be found in Norway and Finland (ZEW (2009)). The case of Thuringia is particular as there is no rationing on the supply side of day care, whereas rationing is a relevant issue in Norway (Schøne (2004)). As a consequence, I can disentangle demand and supply effects and identify a quantitatively large demand effect of the introduction of the cash-for-care subsidy.

As opposed to the literature, I focus on enrolment instead of female labour supply recognising that in countries such as Germany many mothers do not work but still make use of day care (see e.g. Bick (2010)). In particular with respect to the positive impact of early childhood education on skill formation, enrolment plays an important role in the long run. Schøne (2004) rather concentrates on female labour supply finding a small reduction. Moreover, I make us of a natural control group that exists because within the same institutional environment of Germany other states have not introduced the policy. Thus, I do not have to rely on possibly different time periods to apply a DDD design what strengthens my identification.

The Finnish home care allowance as analysed by Ilmakunnas (1997) based on survey data is quite different as the allowance is means-tested and taxable, whereas in Thuringia the amount parents forego by making use of publicly subsidised day care is the same for all children. A policy reform going into the opposite direction is analysed by Sánchez-Mangas and Sánchez-Marcos (2008) for the case of Spain. There, women get a monthly

cash benefit of 100€ per child under the age of three if they work. In a DDD estimation approach, they identify a positive impact of the subsidy on female labour supply. Haan and Wrohlich (2009) simulate the introduction of such a subsidy that is paid conditional on female labour force participation for Germany. In a structural model they predict a positive impact on female labour supply. Baker et al. (2008) study the introduction of subsidies for day care in Quebec and find a large increase in the use of day-care facilities.

This chapter is also related to the large literature that analyses the impact of changes in childcare cost on the mother's labour supply using different reforms. Kornstad and Thoresen (2007) provide a survey of several results in a paper that models rationing in detail. For the US and Canada in particular, the literature finds strongly negative effects of childcare cost on labour supply, see e.g. Averett et al. (1997), Cascio (2009), Kimmel (1995) and Kimmel (1998) for the USA and Lefebvre and Merrigan (2008) and Powell (1997) for Canada. In contrast, the effects found for European countries are comparably low or insignificant, see e.g. Brink et al. (2007) and Lundin et al. (2007) for Sweden, Havnes and Mogstad (2009) for Norway, Wrohlich (2006) for Germany and Wetzels (2005) for the Netherlands. An important factor that explains these differences could be rationing.

Section 4.2 provides an overview of the family policy for young children in Thuringia and the details of the cash-for-care reform. Section 4.3 sketches a theoretical framework and introduces the identification strategy. After the data have been presented in section 4.4, section 4.5 presents the results of the identifying regressions. Section 4.6 provides some robustness checks and extensions, before section 4.7 concludes.

4.2 Institutional background

This section focuses on the institutional environment in the German state of Thuringia where the cash-for-care subsidy has been introduced in 2006. As the subsidy is introduced for two-year olds, I focus on the policies concerning children up to age three. The group of two-year old children that is of special interest for my analysis is particular as it is neither covered by a federally guaranteed legal claim to a child care slot nor by federally paid child-raising allowances. Instead, different German states apply different policies to the parents of two-year olds.

Thuringia: the new scheme

Figure 4.1 gives an overview of the relevant policies for young children in Thuringia. Starting from July 1, 2006 on, Thuringia replaced the existing child-raising allowance that is described below by a cash-for-care subsidy and introduced a legal claim to publicly subsidised childcare for every two-year old child.[4] Under the new scheme, parents dispose

[4]Saxony-Anhalt is the only other German state which guarantees a childcare slot to every two-year old child. In Rhineland-Palatinate a legal claim will be introduced in August 2010. The cash-for-care

of an additional income of 150€ if they do not make use of publicly subsidised day care for their two-year old child. Thus, the cost of day-care increase by 150€ and I refer to this amount as the cash-for-care subsidy.

Practically, the new allowance is a bit more complicated. Parents get between 150€ and 300€ per child and month depending on the number of children older than two they have. If, however, parents make use of a nursery for their two-year old child, 150€ of the allowance go to the day-care facility[5] so that parents dispose of 150€ more once they raise their child at home. Parents who make use of a nursery for less than 9 hours per day, which is the definition of full-time use, only have to pay the proportion of 150€ that corresponds to the the ratio of day-care hours of the child to 9 hours per day. The usual fees for childcare facilities are maintained. As a consequence, parents of two-year old children have to pay the regular nursery fees that apply to all children between ages two and six.

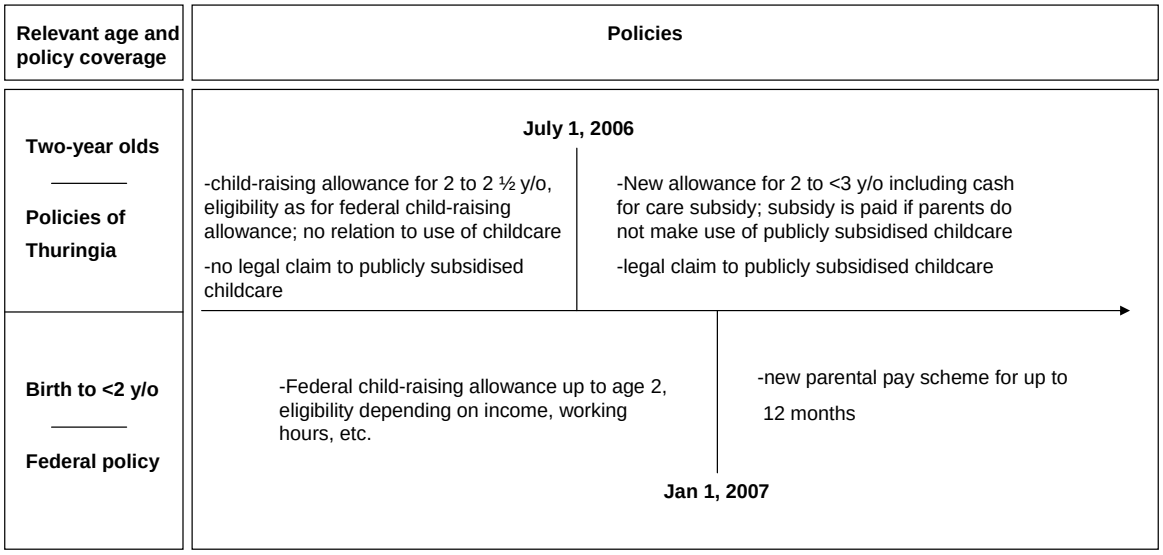


Figure 4.1: State and federal policies affecting children younger than three in Thuringia

The previous policy in Thuringia and federal policies

Parents of young children have been eligible for a federal child-raising allowance (Bundeserziehungsgeld) that depends on certain income and working hours conditions at most up to a child's second birthday. Hence, the eligible parent must not work for more than 30 hours per week and only gets the maximal amount of 300€ for a one-year old child if the annual family income does not exceed 16,500€[6]. Until July 1, 2006,

subsidy is unique in Germany.

[5]Below, I consider possible changes in incentives that the cash-for-care subsidy might cause for providers.

[6]The income threshold is 13,500€ for a single-parent family. For higher incomes, the allowance is reduced.

the state of Thuringia, one of Germany's 16 states, paid a child-raising allowance to all those parents of children between ages 2 and 2.5 who had made use of the federal child-raising allowance up to the child's second birthday. The conditions for eligibility and the amounts paid are the same as for the federal child-raising allowance. Parents with children aged 2.5 to three were not eligible for any sort of child-raising allowance. There was no legal claim to a publicly subsidised childcare slot, but the large majority of parents sent their two-year old child to a publicly subsidised nursery for long hours.

Since January 2007, the federal allowance scheme has been replaced by a parental pay scheme (Elterngeld) for up to 12 months[7] that pays a parent 67% of her previous net income up to at most 1,800€ per month (the minimal amount is 300€). Also, by federal laws, parents get a child benefit of 154€[8] per month and child for every first to third child and have a legal claim on a child care slot from age three on.

In my analysis, I make use of the introduction of the new scheme on July 1, 2006, focusing on the cash-for-care subsidy. While the child-raising allowance for 2 to 2.5 year olds in Thuringia was paid independently of the parents' childcare choices, the cash-for-care subsidy is paid conditional on the children not attending a publicly subsidised nursery. Thus, it introduces an additional fixed cost component for publicly subsidised day care. I exploit the large price increase to analyse the effects on the enrolment of children in publicly subsidised childcare facilities.

Table 4.1 gives an idea of the financial choices parents make before and after July 1, 2006. As an example, I choose a family living in the city of Erfurt with a 2.5 year old child.[9] The median net household income in Thuringia in 2006 is 1,470€. I choose exemplary household income situations to characterise two distinct cases and calculate the total income when childcare cost and child-related transfers are accounted for. In the example, one family has a monthly net income of 800€, so a bit more than half the median income, the other family disposes of a net income of 3,000€, i.e. twice the median income.

As *Table 4.1* shows, the introduction of the cash-for-care subsidy makes it more attractive to care for the child oneself instead of making use of full-time day care. The difference between day care and at-home-care is much larger in 2007 than in 2006. In addition, the example exposes that the cash-for-care subsidy is much more relevant for the total income of low-income households.[10]

[7] 12 months is the maximal period for one parent, if both parents make use of the allowance, the maximal period of payments is 14 months.

[8] This amount was paid for every first, second or third child between 2002 and 2008, from the fourth child on parents got 179€ per fourth and additional child and month. Since then the benefit payments have been raised.

[9] The choice of a specific city is important for determining the nursery fees and board cost.

[10] Of course, the day-care decision often is interrelated to a labour supply decision that affects household income as well. The difference between home care in 2006 and 2007 as well as the relative increase in the attractiveness of home care, however, are unaffected by the labour supply decision.

Table 4.1: Family with one child aged 2.5 in Erfurt, monthly amounts in €

childcare	**income + child benefit**	**nursery fee + board**	**in 2007: subsidy**	**total income**	$\frac{\textbf{subsidy}}{\textbf{tot income}}$
day care	800 + 154	0 + 47	0	907	
at home 06	800 + 154	0	0	954	
at home 07	800 + 154	0	150	1104	13.6%
day care	3000 + 154	132 + 47	0	2975	
at home 06	3000 + 154	0	0	3154	
at home 07	3000 + 154	0	150	3304	4.5%

Availability

When analysing the parents' childcare decision, the literature, see e.g. Kornstad and Thoresen (2007) emphasises three important factors of influence: availability, quality and cost of childcare. In many countries availability of publicly subsidised day care is an important issue because there is rationing as demand exceeds supply, see e.g. Chiuri (2000) and Wrohlich (2006). If the availability of childcare services is rationed, important behavioural effects due to reforms might not be reflected in the capacity utilisation of publicly subsidised nurseries. Rather, the effects take place in private childcare arrangements and the adaptation of female labour supply.

However, evidence suggests that, in contrast to all Western German states, rationing is not an issue for two-year old children in Thuringia (which lies in East Germany). Rationing is neither the case in the number of slots provided nor in the desired time of care. Not only do parents have a legal claim to a child care slot but also already before the introduction of the legal claim to a childcare slot, 80% of the two-year old children in Thuringia went to a publicly subsidised nursery. In addition, in each county the number of slots available in day-care institutions exceeds the number of slots occupied (see *Table 4.8*). Furthermore, the day-care centres in Thuringia have long opening hours. In the city of Erfurt, e.g. 97 out of 98 publicly subsidised nurseries offer 10 or more hours of day care.

Quality

The quality of day-care provision in Thuringia is controlled by the state government. Only if certain criteria concerning the qualification of the personnel, the children-teacher ratio[11], the facilities, cleanliness and foot acreage per child are met, the childcare institution may offer approved slots. The quality standards are decided upon on the state level. Thus, the same standards apply in the whole state and it is well possible to detect changes in the standards by analysing the legislation and decrees. None such changes have taken place between March 2006 and March 2007. Also, the payment of the large

[11]The nursery-school teacher child ratio has not changed between 2006 and 2007 remaining at 10 children between ages two and three per nursery-school teacher.

public subsidies is subject to the compliance with the state-wide standards.

An additional criterion for quality are opening hours. Although, there is no survey of the opening hours of the day-care institutions, it is known how long children attend nurseries and the average time has increased between 2006 and 2009. As a consequence it is very unlikely that opening hours have shortened. Altogether, there is no evidence that the introduction of the cash-for-care subsidy could have had a negative impact on quality. The standards have not been changed and the number of employees has been stable.[12] In addition, under the new regime, childcare centres get the 150€ from the parents. Thus, they have incentives to improve quality to please the parents.

