Essays in Applied Microeconomics

Adrian Lerche

TESI DOCTORAL UPF / ANY 2019

DIRECTOR DE LA TESI
Albrecht Glitz

Departament d'Economia i Empresa

Universitat
Pompeu Fabra
Barcelona

To my parents and grandparents.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor Albrecht Glitz who supported me in all aspects of my PhD. I very much appreciate his guidance in choosing my research topic, his advice for improving my research, his time for reading the many drafts of my papers and his financial support. I am also glad to have had the chance to work on a joint project that gave me profound insights into carrying out research.

The support from the faculty at UPF, in particular Libertad González, Christian Fons-Rosen, Gianmarco León, Giacomo Ponzetto, Ruben Enikolopov and the rest of the LPD research group, was exceptional and I would like to thank them for listening to my research ideas and their constructive feedback. I extend my special thanks to Marta Reynal-Querol for choosing me for her FPI scholarship. I am grateful for the many other people that took their time to talk about my research. Among those, I would like to mention Christian Dustmann, who invited me to spend time at UCL, and Uta Schönberg since both had an undeniable influence on my research direction.

I would also like to thank Marta Araque and Laura Agustí for their administrative support. They were always ready to help, may it be with getting the required documents and signatures for research data, the ever changing reimbursement system, knowing who to contact at UPF, or any other organizational issues.

I could not have imagined a better place to write my thesis than Barcelona. It is an inspiring city, full of life and constant change. I feel fortunate that it was my home for so many years and I am thankful for the thousands of great moments that I have experienced there. My colleagues from UPF are one major reason for such a great time. Paul, Christoph H., Christoph A., Christian, Flavio, Stefan, Adam, Bjarni, Donghai, Flo, Alex, Dijana, Giulia, Korie, Liza, Nick, Cate, Miguel, I am very happy to have met you. The second reason for remembering these years so vividly is the time spent with everybody

else. Laura, Basti, Sven, Robbie, I appreciate the many memories shared with you in Barcelona and afterwards. My rock climbing friends, it was so nice of you to integrate me even though I barely spoke Spanish or Catalan. Martin, Ludwig, Christian, every time I visited Benshausen, I knew that meeting you was as if no time had passed in between.

Finally, I wish to thank my family. They had the longest and biggest impact on who I am today. My mom and dad taught me early on to strive for the best possible outcome, and that hard work and dedication are the best route to success. They always supported me in my educational decisions, emotionally and financially, no matter how far away I moved, and helped me immensely when things did not work as expected. My sister had an important influence in her own right. It was great to spent time with her in Barcelona and during the short stays when traveling to and from Germany. She also made me realize that I should appreciate all the opportunities that I had so far in my life.

Abstract

This thesis consists of three self-contained chapters that empirically evaluate the influence of capital costs and occupational regulation on labor demand. In the first chapter, I study the effects of investment tax credits on firms' input choices in Germany. I find evidence that such a policy has a strong positive direct effect on firm investment and employment, and that positive spillovers between firms lead to sizable further adjustments. In the second chapter, I estimate the firm-level capital-labor elasticity of substitution. I set up a model of firm production with size-dependent capital costs and estimate the model for a German tax policy targeted towards manufacturing firms. The estimated elasticity implies important complementarities between capital and labor in firm production. In the third chapter, I analyze how the formal recognition of immigrants' foreign occupational qualifications affects their subsequent labor market outcomes using novel German data. The results show that access to regulated occupations after recognition is an important driver for faster assimilation of immigrants' earnings.

Resum

Aquesta tesi consta de tres capítols independents que avaluen empíricament la influència dels costos de capital i la regulació ocupacional sobre la demanda laboral. En el primer capítol, estudio els efectes dels crèdits fiscals sobre inversions a la determinació d'input de les empreses a Alemanya. Trobo proves que aquesta política té un fort efecte directe positiu en la inversió i en l'ocupació de les empreses, i que els efectes indirectes entre empreses porten a ajustaments addicionals positius. En el segon capítol, calculo l'elasticitat de substitució entre el capital i el treball a nivell de l'empresa. Estableixo un model de producció d'empreses amb costos de capital discontinu i estimo el model per una política fiscal alemanya dirigida a empreses manufactureres. L'elasticitat estimada implica importants complementarietats entre el capital i el treball en la producció de les empreses. En el tercer capítol, analitzo com el reconeixement formal de les qualificacions ocupacionals estrangeres dels immigrants afecta els seus resultats del mercat laboral posteriors utilitzant noves dades alemanyes. Els resultats mostren que l'accés a ocupacions regulades després del reconeixement és un factor important per a una assimilació més ràpida dels ingressos dels immigrants.

Preface

This thesis consists of three essays on topics related to labor economics and public economics. The research projects are tied together by their focus on the analysis of labor demand and its determinants.

The first chapter estimates separately the direct and spillover effects of investment tax credits using administrative data on firms in Germany. To identify causal effects, I combine difference-in-differences designs at the firm and regional level that exploit a shift in the tax credit rate for manufacturing firms by firm size. I find that firms increase both investment and employment when receiving more generous tax credits, with implied elasticities with respect to capital costs of 2.8 and 1.1, respectively. On top of these direct effects, positive spillovers between firms lead to sizable further adjustments, magnifying the employment effect in labor markets by up to 120%. A heterogeneity analysis reveals that spillovers in the manufacturing sector tend to be stronger within industries and do not lead to benefits for the service sector. Furthermore, firms in industries with higher investment shares into information and communications technology (ICT) are more likely to shift towards highly educated labor and high skill occupations.

The second chapter exploits the same tax policy to estimate the firm-level elasticity of substitution between capital and labor for manufacturing firms in Germany. The policy reduced capital costs by offering size-dependent investment tax credits with firms with up to 250 employees receiving a higher tax credit rate. I incorporate size-dependent capital costs into a firm production framework to analyze distortions in firm size and capital created by the policy. The model predicts that a fraction of firms decrease their firm size and bunch at the employment cutoff to profit from lower capital costs. Firms below the cutoff also move towards a more capital-intensive production. I present descriptive evidence that is in line with these predictions and structurally estimate the model using maximum likelihood and non-linear least squares techniques. I obtain an elasticity of substitution of 0.08 that suggests substantial complementarities between capital and labor.

Finally, the third chapter, co-authored with Herbert Brücker, Albrecht Glitz and Agnese Romiti, analyzes how the formal recognition of immigrants' foreign occupational qualifications affects their subsequent labor market outcomes. The empirical analysis is based on a novel German data set that links respondents' survey information to their administrative records, allowing us to observe immigrants at monthly intervals before, during and after their application for occupational recognition. Our findings show substantial employment and wage gains from occupational recognition. After three years, the full recognition of immigrants' foreign qualifications increases their employment rates by 24.5 percentage points and raises their hourly wages by 19.8 percent relative to immigrants without recognition. We show that the increase in employment is largely driven by a higher propensity to work in regulated occupations. Relating our findings to the economic assimilation of immigrants in Germany, we further document that occupational recognition leads to substantially faster convergence of immigrants' earnings to those of their native counterparts.

Table of Contents

List of Figures

List of Tables

Chapter 1

DIRECT AND SPILLOVER EFFECTS OF INVESTMENT TAX CREDITS

1.1 Introduction

Governments have long used tax policy in an effort to stimulate economic activity. Because the accumulation of capital is thought to be key to the creation of economic growth, there is frequent reliance on investment tax credits and similar tax incentives that reduce investment costs.[1] Proponents of such tax policies argue that a reduction in investment costs encourages additional investments that lead to the expansion of production, higher labor demand and positive spillover effects between firms. Others warn that such tax benefits provide economic rents to firm owners who would have invested anyway, gen-

[1]Investment tax credits in the U.S. on a national level were first introduced with the Revenue Act of 1962 and played a prominent role until its repeal in 1986. In 1985, government expenses for this policy totaled $21 billion (Chirinko, 2000). In 2004, 40% of U.S. states had their own investment tax credit program (Chirinko and Wilson, 2008).

erate real effects only for the initially targeted firms and prompt the substitution of workers with capital.[2]

For the evaluation of the impact of investment tax credits, therefore, it is necessary to analyze not only the investment behavior and input choices of targeted firms but also the additional adjustment processes that may arise throughout the economy at large. In this context, policy-makers often express a particular interest in the role of tax credits for the creation of jobs, both in targeted firms and through spillovers. Because of the difficulty of separately identifying direct and spillover effects and the scarcity of sufficiently detailed micro-level data, there is so far limited empirical evidence on the adjustment behavior of firms.

In this paper, I investigate the effect of investment tax credits on firm investment, employment and workforce composition, and quantify the relevance of spillover effects between firms. To this end, I consider a tax policy in Germany that was introduced in 1991 immediately after reunification to mitigate considerable economic deficiencies in East Germany. The program provided significant support to firms, with reductions in investment costs of up to 27.5% and annual government expenses of €1–2 billion per year. Firms were able to recover investment costs even in excess of their tax liabilities, meaning that the tax credits were *refundable*.

I exploit variation in the tax credit rate by firm size and time. During the period 1995–2004, initially, manufacturing firms with up to 250 employees in a given business year were eligible for a tax credit rate of 10% while those with more than 250 employees were eligible for a rate of only 5%. In 1999 a change in the rates amplified this differential treatment, increasing the tax credit rate to 20% for firms below the cutoff and 10% for firms above the cutoff. In addition to the overall reduction in investment costs, these changes generated a relative decrease for smaller firms.

[2]For example the discussions surrounding Bill Clinton's proposal for investment tax credits in 1992 illustrates both sides.

I focus on the relative decrease in investment costs and implement difference-in-differences estimations for the period 1995–2004. I estimate the direct effects of the policy change by comparing the behavior of a group of firms below and above the firm size cutoff over time.[3] I then augment the regression model by a difference-in-differences approach comparing firms across labor markets according to the regional share of firms below the cutoff to estimate spillover effects. To mitigate concerns about time-varying shocks and distortions around the size cutoff, I include time-varying industry and labor market fixed effects, and exclude observations close to the cutoff.

The estimation strategy is guided by a theoretical firm production framework that includes two types of labor and regional productivity into a CES production function. An increase in investment tax credits leads to an unambiguously positive effect on investment but the effect on employment and workforce composition is uncertain, depending on the degree of substitution between capital and different types of labor. On top, spillovers between firms within local labor markets may generate a positive or negative impact on investment and employment decisions, and this effect is larger within labor markets where a higher share of firms receive a reduction in capital costs.

My first empirical finding is that tax credits have a substantial direct effect on investment. The increase in the relative tax credit rate leads to 23.4 log points higher overall investment and 25.1 log points higher equipment investment at the intensive margin. Given the underlying relative change in capital costs of 8.2%, the overall investment response corresponds to an elasticity with respect to capital costs of 2.8. Considering the dynamic effects, there is an increase in the intensive margin investment response immediately after the policy change. The effect gets stronger in the subsequent year and

[3] I use the term "firm" interchangeably with "establishment" throughout. The main analysis is conducted at establishment level due to the aggregation of the datasets. The policy cutoff, however, is at firm level. I provide evidence that this simplification does not influence the qualitative results by considering single-establishment firms in a robustness test.

stabilizes at a higher level thereafter. There is, however, no discernible effect on the probability of investing, likely because most firms already invest in any given year before the policy change.

I next consider the employment effects of the policy change and find a positive effect of tax credits on overall employment. Based on the preferred estimate, the change in the tax credit rate leads to 8.7 log points higher employment, equivalent to an elasticity with respect to capital costs of 1.1. This effect is similar when considering full-time workers, workers in regular employment or full-time workers in regular employment. Most of the increase materializes in the year after the policy change and there is a slight further increase in subsequent years. I further find that the increase in employment is almost exclusively driven by hiring additional employees rather than fewer separations. Among the additional hires, a share of 49% was unemployed one year prior which is similar to the average share among all hires in the data.

The investment and employment response translate into more output, measured in terms of revenue. Estimates for the revenue effect, however, tend to be more volatile and less precisely estimated. The response to the tax credit rate change for domestic revenue amounts to 8.3 log points. Taking these results together, tax credits seem to be an effective tool to induce higher investment among targeted firms, and the change in investments then translates to more employment and output within the same firms.

When turning to the regression model that estimates both direct and spillover effects, the estimate of the direct employment effect does not change markedly, with the estimate indicating an increase by 10.1 log points. On top of this direct effect, there are positive spillover effects. They depend on the share of firms in a labor market below the cutoff, as predicted by the theoretical model. In a labor market with only treated firms, the spillover effects lead to an additional increase of employment by 12.0 log points, meaning that spillovers amount to 54% of the combined effect. This result suggests that the benefits from tax credits among targeted firms propagate locally and spillovers have important implications when considering the cost-effectiveness of the policy.

These results mask important heterogeneity in the direct and spillover effect. First, firms with larger capital cost shares measured by relating average annual investment costs to the average annual wage bill have a stronger investment and employment response. Second, the spillover effect tends to be stronger for firms within the same industry, suggesting that similarities between firms are important for creating spillovers as is the case for agglomeration economies. In contrast, there is no significant effect on the service industries, which counters the idea that an increase in demand for local goods and services creates spillovers through local multipliers.

Capital-skill complementarity would suggest that highly educated labor profits more relative to less educated labor from tax credits. I consider the ratios of college-educated versus non-college-educated employees, and high skill occupations such as engineers and managers versus low and medium skill occupations like machine operators. All point estimates are close to zero and statistically insignificant. The coefficients are estimated precisely and I reject estimates of all but modest skill composition changes. Thus, the added capital in firms does not shift employment opportunities towards highly educated labor or high-skilled occupations. The literature on technological change considers information and communication technology (ICT) strongly complementary with skill. When analyzing heterogeneous effects across industries, I find that industries with a higher share of investment and capital in ICT are indeed more likely to shift towards college-educated employees and high skill occupations.

I perform various robustness checks to verify the research design. First, plotting the raw data year by year and checking pretreatment year estimates in the dynamic specifications, I find no differential pretreatment behavior between treatment and control group. Second, adjustments in inputs occur directly in the year after the policy change providing a link between the policy and firm behavior. Third, the results are robust to different sample selection procedures. Average effects are stable for the selection of narrower and wider firm size intervals of the treatment and control group, and for the exclusion of different firm size intervals around the cutoff. Finally, in a placebo test for firms below the cutoff, choosing policy-irrelevant firm size cutoffs leads to insignificant estimates.

This paper contributes to several strands of the literature. Starting with Hall and Jorgenson (1967) a long literature has emerged that empirically quantifies the effect of capital cost changes on firm investment. Earlier studies rely on aggregate time-series data and find surprisingly small responses in investment (Abel, 1980; Abel and Blanchard, 1986; Summers, 1981).[4] In an attempt to overcome measurement bias, subsequent studies use firm-level data and cross-sectional variation in tax policies, and generally find larger effects (Cummins et al., 1994; Edgerton, 2010). A survey of such studies by Hassett and Hubbard (2002) concludes that the elasticity of investment with respect to capital costs is between 0.5 and 1.0. Although these studies consider investment tax credits, capital cost reductions are calculated as the total of all available tax incentives at any given time. The recent literature predominantly focuses on the analysis of specific tax incentive programs that introduce cross-sectional variation. House and Shapiro (2008), Maffini et al. (2016) and Zwick and Mahon (2017) examine special depreciation allowances and find large responses in investment behavior with elasticities of investment of around eight.

To the best of my knowledge, I am the first to estimate causal effects of the impact of investment tax credits on firm investment using plausibly exogenous cross-sectional variation. The main advantage of the analysis of investment tax credits is their clear link to capital costs, that in contrast to depreciation allowances do not depend on assumptions for the discount factor and depreciation schedules. Since investment tax credits in the German case were refundable, there is also no influence of the firms' profit situation on capital costs. I find

[4] The elasticities are considered *too* small, since they imply extremely high capital adjustment costs.

an elasticity of investment of 2.8, which is in between the lower values from earlier studies and the large values of recent studies.

I use this result as a starting point for a comprehensive analysis of the impact of investment tax credits on secondary outcomes. As a novelty, I exploit variation within and across labor markets to separately estimate direct effects and indirect spillover effects. On the one hand, this approach adds to the literature by connecting the firm-level evidence on the investment response to related firm outcomes like employment, that are influenced by a capital cost reduction as well. On the other hand, the existence of spillover effects reveals an adjustment mechanism that operates on an aggregate level and their estimation is a step towards the decomposition of the total effect of tax policy found for example in macro-level tax policy studies (Blanchard and Perotti, 2002; Mertens and Ravn, 2013; Romer and Romer, 2010).

With this approach, I also contribute to the literature on spillover effects between firms. My results relate to Greenstone et al. (2010) and Gathmann et al. (2018) that exploit exogenous variation across labor markets in firm openings and closings respectively.[5] These studies focus on relatively specific events for large firms. I add to this literature by analyzing a far-reaching policy that focuses on smaller firms, and I exploit continuous treatment assignment across labor markets. My results suggest that small firms create spillovers as well, as long as they add up to a sufficient share of a labor market. I consider the importance of agglomeration economies and local multipliers as mechanisms for spillovers and in contrast to Moretti and Thulin (2013) do not find evidence of local multipliers.

Furthermore, my paper speaks to the large literature on place-based policies and their effects on regions (Becker et al., 2010, 2013; Busso et al., 2013; Criscuolo et al., 2016; Dettmann et al., 2016; Etzel and Siegloch, 2018; Kline and Moretti, 2014a) and firms (Bronzini and de Blasio, 2006; Cerqua and Pellegrini, 2014). In these studies,

[5]The overall literature on spillovers is much larger. For an overview particularly concerning agglomeration economies see Combes and Gobillon (2015).

there is a stronger focus on employment, but the evidence on the employment effect is inconclusive. This may be because the analyzed policies often combine a mix of regional and firm-specific incentives, including reductions in capital costs. I add to this literature by studying one particular incentive, investment tax incentives, and study adjustments in input and output independent of other influences.