Cost

The last issue are cost. In addition to transferring the cash-for-care subsidy to the childcare facility, parents of two-year old children have to pay the regular nursery fees that apply to all children between ages two and six. These nursery fees are set by the childcare provider to cover the excess of cost over direct public subsidies. Providers are forbidden to make profits. Thus, the fees are not driven by market forces as, e.g. changes in demand or in the parents wage rates. As a consequence of different cost structures, fees vary between counties and within the counties they may vary by provider. There are also different fee structures in place. By law, fees have to take into account either the number of children in publicly subsidised nurseries or the parents' income, they may account for both. To parents with very low income the fees are refunded by the social welfare offices.

The law introducing the cash-for-care subsidy also specifies that public providers of day-care facilities must not change their fees between 2006 and 2007. Although, the share of day-care facilities that are run in public responsibility varies between counties, often a considerable part is public. Given the large number of providers and the local structure, it is difficult to estimate the behaviour of the non-public providers.

To give, nonetheless, an idea of the different fee structures and their changes in the counties of Thuringia, *Table 4.2* presents an overview of five counties. The information stem from the youth offices of the respective counties. As there is no central data collection of fee scales or effective fees paid, the collection reflects the different information provided. Besides giving an idea of the structures in place, *Table 4.2* serves the purpose of showing that no exceptional increases have taken place between March 2006 and March 2007. If there were large changes, one would expect these to have an impact on the parents' childcare decision beyond the impact of the cash-for-care subsidy.

As *Table 4.2* indicates, the scale of fees has been unchanged for the public day-care centres in Erfurt and Eisenach in the relevant period between March 2006 and March 2007. Only in 2008, the scale has been changed in Eisenach where a fixed fee scale has

[12]The composition of the personnel in the nurseries of Thuringia is very stable. There are 8386 full-time equivalents of pedagogic personnel in 2006, 8177 in 2007, 8321 in 2008 and 8764 in 2009.

Table 4.2: Fee structures in different counties in Thuringia

County	**childcare fees for full-time care, all amounts are monthly**
City of Erfurt	city council fee structure, used by almost all providers, valid since April 1, 2001, 98 day-care centres • no fees if family income is below 800€ • 148€ (max) if income is larger than 3,050€ • higher income thresholds, if family has more than one child • board fees (in addition) 47€ per child
City of Eisenach	fee structure in the three day-care centres run by city itself • Jan. 1998 - April 2008, fixed fees: first child: 76.50€, second child: 53.70€, third child: 30.70€ • since May 2008, income-dependent fees: – no fees if family income is below 900€ – 125.50€ (max) if income is larger than 2,500€ – 70% of fees for second child • board fees (in addition): 1.88€ per day and child
City of Jena	almost all providers use city council fee structure, 25 day-care centres, fees depend on income and number of children • 2005-July 2006: min: 80€, max: 260€ • Aug. 2006 - Dec. 2007: min: 0€, max: 244€ • since 2008: min: 0€, 13% of income up to max: 190€
Nordhausen	15 public and 6 private providers in the county, fees depend on income and/or number of children • average fee paid in 2005: 75.79€, in 2008: 86.86€
Kyffhäuserkreis	56 day-care centres, 20 different providers: • for 17 centres fees have not changed between 2005 and 2008, in the others at least one increase • average fee paid in 2005: 67.25€, mid 2006: 75.71€, end 2007: 87.86€ • almost only fixed fees

been replaced with one where fees vary by income. Jena is an example of a county that has reduced the fees for the day-care centres. The income-dependent fees have been between 80€ and 260€ up to July 2006 whereas they are between 0€ and 190€ since 2008. The counties Nordhausen and Kyffhäuserkreis have a very heterogeneous structure of childcare providers which each decide on their own fees scale. Fees have increased in the relevant time period but not by exceptional amounts. In Kyffhäuserkreis the fee changes vary a lot by day-care institution. Whereas for 17 out of 56 facilities no change in fees has occurred between 2005 and 2008, the others have increased their fees at least once. Given the evidence presented, it seems reasonable to assume that no systematic increase in childcare fees has taken place between March 2006 and March 2007.

The section has shown that the introduction of the cash-for-care subsidy of 150€ per month constitutes a potentially important change to a households's childcare decision by its impact on available income. The change in income is the higher, the lower is the family's total income. Concerning the important factors of influence for the parents' choice, availability, quality and cost, no systematic changes have occurred in the relevant period.

4.3 Theory and identification

4.3.1 Theoretical considerations

In a family consisting of a mother, a father and a child, an important question of time use arises. As none of the parents can care for their child at the same time when working, one of the parents, empirically in most cases the mother, has to make a choice between caring for the child or working and making use of some form of childcare by a third person. In deciding on the sort of care within her general leisure consumption choice, the price of public day-care institutions plays an important role because the cost of day care increase the opportunity cost of every working hour.

If a child becomes eligible for a cash-for-care subsidy that pays parents for not using a publicly subsidised day-care facility, the subsidy causes an income and a substitution effect. It increases the mother's income so that she can reduce her own earning efforts. At the same time, the subsidy distorts the relative price between leisure and consumption. As working becomes relatively more expensive, the substitution effect reduces labour supply and demand for institutionalised day-care provision.[13] Even when mothers do not work but still make use of a nursery, as it is the case for many children in Germany (see Bick (2010)), the price effect of introducing a cash-for-care subsidy should reduce the demand for day care as it can be considered a normal good. The effect of introducing the cash-for-care subsidy can be expected to be the more important the lower is the

[13]See e.g. Schøne (2004) for a more extended theoretical framework.

parents regular wage. Hence, the reduction in demand for child care would be larger for parents with smaller incomes. In contrast, parents with a high income would maintain a rather high willingness to pay and thus for them the effect on demand for nurseries would be lower.

When thinking about the effect of a cash-for-care subsidy on low-income households, it is important to bear in mind that Germany is a welfare state. The subsidy increases the welfare payments people get for not working (it is not offset against other transfers) and thus the outside option for employees with low income rises. As a consequence, low-skilled parents may be less inclined to look for a job, others may even decide to stop working. Hence, the cash-for-care subsidy contributes to the already existing poverty trap.

Besides the change in the extensive childcare decision, the introduction of the cash-for-care subsidy could have an impact on the intensive childcare decision. Here, the effects seem theoretically ambiguous. If the subsidy was paid in a rigid design meaning that parents could only decide between 150€ and a full abandonment of publicly subsidised day care or no subsidy payments and any desired day-care use, a polarisation in day-care hours would be a plausible result. Parents who would make use of day care for few hours only could be expected to rather take their child out of a nursery and either work less or pay for private childcare. In contrast, those in need for long day-care hours might extend their use. The observable pattern would be an increase in reported day-care hours per child enroled in a nursery. The specific design of the subsidy in Thuringia, though, allows for a partial use of the subsidy as well as of day-care services. Hence, diverse effects on the intensive and extensive margin can be expected to superimpose each other. Consequently, the effect of the subsidy on childcare intensity seems to be ambiguous and it is not obvious how to interpret a possible effect on observed day-care hours.

An alternative to publicly subsidised childcare other than to care for the child oneself is to make use of private paid or unpaid childcare. Given the good quality of publicly subsidised childcare and the high subsidies paid by the state government and the municipalities, privately paid childcare can be expected to be of negligible relevance in Thuringia. Thus, the realistic alternative seems to be informal childcare. I try to control for it by variables capturing household size the argument being that larger households offer more possibilities for informal care.

4.3.2 Identification strategy

The cash-for-care subsidy is specifically paid to parents of two-year old children in one state of Germany only, Thuringia. Thus, the identification strategy for the causal effect of the introduction of the cash-for-care subsidy onto the enrolment share can make use of the very limited group of eligible parents. There are two natural control groups to

which the treatment group of two-year olds in Thuringia can be compared. Firstly, it is possible to compare the situation in Thuringia with that in other Eastern German states. Secondly, a comparison across age groups can serve to identify the effect. As the treatment is limited to a one-year cohort, the groups of parents of one- or three-year olds can serve as comparable control groups.

The comparisons with a different German state as well as with the control group of younger or older children can each be exploited in a difference-in-difference (DD) approach. For the comparison between states, the estimator looks as follows:

$$DD_{\text{state}} = (E_{cfc} - E_{06})^T - (E_{cfc} - E_{06})^S \tag{4.1}$$

where E is the enrolment share, i.e. the share of children of a specific age that make use of a publicly subsidised day-care facility in a county. Here, only the group of two-year olds is considered. The first difference $(E_{cfc} - E_{06})^T$ indicates the change in the enrolment between the post-introduction period 2007-2009 and the year 2006 in Thuringia whereas the second term $(E_{cfc} - E_{06})^S$ shows the same difference for a different state as e.g. Saxony. Thus, the estimator would be zero if the group of two-year olds was affected by a similar shock in both states implying that the particular legislation in Thuringia that does not have consequences for children in a different German state cannot explain the effect. In contrast, if there is a negative impact of the cash-for-care subsidy for enrolment in Thuringia, DD_{state} will have a negative sign.

In the comparison across age groups in Thuringia, the estimator would be

$$DD_{\text{age}} = (E^2_{cfc} - E^2_{06}) - (E^3_{cfc} - E^3_{06}), \tag{4.2}$$

The first difference $(E^2_{cfc} - E^2_{06})$ measures the change in the enrolment share of two-year olds between the periods where the cash-for-care subsidy was in place and the ex-ante period 2006. The second difference $(E^3_{cfc} - E^3_{06})$ looks at the same time structure for the three-year old cohort. If there was a common effect on the childcare choice in Thuringia, it would show up in both differences and the estimator would not be significantly different from zero. Such a common shock could e.g. be a report on severe grievances in publicly subsidised day care in Thuringia that lets parents look for alternatives or a large increase in fees that may be very heterogeneous and thus difficult to control for. If the effect of the cash-for-care subsidy is such that only the enrolment share of the two-year olds decreases, DD_{age} will have a negative sign.

When one extends the identification approach to a difference-in-difference-in-difference (DDD) estimation, both of these DD designs are combined. Hence, the estimation accounts at the same time for common shocks to children of different age cohorts as well as to children in different states. The estimator is

$$DDD = \underbrace{[(E_{cfc} - E_{06})^T - (E_{cfc} - E_{06})^S]}_{\text{treatment group: 2 year olds}} - \underbrace{[(E_{cfc} - E_{06})^T - (E_{cfc} - E_{06})^S]}_{\text{control group: 3 year olds}} \tag{4.3}$$

If the cash-for-care subsidy reduced the enrolment share of two-year olds in Thuringia beyond common effects on two-year olds and beyond possible common shocks to the enrolment share across age cohorts, the sign of equation (4.3) would be significantly negative.

Still, a concern for the identification strategy could be that a shock occurred between March 2006 and March 2007 that has affected the specific group of two-year olds in Thuringia. Such a shock could be a reduction in the supply of day care for two-year olds, e.g. by a change in the standards to those for one-year olds that are more strict, but I have no hint that any such shock occurred.

The empirical literature has identified several effects that have an impact on the child care decision. As Coneus et al. (2009) show for West Germany, parents are more likely to enrol their child in a nursery if they are better educated and if the mother works[14]. In addition, children are more likely to attend non-parental child care if their parents have German nationality and if they live in a city, the latter effect may be correlated with availability, though. As Wrohlich (2004) shows child care costs also play an important role. Hence, I also take these controls into account.

For my analysis, no individual data are available that would allow for an analysis in the strain of the above-mentioned literature. Instead, I focus on the county level in Thuringia. Consequently, the child care decision is not a zero-one individual choice any more. Rather it reflects itself in the enrolment share of the specific county. Therefore, I make use of average county characteristics to explain the child care decision.

Of course, parents of young children are a different group than, for instance, pensioners and older parents that also contribute to the average county characteristics. However, if parents do not differ from the average of a county in very different ways in the various counties, the average of the county also captures those characteristics that distinguish parents of young children between the counties. As a consequence, looking at the characteristics of a county allows to identify different factors of influence on the child care choice.

For the factors determining the child care choice on the county level I revert to the literature on the determinants on the individual level. Hence, the following factors of influence are of interest: labour supply, education and cost. As labour supply, in particular of women, plays an important role for the individual child care decision, I would expect this also to hold on the county level, having a higher enrolment share of children in nurseries in the counties with higher female labour force participation. On the individual level the problem arises that female labor supply and the child care choice are highly interdependent and assuming exogeneity of either one might bias the estimation result. However, as Coneus et al. (2009) argue, the fact that in Germany a

[14]They instrument the mother's labour supply decision with the labour supply two years prior to the child's birth to overcome the simultaneity problem.

lot of children are enroled in institutional day care although their mothers do not work underestimates the effect of female labour supply on the childcare decision. Thus, the endogeneity bias rather works against significant outcomes.