Finally, by considering heterogeneous effects across labor types, my findings speak to the often voiced concern that tax policies and place-based policies can lead to unwanted redistribution in welfare (Glaeser and Gottlieb, 2008; Kline and Moretti, 2014b; Neumark and Simpson, 2015). The literature on the shift of production technology towards automation (e.g. Acemoglu and Autor, 2011) and capital in general (Krusell et al., 2000) suggests that there can be an advantage for high-skilled over low-skilled labor. The tax credit program does not have adverse effects on the skill composition on average and creates employment opportunities for unemployed individuals. However, I find an influence of ICT on a shift towards high-skilled labor. Similar to Akerman et al. (2015), this result points to potential adverse effects of government programs supporting ICT investments.

The structure of the paper is as follows. Section 1.2 explains the policy intervention in more detail focusing on the relevant regulations for the empirical analysis. Section 1.3 introduces a theoretical framework that provides intuition for expected firm behavior with heterogeneous labor types and spillovers. Section 1.4 explains the estimation strategy. Section 1.5 provides detail on the data including descriptive statistics and sample selection. Section 1.6 presents the main results and Section 1.7 relates the results in a back-of-the-envelope cost-benefit analysis. Section 1.8 concludes.

1.2 Policy Intervention

After the second world war, Germany split into two countries, West Germany and East Germany.[6] While West Germany experienced continued growth with a market-based economy, East Germany faced large war reparations and inefficiencies in its communist economic system. The fall of communism throughout Eastern Europe led to the reunification of Germany in 1990. The diverging prior development however created a country with economically disparate regions. Figure 1.1 shows that over the period 1991 to 1994, East Germany had on average 46% lower GDP per capita, 47% lower capital per worker, 30% lower earnings per worker and an unemployment rate of 13.4% compared to 7.1% in West Germany. To speed up economic convergence, the government provided considerable financial support to regions in East Germany. Besides cash transfers to private households and large infrastructure investments, efforts were focused on increasing the capital stock of firms. The most salient policy in this respect was an investment tax credit program (*Investitionszulagengesetz*) which is at the center of this paper.

The program started immediately after reunification in 1991 and lasted until 2013. It provided tax credits for equipment investments to firms located in East Germany and West Berlin. From 1999, it also covered investments in structures. At the beginning of the program, firms of all industries were eligible for the program but over time access became more restricted and by 1997 coverage applied almost exclusively to manufacturing firms.[7] Tax credits typically reduced investment costs by around 10% but depending on the exact location

[6] During the time of separation the official designation for West Germany was *Federal Republic of Germany* and for East Germany *German Democratic Republic*. I use the common names since they are still used to refer to the respective parts of Germany after reunification.

[7] Retail businesses continued to have limited eligibility until 2001. Manufacturing service businesses like construction design or research gained access to tax credits in 1999. Accommodation businesses were eligible from 2007.

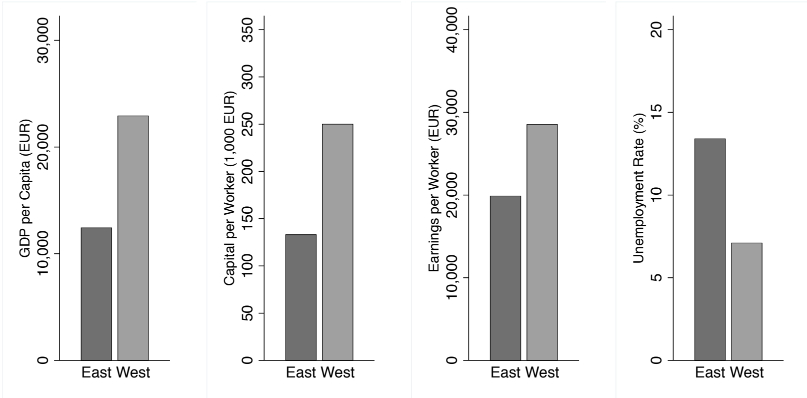


Figure 1.1: Economic Indicators for West and East Germany for 1991–1994

Note: The bars represent averages over the years 1991-1994. Data source for Panel (a)-(c) is the Federal Statistical Office. Data source for Panel (d) is the Federal Employment Institute. The unemployment rate is based on figures of June of each year. All of Berlin is counted towards the statistics of East Germany.

and firm size, the reduction could be as high as 27.5%.[8] Tax credits were fully paid even if they exceeded tax liabilities of a firm and did not depend on the life span of the investment good.[9]

Importantly, all details on firm eligibility and the tax credit rate were precisely defined by law without room for discretion on a case by case basis. This led to an entitlement to tax credits for eligible firms and thus certainty for the planning of long-term investment projects.[10] The tax credits therefore can be considered a pure capital cost reduction for firms. The administrative cost for receiving tax

[8]The highest rate applied to equipment investments of manufacturing firms with at most 250 employees in regions close to the Czech and Polish borders from 2002 to 2009.

[9]This is in contrast to special depreciation allowances where the decrease in cost of capital depends on the profit situation of a firm and on the years of depreciation due to the present value of future tax deductions. House and Shapiro (2008) and Zwick and Mahon (2017) provide detailed explanations.

[10]This is in contrast to various place-based policies that distribute grants to investment projects via a competitive application process with the final outcome

credits was small. Firms filled out a tax credit claim form describing the investment good and the value of investment. Tax officers would check the correctness of the claim after the end of the business year and a positive assessment would trigger the transfer of tax credits. To reduce adverse incentives, a number of further eligibility criteria needed to be satisfied. Assets had to stay within the firm for at least 3 years to prevent East German firms from becoming pass-through companies of buying and reselling fixed assets.[11] Planes, passenger cars and low value assets such as office equipment or basic tools were never eligible because verifying their continued presence within an eligible firm in East Germany (or West Berlin) would entail large monitoring costs.

The program was costly. Figure 1.2 summarizes overall government expenses for tax credits by year based on available information in the official subsidy reports. From 1992 to 1995 expenses totaled around €2 billion per year. After 1993, expenses declined steadily, which can be explained by the reduction in eligible industries. They reached a low of €645 million in 1998 and stabilized thereafter at around €1 billion per year. Starting in 2000, expenses from tax credits for investments in structures contributed to the total and generally made up around 15% each year.

In the empirical analysis, I focus on manufacturing firms as the main recipients of tax credits and consider the time period between 1995 and 2004, comparing their behavior around a sudden change in tax credit rates for equipment investments in 1999. From July 1994, manufacturing firms with up to 250 employees received a tax credit rate of 10% on equipment investments. Firms with more employees instead received 5%. The program defined firm size as the number of employees at the beginning of a business year without differentiating full-time and part-time employment, and excluding vocational trainees since they are employed through special educational con-

based for example on the perceived success of the project or a commitment to hire additional employees.

[11]After 1999 the minimum time period was extended to 5 years.

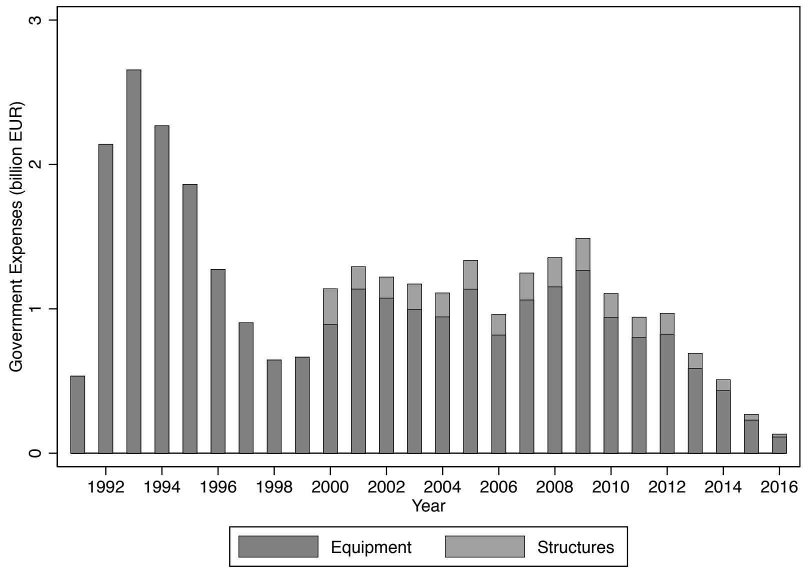


Figure 1.2: Government Expenses for Tax Credits by Year and Asset Type

Note: Information taken from subsidy reports of the German government. Values that are available only in Deutsche Mark were converted to euro using the official conversion rate. Figures on government expenses get revised over time due to additional information. The presented figures are taken from the most recent report available for each year respectively to reflect the most current information status. Expenses after 2016 excluded.

tracts. At the beginning of 1999, tax credits were raised for a broad range of equipment investments, with firms below the employment cutoff now receiving a rate of 20% and firms above the cutoff receiving 10%.[12] The announcement of the policy change was published in August 1997 but because of disputes with EU law, it got approval only by the end of 1998. The adjustments led to a decrease in capital

[12]At the same time an investment limit for receiving the higher tax credit rate of €2.56 million per year for firms below the employment cutoff was eliminated. The additional tax credit rate for investments in structures was independent of firm size and thus, did not lead to relative differences.

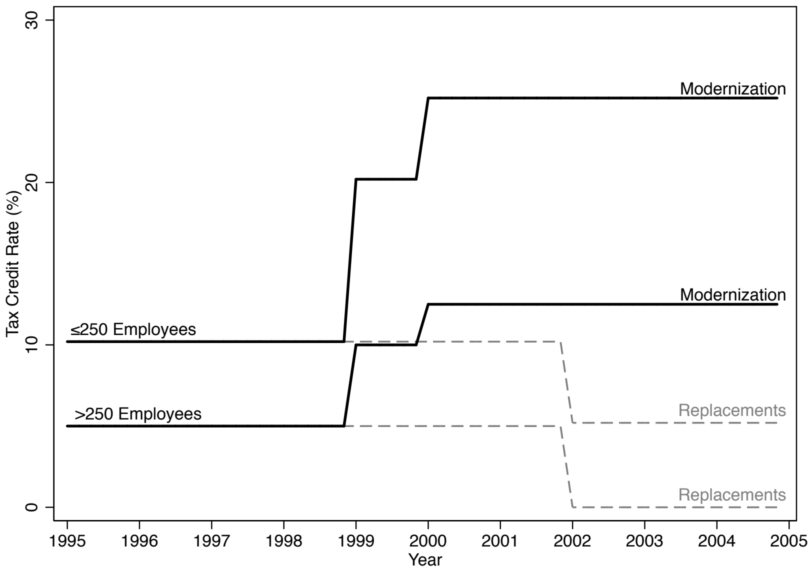


Figure 1.3: Equipment Tax Credit Rates for Manufacturing Firms

Note: The shown tax credit rates apply to manufacturing firms in most parts of East Germany excluding Berlin for equipment investment that start and end at the same day. For Berlin and in some years for areas close to Berlin rates were lower. There was a slight increase in the tax credit rate for modernization investments in regions close to the Polish or Czech border starting in 2001. Changes in tax credit rate were usually accompanied by phase-out periods to allow a constant rate for longer-lasting investment processes.

costs for all manufacturing firms. However, the increase in the tax credit rate was larger for firms below the cutoff and granted them a relative decrease of capital costs of equipment compared to firms above the cutoff.

The change applied to so-called modernization investments, which included among others any investment that could potentially increase production, change the production process or produce different products. Any investment in new equipment that did not directly replace a similar asset fell within this category. Even (high value) office equipment could be part of this category as long as it was bought in connection to a specific modernization investment.

Figure 1.3 summarizes the general tax credit changes for equipment investments of manufacturing firms in East Germany between 1995 and 2004.[13] Apart from the adjustment in 1999, there was another increase for modernization investments in 2000 that further strengthened the relative advantage of firms below the cutoff to those above. Tax credits for non-modernization equipment investment remained unchanged during the policy update in 1999 but were reduced in 2002. However, this reduction did not change the differential treatment and maintained a higher rate of 5 percentage points for firms below the cutoff before and after the change.

Since the definition for firm size changed markedly in 2005, I exclude those years from my analysis. From then on the cutoff value followed the definition of small and medium firms by the European Union that takes ownership structure into account and defines the cutoff with respect to the number of employees, revenue and total assets.[14]

1.3 Theoretical Framework

To understand firm behavior after capital cost changes, it is helpful to outline a simple model of firm production. The literature has already developed detailed models of firm investment behavior in which adjustment costs and corporate taxation play an integral part.[15] Since I rely on a reduced-form approach in the empirical analysis, I abstract from both these issues and focus on a static model with capital costs

[13] Berlin had generally lowered rates throughout the time period. Regions close to the Polish and Czech border received slightly higher rates between 2001 and 2009. For long investment projects, additional rules applied around changes of the tax credit rate.

[14] The definition can be found in the "Commission recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises" (ABl.EU, #L 124 pp36).

[15] Important examples are Hall and Jorgenson (1967), Hayashi (1982) and Abel and Eberly (1994). Hassett and Hubbard (2002) and Bond and Reenen (2007) summarize basic model assumptions and survey further approaches.

consisting of a universal capital rental rate and a firm-specific tax credit rate.[16] I further focus on labor as input to the production process, consider heterogeneous labor types and capture the influence of spillovers in regional firm production with a regional productivity shifter similar to Greenstone et al. (2010) and Gathmann et al. (2018).

The model assumes many firms i within many regions r. Each firm produces one differentiated good according to the nested CES production function

$$F(K_i, U_i, S_i) = Y_i = A_i A_{ir} \left[\left(a_K K_i^{\rho} + a_S S_i^{\rho} \right)^{\frac{\mu}{\rho}} + a_U U_i^{\mu} \right]^{\frac{1}{\mu}}, \qquad (1.1)$$

where output Y_i is produced using capital K_i, low-skilled labor U_i and high-skilled labor S_i as inputs. The nesting of the three input factors follows Krusell et al. (2000) to allow for differential adjustment of the two labor inputs to a change in capital costs. In particular, the elasticity of substitution between low-skilled labor and capital is $\frac{1}{1-\mu}$ and the elasticity of substitution between high-skilled labor and capital is $\frac{1}{1-\rho}$. Production also depends on a firm-specific production parameter A_i and a productivity shifter A_{ir}. Although the productivity shifter is firm-specific, it depends on aggregate outcomes in region r. I consider the behavior of A_{ir} in detail once I turn to the effect of spillovers.

Each firm chooses inputs according to the rental rate of capital r, wage w_U for low-skilled and wage w_S for high-skilled labor. There is fully elastic capital and labor supply which leads to equalization of input prices throughout the economy and firms take the input prices as given. The cost of capital still differs between firms since there is a firm-specific reduction through tax credits with rate τ_i. Firms set the product price p_i facing monopolistic competition with a downward

[16]The qualitative equilibrium results hold true nonetheless. An important difference is that in my model firms choose capital stock and not the investment rate as is the case when including adjustment costs.

sloping inverse demand curve

$$p_i = BY_i^{-\frac{1}{\eta^D}}, \tag{1.2}$$

where the price depends on the elasticity of demand $\eta^D > 1$ and a demand shifter B.[17]

The profit maximization problem for a firm is well-defined and the first-order conditions fully explain the production decisions of a firm. The first-order conditions are

$$(1-\tau_i)r = \left(1-\frac{1}{\eta^D}\right) Ba_K Y_i^{1-\mu-\frac{1}{\eta^D}} X_i^{\frac{\mu-\rho}{\rho}} K_i^{\rho-1}(A_i A_{ir})^{\mu} \tag{1.3}$$

$$w_S = \left(1-\frac{1}{\eta^D}\right) Ba_S Y_i^{1-\mu-\frac{1}{\eta^D}} X_i^{\frac{\mu-\rho}{\rho}} S_i^{\rho-1}(A_i A_{ir})^{\mu} \tag{1.4}$$

$$w_U = \left(1-\frac{1}{\eta^D}\right) Ba_U Y_i^{1-\mu-\frac{1}{\eta^D}} U_i^{\mu-1}(A_i A_{ir})^{\mu} \tag{1.5}$$

where $X_i = a_K K_i^{\rho} + a_S S_i^{\rho}$.

To show the impact of a change in the cost of capital through adjustments in the tax credit rate on optimal capital and labor, I totally differentiate the production function (1.1) and all first order conditions (1.3), (1.4) and (1.5). By reformulating the results, the price elasticity of capital and the cross-price elasticity of high-skilled and low-skilled labor respectively are

$$\begin{aligned}
e_{K_i} &= \underbrace{\left[\eta^D s_{K_i} + \frac{1}{1-\mu} s_{U_i} + \frac{1}{1-\rho} s_{S_i} - \left(\frac{1}{1-\mu} - \frac{1}{1-\rho}\right) s_{U_i} \frac{a_S S_i^{\rho}}{X_i}\right] Z}_{} + (\eta^D - 1) e_{A_{ir}} && (1.6)\\
e_{S_i} &= \left[\left(\eta^D - \frac{1}{1-\rho}\right) s_{K_i} + \left(\frac{1}{1-\mu} - \frac{1}{1-\rho}\right) s_{U_i} \frac{a_K K_i^{\rho}}{X_i}\right] Z + (\eta^D - 1) e_{A_{ir}} && (1.7)\\
e_{U_i} &= \underbrace{\left(\eta^D - \frac{1}{1-\mu}\right) s_{K_i} Z}_{\text{direct effect}} + \underbrace{(\eta^D - 1) e_{A_{ir}}}_{\text{indirect effect}}, && (1.8)
\end{aligned}$$

[17]Monopolistic competition leads to decreasing returns to scale. In this case, there is a unique interior solution. Other approaches introduce a fixed input factor or restrict full elasticity of labor supply.

where each elasticity $e_{\#} = \frac{d\#}{d\tau_i}\frac{1-\tau_i}{\#}$ is with respect to the net of tax rate,[18] $s_{K_i} = \frac{(1-\tau_i)r}{p_i}\frac{K_i}{Y_i}$ is the share of capital cost (after tax credit) in revenue, $s_{U_i} = \frac{w_U}{p_i}\frac{U_i}{Y_i}$ is the share of low-skilled labor cost in revenue, $s_{S_i} = \frac{w_S}{p_i}\frac{S_i}{Y_i}$ is the share of high-skilled labor cost in revenue and $Z = \frac{\eta^D}{\eta^D-1}$ is an additional scaling term.