In addition, education can be expected to be of high relevance. It seems to be reasonable to expect that the enrolment share is higher if more parents have a good education, in particular a college or university degree. A possible channel for education to impact enrolment of children could be that parents with a good education often have a high preference for a good education for their children. Also cost and household income can be accounted for in the county-level approach. Given that household incomes vary across counties, the cash-for-care subsidy (which is the same for all counties) could have a varied impact on the day-care choice as the reservation wages are distinct.

There are additional factors of influence that should be accounted for. Coneus et al. (2009) find that in Germany the probability of sending a child to day care is lower if the child has at least one non-German parent. As potential reasons they give cultural differences or larger families that offer more possibilities for informal care. I include the share of foreign nationals in the county population. A possible city effect is captured by a dummy for whether the county is a city or has a rural structure where evidence suggests that the enrolment is higher in the cities. County-specific preferences could be captured by election outcomes, the share of votes for the conservative CDU reflecting a more traditional view of the distribution of household chores and child care choices. Hence, in strongholds of the CDU, enrolment of children in day care is supposedly lower.

Besides, measures of wealth or poverty could have an influence on the childcare decision, for the former it would be positive, for the latter rather negative. As different empirical works show, the responsiveness to changes in childcare cost declines in income, see e.g. Anderson and Levine (1999), Kimmel (1995) and Haan and Wrohlich (2009). Controls reflecting wealth could be net tax revenues per head. As a measure of poverty one could use the share of households getting welfare or long-term unemployment benefits or the share of children living in a household which benefits from welfare payments (Bedarfsgemeinschaften) to all children in the county.

A feature that helps to identify the effect of the cash-for-care subsidy is the timing of its introduction. The law introducing the cash-for-care subsidy was passed in December of 2005 and one can reasonably assume this policy change and therefore the large cost increase for day care to be exogenous for the relevant age group of two-year olds. This argument is strengthened by the fact that the conception decision for a two-year old in March 2007 has to be taken in the summer of 2004 at the latest. By then the design of the law was not clear. In addition, the cash-for-care subsidy was introduced in a sharp design. This means that parents had no chance to receive transfers under the old state child-raising allowance scheme in July 2006 even if they were in the relevant income categories and their child younger than 2.5 (Thüringen (2005)).

4.4 Data and descriptive statistics

The state of Thuringia with a population of 2.3m is a small state in the centre of Germany (overall German population: 82m in 2008). It consists of 23 counties of which six are single cities. The size of the cities varies between 40,173 (Suhl in 2008) and the capital Erfurt with 203,333 inhabitants. The non-city counties are more similar in size with the smallest county having a population of 61,315 (Sonneberg in 2008), the largest one being Gotha with 140,041 inhabitants. I use county level data so there is in general a cross section of 23 data points for every year in Thuringia.

Parents decide on the use of a publicly subsidised nursery for their child. Although I do not observe the individual parents' characteristics that might influence the childcare decision, the child and youths statistics of Thuringia provide data that reflect aggregated choices on the county level. A good measure for the use of nurseries is the enrolment share which I define as the share of children of a specific age who are enroled in a publicly subsidised day-care facility (Tageseinrichtungen, Tagespflege) to the overall cohort size of the age group in a specified region.

Looking at the different age groups between birth and the age of six in Thuringia for the years 2006 to 2009, the group of two-year olds stands out as it is the only age group with a large reduction in the enrolment share, see *Figure 4.2*. Between March 15, 2006 and March 25, 2007 and thus in the period where the cash-for-care reform that affects the two-year cohort has come into force on July 1, 2006, the enrolment share has decreased by 6.1 percentage points from 79.5% to 73.4%. In contrast, the enrolment share for the three to five year olds is between 93% and 98% showing some variation over the years but no clear trend. For the one year olds, one observes a trend to increasingly higher enrolment.

As the illustration of the enrolment shares by county over the years 2006 to 2009 shows, see *Figure 4.3*, the enrolment share has decreased in 20 out of the 23 counties of Thuringia in 2007 staying constant in one and increasing in the two remaining counties.[15] In the latter two counties, however, there is a decrease between 2007 and 2008. Hence, the reduction between 2006 and 2007 is not a phenomenon of a specific region of Thuringia, it is visible in every single county. Besides, the Figure shows that there is quite some variation in the enrolment share between the different counties and I make use of the variation in my estimations.

It is important to note that constraints on the supply of day-care services are not an issue in Thuringia. In the majority of counties more than 10% of the approved childcare slots are not used. In the few cities where the share of free slots is below 10% it is below 5% only in Erfurt in 2006 and in Jena in 2009. For details see *Figure 4.8* in the Appendix.

[15] A map of Thuringia indicating the geographical location of the different counties can be found in *Figure 4.5* in the Appendix.

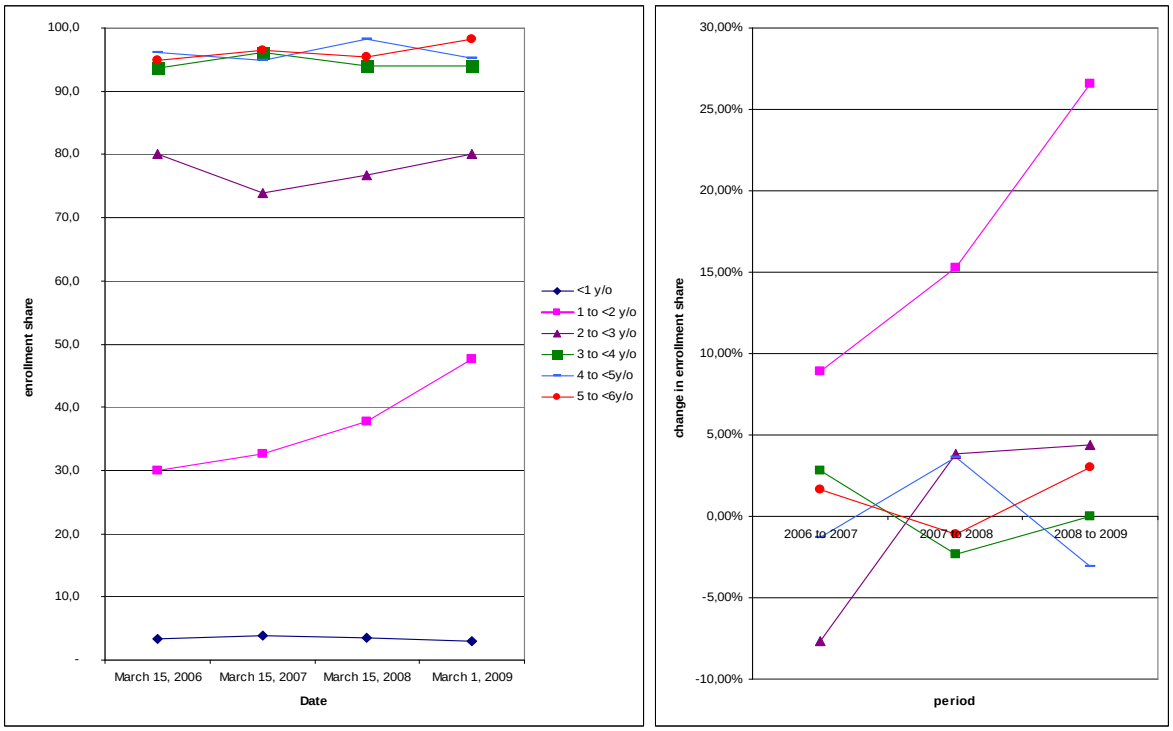


Figure 4.2: Share of children enroled in day-care facility to overall cohort size, different age groups

Besides having an effect on the enrolment share, the introduction of the cash-for-care subsidy could also have an impact on the intensive childcare choice decision of parents, i.e. parents could decide to make use of the day-care services for more or less hours per day. However, if one calculates the average time a two-year old child spends in nursery by weighing the different reported time categories with the number of children in each, there does not appear to be an effect of the introduction of the cash-for-care subsidy in the descriptive statistics. Over the whole period from 2006 to 2009 there seems to be a trend towards longer day-care hours over all counties. For details see *Figure 4.4* in the Appendix.

The share of two-year olds with one parent who has a foreign nationality in the total number of two-year olds in nurseries is in general very small. For the whole of Thuringia it is between 3.5% and 4.2% on average in 2006 and 2009 respectively. There is an important difference between the counties though. In the cities the share is higher with Erfurt and Jena having the highest share (here it is between 7.8% and 9% over the four year period). In contrast, in most rural counties, the share is around 2%.

As the data on children is collected on a yearly basis and not available before 2006, I only have one ex-ante data point in each county and three ex-post points in the time dimension 2006 to 2009. Therefore, my panel data set consists of 23 counties in Thuringia in general over four periods, 2006-2009, thus I have 92 data points. *Table 4.9*

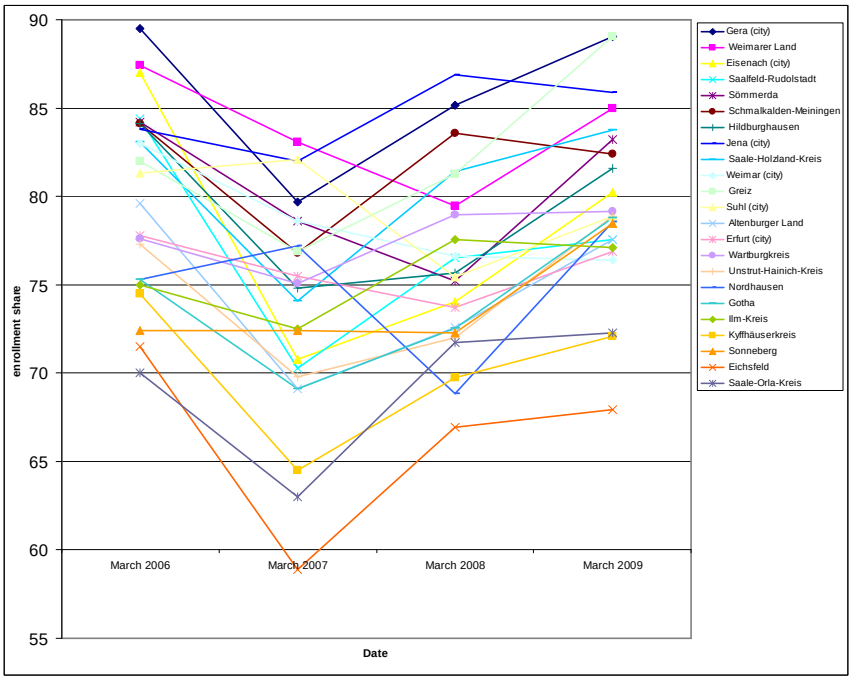


Figure 4.3: Enrolment share of two-year old children by county in Thuringia (2006-2009)

in the Appendix gives the summary statistics of all variables used.

The data concerning children in nurseries originate from the children and youth statistics of the statistical office of Thuringia. Since 2006, they are collected every year in March. They cover the whole population and provide quite detailed information on children in all publicly subsidised day-care facilities.[16] In particular, the following information are collected: For every day-care facility[17], the exact location and provider are reported as well as some characteristics and the number of approved slots and attending children. Concerning the children, every nursery has to report for every single child the month and year of birth, gender, one of four categories indicating the time spent in the nursery as well as whether German is the main language in the family and whether at least one parent has a non-German nationality. In addition, information on the personnel such as gender, age, qualification and working hours are reported. The enrolment share of the two-year olds is collected annually in March.

The main explanatory variable is the cash-for-care subsidy. It takes the value of 150€

[16]Before 2006, only data on the personnel of day-care centres was collected, including qualification, working hours, number of facilities and the like.

[17]There may, of course, be non-publicly subsidised day-care centres as well. However, given the extensive use of the publicly subsidised facilities with an enrolment share of 80% for two-year olds in 2006, it is highly unlikely that there is a relevant private day-care sector in Thuringia. In addition, there are no supply constraints forcing parents to look for alternatives and the price differences between publicly subsidised and private facilities can be expected to be large given the overall subsidies for publicly subsidised day-care centres.

for all years after its introduction, i.e. for 2007 to 2009 and 0 in 2006.