Each elasticity consists of a firm-specific direct effect and an indirect regional effect. Turning first to the direct effect of tax credits, the elasticities mimic those in a model of firm production with just one labor type (e.g. Hamermesh, 1993). The elasticities of capital and labor largely depend on two effects, a scale effect from changes in input prices and a substitution effect between capital and the two types of labor. In the case of capital, these effects work in the same direction. An increase in tax credits leads to more demand for capital because of an expansion of production and a shift from labor towards capital. On the other hand, for each labor type the overall effect is ambiguous since there is higher demand from expansion in production but lower demand from the shift towards capital. The net effect depends on the relative magnitude of the elasticity of product demand and the elasticities of substitution. The effect can be different for each firm because of differences in the capital cost and labor shares. Because of monopolistic competition and the nesting of the CES production function, additional terms show up. First, less than fully elastic product demand amplifies adjustments of all inputs through the term Z. Second, for the elasticity of capital, there is an additional term that reweighs the impact of each elasticity of substitution. Third, for the elasticity of high-skilled labor with respect to the net of tax rate, there is a dependence on the elasticity of substitution of low-skilled labor. Independent of the parameter choices, the elasticity of capital is larger than the elasticity for either labor type.

[18] It is straightforward to show that this elasticity is equivalent to $-\frac{d\#}{dc_K}\frac{c_K}{\#}$, the capital cost elasticity.

To better judge the effect on the composition of labor inputs, it is useful to consider the elasticity of the ratio between high-skilled and low-skilled with respect to the net of tax rate. It is

$$\frac{d\frac{S_i}{U_i}}{d\tau_i}\frac{1-\tau_i}{\frac{S_i}{U_i}} = \left(\frac{1}{1-\mu} - \frac{1}{1-\rho}\right)\frac{a_K K_i^\rho}{X_i}. \tag{1.9}$$

The sign and magnitude of this elasticity depend on the relative magnitude of each elasticity of substitution. If the elasticity of low-skilled labor is higher than the one of high-skilled, there is capital-skill complementarity implying a shift towards high-skilled labor.

The indirect effect changes each elasticity of input by the same additive term consisting of the elasticity of regional productivity and the elasticity of product demand. Since the regional productivity depends on aggregate outcomes, a change for one firm will lead to adjustments for all firms in the same region. By redoing the maximization problem for a firm j, the elasticities with respect to the net of tax rate of firm i are

$$\frac{dK_j}{d\tau_i}\frac{1-\tau_i}{K_j} = \frac{dS_j}{d\tau_i}\frac{1-\tau_i}{S_j} = \frac{dU_j}{d\tau_i}\frac{1-\tau_i}{U_j} = (\eta^D - 1)\frac{dA_{jr}}{d\tau_i}\frac{1-\tau_i}{A_{jr}}, \tag{1.10}$$

which closely resembles the indirect effects of firm i.

The productivity shifter captures spillovers between firms. I follow the literature and assume that the productivity shifter depends on the overall economic activity within a region. I define

$$A_{ir} = \sum_{j \in S_r} Y_j^{\lambda_{ij}}, \tag{1.11}$$

where the set S_r contains all firms in region r and λ_{ij} is the elasticity of agglomeration between firm i and j.[19] This definition encompasses many of the characteristics of spillovers discussed in the literature.

[19] I use output as the measure for economic activity since the according elasticity is unambiguously greater or equal zero.

Aggregate output measures the degree of economic activity in a region, relating to the advantages found from clustering economic activity in proximity such as reduced transportation and communication costs in the supply chain of production, knowledge spillovers and thick labor markets (e.g. Moretti, 2011). These mechanisms suggest that firms may not profit equally from these advantages. By including firm-specific elasticities of agglomeration, I capture differences in the reliance of firms on local production networks. The measure of regional productivity also permits local multiplier effects as an alternative explanation for spillovers (Moretti, 2010). An increase in the employment within one industry may boost the demand for local goods and services and thereby impact employment in the non-tradable sector.

To derive intuitive closed-form solutions, I assume the same productivity shifter for all firms within a region by setting the elasticity of agglomeration $\lambda_{ij} = \lambda$. With this simplification, the elasticity of regional productivity with respect to a change in tax credit rate of firm i is

$$\frac{dA_r}{d\tau_i}\frac{1-\tau_i}{A_r} = \frac{\lambda\eta^D}{1-\lambda\eta^D}\frac{s_{K_i}Y_i^\lambda}{\sum Y_j^\lambda}Z > 0, \tag{1.12}$$

where I assume that $\lambda\eta^D < 1$.

The elasticity importantly depends on the interaction of the elasticity of agglomeration and the elasticity of product demand. Larger values for both elasticities imply a more pronounced impact on regional productivity. Furthermore, firms with a higher share of output in total regional output and larger capital cost shares impact regional productivity more. An initial increase in output of one firm due to tax credits will spread and lead to an unambiguous positive output effect due to spillovers in the whole region. The additional assumption on the magnitude of the elasticities rules out boundary cases in which spillover effects lead to infinite aggregate output.

If several firms in the same region experience a tax credit rate change, then the initial increase in aggregate output will be larger and will magnify spillover effects. Intuitively, the elasticity with respect to the tax credit rate of multiple firms comes about by sequentially calculating the equilibrium adjustments. Considering infinitesimally small tax rate changes, this reduces to summing up all elasticities of regional productivity for firms with changing capital costs. For notational simplicity, I consider the case where firms start out with the same tax credit rate $\tau_i = \tau$ and a subset receives the exact same tax credit rate change $d\tau_i = d\tau$. As result, the elasticity of regional productivity is

$$\frac{dA_r}{d\tau}\frac{\tau}{A_r} = \sum_{i|d\tau_i \neq 0} \frac{dA_r}{d\tau_i}\frac{1-\tau_i}{A_r} = \frac{\lambda\eta^D}{1-\lambda\eta^D} Z \frac{\sum_{i|d\tau_i\neq 0} s_{K_i} Y_i^\lambda}{\sum Y_j^\lambda}. \tag{1.13}$$

The elasticity is comparable to the one before. However, the adjustments of multiple firms lead to a summation of the output of all firms with tax credit rate change weighted by their share of capital costs. This means that there are larger spillover effects in regions where a tax credit rate change affects more firms.

In sum, the theoretical framework predicts the following. An increase in the tax credit rate leads to the use of more capital. The sign of the direct effect on labor depends importantly on the magnitude of the substitution effect between capital and the two labor types. Whether there is a shift towards high-skilled or low-skilled labor depends on the relative magnitude of both elasticities of substitution. On top of these direct effects, spillovers between firms lead to an additional positive effect on capital and employment. The spillover effects are larger in labor markets where more firms experience an increase in the tax credit rate. Finally, the proposed mechanisms for spillover effects suggest that the effect can vary firm by firm.

1.4 Estimation Strategy

The estimation strategy is guided by the theoretical framework and uses the described change of tax credit rates in 1999. I implement a difference-in-differences estimation approach and compare the adjustment of firms below and above the firm size cutoff before and after the change. I start by estimating average effects using the regression model

$$Y_{ibt} = \beta \text{Treated}_b \times \text{Post}_t + X_{it}^{'}\gamma + \psi_i + \psi_{nt} + \psi_{lt} + \epsilon_{ibt}, \qquad (1.14)$$

where Y_{ibt} is the outcome variable for each firm i with treatment b in year t. The variable Treated_b classifies firms into treatment and control group. I consider firms with up to 250 employees in 1998 as treated and those above as untreated to reflect the relative advantage for firms below the cutoff. The classification is fixed over time. I interact this variable with the dummy variable Post_t, which categorizes years 1999 to 2004 as treatment period to reflect the change in tax credit rate. ψ_i, ψ_{nt} and ψ_{lt} are firm, industry-year and labor market-year fixed effects, respectively. Firm fixed effects control for level differences in firm characteristics that stay constant over time such as those correlated with average firm size, industry and location. Industry-year and labor market-year fixed effects can control for the possibility that industry-specific shocks, and labor market-specific policies and economic developments coincide with the update of the tax credit program. I include various control variables $X_{it}^{'}$ depending on the specification and with log average firm wage being used for all main estimations.[20]

The coefficient of interest is β. Without confounding factors, it provides a causal estimate for the effect of the reduction of capital costs caused by the policy change on each outcome variable. The set of fixed effects already controls for many confounding factors. However, the firm size cutoff introduces additional firm incentives. Firms can adjust their size over time and thus are able to cross the cutoff. Such movements imply that these firms receive a different tax credit rate than assigned by the treatment status in the regressions. If this

[20] Using pretreatment wage growth trends instead of log wage leads to qualitatively the same results.

change in firm size is unrelated to the tax credit program, for example because of a general decline of a specific firm, this would not influence the causality claim. It would affect the interpretation of the coefficients though since it partly captures intent to treat effects. To minimize this issue, I exclude firms that are close to the cutoff in 1998 and thus those that are more likely to cross the cutoff in either direction. On the other hand, firms may intentionally move just below the cutoff or delay moving above to take advantage of the higher tax credit rate. To prevent biases of such behavior in the estimation, I further exclude observations in any year for which firm size is close to the cutoff. Since the exclusion of particular observations is directly connected to employment of a firm, it will bias the estimations with employment as outcome variable. In this case I select firms only based on their firm size in 1998.

I study several outcomes. First, I am interested in the effect on capital inputs. As is common in the literature, I use investment as a directly measured variable in the dataset to proxy capital adjustment. I consider the intensive margin, extensive margin and a combined measure. I also specifically consider equipment investment. In a second step, I focus on labor inputs and analyze the effect on the number of total employees and for various subcategories separately. To assess the impact of the policy on the skill composition of a firm's workforce, I also use employment ratios by education and occupation as an outcome variable.

The error term ϵ_{ibt} includes all other omitted factors. I cluster standard errors at the regional level allowing for heteroscedasticity and arbitrary correlation between firms within the same region over time. I consider German *Landkreise* as regions in my analysis. There are 76 regions in East Germany.[21] The regions are then divided into 56 labor markets following a classification used by Dustmann and Glitz (2015).

[21] Various regions merged due to reforms at the state level. I use the regional disaggregation as of 2014 to ensure consistency over time.

Since the average effect can mask interesting adjustment patterns, I estimate the dynamic regression model

$$Y_{it} = \sum_{p=1995}^{1997} \delta_p D_{bp} + \sum_{p=1999}^{2004} \delta_p D_{bp} + X_{it}^{'}\gamma + \psi_i + \psi_{nt} + \psi_{lt} + \epsilon_{ibt} \quad (1.15)$$

where in comparison to specification (1.14) the interaction term is omitted and instead a set of yearly dummies D_{bp} is introduced. Each dummy variable assigns a value of one to firms in the treatment group for the corresponding year p and zero otherwise. The coefficients δ_p for the years 1999 to 2004 capture the dynamic treatment effect. If the treatment effect is indeed causal, then treatment and control group have parallel trends absent the policy change. This would not be the case if anticipation effects from the announcement of the policy or long-term influences from similar policy changes before 1995 influence firm behavior. For this reason, I also examine the pre-treatment effect by including dummies for the years 1995 to 1997. Observing statistically insignificant estimates close to zero before the start of the treatment provides an indication that the identifying parallel trends assumption indeed holds true.

Both regression approaches so far only consider the direct effect of the policy change. For the estimation of spillover effects, I use the regression model

$$\begin{aligned} Y_{iblt} = \beta \text{Treated}_b \times \text{Post}_t + \alpha \text{ShareBelow250}_{l,98} \times \text{Post}_t \\ + X_{it}^{'}\gamma + \psi_i + \psi_t + \epsilon_{iblt}, \quad (1.16) \end{aligned}$$

where $\text{ShareBelow250}_{l,98}$ is the share of employees working in firms with up to 250 employees in labor market l for year 1998. The variable is interacted with the treatment period dummy Post_t to set up a difference-in-differences estimation with continuous treatment status. There are now two coefficients of interest, β, the direct effect of receiving a higher tax credit rate as in specification (1.14) and α. The latter coefficient provides an estimate of the difference in the outcome variable by the share of firms below the cutoff that is due to the change in tax credit rate. This setup mimics the results of the theoretical framework with α corresponding to an estimate of the spillover effects. Since I use variation at labor market level, I cannot control for the same set of fixed effects as before. Instead, I include firm and year fixed effects in the baseline and add industry-year, area (federal state)-year and regional pre-treatment growth trends as robustness tests.

The analysis of dynamic effects is again helpful to better understand adjustment behavior and to check the parallel trends assumption for the estimation of spillovers. The regression model is

$$\begin{aligned} Y_{iblt} = &\sum_{p=1995}^{1997} \delta_p D_{bp} + \sum_{p=1999}^{2004} \delta_p D_{bp} + \sum_{p=1995}^{1997} \theta_p \text{ShareBelow250}_{lp,98} \\ &+ \sum_{p=1999}^{2004} \theta_p \text{ShareBelow250}_{lp,98} + X_{it}^{'}\gamma + \psi_i + \psi_t + \epsilon_{iblt}, \end{aligned} \quad (1.17)$$

where $\text{ShareBelow250}_{lp,98}$ are a set of variables measuring the share of employees working in firms with up to 250 employees in labor market l in year 1998. Each variable takes on this value in year p and is zero otherwise. The set of coefficients δ_p still estimates the direct impact of the policy change over time. The set of coefficients θ_p estimates the dynamic effect of spillovers over time. If the estimation for spillovers is causal, then there should not be any differential effect on firms between labor markets before the policy change. I check for this assumption by including coefficients for pre-treatment periods.

1.5 Data

1.5.1 Data Sources

The empirical analysis relies on two data sets, the AFID Establishment-Panel by the Federal Statistical Office of Germany and the Establishment History Panel (BHP) by the Institute for Employment Research (IAB).

The AFID dataset has a broad coverage of variables for investment, employment and output for the universe of manufacturing and mining firms with more than 20 employees in Germany. With its unusual richness it perfectly fits the needs for the general empirical analysis. Firm variables are collected through various administrative surveys and are used to inform the government and the public about key economic statistics like aggregate output and investment. Because of the importance of these statistics, firms are required by law to provide truthful information. The AFID dataset merges the underlying surveys through a unique firm identifier and aggregates information from monthly and quarterly surveys by year. Information is available since 1995 and new waves are continuously added. The dataset is especially suited for the investment analysis since there is separate information for equipment. Further subcategories distinguish different modes of acquisition such as self-production, leasing and purchase. For measures of output, there are revenue, production value, orders and the number of distinct products. Revenue is divided into domestic and foreign. There is however only limited information on labor inputs, with total employment and wage bill being most informative.[22] For each firm the 4-digit industry code and location at regional level is provided as well.

For a more detailed employment analysis, I use the BHP dataset which provides information on overall employment, employee composition, employee inflows and outflows, and wages. The data are based on the employment histories of the entire labor force covered by social security. They are collected from mandatory communication between firms and the Federal Employment Agency on changes in employment. The BHP aggregates this information at firm level for 30 June of every year for West Germany since 1975 and for East Germany since 1992. I focus on the years 1995 and 2004. The final

[22]There is a distinction of employees by contract type. I do not use this information since it does not translate well to other economic concepts.

sample consists of a 50% random draw of firms and all available years of selected firms are included. For employment there are counts for the total and for subcategories of education, occupation types and age and full-time employment and vocational training. For a better understanding of changes over time, inflow and outflow information provides the number of employees that did not work in the same firm one year before and one year after, respectively. These flows are again divided into subcategories. I use information for incoming and leaving employees that were unemployed one year before and after, respectively. Average wage is based on full-time employees and available for quartiles and by education. The dataset includes firms of all industries. Information at 3-digit industry level and a region variable allows the selection of relevant firms.

The AFID and BHP dataset are distinct and it is not possible to link them. Therefore, it is necessary to calculate key policy variables separately and take into account changes in the data collection over time independently. One important variable in the analysis is firm size, since it is the basis for classifying treatment and control group. The program definition requires information on the head count of overall firm employment and vocational trainees.[23] The AFID dataset lacks information on the latter. To address this issue, I match the number of vocational trainees at firm level for the years 1999 to 2001 from the cost structural panel (KSE) by the Federal Statistical Office.[24] For observations that are unmatched, I impute values assuming a constant share of trainees within firms or if unknown within industries. For the BHP dataset, on the other hand, information on the marginally employed is missing for years before 1999. I impute missing observations by assuming a constant share

[23]The definition considers the business year start for calculations. Employment figures in the AFID and BHP dataset do not have this same timing, however, an auxiliary analysis using the IAB establishment survey does not show a systematic difference of employment levels in manufacturing firms within a given year.

[24]The KSE is a yearly firm survey of a stratified random sample in the manufacturing and mining industry and focuses on the production process.

of marginally employed within firms or if unknown predict the share within industries for different firm size.[25] The change in reporting of marginally employed unfortunately coincides with the policy change. To reduce the risk that the imputation of marginally employed influences the estimation, I drop firms with an average share of more than 25%, which is above the 95$^{\text{th}}$ percentile within the manufacturing industry.

Another issue concerns the continuity of firm identifiers. In the AFID data, firm identifiers are constant even if the ownership structure of a firm changes, in the BHP this is not the case. However, in the BHP the firm identifier can change for relatively simple reasons such as changes in the legal structure. In the analysis, firm structure possibly influences decision-making processes and thus leading to abrupt changes in production behavior. To have firm identifiers that exclude such changes, I separate firm identifiers in the AFID dataset when there are changes in overall firm structure and I exclude firms with more than one-hundred establishments in a given year. For the BHP, I follow Hethey-Maier and Schmieder (2013) and create unique identifiers for firms that are connected through employee flows.

Further adjustments are the reclassification of regions as of 2014, the classification of regions into labor markets according to Dustmann and Glitz (2015) and a classification of 2-digit industries that further aggregates uncommon industries.