All data on Thuringia are provided by the statistical office of Thuringia and stem from three basic data sources, the labour market data, the national accounts and the Mikrozensus. The labour market data are collected by the regional branch of the Federal Employment Agency (Bundesagentur für Arbeit) from employers. In the labour market statistics, one can find a lot of data on the county level, including monthly unemployment data by gender. Even more detailed, the labour market data are available on a quarterly basis and, to be close to the childcare data, I take the data from the end of March for each year 2006-2009.

The baseline for the labour market statistics is the total number of employees (sozialversicherungspflichtig Beschäftigte). There are many employees who work in a different county than they live in. As it seems to be very likely that the large majority of parents enrol their child in a nursery in the county they live in, I always use the employee data that are sorted based on the county of residence of the employee.

I define female labour supply as the share of female employees to total working-age women, i.e. all women between ages 15 and 65. To take account of part-time employment, I calculate the share of women working part-time in all female employees.[18] The level of education is registered in different categories. I use the lowest category and combine the highest two. Hence, I include the share of employees without any finished vocational training to the total number of employees and the share of employees with a degree from a college of higher education or a university in my regressions.

In addition, the labour market statistics provide information about the recipients of welfare payments. Based on these data, I use the share of the number of employable long-term unemployed who receive welfare payments (Arbeitslosengeld II) to the civil labour force. In a broader measure of the recipients of welfare payments, I add beneficiaries of the so called 'Sozialgeld', paid to those who are not able to work, to the number of recipients of Arbeitslosengeld II (*welfare recipients*). An alternative measure of poverty that focuses on *poor children* is the share of families with children which get welfare payments (Bedarfsgemeinschaften mit Kindern) to the total number of families[19] in a county. The national accounts data procure the disposable income per head which is calculated on an annual basis. These data are only available for the years 2005 until 2007 on the county level.

The Mikrozensus is a representative national survey in which 1% of the German population takes part every year. Individuals are legally obliged to participate, thus there is no participation bias. Data are collected in personal interviews and the collection process is spread equally over all weeks of the year. The Mikrozensus provides a lot of information about employees, family structures, etc. Also, it contains for each county

[18] I use shares for all these variables to make the data comparable over the counties.

[19] See below for the definition of a family.

the median net household income and the number of residents of different age groups. As an alternative to disposable income, I use the median net household income in a county as these data are available for four years (2005 to 2008).

In regressions using such data, I recur on the previous year, i.e. take the 2005 number in the regression on the 2006 enrolment share because the child care statistics are collected in March. If income is to influence this decision, it seems to be reasonable to base the child care choice on realised income rather than on expected income as it would be the case when taking the same year data.

Moreover, the Mikrozensus registers the number of families in a county where a family is defined as a household where at least one non-married child lives with one or both parents. In the regressions, I use the variable *big households* that is composed as the share of households with three or more members to the total number of households[20] to account for family structures and the possibility of informal childcare.

In addition, the statistical office provides data on the election outcomes in the different counties. Thus, I can include the votes for the CDU specified as the average deviation of the share of votes for the conservative party (CDU) in a specific county from the average of Thuringia of the two state elections 2004 and 2009 and the federal elections 2005 and 2009. $City$ is a dummy for city as opposed to non-city county, $BorderWest$ a dummy taking the value of 1 for every county along the border to a Western German state. Besides, I use some variables that reflect differences in fertility between the counties. One measure is the share of under 6 year-old children to the number of women aged 18 to 40, another is the share of children below the age of 18 in the total population of a county. An additional population variable is the share of foreign nationals in the total population. Last, there are the communal excise tax revenues aggregated on the county level. Together the different control variables are supposed to capture the heterogeneity of the counties in particular concerning educational, social, economic and demographic factors.

4.5 Results

To get a first impression of the effect of the cash-for-care subsidy, I regress the enrolment share of the two-year olds on the cash-for-care subsidy and different controls such as female labour supply and education. I also incorporate additional county characteristics. In columns 1 and 2, *Table 4.3* presents the results of the two pooled OLS regressions where standard errors are clustered on the county level to adjust for the serial correlation of the county characteristics over time. The regressions also include time dummies to account for specific year effects and a linear Trend. As these regressions use current

[20] A household is defined in the Mikrozensus as a group of persons who live together and form an economical unit as well as persons who live on their own.

accounts data, the number of observations is 69, i.e. only the years 2006 to 2008 are covered.

Both columns show that the cash-for-care subsidy has a highly significant negative effect on the enrolment share. Although the quantitative effect has to be interpreted with caution in this rather crude first regression, the introduction of the cash-for-care subsidy results in a reduction of the enrolment share by 9.4 percentage points and thus a higher reduction than the descriptive statistics would indicate. In addition, the following variables have a significant impact. The higher is the share of college/university graduates and female labour supply[21] in a county, the higher is the enrolment share. In contrast, the share of employees without a vocational training and the share of part-time working females have a negative impact on the enrolment share. Hence, these controls have the expected sign. In the second-column regression, controls such as the proximity to the border to Western Germany or the share of foreign nationals are significant, whereas other controls such as the share of votes for the CDU, the excise tax revenues, the share of welfare recipients or the share of children do not have a significant impact.

The OLS estimates may be biased due to omitted variables that systematically differ between counties and might have an impact on the enrolment share. The different day-care fee levels and structures could, for example, be such a variable that I cannot control for directly because of missing data. To capture county-specific effects that are difficult to control for by covariates, in the following, I introduce county fixed effects and make use of the panel structure of the data thus accounting for county-specific effects. As the time-dimension is important for identification in fixed-effects panel regressions, I refrain from using the national accounts data and use the net median household income instead of the disposable income because then I dispose of a panel with 92 observations (instead of 69). All regressions are run with robust standard errors to account for the fact that the controls within a county are not independent over time.

Columns 3 to 5 in *Table 4.3* present the results of the fixed-effects panel regressions.[22] The third column presents the results of a minimal fixed-effects regression of the cash-for-care subsidy, a Trend and time dummies on the enrolment share. The highly significant, negative impact of the cash-for-care subsidy on the enrolment share is quantitatively very similar to the effect in the OLS regressions. The *Trend* variable, also highly significant, reflects that there seems to be a tendency towards higher enrolment for two-year olds in Thuringia. Based on using the net median household income to account for income differences between counties, in columns 4 and 5 I add more controls to the regression. The effect of the cash-for-care subsidy onto the enrolment share as the dependent variable in all regressions is always highly statistically significant and negative. The quantitative

[21] Despite the previously discussed endogeneity bias in female labour supply, it turns out to be significant.

[22] *rho* indicates the variance not explained by differences across entities. R_i^2 with $i = o, b, w$ is the overall, between county and within county variation explained by the regression.

Table 4.3: Clustered OLS and fixed-effects panel regressions, two-year olds in Thuringia

	(1) enrolsha OLS	(2) enrolsha OLS	(3) enrolsha FE panel	(4) enrolsha FE panel	(5) enrolsha FE panel
Cash-for-care	-9.3556***	-9.4109***	-9.3913***	-8.7170***	-8.6516***
	(1.983)	(1.918)	(1.281)	(1.372)	(1.697)
Disposable income	-0.0005	-0.0006			
	(0.001)	(0.001)			
Median income				0.0039	0.0033
				(0.004)	(0.004)
High education	41.8859**	103.2416**		327.1830***	294.7657**
	(15.728)	(42.679)		(112.866)	(132.432)
No training	-247.3391***	-196.9449**			-285.1250*
	(79.193)	(81.178)			(162.444)
Female labour	81.1008***	86.5631*		106.8771	182.1993***
	(19.343)	(44.434)		(64.854)	(60.891)
Female part-time	-48.8968	-42.2756**			
	(30.463)	(19.554)			
Fertil under18	-124.9382	-139.0918			
	(113.788)	(89.673)			
Votes CDU´	0.0404				
	(0.205)				
Excise tax	-0.0125				
	(0.011)				
Border West	1.2097	3.2341*			
	(2.045)	(1.735)			
Foreigners		-2.8843*		-1.8541	
		(1.466)		(2.897)	
Welfare recip.		17.2752			76.7712*
		(32.657)			(40.523)
Trend	0.8527	1.0902	3.0217***	1.0011	
	(1.248)	(1.198)	(0.413)	(0.914)	
Constant	101.1106***	86.4011*	77.5174***	-6.0786	-33.9846
	(28.201)	(44.930)	(0.668)	(32.739)	(30.634)
Time dummies	no	no	yes	yes	yes
N	69	69	92	92	92
F	20.319	29.635	28.567	16.804	16.882
R^2	.6361158	.6679346	.5335155	.5654748	.5984661
R^2_o			.1781604	.3505638	.3835561
R^2_b			3.29e-30	.3982368	.4514695
R^2_w			.5335155	.5654748	.5984661
rho			.762291	.934652	.9543304

Clustered standard errors (columns 1 and 2) in parentheses
Robust standard errors (columns 3-5) in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

impact varies a bit, but is in the same range as in all previous regressions and important. The signs of the control variables are as expected and do not change as compared to the OLS regressions.

Quantitatively, the interpretation is as follows. If one increases the share of college or university graduates by one percentage point, the enrolment share increases by 2.9 to 3.3 percentage points depending on the column considered. This effects seems to be very high. A reason for caution in the interpretation of the quantitative impact of the variables is that there is limited time variation in a single variable in a county. As a consequence, fixed effects regressions that rely on changes over time for identification are not very precise in estimating the impact.

Difference-in-difference over states

So far, I have focused on effects that can be controlled for by including the differences between the counties. A concern, though, could be the existence of more general shocks that are common to a broad region and that have had an impact on the enrolment share of two-year olds between 2006 and 2007. In case of e.g., a large labour market shock, a reduction in the enrolment share might have occurred in other German states as well. However, as I have only concentrated on the state of Thuringia so far, I would wrongly attribute an effect of a state law to the observed pattern, although the effect is not particular to the state. To shed light on this concern, I use the German state of Saxony as a control group.

There are several reasons for the choice of Saxony as opposed to other German states for a comparison. Saxony is a neighbouring state of Thuringia and thus the two states are geographically close to each other. Both states are located in East Germany and have a similar tradition of institutionalised early childhood care which is not the case in any one of the three Western German states that share a common border with Thuringia. An important implication of this tradition is a large availability of childcare services. As the lower part of *Table 4.8* in the Appendix indicates, one also finds an excess supply of approved slots in each county in Saxony.

In addition, the states of Thuringia and Saxony are the economically most advanced in East Germany. Hence, in many characteristics the two states look rather similar as can be seen from *Tables 4.9* and *4.10* in the Appendix that provide the summary statistics for both states. The largest differences between Thuringia and Saxony can be found in the voting behaviour, in the revenues from the communal excise tax and in the size of the counties as indicated by *population*. After a reform of the counties, the 13 counties in Saxony are much larger.[23] Whereas the characteristics of the counties look quite alike, there seems to be a level difference in the enrolment share of two-year olds.

To implement the difference-in-difference approach that compares the group of two-

[23]The data on Saxony are provided by the Statistical Office of Saxony. The principal data sources correspond to those in Thuringia.

year olds with the analogous cohort in Saxony as described in equation (4.1), I use the following set-up:

$$\begin{aligned} E_{ist} &= \alpha_1 + \alpha_2 Z_{ist} + \alpha_3 CFC_{it} + \alpha_4 TH_{is} + \alpha_5 (CFC_{it} \times TH_{is}) \\ &\quad + \alpha_6 Trend + \alpha_7 (Trend \times TH_{is}) + \epsilon_{ist} \end{aligned} \tag{4.4}$$

where i indexes individuals, t indexes time, being 1 after 2006 and 0 before the introduction of the subsidy and s indexes the state (1 for Thuringia and 0 for Saxony). E is the enrolment share of two- or three-year olds respectively. Z is a vector with different control variables that might affect the enrolment share. CFC is a dummy variable that captures whether the ex-ante period 2006 or the periods after the introduction of the cash-for-care subsidy are considered. Its value is 0 in 2006 and 1 for the period 2007 to 2009. TH is a dummy variable that takes on a value of one for the state of Thuringia and 0 else. The effect of interest, i.e. the impact of the introduction of the cash-for-care subsidy onto the enrolment share in Thuringia would show up in α_5. α_3 captures whether there is a common shock to the enrolment of two-year olds between March 2006 and March 2007. The coefficients α_6 and α_7 are important to clarify the interpretation of the Trend as they potentially allow to distinguish between a common trend for the group of two-year olds and a phenomenon that is specific to Thuringia.