1.5.2 Sample Selection

The sample selection follows the eligibility criteria of the program. First, I select firms in manufacturing industries[26] for years 1995 to

[25]Since I rely on broad firm size intervals in the analysis, and vocational trainees and marginally employed constitute a small share of a firm, measurement error from both imputations should lead to relatively few misclassifications into treatment and control group.

[26]These include all industries with a WZ 1993 classification of D.

2004 located in East Germany excluding Berlin. To avoid peculiar behavior of entering and exiting firms, I also condition on them being economically active throughout the period of the analysis.

Second, I check for bunching behavior around the cutoff. Figure 1.4 displays the size distribution of manufacturing firms in East Germany and West Germany around the cutoff for 1999 to 2004. For West Germany a decrease in the density with increasing firm size is apparent which is as expected (e.g. Axtell, 2001). For East Germany this pattern is generally true as well, however, just below the cutoff there is excess mass. This points to bunching of East German manufacturing firms. I therefore exclude all firms with a size of between 226 and 274 employees in 1998.[27] I further exclude observations within that size interval in any other year.[28]

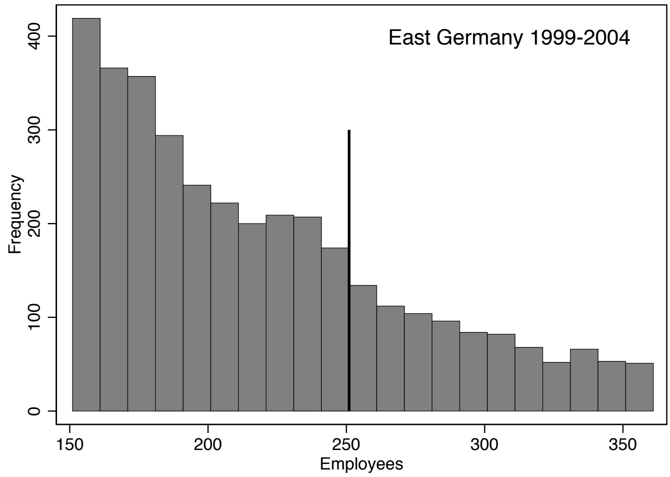
As a last step, I include only firms with at least 40 employees and a maximum of 1,500 in 1998 and observations that lie within the same interval for any other year. This reduces the problem from biases due to heterogeneous effects among observationally different firms (Heckman et al., 1997).[29] To check whether the choice of the excluded interval and the size interval has an impact on estimation results, I run robustness test that vary these interval boundaries.

Figure 1.5 considers the impact of the exclusion around the cutoff on firms moving outside their assigned treatment status. As a comparison, I include the case without restriction. The specific sample selection has little impact on the treatment group. At most 3.5% of firms move above the cutoff and once observations around the cutoff are excluded this share reduces to a maximum of 1.5%. There is more movement within the control group. Without the exclusion of firms

[27]I determine this cutoff by implementing a structural approach that adapts ideas from Garicano et al. (2016). The maximum likelihood estimation leads to a value of 274.6. Implementation details are available upon request.

[28]The exclusion of single observations is not appropriate for estimating the effect on employment since then there would be a selection based on the dependent variable. In this case, I only condition on firm size in 1998.

[29]I implemented a propensity score matching approach between treatment and control group and found qualitatively similar results to those in the main text.



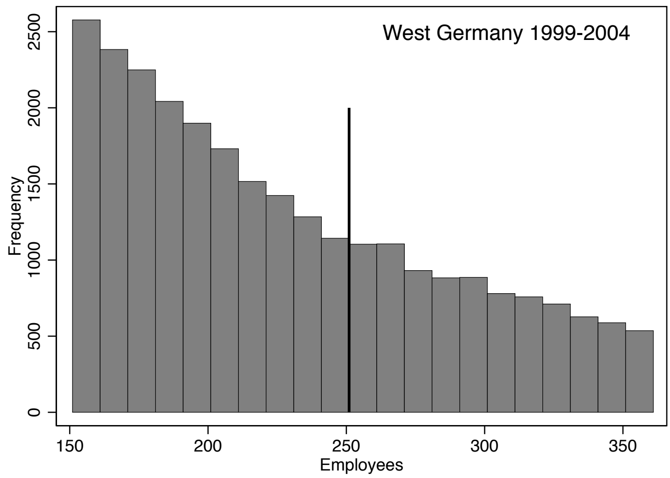


Figure 1.4: Firm Size Distribution of Manufacturing Industry

Note. Data from AFID establishment panel. The sample consist of all observations of manufacturing firms for East and West Germany excluding Berlin between 1999 and 2004. Firm size is based on total number of all employees but excludes vocational trainees.

around the cutoff, 8.6% of firms in the control group fall below the cutoff already in 1999. In 2003 this share is at 26.4%. It is likely that some of these firms move on purpose to take advantage of higher tax credit rates. For the sample with excluded bandwidth, the share of firms moving below is considerably lower. After the policy change the share increases slowly over time, reaching 6.8% in 2001 and is highest in 2004 with 16.9%. Even though the share in 2004 is non-negligible, in this case, a movement below the cutoff is less problematic since these observations are not affected by bunching behavior.

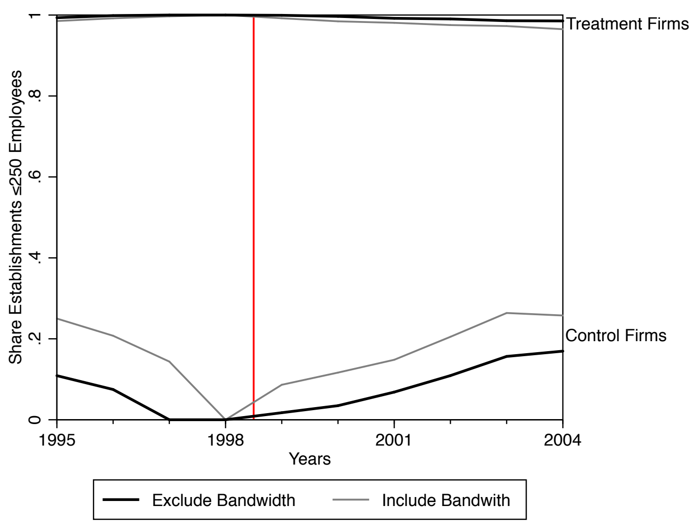


Figure 1.5: Movement of Treatment and Control Group Around Firm Size Cutoff

Note: Data from AFID establishment panel. The figure shows the share of firms in the treatment and control group below the firm size cutoff. Firm size is defined as total employee head count minus vocational trainees. The treatment group consists of firms with 250 or fewer employees. The sample selection for 'Exclude Bandwidth' is according to main text and specifies the exclusion of firms and observations around the firm size cutoff. The case of 'Include Bandwidth' follows the same sample selection, but keeps firms and observations around the size cutoff.

1.5.3 Descriptive Statistics

Table 1.1 presents a selection of firm variables from the AFID dataset in Panel A and the BHP dataset in Panel B for the years 1995 to 2004. I show descriptive statistics for all manufacturing firms in West and East Germany (excluding Berlin), and for treatment and control group of the empirical analysis. Overall, firms in West Germany are larger in many respects compared to those in East Germany. The average number of employees in the AFID dataset is 155.19 in West Germany and 89.42 in East Germany. They have the same likelihood of investing in any given year but in West Germany the investment value is larger. These differences in input factors translate to higher revenue with €30.75 million compared to €13.03 million. Panel B reports very similar differences for the number of employees. On top, it shows that full-time employees earn considerably more in West Germany. In terms of employee composition, there are actually more employees with college degree, more high-skilled occupations and more vocational trainees in East German manufacturing firms.

Turning to the estimation sample, there are again clear differences in size. This is not surprising given the definition of treatment and control group. The treatment group is similar to the average East German firm when comparing means, but the control group is far larger in every respect.[30] The groups have similar employee composition. For example the share of college graduates is 11.16% in the treatment group compared to 14.82% in the control group.

As a step towards the actual estimation, Figure 1.6 presents the raw means for treatment and control group over time for different investment and employment outcomes. Encouragingly, in each plot treatment and control group have a similar pre-treatment trend. In the upper left panel, there is a continuous decrease in the log of total investment that continues after the policy change in 1999. For subsequent years the investment level stays higher for the treatment

[30] This does not mean that there is no overlap. Standard deviations are usually large, especially for investment and employment composition.

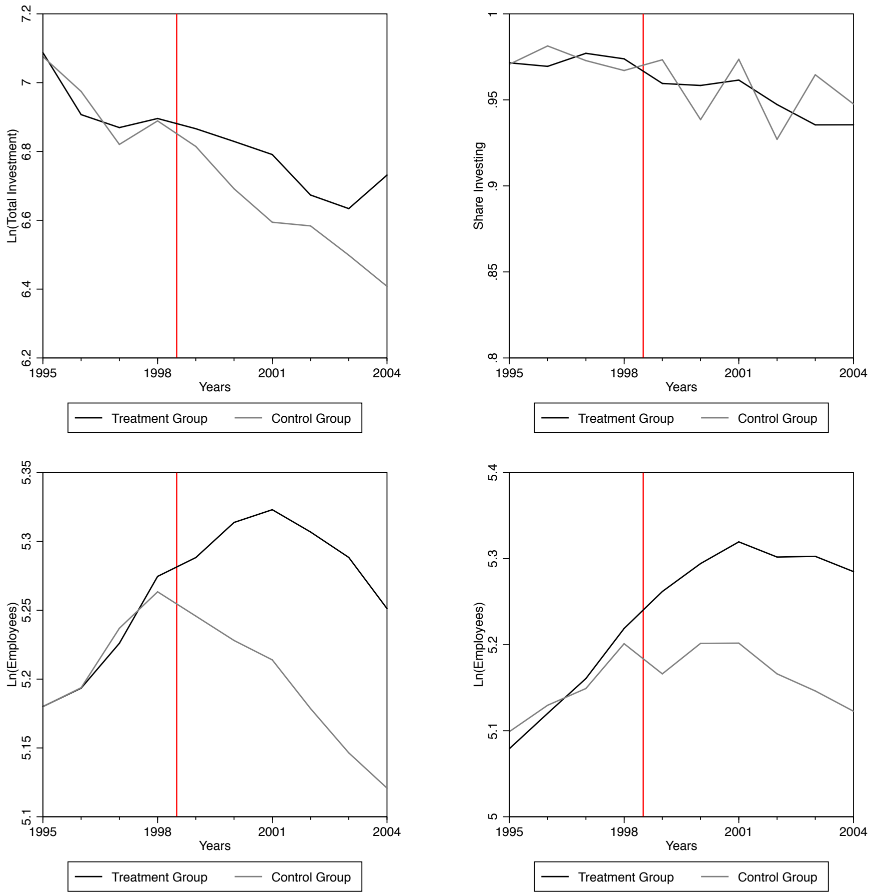


Figure 1.6: Average Firm Investment and Employment By Treatment Status

Note. The figures report raw means for each outcome by treatment and control group over the period 1995 to 2004. Outcome variables are the log of total investment in the upper left panel, share of firms with positive total investment in the upper right panel, log of employment based on AFID data in the lower left panel and log of employment based on BHP data in lower right panel. For each data point, I subtract the group mean for the pre-treatment period (1995–1998) and add the pooled mean for the same period for facilitating a comparison of trends.

Table 1.1: Descriptive Statistics for Manufacturing Firms

Panel A. AFID Data

	West Germany	East Germany		
	All	All	Treatment	Control
No. of Employees (mean)	155.19	89.42	91.88	508.19
No. of Employees (median)	57	47	75	414
Investing (%)	86.0	86.0	94.0	97.0
Investments (thousand EUR)	1,267.50	1,172.35	805.96	8,605.11
Revenue (million EUR)	30.75	13.03	11.39	101.73
Part of Multi-Establishment Firm	22.0	22.0	9.98	15.76
Observations	357,662	64,154	16,374	1,266

Panel B. BHP Data

	West Germany	East Germany		
	All	All	Treatment	Control
No. of Employees (mean)	129.24	81.09	89.89	468.66
No. of Employees (median)	47	41	72	404.5
Full-time employees (mean)	111.63	70.95	80.56	424.02
Average Daily Wage Full-time	85.86	58.16	59.54	78.23
Share College Degree (%)	6.43	10.58	11.27	14.70
Share High-Skilled Occupation (%)	12.63	13.06	13.08	16.87
Share Vocational Trainees (%)	3.89	5.30	4.33	4.13
Observations	212,924	37,239	9,180	1,000

Note: Statistics are based on firms in the manufacturing sector for the years 1995–2004 excluding Berlin. The number of observations is based on the according statistic for number of employees. The group of all manufacturing firms includes those with more than 20 employees to allow for a better comparison between AFID and BHP data. Treatment and control group are according to the estimation sample.

group reflecting a positive investment response. The decision to invest, shown in the upper right panel, is relatively stable with a value close to 100% and there is little differential movement before or after 1999. I plot log employment using AFID data in the lower left panel and BHP data in the lower right panel. The evolution in both graphs is quite similar. This is remarkable given that the actual data collection was independent of each other and speaks to the quality of both datasets. Until 1999, employment increases for the treatment and control group. Subsequently, employment stays constant or decreases in the control group whereas there is continued growth for the treatment group until 2001. After 2001, employment decreases for the treatment group as well at similar levels as the control group. This pattern indicates a positive employment response to the relative reduction in capital costs.

1.6 Results

1.6.1 Investment

As a first outcome, I study firm investment. Table 1.2 summarizes the average effect for various investment measures. In column (1), I consider the log of total investment at the intensive margin and find that the policy change leads to 23.4 log points higher investment for the treatment group compared to the control group. Since the policy change only affected tax credits on equipment investment, column (2) reports the estimate for this subcategory. The estimate is slightly higher with 25.1 log points and statistically significant. Both outcomes measure only the intensive margin. In column (3), I check for differences in the probability of investing. However, there does not seem to be any response with a point estimate of zero and small standard errors. This is likely the case because of the high rate of firms that invest in any given year independent of treatment. In the literature, one outcome of interest is the investment rate (I_t/K_{t-1}) which combines intensive and extensive margin. I do not observe capital in my datasets. As proxy, I consider investment divided by the average total investment during the period of analysis. I find a positive and statistically significant response of 0.171 for total investment and 0.167 for equipment investment.

Taken together, these results show a positive investment response to tax credits. To compare the effect to previous findings in the literature, I calculate the elasticity with respect to capital costs. The change in capital costs for treatment and control group is equal to $\Delta t_i/(1-t_i)$. Taking into account the changes in tax credit rate in 1999 and 2000, and assuming that all investments are modernization investments, capital costs decreased by 15.74% for the treatment group

Table 1.2: Difference-in-Differences – Investment

	Dependent Variable: Log Total Investment (1)	Log Equipment Investment (2)	Investing (1/0) (3)
Direct Effect	0.234**	0.251***	-0.000
	(0.089)	(0.092)	(0.008)
Observations	15,275	15,071	15,900

Note. Data from AFID establishment panel. Each coefficient is estimated from different regression following main specification (1.14). The dependent variables are the log of total investment in column (1), log of equipment investment in column (2), a dummy for having positive total investment in column (3), total investment over the average total investment between 1995–2004 in column (4) and equipment investment over the average total investment between 1995–2004 in column (5). Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. Standard errors in parentheses are clustered at the regional level: * p<0.10, ** p<0.05, *** p<0.01

and 7.46% for the control group. This leads to an intensive margin elasticity of total investment of 2.825.[31] The consensus range for the elasticity of investment proposed by Hassett and Hubbard (2002) is 0.5 to 1.0 although recent studies by House and Shapiro (2008), Maffini et al. (2016) and Zwick and Mahon (2017) find much larger elasticities of around 8. Zwick and Mahon (2017) provide evidence that smaller firms have larger elasticities which could explain the rather large elasticity in my case as well.

Table 1.A1 in the appendix considers a few robustness tests for the investment specification. Column (1) reproduces the results from before. Given that investment is highly volatile, in column (2), I consider a sample where I exclude firms with investment growth in 1997 above the 95$^{\text{th}}$ percentile. The response is slightly lower for log of total investment and log of equipment investment but there is no

[31]For cases where there are structures and replacement investments, the calculated elasticity is a lower bound. The extensive margin elasticity of total investment with respect to the tax credit rate is zero.

response for the probability of investing. In column (3), I only control for average firm wage, firm fixed effects and year fixed effects. The coefficients are again smaller but lead to qualitatively similar results. Finally, in column (4), I only consider single-establishment firms to eliminate inconsistencies between tax credit eligibility and the level at which production decisions are taken. I again exclude firms with high volatility in their investments. In this case, the average effect becomes slightly larger although there is still no response in the extensive margin. Overall, the investment response is similar throughout distinct specifications and for different sample selections.

For a better understanding of investment behavior over time, I study the dynamic specification. Figure 1.7 presents the coefficients and their 95% confidence intervals for the log of investment, the log of equipment investment and the probability of investing. For both measures of the intensive margin, there is an upturn directly in the year after the policy change. In the subsequent year, investment further increases with coefficients of 20.6 log points for total investment and 23.2 log point for equipment investment. Afterwards, there is notable fluctuation in the effect size with a short period of smaller coefficients followed by increases in 2003 and 2004. Standard errors are relatively large so that not all coefficients after the policy change are statistically significant. For pretreatment periods, coefficients are close to zero which suggests that treatment and control group follow the same trend. For the extensive margin, there is not any clear dynamic pattern. Coefficients move around zero for periods before and after the policy change leading to the zero result on average.