Table 4.4 shows the outcome of a fixed effects panel regression that implements equation (4.4).[24] There are 144 observations covering the 36 overall counties, 23 in Thuringia and 13 in Saxony over the four-year period. The cash-for-care subsidy in Thuringia shows a highly significant negative impact on the enrolment share of in-state two-year olds. Furthermore, a common positive trend for the age group seems to exist that is not a particular phenomenon of Thuringia (see columns 1 and 2). As the DD estimation shows, the negative impact of the cash-for-care subsidy on the two-year olds in Thuringia can be confirmed. It is clearly distinguishable from general changes in the enrolment share of two-year olds across states and quantitatively important.

Difference-in-difference-in-difference estimation

Up to now, I have only considered the age group of the two-year olds in my regressions. Although the cash-for-care subsidy is only paid to parents of these children in Thuringia, the law that introduced it has also changed the financing for publicly subsidised nurseries focusing the subsidies on the slots occupied rather than on the number of slots approved in a facility. In addition, the law makes it easier for parents to choose the nursery for their child independent of the particular town or village they live in. Both of these changes might have had a common effect on all children in Thuringia. If this was the case, the regressions presented so far would wrongly attribute the effect that occurred between March 2006 and March 2007 in Thuringia to the cash-for-care subsidy because only the group of two-year olds has been considered. As a consequence, I cannot distinguish

[24] As I run a fixed effects regression, α_4 does not appear in *Table 4.4* as all effects that are common to Thuringia are part of the county fixed effects.

Table 4.4: Diff-in-diff estimation for two-year olds in Thuringia and Saxony (FE)

	(1)	(2)	(3)
	enrolsha	enrolsha	enrolsha
Cfc*Thuringia	-7.1713***	-7.7617***	-8.4507***
	(0.947)	(1.449)	(1.486)
Cash-for-care	-2.2039**	-1.8267**	-0.7748
	(0.821)	(0.853)	(1.074)
Trend	2.9151***	2.7265***	-0.4719
	(0.280)	(0.278)	(1.182)
Trend*Thuringia		0.2952	0.5997
		(0.494)	(0.561)
Median income			0.0025
			(0.004)
Female labour			120.4490**
			(57.524)
No training			-202.5692
			(132.142)
High education			281.1495**
			(115.325)
Constant	71.3414***	71.3414***	3.2486
	(0.456)	(0.457)	(35.016)
Time dummies	yes	yes	yes
N	144	144	144
F	56.095	56.270	32.559
R_o^2	.0715402	.0611461	.0053771
R_b^2	.5510947	.5510947	.0003745
R_w^2	.5881756	.5890656	.6243041
rho	.9398835	.9380536	.9783475

Robust standard errors in parentheses

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

a general effect from the childcare reform or some other shock that has influenced the parents' childcare decision in Thuringia in 2006 from the particular cash-for-care subsidy.

To overcome this concern, I add the group of three-year olds to the sample so that I regress the county characteristics of Thuringia and Saxony on the enrolment share of two- and three-year olds respectively. In effect, I implement a difference-in-difference-in-difference (DDD) design that compares the two cohorts across the two states. The DDD design can therefore control for general schocks by comparing cohorts across states as well as for state specific shocks that go beyond the specific cohort of two-year olds by comparing different cohorts.

The group of three-year olds is chosen for several reasons. Firstly, the group of three-year olds is close to the two-year olds because of the similar age. In contrast to the one-year olds, rationing is not problematic as there is a legal claim to childcare for three-year olds as well, whereas parents of one-year olds have to justify their wish for a day-care slot and may not be granted one. Secondly, in Thuringia there is an institutional difference between day care for children younger than two and those with two years of age and older. For instance, the nursery teacher child ratios are higher in the facilities for one-year old children (Kindertagespflege) and usually fees are higher as well. Thirdly, the introduction of the new parental pay scheme on January 1, 2007 may have an effect on the enrolment of one-year olds in nurseries in the period considered. Hence, the group of three-year olds seems to be the best comparison group.

The implementation of the DDD approach as presented in equation (4.3) is carried out by

$$\begin{aligned} E_{iast} &= \alpha_1 + \alpha_2 Z_{iast} + \alpha_3 CFC_{iat} + \alpha_4 2YO_{ia} + \alpha_5 TH_{is} \\ &\quad + \alpha_6(CFC_{iat} \times TH_{is}) + \alpha_7(CFC_{iat} \times 2YO_{ia}) + \alpha_8(TH_{is} \times 2YO_{ia}) \\ &\quad + \alpha_9(CFC_{iat} \times 2YO_{ia} \times TH_{is}) \\ &\quad + \alpha_{10} Trend + \alpha_{11}(Trend \times 2YO_{ia}) + \alpha_{12}(Trend \times TH_{is}) \\ &\quad + \alpha_{13}(Trend \times TH_{is} \times 2YO_{ia}) + \epsilon_{iast} \end{aligned} \tag{4.5}$$

where a indexes the age or treatment group (1 for two-year olds and 0 for three-year olds). The names of the variables are the same as in equation 4.1, new is $2YO$ which is a dummy variable for the age group and distinguishes the treated two-year olds from the control group of three-year olds being of value 1 if the child is two years of age.

Here, α_9 captures the essential effect of the cash-for-care subsidy onto the enrolment share of two-year olds in Thuringia, α_3 captures whether there is a common shock to the child care choice between March 2006 and 2007, α_4 contains possible systematic differences between the two age groups and α_5 accounts for systematic differences between the states as e.g. the level difference in the enrolment share, see Tables *4.9* and *4.10*. A general trend in the enrolment shares would be captured by α_{10}, whereas a possible common trend for two-year olds is caught by α_{11}.

In using second-order interactions, it is important to also include all first-order interactions into the regression. This assures that the second-order interactions do not capture the effect of the omitted first-order interaction which would result in biases. *Table 4.5* summarises the results of the DDD estimation.

As I regress the same county characteristics on the enrolment share of the two- and three-year olds, the number of observations doubles to 288. I use an OLS regression with clustered standard errors where clustering takes place on the county level to adjust for the autocorrelation of covariables within a county over time. Time and county fixed effects are introduced to capture the particularities of the different counties and avoid omitted variable biases.

In all three regressions presented in *Table 4.5*, the introduction of the cash-for-care subsidy has a significantly negative impact on the enrolment share of the two-year olds thus confirming the qualitative result obtained in previous regressions. The negative impact of the cash-for-care subsidy can be explained neither by a general effect on the enrolment share in Thuringia nor by a possible common negative shock on the enrolment share in general. On the contrary, an additional insight of the DDD design is that there seems to be a positive shock on enrolment in 2006 which is indicated by the positive sign of the *Cash-for-care* variable itself.[25] The DDD estimation strongly confirms the existence of an underlying trend in the enrolment share of two-year olds. Because of the short time dimension of the data, it is essential to identify the trend across states. Controlling for the systematic differences between Thuringia and Saxony as well as for the two-year cohort in both states is important because of level differences.

Taking all the evidence provided by the different regressions into account, the results hint towards a strong negative effect of the cash-for-care subsidy in Thuringia and a general positive trend in the enrolment rate of two-year olds that is common to Saxony and Thuringia. Given the inclusion of the general trend for two-year olds, the quantitative effect of the introduction of the cash-for-care subsidy onto the enrolment share of two-year olds in Thuringia can be calculated by adding the significant coefficients of all variables that contain 'CFC', i.e. α_3, α_7 and α_9[26].[27] The reduction in the enrolment share is larger than 9 percentage points (change of -9.63 for the column 2 regression, -9.12 for the column 3 regression).

[25]The occurrence of the positive effect is driven by an increase in the enrolment share of the three-year olds. When I regress the enrolment share for the three-year olds onto the post-2006 variable, a Trend and time dummies in a fixed effect panel regression, there is a significantly positive effect between 2006 and 2007.

[26]α_6 is not significantly different from zero and thus excluded.

[27]Note that $Thuringia$ and $2 - year - old$ are dummy variables.

Table 4.5: DDD estimation, OLS with county-clustered standard errors

	(1)	(2)	(3)
	enrolsha	enrolsha	enrolsha
Cfc*2-year-old*Thuringia	-5.7483***	-7.8904***	-7.8904***
	(1.329)	(2.250)	(2.268)
Cash-for-care	1.2324***	2.8146***	3.3230***
	(0.441)	(0.888)	(0.959)
2-year-old	-26.8864***	-29.8612***	-29.8612***
	(1.498)	(1.506)	(1.518)
Thuringia	13.3097***	8.2872***	15.3042**
	(1.130)	(1.180)	(6.544)
Cfc*Thuringia	-1.4230*	0.1287	-0.0674
	(0.823)	(1.376)	(1.424)
Cfc*2-year-old	1.3965*	-4.5531***	-4.5531***
	(0.720)	(1.334)	(1.345)
Thuringia*2-year-old	13.4203***	12.3492***	12.3492***
	(2.091)	(2.066)	(2.083)
Trend		-0.2483	-0.9458
		(0.369)	(1.153)
Trend*2-year-old		2.9748***	2.9748***
		(0.432)	(0.436)
Trend*Thuringia		-0.7758	-0.7051
		(0.521)	(0.577)
Trend*2-year-old*Thuringia		1.0710	1.0710
		(0.651)	(0.657)
Median income			0.0012
			(0.003)
High education			131.4969
			(116.774)
No training			-44.1153
			(102.892)
Female labour			19.6072
			(50.553)
Constant	87.7404***	93.7870***	65.1530*
	(0.704)	(0.725)	(33.362)
County dummies	yes	yes	yes
Time dummies	yes	yes	yes
N	288	288	288
F	.	.	.
R^2	.9201917	.9323135	.9326174
R^2	.9201917	.9323135	.9326174

Clustered standard errors in parentheses

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

4.6 Robustness

A possible concern is that the counties in Thuringia could be very different from each other in the ex-ante situation. If counties with a high enrolment share of two-year olds systematically vary in their characteristics from counties with a low enrolment share, changes attributed to the cash-for-care subsidy could rather reflect the differences that already existed before-hand.

Table 4.12 in the Appendix shows the summary statistics for the 12 counties of Thuringia that have an above-average enrolment share in 2006 whereas *Table 4.13* summarises the county characteristics of the 11 counties that have ex-ante a below-average enrolment share. As can be seen, the two groups of counties look very much alike in most characteristics. The biggest difference exists in the number of city counties most of which are in the category with the higher enrolment shares. The regressions account for this in the fixed effects. Also the voting behaviour differs between the groups of counties and in the counties with the higher ex-ante enrolment share, there is a larger share of college and university graduates.

It would be interesting to identify which characteristics have had an impact on the extent of the change due to the introduction of the cash-for-care subsidy. Therefore, I make use of the county heterogeneity to identify such characteristics. In accordance to the arguments outlined in section 4.3, I would expect certain characteristics to decrease the reduction in the enrolment share. These are a high income, high education and a large labour force participation as the variables capture higher opportunity cost of care at home and preferences for early childhood education.

In contrast, other characteristics can be expected to increase the reduction due to the subsidy. If there are more large households, informal care may be a more relevant alternative. Caring for a child informally might be more attractive if a family has several children or if one can share the care between neighbours. Thus, a higher fertility could increase the reduction. Measures of poverty could capture that the cash-for-care subsidy is more relevant for low household incomes. In addition, high election results for the CDU could reflect underlying preferences such as a more traditional view of the family. Then, parents would make more use of the subsidy. To implement the analysis, I introduce interaction terms into the fixed-effect panel regressions as found in columns 4 and 5 of *Table 4.3*. The results can be seen in *Table 4.6*[28]

As column 2 of *Table 4.6* indicates, the reduction in the enrolment share due to the cash-for-care reform is significantly lower if more highly educated employees live in the county. In contrast, the negative impact of the cash-for-care subsidy is higher if the conservative CDU scores above state average voting shares in the state and federal elections (column 3). Also, the impact is larger if there are many large households in the county (column

[28]The same results can be obtained, if the interaction terms are introduced in the across states difference-in-difference estimation.