1.6.2 Employment

Given the positive response in investment, I then study adjustments in the use of labor inputs. Table 1.3 presents the estimation results for the effect of the tax credit rate change on employment. For the regression of column (1), I use employment information from the AFID dataset. I find a positive and statistically significant effect of 11.3 log

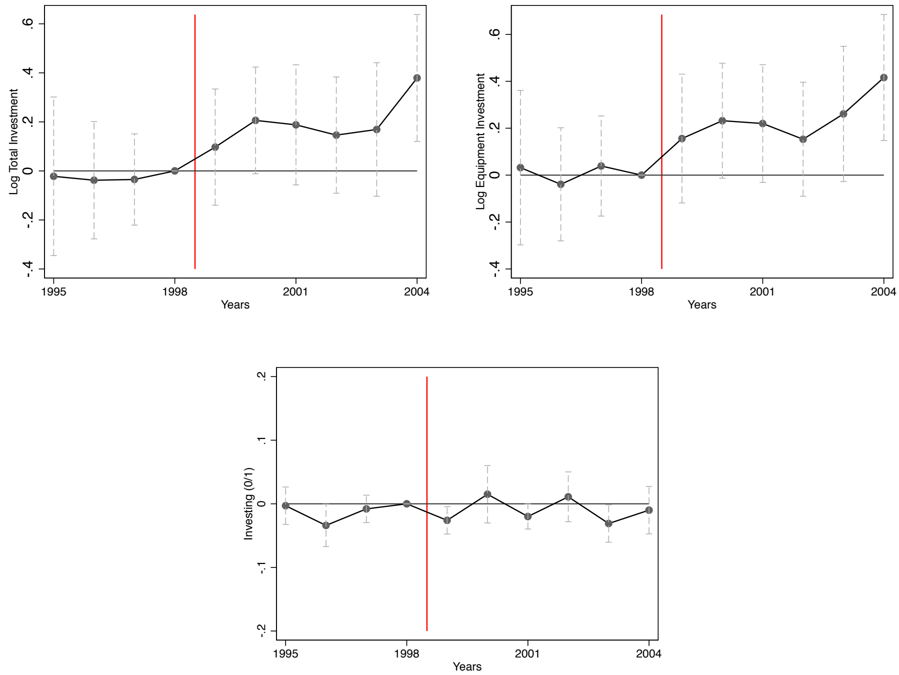


Figure 1.7: Dynamic Effect of Policy Change on Investment Behavior

Note. Coefficients from estimation of dynamic specification (1.15). The dependent variables are the log of total investment in the upper panel, log of equipment investment in the middle panel and a dummy for having positive total investment in the lower panel. Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. Firms with volatile investment, measured by growth of total investment in 1997 above the 95^{th} percentile, are excluded. 95% confidence intervals are displayed for each coefficient using clustered standard errors at the regional level.

points on total employment. The response is smaller than for investment but still sizable. In column (2), I consider the same employment measure based on the BHP data. In this case the coefficient is 8.7 log points which is similar in magnitude and confirms the positive employment response. The corresponding elasticity is 1.051. To analyze whether the employment effect applies to different subgroups of workers, in column (3) to (5), I consider regular employees, full-time

Table 1.3: Difference-in-Differences – Employment

	Dependent Variable: Log Employment of				
	Total (AFID)	Total (BHP)	Regular	Full-Time	Full-Time-Regular
	(1)	(2)	(3)	(4)	(5)
Direct Effect	0.113***	0.087**	0.076**	0.088**	0.087**
	(0.031)	(0.037)	(0.036)	(0.038)	(0.038)
Observations	17,637	10,116	10,116	10,116	10,116

Note. Each coefficient is estimated from different regression following main specification (1.14). The dependent variables are the log of total employment in column (1) and column (2), log of regular employees in column (3), log of full-time employees in column (4) and log of regular full-time employees in column (5). Data from AFID in column (1) and BHP in column (2)-(5). Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. Standard errors in parentheses are clustered at the regional level: * p<0.10, ** p<0.05, *** p<0.01

employees and full-time regular employees.[32] The average effects are 7.6 log points on regular employees, 8.8 log points on full-time employees and 8.7 log points on full-time regular employees which is very close to the effect on total employment.

Figure 1.8 presents the according dynamic effects for total employment and full-time employment. For the case of the AFID data, there is a continuous increase in total employment during the treatment period. In the first year after the policy change employment is 3.3 log points higher. After three years it reaches 9.9 log points and then levels off. For the case of BHP data, the effect on total employment is more immediate. After one year, employment is 7.7 log points higher. Subsequently, there is a small drop, that is followed by a slow increase over time reaching 10.9 log points in 2004. When considering only full-time employees, there is again an immediate re-

[32]Since these measures mostly exclude marginally employed workers, using them as outcome variable can check whether reporting issues of the marginally employed in the BHP dataset are driving the results on total employment.

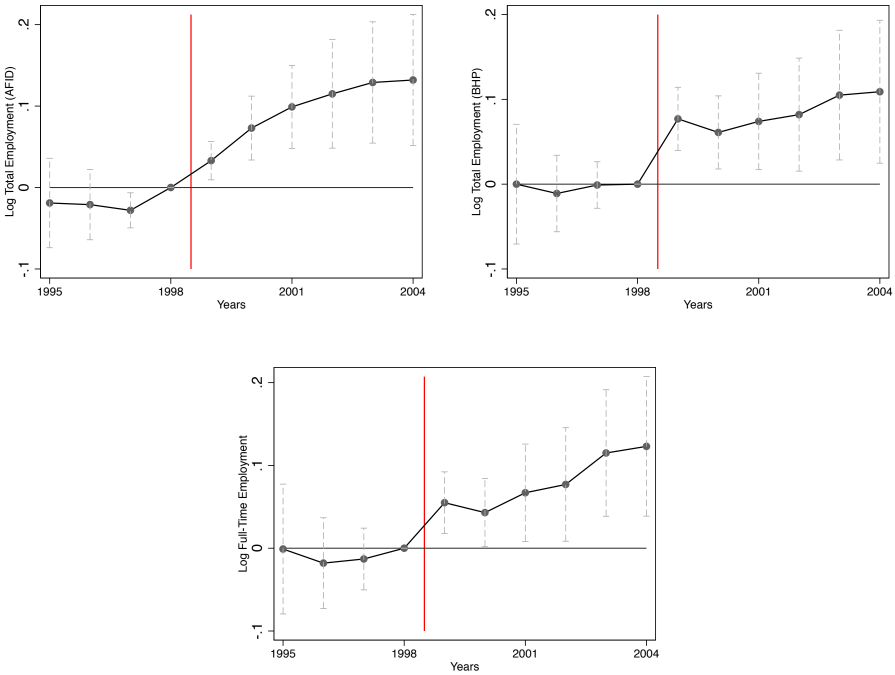


Figure 1.8: Dynamic Effect of Policy Change on Firm Employment

Note. Coefficients from estimation of dynamic specification (1.15). The dependent variables are the log of total employment. Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. 95% confidence intervals are displayed for each coefficient using clustered standard errors at the regional level.

sponse of 5.5 log points. Even though there is again a drop in the subsequent year, there is a stronger increase in the effect over time, reaching 12.3 log points in 2004. When checking for parallel trends in employment for pretreatment periods, coefficients are again close to zero and mostly statistically insignificant. For total employment with the AFID data there seems to be some movement already in 1998, however, it is small in magnitude.

For these employment regressions, I exclude firms that are close to the firm size cutoff in 1998, but I do not exclude observations for firms that moved close to the cutoff in prior or subsequent years. Thus,

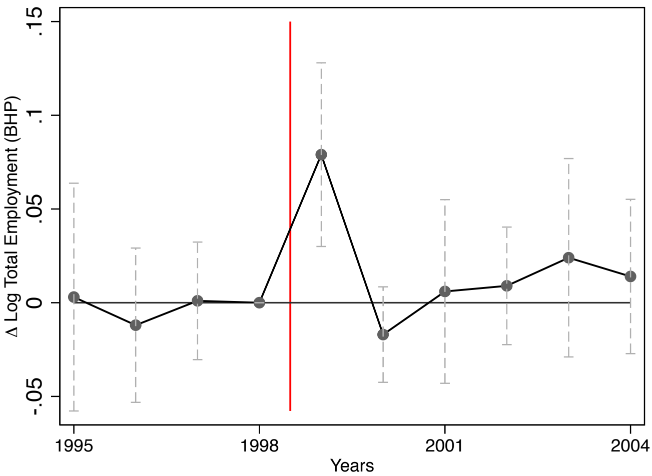


Figure 1.9: Dynamic Effect of Policy Change on Firm Employment Growth

Note. Coefficients from estimation of dynamic specification (1.15). The dependent variables are the log of total employment. Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. 95% confidence intervals are displayed for each coefficient using clustered standard errors at the regional level.

my estimates potentially pick up bunching at the cutoff. As robustness test, I estimate the dynamic specification with the yearly change in the log of total employment as outcome variable and present the results using BHP data in Figure 1.9. For this regression, I exclude firms around the cutoff in 1998 and observations close to the cutoff in any other year, thereby eliminating observations with bunching. The effect is close to zero and statistically insignificant for all years except 1999. For 1999, the effect on the log growth rate is 7.9 log points. Even though statistically insignificant, the coefficients in subsequent years are still in line with the level results suggesting a decrease in employment in 2000 and small positive growth over time in subsequent years.

1.6.3 Revenue

In the previous subsections, I find that firms increase investments and employment due to tax credits. A natural extension to these results is the analysis of output. Table 1.4 provides estimation results for various revenue measures and the number of distinct products as proxy for output.[33] In column (1), I use total revenue as outcome variable and find a small effect of 1.1 log points. However, the effect on domestic revenue in column (2) is 8.3 log points and on domestic revenue of manufacturing-specific goods in column (3) is 8.0 log points. This is puzzling since this implies a reduction in exports to compensate for changes in domestic production, although exports equal a small share of production on average. I therefore consider the possibility that the results for total revenue are driven by outliers among exporting firms. In column (4), I restrict the sample to firms with an average export share of not more than 15%. In this case, the effect is 7.6 log points which is close to the effect on domestic revenue (though statistically insignificant). Thus, the discrepancy between the effect on total revenue and domestic revenue seems to be confined to those firms that already export sizable amounts. As additional outcome, I include the number of distinct products in column (5) but I do not find an effect of tax credits on the number of products.[34]

1.6.4 Robustness Tests

Taken together, these results provide a positive assessment for investment tax credits. Not only do they increase firm investment, but they also lead to more employment overall, and thereby to higher revenue. The results are all based on a specific selection of the firm size sample. I check the robustness of these results to the exclusion of different firm size intervals around the cutoff and the selection of

[33] The revenue measures will elicit the same effects if there is no change in output prices.

[34] It is possible that this result reflects the introduction of novel products and a discontinuation of old ones at the same time.

Table 1.4: Difference-in-Differences – Effect on Output

	Dependent Variable:				
	Log Revenue	Log Domestic Revenue	Log Domestic Manuf Revenue	Log Revenue (low exporting)	Log Products
	(1)	(2)	(3)	(4)	(5)
Direct Effect	0.011	0.083**	0.080*	0.076	0.022
	(0.037)	(0.038)	(0.040)	(0.050)	(0.032)
Obs	15,906	15,898	15,897	11,689	15,547

Note. Each coefficient is estimated from different regression following Specification 1.14. The dependent variables are log of total revenue in column (1), log of domestic revenue in column (2), log of domestic revenue for manufacturing output in column (3), log of revenue in column (4) and log of the number of distinct products in column (5). The sample in column (4) is conditional on an exporting share of below 0.15. Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. Standard errors in parentheses are clustered at the regional level: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

different lower and upper firm size bounds. Additionally, I implement a placebo test by estimating the effect on the sample of the treated firms and selecting several policy irrelevant cutoffs.

Table 1.A2 in the appendix reports coefficients for the most relevant measures of investment, labor and revenue. The interval of excluded firms around the cutoff varies in each column. In column (1), I do not exclude any firm. Column (3) follows the main specification with an excluded interval of 225 to 275 employees. In addition, I report an intermediate case in column (2) and larger intervals in column (4) and (5). Reassuringly, estimates for all outcomes are robust throughout the different intervals. This applies especially to log overall investment, the extensive margin investment decision and employment. For log equipment investment there are slight differences although there is no clear pattern as a function of the size of the excluded interval. For log domestic production, coefficients grow somewhat by the size of the excluded interval, but qualitatively, the results are the same. Taken together, there is little evidence that the exclusion of firms around the cutoff has a notable impact, possi-

bly because the number of firms that bunch at the cutoff is small in comparison to the whole sample.

Table 1.A3 keeps the excluded bandwidth constant but instead varies the smallest and largest included firm size. In this table, column (4) reproduces the main specification and columns to the left use smaller intervals whereas columns to the right use larger intervals. Again, independent of the chosen interval, estimates lead to qualitatively similar results with limited fluctuations in the main coefficient of interest. The largest differences apply to the investment variables for which the coefficients in column (1) are somewhat smaller and those in column (6) larger than the rest. It should be noted that for the smallest firm interval, the number of observations is considerably smaller which may lead to higher susceptibility to outliers.

Table 1.A4 reports the estimation result for the placebo cutoffs. The sample only consists of firms in the treatment group. I focus on the cutoffs 80, 100 and 125 to have a reasonable number of firms below and above these new cutoffs. I still exclude firms around the cutoffs as in the main specification. For the majority of outcomes, coefficients are close to zero and statistically insignificant. This holds true for all placebo cutoffs. For the extensive margin investment decision, coefficients are negative and statistically significant. However, given the high probability of firms investing in any given year, the coefficient is not economically significant.

1.6.5 Spillover Effects

All results so far speak to the direct effect of investment tax credits. In this subsection, I exploit the labor market-level variation in the number of employees working in firms below the cutoff. I start with the analysis of specification (1.16) where the treatment intensity is equal for all firms in the same labor market. Since this specification estimates direct and indirect effects simultaneously, I still exclude firms close to the cutoff in 1998. However, compared to previous estimations, I additionally include firms with 20-40 employees to have a larger sample size for more precision in the spillover analysis. The outcome variable is the log of total employment in all estimations.

Table 1.5 reports the according results. In the baseline in column (1), the direct effect on treated firms is 10.7 log points which is close to the previously estimated employment effect. The coefficient for the spillover effect is 0.118. Both the direct and indirect effect is stable

Table 1.5: Spillover Effects on Labor Inputs

	Dependent Variable: Log Total Employment				
	(1)	(2)	(3)	(4)	(5)
Direct Effect	0.107***	0.093**	0.112***	0.113***	0.101***
	(0.037)	(0.036)	(0.038)	(0.038)	(0.038)
Regional Share Firms Below Cutoff	0.118	0.116*	0.126*	0.117*	0.120**
	(0.075)	(0.067)	(0.072)	(0.062)	(0.056)
Observations	19,328	19,328	19,328	19,328	19,328
Controls					
Growth trends	-	X	-	-	X
Industry-Year FE	-	-	X	X	X
Area-Year FE	-	-	-	X	X

Note. Each column is estimated from different regression following specification 1.16. The dependent variable is log total employment. Additional controls are log average firm wage, firm and year fixed effects, and the controls specified in each column. The sample includes firms within firm size of [20,225],[275,1500] in 1998. Standard errors in parentheses are clustered at the regional level: * p<0.10, ** p<0.05, *** p<0.01

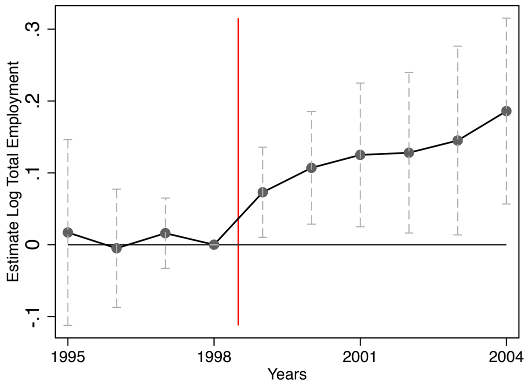


Figure 1.10: Dynamic Effect of Spillovers on Labor Inputs

Note. Coefficients θ_p from estimation of dynamic specification (1.17). The dependent variable is the log of total employment. The firm size interval is [20,225],[275,1500]. Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and area (Bundesland)-year fixed effects. 95% confidence intervals are displayed for each coefficient using clustered standard errors at the regional level.

to the inclusion of additional control variables. I include pretreatment growth trends in average firm size at the labor market level in column (2), industry-year fixed effects in column (3), industry-year and area (federal state)-year fixed effects in column (4) and all of the previous controls together in column (5). The direct effect fluctuates between 9.3 and 11.3 log points and the indirect effect is between 0.116 and 0.126. The average firm is in a labor market where 78% of employees work in a firm below the cutoff, translating to spillover effects on employment of 9.0 to 9.8 log points. In contrast to the direct effect, the spillover effects boost employment for firms in the treatment and control group.

In Figure 1.10, I plot the dynamic spillover effects estimated from specification (1.17). I find an immediate response in employment after the policy change. However, the effect gets larger over time and has not stabilized by 2004 which suggests that the complete propagation of spillovers takes extended time. For the pretreatment periods, the spillover effect is close to zero and statistically insignificant, suggesting parallel trends between firms in different labor markets.

1.6.6 Flow Information

I continue with an analysis of employee flow information. The BHP dataset reports the number of employees within each firm that did not work there one year prior or that do not work there one year after. For my empirical analysis, I first relate the inflow and outflow information to overall employment one year prior. This is equal to the yearly growth rate for net flows and the hypothetical growth rate for inflows and outflows assuming the other flow value to be zero. I then accumulate these rates over time and take logs to have a measure which is similar to log employment.[35] I apply the same procedure to flow information of employees within firms that were unemployed one year prior or are unemployed one year after.

Table 1.6 reports the results for each of the accumulated flow variables. In column (1) to (3), the dependent variables include flows from all employees. The result for net flows in column (1) is nearly identical to the result using log employment with an increase of cumulated net flows by 8.6 log points. This serves as a check on the validity of the analysis of flow information. In column (2), I focus on the inflow rate and find an increase by 10.2 log points. Since the additional inflows in treated firms are larger than net flows, there must be an increase in outflows as well. This is confirmed in column (3), although the coefficient is small and statistically insignificant. These results show that firms that received relatively more tax credits after the policy change increased firm size predominantly through hiring additional employees. Importantly, the point estimate does not suggest that the control group has higher outflows which counters the concern that the employment effect is due to a shift of employees from the control group to the treatment group. If anything, the

[35] This is the case since firm fixed effects control for average firm employment.