Table 4.6: Fixed effects panel regressions with interaction terms, two-year olds Thuringia

	(1) enrolsha	(2) enrolsha	(3) enrolsha	(4) enrolsha	(5) enrolsha	(6) enrolsha
Cash-for-care	-8.3635	-12.0016***	-8.6697***	-15.0366***	-0.2462	17.9331
	(4.950)	(2.180)	(2.010)	(2.768)	(4.629)	(14.615)
Median income	0.0032	0.0033	0.0045	0.0058*	0.0046	0.0030
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)
High education	292.2498*	223.2453	346.8224**	269.0493*	326.9315***	300.1561**
	(142.466)	(143.359)	(126.063)	(155.561)	(109.572)	(137.516)
No training	-282.9357	-315.4503*	-288.9469	-379.8726**	-320.6832	-232.4833
	(167.301)	(172.715)	(205.437)	(165.341)	(196.998)	(204.318)
Female labour	180.3552**	116.4896*	101.6307	141.1423**	93.3025	91.9697
	(70.158)	(65.081)	(77.117)	(65.024)	(73.032)	(78.704)
Welfare recip.	77.3128**					
	(36.505)					
Cfc*Welfare recip.	-0.8368					
	(15.152)					
Cfc*High education		25.0251**		24.6295**		
		(10.229)		(10.822)		
Votes CDU			0.0000			
			(0.000)			
Cfc*Votes CDU			-0.2883***			
			(0.085)			
Cfc*City			-1.0588		-1.8041	-1.7760
			(2.250)		(2.622)	(2.821)
Poor children				29.2062		
				(27.413)		
Cfc*Poor children				18.3710		
				(18.768)		
Big household					16.8807	
					(16.725)	
Cfc*Big household					-28.0155**	
					(11.123)	
Fertil under6						-28.5538
						(63.287)
Cfc*Fertil under6						-75.9450*
						(39.312)
Constant	-33.2225	30.8366	20.7098	14.3117	24.8011	35.8630
	(35.813)	(40.952)	(39.822)	(40.719)	(39.739)	(37.897)
Time dummies	yes	yes	yes	yes	yes	yes
N	92	92	92	92	92	92
F	16.002	17.475	45.391	20.186	23.214	20.883
R_o^2	.3838274	.4085323	.3612926	.3897464	.3639953	.3597902
R_b^2	.4510238	.4706617	.4380911	.4621697	.4325674	.4236174
R_w^2	.5984876	.5910167	.5965598	.6085745	.61013	.6096953
rho	.9527844	.9338482	.968761	.9578664	.9637654	.9610025

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5) and if the share of children below the age of six is high (column 6). The result of the votes could reflect the underlying preferences of the population and more traditional family structures. The effect of the large households could suggest that single-parent households react less to the subsidy, although the effect has to be interpreted prudently.

I include an interaction with the city dummy in several of these regressions to control for systematic differences. As the descriptive statistics show, cities have ex-ante higher enrolment shares. Consequently, I want to avoid that the effects found in election results, household size and fertility are due to the differences between city and rural counties.[29] The significance of the effects does not depend on the inclusion of the city dummy interaction term, though. The poverty measure presented in column 1 does not have a significant impact on the size of the change in the enrolment share due to the introduction of the cash-for-care subsidy. I have also incorporated several other controls as interaction terms into these panel regressions. However, neither female labour supply nor any one of the other welfare or poverty covariates have a significant impact on the size of the change in the enrolment share.

Lastly, I have a look at the impact of the cash-for-care subsidy on the intensive childcare decision. Here, the dependent variable is the number of daily hours, two-year olds spend in publicly subsidised day care in Thuringia on a county level.[30] The hours are a weighted average over four categories of day-care time.[31] *Table 4.7* presents the result of the regressions. For column 1, I have run an OLS regression, for columns 2 to 4 difference-in-difference regressions that compare the two- and three-year olds in Thuringia. The OLS regression finds no effect of the cash-for-care subsidy, in the DD regressions a positive effect can be found in columns 2 and 4. The relevant variable is the interaction term between the two-year cohort and the cash-for-care subsidy as the children of age two are the treatment group. However, once I account for a possible unobserved trend, the introduction of the subsidy has no significant effect any more and this also holds if one introduced the trend variable into the column 2 regressions. Instead, there seems to be a positive trend in the day-care intensity of two-year olds. See *Figure 4.4* in the Appendix for a graphical presentation of the county data. On the basis of the data I have, no clear effect of the introduction of the cash-for-care subsidy onto the intensive childcare choice can be identified.

[29] As I use fixed effect regressions, all potential city effects are accounted for in the basic setup. Hence, I have not included the city dummy into the basic regression.

[30] I also have the number of hours for the three-year olds in Thuringia and use them in the DD estimation.

[31] The data provide the number of children in day care that attend for ≤ 5 hours, between five and seven hours, between seven and ten hours and for more than ten hours a day.

Table 4.7: OLS and DD estimation of hours in day care, two-year versus three-year cohorts, Thuringia

	(1) hours 2 y/o	(2) hours DD	(3) hours DD	(4) hours DD
Cfc*2-year-old		0.0922**	0.0304	0.0922**
		(0.033)	(0.049)	(0.033)
Cash-for-care	0.3994	0.1110***	-0.0427	0.1007
	(0.286)	(0.035)	(0.052)	(0.220)
Median income	-0.0003*		-0.0002	-0.0002
	(0.000)		(0.000)	(0.000)
High education	-4.7988		-0.1262	-0.1262
	(9.910)		(7.911)	(7.885)
No training	4.7374		0.5586	0.5586
	(7.850)		(5.973)	(5.954)
Female labour	-2.0072		0.7414	0.7414
	(5.127)		(4.132)	(4.119)
2-year-old		-0.2236***	-0.2545***	-0.2236***
		(0.036)	(0.039)	(0.036)
Trend			0.0426	
			(0.070)	
Trend*2-year-old			0.0309*	
			(0.015)	
Constant	10.5330***	8.4001***	8.4390***	8.4817***
	(3.488)	(0.023)	(2.302)	(2.344)
County dummies	yes	yes	yes	yes
Time dummies	yes	yes	yes	yes
N	92	184	184	184
F	.	.	.	.
R^2	.904236	.8642219	.8687839	.8672963

Clustered standard errors in parentheses

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

4.7 Conclusions

Many industrialised countries are confronted with the challenge of demographic change and ageing societies. As a consequence, family policy gains importance as it has the potential to alleviate the lack in qualified employees in at least two aspects. Firstly, a sufficient supply of day-care is expected to increase female labour force participation and reduce gender wage inequality. Secondly, good quality childcare is a solid foundation for equal chances at school and successful educational careers (see e.g. Temple and Reynolds (2007) and Currie (2001)). Thus, it contributes to future economic growth.

This paper analyses the impact of the introduction of a cash-for-care subsidy in the German state of Thuringia. Although specific to the single state, the analysis is of more general interest because the federal German government envisages to introduce a similar policy for the whole country in 2013 at the same time when every parent gets a legal claim to a child care slot for her under three year old. In addition, the specific environment of Thuringia is distinct from other regions for which a similar policy has been analysed in that there are no constraints on the supply side. As this component is an important factor influencing the outcome of a similar cash-for-care subsidy in Norway (see Schøne (2004)), the analysis contributes additional insights.

Parents of young children have the choice between either staying at home and caring for their child or working and making use of some sort of day-care provision. In Germany also quite some non-working mothers send their child to a nursery. The parents' decision is influenced by the price of childcare as it determines the opportunity cost of working or other activities. I make use of a large exogenous price increase for the childcare of two-year old children in publicly subsidised nurseries in Thuringia to analyse its consequences on the parents' childcare decision. The effect is identified in a DDD design that utilises that the reform only concerns a very particular age group in one German state. Hence, the age group of three-year olds as well as the comparable cohorts in the similar state Saxony are used as control groups.

The introduction of the cash-for-care subsidy has a quantitatively important negative impact on the enrolment share of children in the respective group of two-year olds in Thuringia reducing the enrolment share by more than 9 percentage points based on an initial level of 80%. Given a median net household income in Thuringia of $1,470$€, the amount of the subsidy of 150€ per two-year old and month is quantitatively important. Theory would suggest that in particular households with low incomes make use of the subsidy.

For my analysis I do not dispose of individual level data. However, the results obtained on the basis of county-level data seem to confirm the prediction. The reduction in the enrolment of two-year olds due to the introduction of the cash-for-care subsidy in Thuringia is less pronounced if many highly educated parents live in the county. In contrast, the reduction is larger if parents have a more traditional view of the family

and if there are more informal alternatives to publicly subsidised childcare. Besides, I identify a positive trend in the enrolment share of two-year old children across the states of Thuringia and Saxony.

The cash-for-care subsidy seems to influence the child care choice mostly on the extensive margin, i.e. parents that would have sent their child to publicly subsidised nurseries before July 2006 refrain from doing so afterwards. In contrast, on the intensive margin no significant effect takes place.

In Thuringia, the introduction of the cash-for-care subsidy causes a large reduction in the enrolment share and a bias in the reduction as it is larger for children of less-educated parents. As the other Eastern German states are rather comparable to Thuringia, a similar pattern could be expected from a general introduction of a cash-for-care subsidy as well. Hence, one would expect a reduction in the enrolment that does not stop the general trend towards higher enrolment shares, though. In the Western German states, rationing is a severe problem. Therefore, it is unlikely that one will observe a reduction in the enrolment shares as a consequence of the introduction of a cash-for-care subsidy. However, the selection effect might occur which would result in more children of highly educated parents making use of the limited slots.

From a budgetary perspective it may seem to be cheaper in the short-run to subsidise parents with 150€ than to invest into childcare and subsidise childcare facilities. However, there are two important arguments that let savings in the budget seem unlikely. Firstly, there are many parents who do not intend to make use of institutionalised child care even if one is available. Thus, subsidising parents by a cash-for-care subsidy implies that a lot of money is spent on parents who would not be subsidised by offering childcare slots what makes the subsidy rather expensive. Secondly, the positive trend in demand for childcare has neither been reversed nor stopped by the introduction of the cash-for-care subsidy in Thuringia. Hence, when the government guarantees a legal claim to publicly subsidised childcare, it has to invest into new facilities. It cannot avoid the investment by a cash-for-care subsidy.

4.8 Appendix

Table 4.8: Supply of day-care slots for children of all ages in Thuringia and Saxony

county	(1)	(2)	(3)	(4)	(5)	(6)
THURINGIA	79,454	93,159	14.7%	31,093	94,194	13.9%
Stadt Erfurt	7,249	7,500	3.3%	7,353	7,925	7.2%
Stadt Gera	3,563	4,563	21.9%	3,639	4,380	16.9%
Stadt Jena	3,804	4,228	10.0%	4,314	4,529	4.7%
Stadt Suhl	1,118	1,287	13.1%	1,068	1,377	22.4%
Stadt Weimar	2,554	2,863	10.8%	2,692	2,921	7.8%
Stadt Eisenach	1,542	1,691	8.8%	1,607	1,753	8.3%
Eichsfeld	4,083	5,291	22.8%	4,139	5,090	18.7%
Nordhausen	3,052	3,591	15.0%	3,132	3,514	10.9%
Wartburgkreis	4,843	5,802	16.5%	4,793	6,032	20.5%
Unstrut-Hainich-Kreis	4,216	5,250	19.7%	4,391	5,122	14.3%
Kyffhäuserkreis	2,855	3,335	14.4%	2,798	3,278	14.6%
Schmalkalden-Meiningen	4,353	5,285	17.6%	4,622	5,301	12.8%
Gotha	5,104	5,853	12.8%	5,118	6,012	14.9%
Sömmerda	2,754	3,132	12.1%	2,853	3,312	13.9%
Hildburghausen	2,238	2,596	13.8%	2,359	2,708	12.9%
Ilm-Kreis	3,657	4,473	18.2%	3,762	4,372	14.0%
Weimarer Land	3,205	3,734	14.2%	3,356	3,766	10.9%
Sonneberg	1,987	2,321	14.4%	1,984	2,446	18.9%
Saalfeld-Rudolstadt	3,950	4,420	10.6%	3,897	4,535	14.1%
Saale-Holzland-Kreis	3,206	3,921	18.2%	3,146	3,822	17.7%
Saale-Orla-Kreis	3,007	3,702	18.8%	2,995	3,649	17.9%
Greiz	3,933	4,717	16.6%	3,935	4,710	16.5%
Altenburger Land	3,181	3,604	11.7%	3,140	3,640	13.7%
SAXONY	236,757	214,361	9.5%	267,514	244,853	8.5%
Stadt Chemnitz	14,150	12,633	10.7%	15,375	14,439	6.1%
Erzgebirgskreis	19,213	17,099	11.0%	21,615	19,404	10.2%
Mittelsachsen	18,464	16,831	8.8%	20,689	18,740	9.4%
Vogtlandkreis	12,961	11,664	10.0%	14,355	13,028	9.2%
Zwickau	17,894	16,349	8.6%	19,618	18,046	8.0%
Stadt Dresden	29,923	27,454	8.3%	36,146	34,168	5.5%
Bautzen	18,904	17,387	8.0%	21,424	19,450	9.2%
Görlitz	16,209	14,386	11.2%	17,307	15,314	11.5%
Meißen	15,440	13,760	10.9%	17,750	16,236	8.5%
Sächsische Schweiz-Osterzgebirge	14,897	13,603	8.7%	16,993	15,685	7.7%
Stadt Leipzig	28,849	26,581	7.9%	33,173	30,745	7.3%
Leipzig	16,510	14,865	10.0%	18,815	16,835	10.5%
Nordsachsen	13,343	11,749	11.9%	14,254	12,763	10.5%

columns 1(4): number of children in in publicly subsidised day care 2006 (2009)
columns 2(5): number of approved slots in publicly subsidised day care 2006 (2009)
columns 3(6): excess of slots over enrolment as a share of approved slots 2006 (2009)

Table 4.9: Summary statistics Thuringia (23 counties)

Variable	Mean	Std. Dev.	Min.	Max.	Obs
Enrolment 2yo	77.88	(6.294)	59.1	90.400	92
Enrolment 3yo	94.61	(3.204)	87.918	102.752	92
Cash-for-care	112.5	(65.308)	0	150	92
Disposable inc.	14678.61	(846.358)	13175	17392	69
Median income	1516.95	(139.976)	1200	1950	92
Female labour	0.497	(0.038)	0.405	0.593	92
Female part-time	0.277	(0.031)	0.216	0.375	92
No training	0.098	(0.014)	0.075	0.143	92
High education	0.101	(0.049)	0.06	0.283	92
Foreigners	2.097	(1.109)	0.700	6	92
Poor children	0.111	(0.035)	0.034	0.244	92
Big household	0.28	(0.071)	0.129	0.488	92
Welfare recip.	0.22	(0.057)	0.106	0.334	92
ALOIIsha	0.17	(0.044)	0.086	0.263	92
Fertil under6	0.346	(0.025)	0.274	0.4	92
Fertil under18	0.132	(0.011)	0.102	0.166	92
Votes CDU	-0.255	(5.044)	-8.25	18.625	92
Excise tax	188.49	(77.243)	91	512.331	92
Border West	0.435	(0.498)	0	1	92
Trend	2.5	(1.124)	1	4	92
City	0.261	(0.442)	0	1	92
populat	100030.1	(34736.22)	40173	203333	92

Table 4.10: Summary statistics Saxony (13 counties)

Variable	Mean	Std. Dev.	Min.	Max.	Obs
Enrolment 2yo	65.861	(6.155)	49.46	77.863	52
Enrolment 3yo	91.7	(2.425)	86.932	97.428	52
Cash-for-care	112.5	(65.586)	0	150	52
Disposable inc.	15084.31	(485.344)	13919	16057	39
Median income	1493.23	(79.091)	1288	1617	52
Female labour	0.5	(0.029)	0.423	0.560	52
Female part-time	0.303	(0.03)	0.235	0.383	52
No training	0.091	(0.006)	0.079	0.106	52
High education	0.125	(0.039)	0.088	0.229	52
Foreigners	2.53	(1.634)	1.049	6.504	52
Poor children	0.13	(0.027)	0.089	0.206	52
Big household	0.237	(0.04)	0.15	0.303	52
Welfare recip.	0.252	(0.04)	0.2	0.349	52
ALOIIsha	0.197	(0.031)	0.153	0.273	52
Fertil under6	0.365	(0.027)	0.3	0.412	52
Fertil under18	0.131	(0.008)	0.115	0.147	52
Votes CDU	0.331	(3.357)	-7.62	4.553	52
Excise tax	224.145	(71.812)	118.055	406.513	52
Border West	0.077	(0.269)	0	1	52
Trend	2.5	(1.129)	1	4	52
City	0.231	(0.425)	0	1	52
population	327955.1	(88963.72)	215542	507954	52

Table 4.11: Counties of Thuringia sorted according to their enrolment share in the year 2006

Table 4.12: Above average enrolment share 2006

Variable	Mean	Std. Dev.	Min.	Max.	Obs
Disposable inc.	14524.917	(788.348)	13277	16469	12
Median income	1486.333	(182.81)	1200	1750	12
Female labour	0.475	(0.032)	0.413	0.538	12
Female part-time	0.25	(0.014)	0.235	0.289	12
No training	0.098	(0.012)	0.079	0.115	12
High education	0.119	(0.058)	0.066	0.267	12
Foreigners	2.208	(1.328)	0.700	5.4	12
Poor children	0.108	(0.045)	0.07	0.244	12
Big household	0.295	(0.088)	0.129	0.419	12
Welfare recip.	0.295	(0.042)	0.233	0.365	12
ALOsha	0.246	(0.033)	0.2	0.304	12
ALOIIsha	0.177	(0.035)	0.122	0.239	12
Fertil under6	0.324	(0.025)	0.274	0.351	12
Fertil under18	0.136	(0.009)	0.12	0.15	12
Votes CDU	-1.91	(3.7)	-8.25	2.95	12
Excise tax	154.917	(55.31)	92	284	12
Border West	0.333	(0.492)	0	1	12
Trend	1	(0)	1	1	12
Cityumm	0.417	(0.515)	0	1	12

Table 4.13: Below average enrolment share in 2006

Variable	Mean	Std. Dev.	Min.	Max.	Obs
Disposable inc.	14243.545	(745.472)	13175	15653	11
Median income	1474.727	(86.461)	1333	1659	11
Female labour	0.467	(0.032)	0.405	0.512	11
Female part-time	0.263	(0.033)	0.216	0.334	11
No training	0.111	(0.015)	0.095	0.143	11
High education	0.083	(0.027)	0.061	0.153	11
Foreigners	1.936	(0.677)	1.2	3.2	11
Poor children	0.112	(0.032)	0.051	0.171	11
Big household	0.305	(0.074)	0.214	0.488	11
Welfare recip.	0.318	(0.061)	0.23	0.404	11
ALOsha	0.265	(0.048)	0.199	0.338	11
ALOIIsha	0.188	(0.051)	0.111	0.255	11
Fertil under6	0.346	(0.017)	0.319	0.375	11
Fertil under18	0.143	(0.01)	0.131	0.166	11
Votes CDU	1.55	(5.996)	-4.925	18.625	11
Excise tax	140.364	(55.142)	91	272	11
Border West	0.545	(0.522)	0	1	11
Trend	1	(0)	1	1	11
Cityumm	0.091	(0.302)	0	1	11

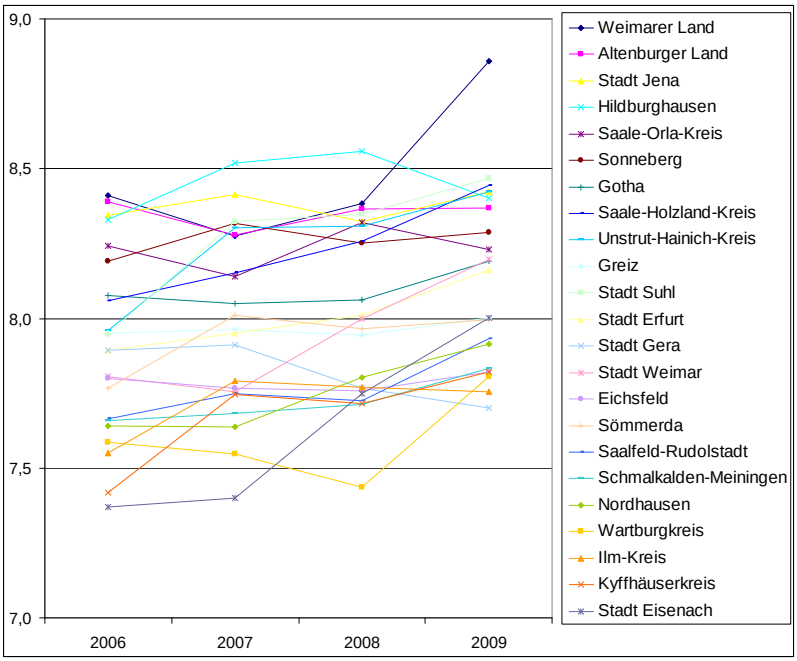
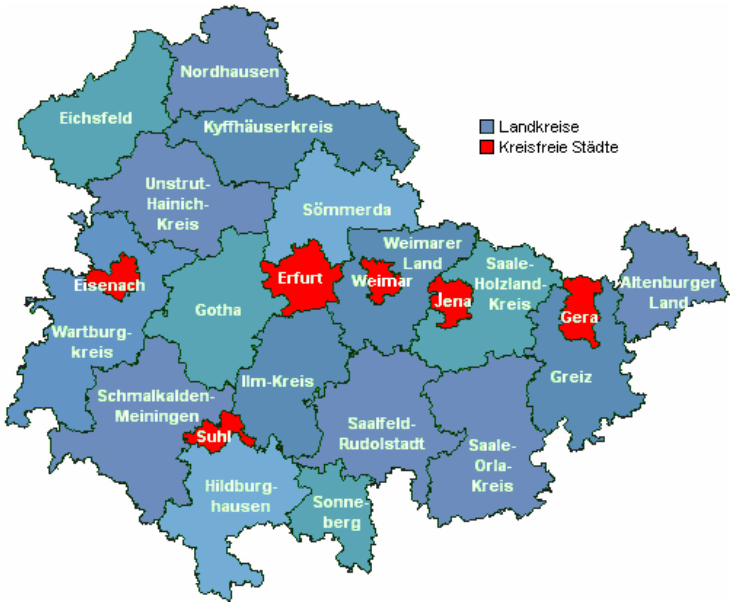


Figure 4.4: Average weighted time two-year old children spend in day care by county of Thuringia (2006-2009)



http://portal.thueringen.de/zufimap/htdocs/zufimap/index.html

Figure 4.5: Map of Thuringia and its counties

Bibliography

Abel-Koch, J. (2010). Endogenous trade policy with heterogeneous firms. mimeo, University of Mannheim.

Alexeev, M., Janeba, E., and Osborne, S. (2004). Taxation and evasion in the presence of extortion by organized crime. *Journal of Comparative Economics*, 32:375–387.

Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1(3-4):323–338.

Alm, J., Martinez-Vazquez, J., and Schneider, F. (2003). 'Sizing' the problem of the hard-to-tax. mimeo, AYSPS Conference.

Almeida, R. and Carneiro, P. (2009). Enforcement of labor regulation and firm size. *Journal of Comparative Economics*, 37:28–46.

Anderson, P. M. and Levine, P. B. (1999). Child care and mothers´ employment decisions. Working Paper 7058, NBER.

Arnott, R. J., Hosios, A. J., and Stiglitz, J. E. (1988). Implicit contracts, labor mobility, and unemployment. *American Economic Review*, 78:1046–1066.

Averett, S. L., Peters, H. E., and Waldman, D. M. (1997). Tax credits, labor supply, and child care. *The Review of Economics and Statistics*, 79(1):125–135.

Axtell, R. L. (2001). Zipf distribution of U.S. firm sizes. *Science*, 293:1818–1820.

Baker, M., Gruber, J., and Milligan, K. (2008). Universal childcare, maternal labor supply, and family well-being. *Journal of Political Economy*, 116(4):709–745.

Bernasconi, M. (1998). Tax evasion and orders of risk aversion. *Journal of Public Economics*, 67:123–134.

Bick, A. (2010). The quantitative role of child care for fertility and female labor force participation. mimeo, Goethe University, Frankfurt.

Borjas, G. J. (1999). The economic analysis of immigration. In *Handbook of Labor Economics, Volume 3*, chapter 28, pages 1697–1760. O. Ashenfelter and D. Card.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics*, 118(4):1335–1374.

Borjas, G. J. (2006). Native internal migration and the labor market impact of immigration. *Journal of Human Resources*, 42(2):221–258.

Borjas, G. J., Grogger, J., and Hanson, G. H. (2010). Immigration and the economic status of African-American men. *Economica*, 77:255–282.

Borjas, G. J. and Katz, L. F. (2005). The evolution of the Mexican-born workforce in the United States. Working Paper 11281, NBER.

Brink, A., Nordblom, K., and Wahlberg, R. (2007). Maximum fee versus child benefit: a welfare analysis of Swedish child-care fee reform. *International Tax and Public Finance*, 14:457–480.

Bucovetsky, S. (2003). Efficient migration and income tax competition. *Journal of Public Economy Theory*, 5:249–278.

Bühn, A., Karmann, A., and Schneider, F. (2007). Size and development of the shadow economy and of do-it-yourself activities in Germany. Working Paper 2021, CESifo.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64.

Card, D. (2007). How immigration affects U.S. cities. Discussion Paper 11/07, CReAM.