Table 1.6: Difference-in-Differences – Analysis of Flow Data

	Dependent Variable:					
	Flows All Employees			Flows Unemployment		
	Net	In	Out	Net	In	Out
	(1)	(2)	(3)	(4)	(5)	(6)
Direct Effect	0.086**	0.102**	0.022	0.031	0.050	0.018
	(0.037)	(0.041)	(0.032)	(0.019)	(0.034)	(0.026)
Observations	10,099	10,099	10,099	10,099	10,099	10,099

Note. Each coefficient is estimated from different regression following main specification (1.14). The dependent variables are the log of cumulated net flows growth in column (1), the log of cumulated inflow growth in column (2), the log of cumulated outflow growth in column (3), the log of cumulated net flow growth from or to unemployment in column (4), the log of cumulated inflow growth from or to unemployment in column (5) and the log of cumulated outflow growth from or to unemployment in column (6). Growth rate is defined as a flow over past year total employment. Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. Standard errors in parentheses are clustered at the regional level: * p<0.10, ** p<0.05, *** p<0.01

higher cumulated outflow among treated firms suggests changes in employment structure.

From a policy perspective, it is of interest whether unemployed individuals gain from tax policies. I analyze the flows related to unemployment in column (4) to (6). Even though the coefficients are statistically insignificant, their magnitudes are economically significant. The effect of inflows from unemployment in column (5) is 5.0 log points. When comparing this estimate to the one for overall inflows, the hiring of unemployed people constitutes 49% among the additionally hired employees. This is sizable and similar to the share of all hires of 60%. The effect on the accumulated outflow rate into unemployment is 1.8 log points which is almost as high as for overall outflows. Thus, it seems that the additional separations are of employees who have problems finding re-employment.

I consider the dynamic effects of the flow variables in Figure 1.A1 in the appendix. Taken together, the dynamic results confirm the previous findings. It is of interest that the net flow and inflow for unemployment are statistically significant in this specification. It is also the case that the share of inflows from unemployment is larger at the beginning with 71.4% and reduces over time. Furthermore, the increase in outflows, although still statistically insignificant, slowly increases over time which suggest that firms first hire additional employees and only let go of unnecessary employees over time.

1.6.7 Heterogeneous Effects

Based on the theoretical model, there is reason to believe that the direct effect of investment tax credits varies by the capital cost share of firms. Even though the theoretical model suggests a measure relating capital costs to revenue, in practice such a measure is biased by the more volatile reaction of revenue to economic changes and exceeds one in many cases. Therefore, I relate the average yearly investment costs to the average yearly wage bill to have a measure between zero and one, which still reflects differences in capital costs. Table 1.7 provides evidence that there are indeed differences empirically. The coefficients for the direct effect are based on firms with zero capital costs. Although this is an unrealistic boundary case, coefficients are close to zero and statistically insignificant as expected. The larger the capital cost share becomes the larger the response for investment and employment. The coefficients mirror the average findings from before, that there is a stronger response for investment rather than employment.

To learn about the underlying mechanisms of spillovers, I examine two additional specifications. First, if spillover occur due to advantages in the production network or input sharing, then the share of firms below the cutoff in related industries should have a larger effect on employment than the share in unrelated industries. Aggregated input and output statistics and job to job movements indicate a strong dependence of firms within industry classes. I therefore split the share of employees in firms below the cutoff into a within industry

Table 1.7: Differences-in-Differences – Effect by Capital Cost Share

	Log Total Investments (1)	Investing (0/1) (2)	Log Employment (3)
Direct effect	0.020	0.004	0.041
	(0.145)	(0.017)	(0.055)
By capital cost share	0.880*	-0.015	0.301*
	(0.485)	(0.049)	(0.156)
Obs	14,453	15,792	17,412

Note. Each column is estimated from different regression following main specification (1.14) where the main interaction term is further interacted with a firms average capital cost share measured as average yearly investment costs over average yearly investment costs and average yearly wage bill. The dependent variables are the log of total investment in column (1), dummy of investing in column (2) and log of employees in column (3). Data is from AFID dataset. Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. Standard errors in parentheses are clustered at the regional level: * p<0.10, ** p<0.05, *** p<0.01

share and an across industry share:

$$\underbrace{\frac{\sum_{i \in S_l, L_i \leq 250} L_i}{\sum_{i \in S_l} L_i}}_{\text{Whole labor market}} = \underbrace{\frac{\sum_{i \in S_l, i \in S_n, L_i \leq 250} L_i}{\sum_{i \in S_l} L_i}}_{\text{Within industry}} + \underbrace{\frac{\sum_{i \in S_l, i \notin S_n, L_i \leq 250} L_i}{\sum_{i \in S_l} L_i}}_{\text{Across industry}} \tag{1.18}$$

Table 1.8 reports the results of estimating direct effects, spillover effects within an industry and spillover effects across industries. As before, I report results for a baseline specification and for specifications with various additional control variables. Since I just split the previous share variable into two, there is no change on the direct effect. For the spillover effect, it is the case that the coefficient for the within industry share is larger than for the across industry share. In column (1) and (4) the value is nearly double. This suggests that spillovers are stronger between firms in the same industry, possibly because of their links in the production network. It should be noted that the standard errors are too large to statistically test for the differences in coefficients.

Table 1.8: Spillover Effects on Labor Inputs By Industry

	Dependent Variable: Log Total Employment				
	(1)	(2)	(3)	(4)	(5)
Direct Effect	0.106***	0.093**	0.112***	0.113***	0.101***
	(0.037)	(0.036)	(0.038)	(0.038)	(0.038)
Regional Share Same Industry Below Cutoff	0.201	0.116	0.132	0.200	0.178
	(0.151)	(0.137)	(0.193)	(0.178)	(0.172)
Regional Share Other Industry Below Cutoff	0.104	0.116*	0.125*	0.105*	0.112**
	(0.074)	(0.067)	(0.071)	(0.062)	(0.056)
Observations	19,328	19,328	19,328	19,328	19,328
Controls					
Growth trends	-	X	-	-	X
Industry-Year FE	-	-	X	X	X
Area-Year FE	-	-	-	X	X

Note. Each column is estimated from different regression following specification 1.16. The dependent variable is log total employment. Additional controls are log average firm wage, firm and year fixed effects, and the controls specified in each column. The sample includes firms within firm size of [20,225],[275,1500] in 1998. Standard errors in parentheses are clustered at the regional level: * p<0.10, ** p<0.05, *** p<0.01

The policy applied mainly to manufacturing firms. Local multipliers can create spillovers to the service industry because of changes in local demand for goods and services. To check for such an effect, I use a regression sample consisting of firms in the service industry with employment between 20 and 1,500 employees in 1998 and analyze the effect of the policy by the share of manufacturing workers in firms below the cutoff. Table 1.9 presents the results. Independent of the chosen control variables, estimates are close to zero, ranging from -0.028 to 0.012, and statistically insignificant. These results suggest that spillover effects are confined to firms in the manufacturing industry. The investment tax credits do not seem to create multiplier effects.

Table 1.9: Spillover Effects to Service Industry

	Dependent Variable: Log Total Employment				
	(1)	(2)	(3)	(4)	(5)
Regional Share Manuf Firms Below Cutoff	-0.028	0.004	-0.000	0.001	0.012
	(0.031)	(0.029)	(0.028)	(0.022)	(0.022)
Observations	75,965	75,965	75,965	75,965	75,965
Controls					
Growth trends	-	X	-	-	X
Industry-Year FE	-	-	X	X	X
Area-Year FE	-	-	-	X	X

Note. Each coefficient is estimated from different regression following specification 1.16 excluding the interaction term of the firm size cutoff on a sample of firms in the service industry. The dependent variable is log total employment. Additional controls are firm-level pre-treatment wage growth trends and the fixed effects specified in each column. Standard errors in parentheses are clustered at the regional level: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

1.6.8 Capital-Skill Complementarity

The economic literature extensively discusses the importance of capital-skill complementarity. I examine this adjustment mechanism in the context of tax credits by analyzing various skill ratios. First, I consider the education level of workers comparing shifts in the ratio of the college educated vs. the non-college educated. I use this measure based on all employees and conditional on being a full-time employee. Second, I analyze the skill level of occupations. I build on the categories provided in the BHP dataset and use the ratio of high-skilled vs low- and medium-skilled occupations. In this definition, technicians, engineers, semi professions, professions and managers count as high-skilled occupations. Manual, service and sales occupations count as low- and medium-skilled occupations. As another dimension, I explore shifts to or away from manual occupations. I build the ratio comparing manual to service and sales occupations. Finally, within manual labor, I analyze the ratio of medium-skilled (qualified) vs. low-skilled occupations.

Table 1.10: Differences-in-Differences – Skill Ratios

	Dependent Variable: Log Ratio of				
	College Educated	Full-time College Educ	High-Skilled	Manual Labor	Qualified Within Manual Labor
	(1)	(2)	(3)	(4)	(5)
Direct Effect	-0.011	0.014	-0.027	-0.019	0.004
	(0.032)	(0.031)	(0.043)	(0.045)	(0.043)
Average Share (%)	11.6	12.3	13.4	74.2	42.2
Obs	8,521	8,466	8,623	8,968	8,008

Note. Each coefficient is estimated from different regression following main specification (1.14). The dependent variables are the log of college educated vs. non-college educated in column (1), the log of full-time college educated vs. full-time non-college educated in column (2), the log of high-skilled occupations vs. low-skilled occupations in column (3), the log of manual occupations vs. service occupations in column (4) and the log of qualified manual occupations vs. unqualified manual occupations in column (5). Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. Standard errors in parentheses are clustered at the regional level: * p<0.10, ** p<0.05, *** p<0.01

Table 1.10 reports the average effect on the log of each ratio. All regression coefficients are close to zero and statistically insignificant. I can reject the null hypothesis that there are large shifts in employment composition. For example, for the ratio of the college educated vs non-college educated, the upper bound of the 95% confidence interval includes a coefficient of 0.052, which translates to a shift of the ratio from 0.131 to just 0.138.

This zero result is surprising. The literature considers ICT as an important mechanism for these shifts, which became an important driver of production technology changes in the 1990s. It is possible that in my setting, ICT does not play a big enough role and therefore does not influence the average employment composition. This however does not mean that there is no influence in particular circumstances. To explore this point in more detail, I study heterogeneous effects of tax credits by the intensity of ICT usage. Since the main datasets do not have information on this type of capital, I match information of ICT usage at the industry level from several data sources. First, Sauer and Strobel (2015) provide investment information for 2014 based on data by the Federal Statistical Office and a firm survey by the Ifo Institute. Second, the EU KLEMS project, in collaboration with the DIW, provides various capital input measures since 1991 for Germany. I rely on the real fixed capital stock in 1998. Third, the IAB establishment survey includes extensive margin investment decisions at firm level for 1993-2014. I aggregate this information to 2-digit industries after controlling for broad regional areas, firm size, average wage and year as possible confounding factors. Finally, the Economic Census collects capital expenditure information for U.S. manufacturing firms. I select the year 2002 as it is the earliest publicly available one. Using information at industry level from several data sources is advantageous for thinking about the heterogeneous results as causal. Because of the aggregation, within industry correlations of firm characteristics and ICT usage are excluded. Using information from data sources of different time periods or countries helps to exclude temporary influences and permanent region specific correlations.

Table 1.A5 in the appendix presents the estimation results using each measure of ICT usage. I find that firms in industries with more intensive use of ICT change their employment composition more towards high-skilled labor and high-skilled occupations when receiving investment tax credits. This heterogeneous response is remarkably stable for each ICT measure, even though they are from quite distinct sources. This speaks to the fact that the effect may be indeed causally related to ICT itself. For the ratio of college-educated vs. non-college educated and using the IAB establishment survey, the coefficient for the heterogeneous effect is 0.793 with a standard deviation of the ICT measure of 0.056. The effect size does not seem large, however, it should be noted that there is larger variation at firm level and the industry measure introduces measurement error on the actual firm-level ICT usage.

The effects on shifts of manual occupations and on shifts within manual occupations are noisy. Only when using information from the Economic Census, I find a statistically significant effect for a more intensive use of medium-skilled compared to low-skilled manual labor among firms in industries with higher ICT shares.

1.7 Discussion

The results show that investment tax credits increase investment and employment. For policy decisions, it is crucial to relate these benefits to the incurred government expenses. Given that one stated goal is the increase in employment, I focus on a measure that relates the government expenses needed for increasing employment among the average manufacturing firm in East Germany. To highlight the contribution of spillovers, I separately calculate this measure for only the direct effects and including spillover effects.

I first provide a back-of-the-envelope calculation of government expenses when considering only the direct effect. Based on the full sample of East German manufacturing firms in Table 1.1, the average firm invests €1.17 million per year including zero investment among 14% of firms and employs 81.1 employees. The estimation results in Table 1.2 and 1.3 elicit an elasticity of investment of 2.825 and an elasticity of employment of 1.051.

Assuming a tax credit rate of 10% for all firms independent of size has the following effect. The average firm increases the intensive margin of investment by 28.3 log points which translates to €381.9 thousand in additional average investments reaching €1.552 million per year. The government therefore spends €155.2 thousand in government expenses on tax credits. At the same time, the average firm increases employment by 8.99 employees. Taken together, government expenses of €17,264 lead to one additional employee.[36]

[36]This calculation does not take into account that government receive additional income tax or reduce expenses for unemployment benefits.

This result changes when including spillover effects. Based on Table 1.5 column (5), first, there is a slight increase in the direct effect implying an elasticity of employment of 1.219. Adopting the same investment response as before, this implies government expenses of €14,761 per additional employee. The difference between both numbers reflects partly the variation of the estimated employment effects.

Second, if all firms receive the tax credit of 10%, then the share of employees working in such firms is by construction one. Thus, the estimate of 0.120 translates to an additional employment effect of 12.0 log points and to an increase of the elasticity of employment to 2.668. Spillovers account for 54.3% of this overall employment response. Due to spillovers, government expenses of only €6,258 per additional employee are needed.

1.8 Conclusion

In this paper, I empirically assess an investment tax credit program in Germany to estimate the causal impact on firm input choices. To evaluate the success of this program, I go beyond a firm investment analysis and study the effect on employment of different labor types and spillover effects.

I find that firms increase both their investment and employment substantially and that this translates to higher revenue. For employment, the effect occurs through hiring additional employees, including a sizable share that were unemployed before. I do not find a shift in the employment composition by skill types on average. Nevertheless, industries with higher dependence on ICT technology are more likely to shift to high-skilled labor. Finally, spillovers further increase the employment effect.

These results are encouraging for the use of investment incentive programs in fiscal policies. The fact that there is a benefit for unemployed and low-skilled individuals is important from a welfare perspective since it is believed that these individuals profit the most from support policies. The existence of spillovers are further important for the cost-effectiveness of the policy. However, more research is needed to assess whether investment incentives are the most efficient way of improving economic outcomes and how programs directly targeted to employees compare during the short-run and the long-run.

The influence of ICT on the workforce composition provides a cautionary tale. Given that ICT has become pervasive in the production process and robots are starting to take over many simple production tasks, incentivizing investments in the future may lead to less beneficial outcomes for unemployed and low-skilled individuals.

1.A Appendix

Table 1.A1: Difference-in-Differences – Specification Robustness Investment

	(1)	(2)	(3)	(4)
	Dependent Variable: Log(Total Investment)			
Direct Effect	0.234**	0.218**	0.201**	0.270***
	(0.089)	(0.090)	(0.090)	(0.094)
Observations	15,275	14,547	15,275	13,047
	Dependent Variable: Log(Equipment Investment)			
Direct Effect	0.251***	0.231**	0.227**	0.285***
	(0.092)	(0.093)	(0.092)	(0.099)
Observations	15,071	14,347	15,071	12,880
	Dependent Variable: Investing (1/0)			
Direct Effect	-0.000	-0.000	-0.004	-0.002
	(0.008)	(0.008)	(0.007)	(0.009)
Observations	15,900	15,148	15,900	13,555

Note. Data from AFID establishment panel. Each coefficient is estimated from different regression following main specification (1.14). The dependent variables are the log of total investment in the upper panel and the log of equipment investment in the lower panel. Column (1) reproduces the coefficient from Table 1.2. Column (2) excludes firms with volatile investment measured by growth of total investment in 1997 above the 95$^{\text{th}}$ percentile, column (3) only includes log average firm wage, firm fixed effects and year fixed effects as control and column (4) is conditional on a a sample of single-establishment firms and excluding firms with volatile investment measured by growth of total investment in 1997 above the 95$^{\text{th}}$ percentile. Standard errors in parentheses are clustered at the regional level: * p<0.10, ** p<0.05, *** p<0.01

Table 1.A2: Impact of Excluded Bandwidth Around Cutoff

	Excluded Bandwidth:				
	None	[238,262]	[225,275]	[213,287]	[200,300]
	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: Log(Total Investment)				
Direct Effect	0.209**	0.208**	0.218**	0.212**	0.210**
	(0.082)	(0.085)	(0.090)	(0.096)	(0.098)
Observations	15,559	15,014	14,547	14,168	13,862
	Dependent Variable: Log(Equipment Investment)				
Direct Effect	0.207**	0.189**	0.231**	0.218**	0.215**
	(0.087)	(0.090)	(0.093)	(0.102)	(0.103)
Observations	15,354	14,813	14,347	13970	13,674
	Dependent Variable: Investing (0/1)				
Direct Effect	-0.000	-0.004	-0.000	-0.002	-0.001
	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)
Observations	16,186	15,625	15,071	14,763	14,446
	Dependent Variable: Log(Employees) AFID				
Direct Effect	0.091***	0.105***	0.113***	0.108***	0.109***
	(0.030)	(0.031)	(0.031)	(0.033)	(0.034)
Observations	18,377	18,007	17,637	17,287	16,997
	Dependent Variable: Log(Employees) BHP				
Direct Effect	0.080**	0.080**	0.087**	0.107**	0.114**
	(0.036)	(0.036)	(0.037)	(0.044)	(0.046)
Observations	10,406	10,256	10,116	9,876	9,706
	Dependent Variable: Log(Domestic Revenue)				
Direct Effect	0.058*	0.074*	0.083**	0.088**	0.089**
	(0.032)	(0.038)	(0.038)	(0.040)	(0.039)
Observations	16,988	16,382	15,898	15,510	15,191