Carneiro, P. and Heckman, J. J. (2003). Human capital policy. In Heckman, J. J. and Krueger, A. B., editors, *Inequality in America: What Role for Human Capital Policy?* MIT Press.

Cascio, E. U. (2009). Maternal labor supply and the introduction of kindergartens into American public schools. *Journal of Human Resources.*, 44(1):140–170.

Cebula, R. J. (1997). An empirical analysis of the impact of government tax and auditing policies on the size of the underground economy: The case of the United States, 1973-94. *American Journal of Economics and Sociology*, 56(2):173–185.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review*, 98(4):1707–1721.

Chiswick, B. R. (1982). The impact of immigration on the level and distribution of economic well-being. In *The Gateway: U.S. Immigration Issues and Policies*, pages 289–313. American Enterprise Institute.

Chiuri, M. C. (2000). Quality and demand of child care and female labour supply in Italy. *Labour*, 14(1):97–118.

Coneus, K., Goeggel, K., and Muehler, G. (2009). Maternal employment and child care decision. *Oxford Economic Papers*, 61:i172–i188.

Currie, J. (2001). Early childhood education programs. *The Journal of Economic Perspectives*, 15(2):213–238.

Dabla-Norris, E., Gradstein, M., and Inchauste, G. (2008). What causes firms to hide output? The determinants of informality. *Journal of Development Economics*, 85:1–27.

Davis, D. R. and Weinstein, D. E. (2002). Technological superiority and the losses from migration. Working Paper 8971, NBER.

De Paula, Á. and Scheinkman, J. A. (2008). The informal sector. Working Paper 08-018, Penn Institute for Economic Research.

Dell'Anno, R. (2003). Estimating the shadow economy in Italy: a structural equation approach. mimeo, Department of Economics and Statistics, University of Salerno.

Dharmapala, D., Slemrod, J., and Wilson, J. D. (2008). Tax policy and the missing middle: Optimal tax remittance with firm-level administrative costs. mimeo, SSRN.

Djankov, S., La Porta, R., Lopez-de-Silanes, F., and Shleifer, A. (2002). The regulation of entry. *Quarterly Journal of Economics*, 117(1):1–37.

Dolmas, J. and Huffman, G. W. (2004). On the political economy of immigration and income redistribution. *International Economic Review*, 45:1129–1168.

Epstein, G. S. and Hillman, A. L. (2003). Unemployed immigrants and voter sentiment in the welfare state. *Journal of Public Economics*, 87:1641 – 1655.

Felbermayr, G. J. and Kohler, W. (2007). Immigration and native welfare. *International Economic Review*, 48(3):731–758.

Fortin, B., Marceau, N., and Savard, L. (1997). Taxation, wage controls and the informal sector. *Journal of Public Economies*, 66:293–312.

Freeman, R. B. (2006). People flows in globalization. *Journal of Economic Perspectives*, 20(2):145–170.

Ghironi, F. and Melitz, M. J. (2005). International trade and macroeconomic dynamics with heterogeneous firms. *Quarterly Journal of Economics*, 120(3):865–915.

Giles, D. E. A. and Tedds, L. M. (2002). Taxes and the Canadian underground economy. Canadian Tax Paper 106, Canadian Tax Foundation, Toronto/Ontario.

Goerke, L. and Runkel, M. (2006). Profit tax evasion under oligopoly with endogenous market structure. *National Tax Journal*, 59(4):851–857.

Haan, P. and Wrohlich, K. (2009). Can child care policy encourage employment and fertility? – evidence from a structural model. Discussion Paper 935, DIW Berlin.

Hamilton, J. and Pestieau, P. (2005). Optimal income taxation and the ability distribution: Implications for migration equilibria. *International Tax and Public Finance*, 12:29–45.

Hansen, J. D. (2003). Immigration and income redistribution in welfare states. *European Journal of Political Economy*, 19:735–746.

Havnes, T. and Mogstad, M. (2009). Money for nothing? Universal child care and maternal employment. Discussion Paper 4504, IZA.

Hill, R. and Kabir, M. (1996). Tax rates, the tax mix, and the growth of the underground economy in Canada: What can we infer? *Canadian Tax Journal/Revue Fiscale Canadienne*, 44:1552–1583.

Hopenhayn, H. A. (1992a). Entry, exit, and firm dynamic in long run equilibrium. *Econometrica*, 60(5):1127–1150.

Hopenhayn, H. A. (1992b). Exit, selection, and the value of firms. *Journal of Economic Dynamics and Control*, 16:621–653.

IAW (2009). Wirtschaftskrise führt wieder zum Anstieg der Schattenwirtschaft. Pressemitteilung.

Ilmakunnas, S. (1997). Public policy and child care choices. In Persson, I. and Jonung, C., editors, *Economics of the Family and Family Policies*, pages 178–193. Routledge, London.

Johnson, S., Kaufmann, D., and Zoido-Lobatón, P. (1998). Government in transition - regulatory discretion and the unofficial economy. *American Economic Review Papers and Proceedings*, 88(2):387–392.

Kimmel, J. (1995). The effectiveness of child-care subsidies in encouraging the welfare-to-work transition of low-income single mothers. *The American Economic Review*, 85(2):271–275.

Kimmel, J. (1998). Child care costs as a barrier to employment for single and married mothers. *The Review of Economics and Statistics*, 80(2):287–299.

Kohler, W. and Felbermayr, G. (2009). Can international migration ever be made a Pareto improvement. Research Paper 2009/01, University of Nottingham.

Kornstad, T. and Thoresen, T. (2007). A discrete choice model for labor supply and childcare. *Journal of Population Economics*, 20:781–803.

Lefebvre, P. and Merrigan, P. (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from Canada. *Journal of Labor Economics*, 26(3):519–548.

Loayza, N. V. (1996). The economics of the informal sector: a simple model and some empirical evidence form Latin America. *Carnegie-Rochester Conference Series on Public Policy*, 45:129–162.

Longhi, S., Nijkamp, P., and Poot, J. (2005). A meta-analytic assessment of the effect of immigration on wages. *Journal of Economic Surveys*, 19(3):451–477.

Love, J. M., Eliason, E., Ross, C., Raikes, H., Constantine, J., and Boller, K. (2005). The effectiveness of early head start for 3-year-old children and their parents: Lessons for policy and programs. *Developmental Psychology*, 41(6):885–901.

Lundin, D., Mörk, E., and Öckert, B. (2007). Do reduced child care prices make parents work more? Working Paper 2007:2, Institute for Labour Market Policy Evaluation.

Marrelli, M. (1984). On indirect tax evasion. *Journal of Public Economics*, 25:181–196.

Marrelli, M. and Martina, R. (1988). Tax evasion and strategic behaviour of the firms. *Journal of Public Economics*, 37:55–69.

Michael, M. S. (2003). International migration, income taxes and transfers: a welfare analysis. *Journal of Development Economics*, 72:401–411.

Michael, M. S. (2006). Are migration policies that induce skilled (unskilled) migration beneficial (harmful) for the host country? Working Paper 1814, CESifo.

Mirrlees, J. A. (1982). Migration and optimal income taxes. *Journal of Public Economics*, 18:319–341.

Myers, G. M. and Papageorgiou, Y. Y. (2000). Immigration control and the welfare state. *Journal of Public Economics*, 75:183–207.

Ottaviano, G. I. and Peri, G. (2008). Immigration and national wages: Clarifying the theory and the empirics. Working Paper 14188, NBER.

Peri, G. and Sparber, C. (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, 1(3):135–169.

Powell, L. M. (1997). The impact of child care costs on the labour supply of married mothers: Evidence from Canada. *The Canadian Journal of Economics*, 30(3):577–594.

Prado, J. M. (2008). Government policy in the formal and informal sectors. mimeo, IMT Lucca Institute for Advanced Studies.

Richter, W. E. (2004). Delaying integration of immigrant labor for the purpose of taxation. *Journal of Urban Economics*, 55:597 – 613.

Sánchez-Mangas, R. and Sánchez-Marcos, V. (2008). Balancing family and work: The effect of cash benefits for working mothers. *Labour Economics*, 15(6):1127–1142.

Schneider, F. (2000). The increase of the size of the shadow economy of 18 OECD countries: Some preliminary explanations. mimeo, Paper presented at the Annual Public Choice Meeting, March 2000, Charleston,S.C.

Schneider, F. (2005). Shadow economies around the world: what do we really know. *European Journal of Political Economy*, 21:598–642.

Schneider, F. and Enste, D. H. (2000). Shadow economies: Size, causes, and consequences. *Journal of Economic Literature*, 38:77–114.

Schneider, F. and Klinglmair, R. (2004). Shadow economies around the world: What do we know? Working Paper 1167, CESifo.

Schøne, P. (2004). Labour supply effects of a cash-for-care subsidy. *Journal of Population Economics*, 17:703–727.

Shaw, J., Slemrod, J., and Whiting, J. (2008). Administration and compliance. In *Reforming the Tax System for the 21st Century: The Mirrlees Review*, chapter 12. Institute for Fiscal Studies.

Sinn, H.-W. (2005). Migration and social replacement incomes: How to protect low-income workers in the industrialized countries against the forces of globalization and market integration. *International Tax and Public Finance*, 12:375–393.

Slemrod, J. and Bakija, J. (2004). *Taxing ourselves - A Citizen's Guide to the Debate over Taxes*. MIT Press, 3rd edition.

Stiglitz, J. E. (1982). Self-selection and Pareto efficient taxation. *Journal of Public Economics*, 17:213–240.

Temple, J. A. and Reynolds, A. J. (2007). Benefits and costs of investments in preschool education: Evidence from the child–parent centers and related programs. *Economics of Education Review*, 26:126–144.

Thüringen (2005). Thüringer Familienfördergesetz, Article 3. Gesetz- und Verordnungsblatt für den Freistaat Thüringen.

Traxler, C. (2010). Social norms and conditional cooperative taxpayers. *European Journal of Political Economy*, 26:89–103.

Tuzova, Y. (2008). A model of tax evasion with heterogeneous firms. Technical report, University of Minnesota.

Vannoorenberghe, G. (2008). Trade between symmetric countries, heterogeneous firms and the skill premium. Discussion Paper 43, University of Mannheim, CDSE.

Wellisch, D. and Walz, U. (1998). Why do rich countries prefer free trade over free migration? The role of the modern welfare state. *European Economic Review*, 42:1595 – 1612.

Wellisch, D. and Wildasin, D. E. (1996). Decentralized income redistribution and immigration. *European Economic Review*, 40:187–217.

Wetzels, C. (2005). Supply and price of childcare and female labour force participation in the Netherlands. *Labour*, 19:171–209.

Wilson, J. D. (1980). The effect of potential emigration on the optimal linear income tax. *Journal of Public Economics*, 14:339–353.

Wilson, J. D. (1992). Optimal income taxation and international personal mobility. *American Economic Review Papers and Proceedings*, 82(2):191–196.

Wilson, J. D. (2006). Income taxation and skilled migration: The analytical issues. mimeo, Michigan State University.

Wilson, J. D. (2007). The welfare state versus the common labor market: Which to dismantle? *CESifo Economic Studies*, 53(4):618–636.

Wrohlich, K. (2004). Child care costs and mothers' labor supply: An empirical analysis for Germany. Discussion Paper 412, DIW.

Wrohlich, K. (2006). Labor supply and child care choices in a rationed child care market. Discussion Paper 2053, IZA.

ZEW (2009). Fiskalische Auswirkungen sowie arbeitsmarkt- und verteilungspolitische Effekte einer Einführung eines Betreuungsgeldes für Kinder unter 3 Jahren. Studie im Auftrag des Bundesministeriums der Finanzen.

LEBENSLAUF

Björn Uwe Saß

10/2004 - 06/2010	Promotion in Volkswirtschaftslehre an der Universität Mannheim Lehrstuhl für Volkswirtschaftslehre, Finanzwissenschaft und Wirtschaftspolitik
10/2004	Abschluss als Diplom-Volkswirt, Otto-Friedrich-Universität Bamberg
09/2001 - 06/2002	Auslandsstudienjahr an der Université Jean-Moulin, Lyon III, Lyon, Frankreich
10/1999 - 10/2004	Studium der Europäischen Wirtschaft an der Otto-Friedrich-Universität Bamberg
05/1999	Abitur am Helmholtz-Gymnasium, Essen

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig angefertigt und mich keiner anderen als der in ihr angegebenen Hilfsmittel bedient zu haben. Insbesondere sind sämtliche Zitate aus anderen Quellen als solche gekennzeichnet und mit Quellenangaben versehen.

Mannheim, 30.04.2010

Björn Saß