Note. Data from AFID establishment panel and BHP. Each coefficient is estimated from different regression following main specification (1.14). The excluded firm and observations (for employment only firm) are according to the column titles. Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. Standard errors in parentheses are clustered at the regional level: * p<0.10, ** p<0.05, *** p<0.01

Table 1.A3: Impact of Firm Size Interval

	Firm Size Interval					
	[80,225], [275,750]	[54,225], [275,1125]	[45,225], [275,1350]	[40,225], [275,1500]	[34,225], [275,1800]	[20,225], [275,3000]
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Log(Total Investment)					
Direct Effect	0.165*	0.244***	0.211**	0.218**	0.284***	0.327***
	(0.093)	(0.091)	(0.091)	(0.090)	(0.089)	(0.096)
Observations	6,307	10,459	12,821	14,547	17,043	23,945
	Dependent Variable: Log(Machinery Investment)					
Direct Effect	0.163*	0.262***	0.224**	0.231**	0.273***	0.300***
	(0.092)	(0.093)	(0.092)	(0.093)	(0.092)	(0.095)
Observations	6,218	10,332	12,657	14,347	16,799	23,566
	Dependent Variable: Investing (0/1)					
Direct Effect	0.008	0.013	0.003	-0.000	-0.003	-0.020***
	(0.010)	(0.010)	(0.008)	(0.008)	(0.008)	(0.007)
Observations	6,464	15,625	13,308	15,148	17,819	25,494
	Dependent Variable: Log(Employees) AFID					
Direct Effect	0.091**	0.101***	0.105***	0.113***	0.135***	0.128***
	(0.039)	(0.033)	(0.032)	(0.031)	(0.034)	(0.033)
Observations	7,964	12,968	15,626	17,637	20,455	27,717
	Dependent Variable: Log(Employees) BHP					
Direct Effect	0.097**	0.097**	0.089**	0.087**	0.101***	0.115***
	(0.045)	(0.040)	(0.037)	(0.037)	(0.035)	(0.034)
Observations	4,540	7,406	9,086	10,116	11,826	19,244
	Dependent Variable: Log(Domestic Revenue)					
Direct Effect	0.090**	0.096**	0.082**	0.083**	0.111***	0.121***
	(0.045)	(0.041)	(0.040)	(0.038)	(0.041)	(0.042)
Observations	6,731	11363	13,987	15,898	18,701	26,796

Note. Data from AFID establishment panel and BHP. Each coefficient is estimated from different regression following main specification (1.14). The sample consists of firms according to the size intervals given in the column titles. Except for log employment these intervals apply to observations as well. Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. Standard errors in parentheses are clustered at the regional level: * p<0.10, ** p<0.05, *** p<0.01

Table 1.A4: Placebo Cutoffs Within Treatment Group

	Dependent Variable				
	Log Total Investment	Log Machinery Investment	Investing (1/0)	Log Employees (AFID)	Log Employees (BHP)
	(1)	(2)	(3)	(4)	(5)
Cutoff: 80					
Direct Effect	-0.043	-0.038	-0.030*	0.033	0.006
	(0.078)	(0.076)	(0.016)	(0.026)	(0.030)
Observations	6,552	6,448	6,803	9,939	5,410
Cutoff: 100					
Direct Effect	-0.027	0.023	-0.024**	0.039	0.009
	(0.068)	(0.072)	(0.012)	(0.025)	(0.027)
Observations	8,623	8,510	9,011	12,237	7,036
Cutoff: 125					
Direct Effect	0.010	0.040	-0.020*	0.020	0.026
	(0.079)	(0.076)	(0.011)	(0.027)	(0.036)
Observations	10,385	10,247	10,886	14,025	7,916

Note. Data from AFID establishment panel and BHP. Each coefficient is estimated from different regression following a modified version of specification (1.14). I change the firm size interval to between 50 and 250 employees. I set a cutoff for treatment and control group to 80 in the first panel, 100 in the second panel and 125 in the third panel. I exclude firms in an interval between -24 and +24 of the cutoff. Except for log employment these intervals apply to observations as well. Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. Standard errors in parentheses are clustered at the regional level: * p<0.10, ** p<0.05, *** p<0.01

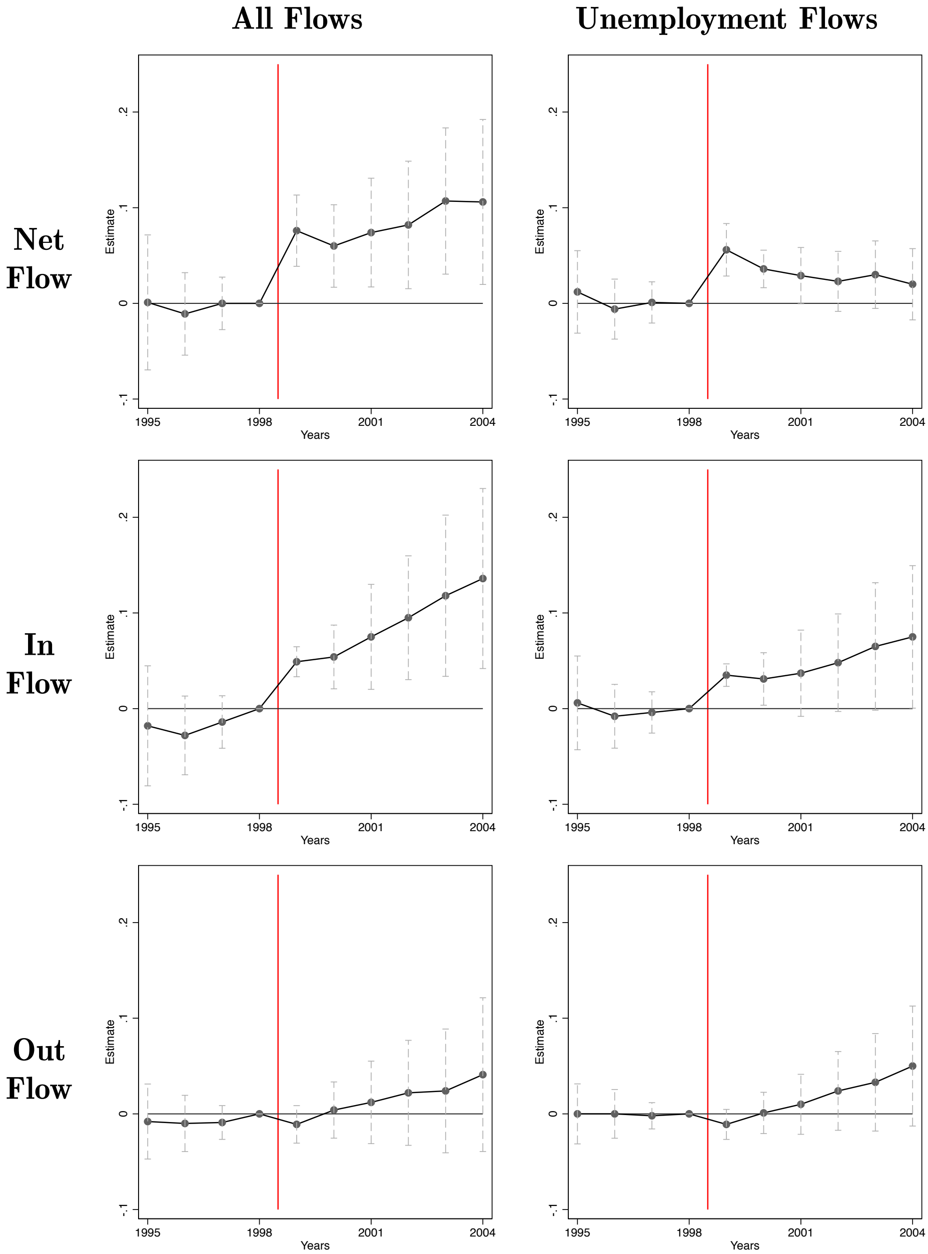


Figure 1.A1: Dynamic Effect By Flow Type

Note. Coefficients from estimation of dynamic specification (1.15). The dependent variables are the log of cumulated flow growth. Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. 95% confidence intervals are displayed for each coefficient using clustered standard errors at the regional level.

Table 1.A5: Differences-in-Differences – Employment Shares By Skill Level

	Dependent Variable: Log Ratio of				
	College to Non-college	College to Non-college (Full-time)	High- to Low-/ Medium-skill	Manual to Service	Qualified vs Non-qualified (Manual)
	(1)	(2)	(3)	(4)	(5)
Direct Effect	-0.021	0.005	-0.041	-0.009	-0.006
	(0.029)	(0.028)	(0.041)	(0.045)	(0.043)
Direct Effect x ICT Invest. Share IAB-LIAB (demeaned)	0.793**	0.710*	0.880**	-0.358	0.439
	(0.374)	(0.389)	(0.439)	(0.582)	(0.46)
Observations	8,521	8,466	8,623	8,968	8,008
Direct Effect	-0.013	0.014	-0.032	-0.022	0.002
	(0.030)	(0.029)	(0.041)	(0.045)	(0.045)
Direct Effect x ICT Invest. Share ifo (demeaned)	0.745*	0.800*	1.081**	0.390	-0.322
	(0.435)	(0.418)	(0.426)	(0.682)	(0.720)
Observations	8,521	8,466	8,623	8,968	8,008
Direct Effect	-0.015	0.011	-0.036	-0.012	-0.010
	(0.030)	(0.029)	(0.042)	(0.045)	(0.043)
Direct Effect x ICT Capital Share EU KLEMS (demeaned)	1.984	1.775	2.877*	-1.084	1.591
	(1.371)	(1.420)	(1.677)	(2.201)	(2.082)
Observations	8,521	8,466	8,623	8,968	8,008
Direct Effect	-0.001	0.024	-0.019	-0.030	0.015
	(0.033)	(0.031)	(0.044)	(0.044)	(0.048)
Direct Effect x ICT Share CapEx U.S. EC (demeaned)	1.378	1.113	1.514	-2.525	3.568*
	(1.388)	(1.276)	(1.378)	(1.988)	(2.142)
Observations	8,521	8,466	8,623	8,968	8,008

Note. Each coefficient is estimated from different regression following main specification (1.14) including interaction term. The dependent variables are the log of college educated vs. non-college educated in column (1), the log of full-time college educated vs. full-time non-college educated in column (2), the log of high-skilled occupations vs. low-skilled occupations in column (3), the log of manual occupations vs. service occupations in column (4) and the log of qualified manual occupations vs. unqualified manual occupations in column (5). The interaction term is the share of firms investing ICT in a given year based on the IAB-LIAB dataset in panel 1, the share of ICT investment based on Sauer, Strobel (2015) in panel 2, the ICT capital share for Germany based on EU KLEMS data and the ICT capital expenditure share based on U.S. Economic Census. Interaction terms are demeaned. Additional controls are log average firm wage, firm fixed effects, industry-year fixed effects and labor market-year fixed effects. Standard errors in parentheses are clustered at the regional level: * p<0.10, ** p<0.05, *** p<0.01

Chapter 2

THE CAPITAL-LABOR ELASTICITY OF SUBSTITUTION: EVIDENCE FROM INVESTMENT TAX CREDITS

2.1 Introduction

The elasticity of substitution between capital and labor is key in determining a vast set of economic outcomes. At the macroeconomic level, the aggregate elasticity influences for example long-term growth and income per capita (Klump and Grandville, 2000), fluctuations in the real business cycle (Cantore et al., 2014; Koh and Santaeulàlia-Llopis, 2017), the effectiveness of monetary and fiscal policy (Chirinko, 2002), the impact of trade openings (Dornbusch et al., 1980) and unemployment levels (Rowthorn, 1999).

Given its relevance, there is a long-standing interest in empirically determining the elasticity of substitution with research proceeding already for close to a century. It is however difficult to identify an unbiased elasticity using time-series data (Diamond et al., 1978; León-Ledesma et al., 2010) and as a consequence no agreement on a consensus value has been reached. Estimates fluctuate widely with typical values falling between 0.5 to 1.5 that permit opposing economic predictions. For example, Karabarbounis and Neiman (2013) argue based on an estimated elasticity of substitution of 1.25 that a reduction in capital prices can explain the recent decline in the labor share. For the case of an elasticity smaller one, the opposite would hold true and the labor share would actually increase.

With the recent availability of detailed firm data, an alternative approach is the estimation of the elasticity of substitution at the firm-level. In its own right, the firm-level elasticity is important as a factor in the response of firms to tax incentives (Hall and Jorgenson, 1967) and minimum wage (Aaronson and French, 2007), and for the decision of firms to innovate (Acemoglu, 2010). The firm-level elasticity is also informative about the macroeconomic one. Oberfield and Raval (2014) show that adjustments at the firm level explain two thirds of the aggregate elasticity in the context of the U.S.

In this paper, I provide novel evidence on the firm-level elasticity of substitution between capital and labor using quasi-experimental variation from an investment tax policy targeted towards German manufacturing firms. The policy provided size-dependent investment tax credits that induced a larger reduction in capital costs for firms with up to 250 employees.

I analyze the firm behavior around the cutoff by setting up a theoretical framework that introduces size-dependent capital costs into the span of control firm production model by Lucas (1978). The model demonstrates that firms have an incentive to decrease their firm size below the cutoff to benefit from lower capital costs. This leads to bunching of firms at the cutoff and missing mass of firms just above the cutoff. Additionally, firms that choose to produce at a firm size below the cutoff move towards a more capital-intensive production and as a result have higher capital-labor ratios. The extent of these distortions importantly depends on the production function parameters and in particular the capital-labor elasticity of substitution.

The descriptive evidence confirms the predictions of the model. There is excess mass at the firm size cutoff and firms change their investment behavior according to their capital costs. To account for measurement error in firm size, I augment the theoretical framework by including an error term to the firm size measure. This adjustment results in a smooth firm-size probability and capital-labor function that closely match the descriptives.

I then estimate the model structurally, using maximum likelihood and non-linear least squares. I use a two-step procedure that first estimates the firm-size density and given the coefficients in a subsequent step the capital-labor relation. With this approach, I recover a capital-labor elasticity of substitution of 0.083. As such, my results are in line with the contemporaneous findings on the micro elasticity. The estimates are far lower than one and provide strong evidence that firms do not have Cobb-Douglas production functions. At the same time applying the results of Oberfield and Raval (2014), these values suggest that the aggregate production function does not have a unitary elasticity of substitution either. Stated differently, if an unitary elasticity were indeed correct at the aggregate level, large changes in the sector composition through entry and exit of firms would need to shift capital-labor ratios since adjustments of the ratio at the firm level are small.

This article is related to the estimation of firm production functions. Earlier research focuses on firm productivity and assumes Cobb-Douglas production functions (e.g Levinsohn and Petrin, 2003; Olley and Pakes, 1996). I instead focus on the elasticity of substitution. In my setting, productivity of firms around the cutoff is on average the same without the tax policy. Instead, firm behavior is driven by the remaining production parameters that are conditional on the productivity of firms at the cutoff.

Raval (2019) estimates the elasticity of substitution for U.S. manufacturing firms. He combines cross-sectional variation in local wages with an instrumental variable approach and finds elasticities in a range between 0.3 and 0.5. Although many robustness tests address the concern that regional differences are driving the results, there is still the possibility that labor costs influence firm location choices and thereby introduce a bias in the estimation. My quasi-experimental setup relies on the sharp difference in capital costs at the firm size cutoff. Firm characteristics would be the same for firms directly to the left and right of the cutoff without the tax policy. This includes regional characteristics. The distortions introduced by the policy are directly modeled and are part of the estimation. Therefore, the estimation controls for biases from firm characteristics.

In contemporaneous and independent research, two papers follow a similar strategy to mine and use size-dependent policies. Moreau (2019) uses labor regulations in France that bind from a firm size of 50 employees. Exploiting the distortions created by such regulations, he estimates a firm-level elasticity for manufacturing firms of 0.12. His theoretical framework assumes a size-dependent profit tax rate to capture the labor regulations[1] and he structurally estimates the degree of bunching and capital distortion at the cutoff. My approach instead relies on a simple investment tax credit program, for which the policy characteristics have clear implications for the modeling approach. With this setting, I reduce the bias from misspecification in the estimation model. As a further difference, the size-dependent input costs lead to shifts in capital-labor ratios and firm density away from the cutoff that I include in my theoretical framework and exploit in the estimation of the model.

Benzarti and Harju (2019) study a tax policy in Finland that sets different payroll tax rates according to a capital depreciation cutoff. They rely on a donut-hole regression discontinuity design to study shifts in capital and labor around the cutoff and show that higher payroll tax rates lead to a decrease both in labor and capital of firms.

[1] This is in contrast to Garicano et al. (2016) and Gourio and Roys (2014), who consider the same policy and model the regulations as labor costs and fixed costs.

They argue that this is consistent with an elasticity of substitution of 0.0. I instead include observations at the cutoff and make use of the additional information contained in the bunching behavior of firms. With the additional information I consider not only the elasticity of substitution but also scale effects in production. This is important since adjustments in labor and capital of firms are influenced by both these effects. I find important scale effects in my setting and estimate the according elasticity of substitution.

The remainder of the paper is organized as follows. Section 2.2 sets up the theoretical framework, Section 2.3 details the investment tax policy, Section 2.4 provides information on the data and the variables used in the analysis, Section 2.5 presents descriptive evidence, estimates an adjusted version of the theoretical framework, and backs out the elasticity of substitution and the remaining production function parameters. Section 2.6 concludes.

2.2 Firm Production Model with Discontinuous Capital Costs

I set up a simple firm production model following Lucas (1978) by assuming that the allocation of productive factors is over managers with varying ability levels. I add to the model by including a size-dependent change in capital costs to capture the regulations of the tax policy. The model provides not only solutions for optimal production of each firm, but through a managerial ability distribution also for aggregate production along the firm size distribution. I first describe the decision problem of each firm separately and then consider the aggregate.

Each firm maximizes profit according to

$$\pi(\alpha_i) = \max_{K_i, L_i} \begin{cases} p_i \alpha_i f(L_i, K_i) - wL_i - \tau r K_i & \text{if } L_i \leq N \\ p_i \alpha_i f(L_i, K_i) - wL_i - rK_i & \text{if } L_i > N, \end{cases} \tag{2.1}$$

where each firm i chooses labor L_i and capital K_i to produce a firm-specific output Y_i. All firms have the same production function $Y_i = f(L_i, K_i)$ but they differ in the ability of the manager that determines productivity $\alpha_i > 0$. Firms pay their inputs the equilibrium wage w and capital rental rate r, respectively, and receive price p_i for each unit of output sold. Firms also receive a tax credit $\tau < 1$ that reduces the capital rental rate, as long as they do not exceed the employment cutoff N. Lower values of the tax credit mean a larger reduction in capital costs.[2]

I focus on the particular case of constant elasticity of substitution (CES) with the production function

$$f(L_i, K_i) = (\lambda L_i^{\rho} + (1-\lambda) K_i^{\rho})^{\frac{1}{\rho}}, \tag{2.2}$$

where λ and ρ govern the production process. In particular, $\rho = \frac{e_{kl}-1}{e_{kl}} \in [-\infty, 1]$, where e_{kl} is the elasticity of substitution between capital and labor. The parameter $\lambda \in [0,1]$ influences the cost ratio between the input factors. The CES production function is widely used in empirical applications due to its flexible functional form and nests for example the Leontieff production function ($e_{kl} = 0$) and the Cobb-Douglas production function ($e_{kl} = 1$) as special cases.

I further assume a downward sloping inverse product demand curve

$$p_i = B Y_i^{-\frac{1}{\eta^D}}, \tag{2.3}$$

where B is a product demand shifter and $\eta^D > 1$ the elasticity of product demand. This leads to decreasing returns to scale and a unique solution for each firm.[3]

[2] In the context of the policy, all firms received tax credits. The model setup assumes that a baseline rate is subsumed in the capital rental rate r and that τ identifies the difference in the tax credit rate of firms below and above the cutoff.

[3] Decreasing returns to scale can also be the result of fixed factors in the production process, for example due to limits in managerial time. Both approaches lead to identical predictions in observed firm behavior in my context, however, the underlying mechanism with respect to the scaling coefficient would be distinct.

The profit function is differentiable for all $L_i \neq N$ and the first order conditions lead to the locally optimal input decision of

$$K^* = B^{\frac{1}{1-\theta}} \alpha^{\frac{1}{1-\theta}} \left(\frac{\theta(1-\lambda)}{r\bar{\tau}} \right)^{\frac{1}{1-\theta}} \left(\lambda(\gamma\bar{\tau})^{\frac{\rho}{1-\rho}} + (1-\lambda) \right)^{\frac{\theta-\rho}{\rho(1-\theta)}} \tag{2.4}$$

$$L^* = B^{\frac{1}{1-\theta}} \alpha^{\frac{1}{1-\theta}} \left(\frac{\theta\lambda}{w} \right)^{\frac{1}{1-\theta}} \left(\lambda + (1-\lambda) \left(\frac{1}{\gamma\bar{\tau}} \right)^{\frac{\rho}{1-\rho}} \right)^{\frac{\theta-\rho}{\rho(1-\theta)}}, \tag{2.5}$$

where $\gamma = \frac{r}{w} \frac{\lambda}{1-\lambda}$, $\theta = 1 - \frac{1}{\eta^D}$, and $\bar{\tau}$ is defined as $\bar{\tau} = \tau$ if $L_i \leq N$ and $\bar{\tau} = 1$ if $L_i > N$ to simplify notation. θ is the scaling factor and is one for constant returns to scale. Due to imperfectly elastic product demand, there is decreasing returns to scale $\theta < 1$ in the model. The solutions are defined implicitly, since the optimal number of employees influences and is influenced by the tax credit rate.

The optimal capital and labor choice implies a capital-labor ratio of

$$\frac{K^*}{L^*} = (\gamma\bar{\tau})^{\frac{1}{\rho-1}}. \tag{2.6}$$

The ratio is inversely related to the value of the tax credit. This is as expected, since a reduction in capital costs leads to an increase in the relative use of capital. The magnitude of such an adjustment depends on the elasticity of substitution with higher elasticities leading to larger relative adjustments.

These results so far exclude the additional incentives introduced by the shift in capital costs at the cutoff. To understand the influence of these incentives on firm behavior, I first consider a firm that chooses optimal labor right at the cutoff $L^* = N$ and define the corresponding productivity as α_N. This firm still receives a reduction in capital costs due to tax credits. A firm with just slightly higher productivity would be located above the cutoff and therefore be ineligible for tax credits. The firm is better off by reducing their firm size to the cutoff value. Such a decision is optimal since the firm chooses a negligible smaller firm size which translates to insignificant changes in revenue, but they benefit from the reduction in capital costs. A similar argument can be made for firms with productivity levels further above α_N. They still can gain from bunching at the cutoff due to the benefits from the reduction in capital costs. However, the larger is the decrease in firm size the larger is the reduction in revenue. There exists a firm productivity α_r for which benefits and costs of bunching cancel out. These firms are the so called marginal bunchers and their firm size is implicitly defined by

$$\max_{K_i} p_N \alpha_r f(N, K_i) - wN - \tau r K_i = p_r \alpha_r f(L_r, K_r) - wL_r - rK_r, \quad (2.7)$$

where p_N is the price given the firm's output at the cutoff and p_r, L_r and K_r are the output price and inputs when maximizing profits without tax credits.

All firms that bunch at the cutoff still optimize profits. Conditional on firm size, this leads to adjustments in the capital inputs. They choose capital at the cutoff according to

$$K_{N_i} = \arg\max_{K_i} p_{N_i} \alpha_{N_i} f(N, K_i) - wN - \tau r K_i, \quad (2.8)$$

where α_i is the productivity of firm i and p_{N_i} is the price given firm i's output at the cutoff. The exact capital value importantly depends on firm's productivity with higher productivity leading to more capital-intensive production at the cutoff.

In the model, the decreasing returns to scale lead to non-zero profits for firms and would induce firm entry. As in Lucas (1978), I establish an equilibrium by assuming that each individual i chooses between earning wage w as a worker or earning the profit $\pi(\alpha_i)$ as a firm owner. The productivity α_i is directly linked to the individual and can be thought of as their managerial ability. Each individual has a fixed managerial ability that is determined by a random draw from the power law distribution

$$\phi(\alpha) = c_\alpha \alpha^{-\beta_\alpha}, \quad (2.9)$$

where $c_\alpha > 0$ and $\beta_\alpha > 0$.

In equilibrium there is a minimum ability α_{min} defined as

$$\pi(\alpha_{min}) = w, \tag{2.10}$$

for which the individual is marginal between being a worker and a firm owner. All individuals with lower managerial ability become workers and all those with higher ability become firm owners. Equilibrium wage works as an opportunity cost that equalizes the number of workers with the aggregate labor demand of the firm owners.

This then allows to describe the firm size distribution as follows. Firms in the productivity interval of $[\alpha_{min}, \alpha_N)$ have a firm size below the cutoff and therefore receive tax credits. All firms with productivity of $[\alpha_N, \alpha_r)$ bunch at the firm size cutoff, creating excess mass at the cutoff and missing mass to the right of the cutoff. Firms with productivity of $[\alpha_r, \infty)$ have a firm size above the cutoff and therefore produce without getting tax credits.

Given the distribution of managerial ability the firm size distribution is[4]

$$\chi(L^*) = \begin{cases} (\beta-1)L_{min}^{\beta-1}L^{*-\beta} & \text{if } L_{min} \leq L^* < N \\ L_{min}^{\beta-1}\left[N^{1-\beta} - L_r^{1-\beta}\left[\frac{T(1)}{T(\tau)}\right]^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}}\right] & \text{if } L^* = N \\ 0 & \text{if } N < L^* < L_r \\ (\beta-1)L_{min}^{\beta-1}\left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} L^{*-\beta} & \text{if } L_r \leq L^* \end{cases} \tag{2.11}$$

where $\beta = \beta_\alpha(1-\theta) + \theta$ and $T(\bar{\tau}) = \lambda + (1-\lambda)\left(\frac{1}{\gamma\bar{\tau}}\right)^{\frac{\rho}{1-\rho}}$.

The corresponding optimal capital choice is

$$k(L_i^*) = \begin{cases} (\gamma\tau)^{\frac{1}{\rho-1}}L_i^* & \text{if } L_{min} \leq L_i^* < N \\ [(\gamma\tau)^{\frac{1}{\rho-1}}N, K_{N_r}] & \text{if } L_i^* = N \\ \gamma^{\frac{1}{\rho-1}}L_i^* & \text{if } L_r \leq L_i^*, \end{cases} \tag{2.12}$$

[4]See Appendix 2.A.1 for the computational steps.

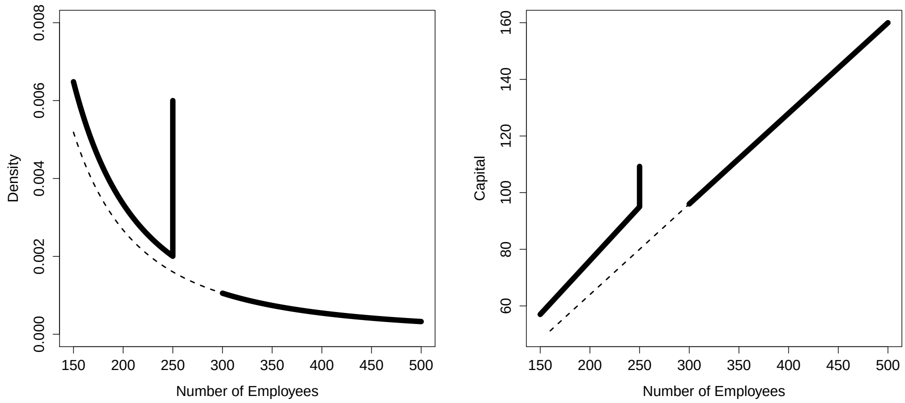


Figure 2.1: Theoretical Firm Size Distribution and Capital-Labor Relation

Note. The graphs show a parametrization for which $L_{min} = 120$, $L_r = 300$, $\beta = 2.3$ and $T = \left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} = 0.8$. The density at the cutoff is capped at 0.006. The actual value is 0.142.

where the capital level at the cutoff is implicitly defined according to the constraint profit maximization in equation (2.8). Firms compensate their reduction in firm size with an increase in capital levels. The exact amount depends on each firm's productivity, with higher productivity leading to higher capital levels. In aggregate this leads to an interval of capital levels at the cutoff with the marginal buncher choosing the highest capital level K_{N_r}. Since there are no firms in the interval (N, L_r), the capital level is not defined for these values.

Figure 2.1 illustrates the theoretical predictions graphically. The left panel shows the firm size distribution. Away from the cutoff, there is a decrease of the density in firm size according to a power law. This is in line with the literature (e.g. Axtell, 2001) that find evidence for such a relation for various countries. Due to the discontinuous drop in the tax credit rate, there are two additional distinct characteristics. First, there is excess mass right at the cutoff. These firms have productivity between α_N and α_r and obtain the highest profit when choosing firm size N thereby creating bunching at the cutoff and missing mass to the right of the cutoff. Second, there can be a shift in the distribution. The figure shows the case of a upward shift of firms below the cutoff that materializes due to firms choosing more employees conditional on their productivity level. This is the result of the reduction in capital costs due to investment tax credits that leads to an expansion of firm production and therefore higher labor demand. The model allows for an downward shift instead as well since there is a counteracting substitution away from labor and towards capital. The direction of the shift depends on the relative magnitude of the elasticity of substitution σ and the scaling in production θ, with higher substitution leading to a downward shift.

The right panel depicts the capital-labor relation. The relation is linear away from the cutoff with higher capital levels in larger firms. Due to the increase in capital costs, there is an unambiguous drop in the capital-labor ratio of the firms above the cutoff. At the cutoff, where the bunching firms locate, each firm chooses a different capital level according to their productivity leading to an interval of predicted values.

The theoretical framework is the foundation for the empirical analysis. As a next step, I consider how well the model predictions correspond to the empirical evidence of the policy found in the data.

2.3 Policy

I apply the theoretical framework to an investment tax credit policy in Germany, the so called *Investitionszulagen* program, that started in 1991 and lasted until 2013. This is the same policy as in Lerche (2018) and I, therefore, outline only the main characteristics relevant for the estimation.

The tax policy provided a reduction in investment costs to manufacturing firms in East Germany through tax credits. The federal government implemented the policy as a place-based support program to alleviate large regional differences. Through an official government act, eligible firms had the right to file a short report and get back a share of their investment costs on fixed assets at the end of each business year according to the tax credit rate. The exact tax credit rate varied over time and according to firm characteristics. The rate was typically around 10% but reached up to 27.5% for small firms in border regions. Although the tax credits were paid within the federal tax system, the individual payments to each firm were irrespective of the firm's tax burden, leading to reduced investment costs even for firms with financial losses. The program implied substantial government expenses that were around $1.2 billion per year during the time of analysis.

There were regular changes in the policy that led to adjustments in the eligibility criteria and tax credit rates. In July 1994 the policy introduced firm size dependent tax credit rates in the spirit of the widely used approach to support small and medium-sized firms (Guner et al., 2008). Between 1994 and the end of 1998 firms with up to 250 employees received a tax credit rate of 10% and those above the cutoff only 5%. In 1999 adjustments in the policy increased the investment tax credit rates. Between 2000 and 2004 firms below the cutoff received 25% and those above received 12.5%. During this time period, the definition of firm size considered the head count of employees at the beginning of a business year without differentiating full-time and part-time employees. Vocational trainees did not count due to their special employment contracts under German law. In the empirical analysis, I focus on the period 1999–2004 to benefit from the large difference in investment costs at the cutoff that allows for a more robust estimation of the model.

In 2005 changes in the policy led to new definitions of the firm size and the cutoff value by applying the definition of small and medium-sized firms of the European Union. From then on, firms of a corporate group had to add up all employees within the group. Furthermore, part-time employees counted according to hours worked. The cutoff criterion expanded to not only include firm size but also revenue and total assets. Due to these significant changes in definitions I disregard these years in the empirical analysis.

2.4 Data

The empirical analysis relies on the years 1999 to 2004 of the AFID Establishment Panel by the Federal Statistical Office of Germany. The dataset combines administrative data from mandatory firm reports used for official economic statistics such as aggregate output and investment. By law all establishments of manufacturing and mining firms with more than 20 employees in Germany have to report on their employment, investment and revenue in regular intervals. Through unique establishment identifiers, the final dataset combines the reports and connects them over time while aggregating monthly and quarterly information by year.

The reporting in the dataset is at establishment level. Since the policy-relevant cutoff applies to the firm level I aggregate all information at firm level using firm identifiers. I choose the firm level also for the empirical analysis since 90.9% of firms are single-establishment firms for which a differentiation between establishment and firm does not occur. For the case of multi-establishment firms, the fact that they are not split in separate single-establishment firms suggests that production is closely linked and that decision making happens as a unit. Even if this is not the case, it seems reasonable to assume that establishments of the same firm have aligned incentives and that an aggregation masks little heterogeneity. I exclude firms with establishments in East and West Germany to avoid the possibility of shifting of economic activity within such firms. I also exclude firms with establishments in Berlin, since a similar argument applies.

For the firm size measure in the empirical analysis, I rely on information on total employment. The data only provide the cumulated number of employees over the reports made by each establishment within a year. It is therefore not possible to exactly measure firm size at the beginning of the year. As a proxy I compute the yearly average by dividing the cumulated employment with the number of reports made within a year. For the majority of the analysis, I exclude vocational trainees from the firm size measure to match the policy definition. For this, I match information on vocational trainees at firm level for the years 1999 to 2001 from the cost structural panel (KSE) by the Federal Statistical Office of Germany. Since there is missing information for a sizable number of observations, I impute values assuming a constant share of trainees within firms, then within 3-digit industry classifiers and finally within the overall manufacturing sector. The timing problem and imputation likely introduces classical measurement error and it is important to consider resulting deviations between theory and the empirical evidence.

The dataset does not provide information on capital. The only possibility to get to a capital measure would be through computations using the perpetual inventory method. It is well-known that applying this method is non-trivial (Collard-Wexler and De Loecker, 2016) and there is agreement that measurement error is unavoidable. Investment, as the flow measure of capital, is much less prone to such concerns. In the empirical analysis I therefore focus on investment and relate it to capital as follows.

I start out with the law of motion of capital:

$$K_{it} = (1 - \delta)K_{it-1} + I_{it} \tag{2.13}$$

where K_{it} is the capital stock of firm i at time t and capital depreciates according to the depreciation rate δ.

By rearranging, I obtain

$$I_{it} = \Delta K_{it} + \delta K_{it-1} = K_{it} + (\delta - 1)K_{it-1}. \tag{2.14}$$

Using the result for the capital-labor ratio from equation (2.6) the investment equation becomes

$$I_{it} = (\gamma\bar{\tau}_{it})^{\frac{1}{\rho-1}} L^*_{it} + (\delta - 1)(\gamma\bar{\tau}_{it-1})^{\frac{1}{\rho-1}} L^*_{it-1}, \tag{2.15}$$

where changes in employment over time can occur for example due to productivity shocks. The equation implies that investments to one part recover depreciations over the business year and to another part facilitate changes in capital driven by transitions in firm size. The equation further details that firms below the cutoff have larger capital stocks and therefore larger investments to recover depreciations compared to firms above the cutoff. Additionally, a move across the threshold entails further adjustments to adapt to the change in capital costs.

For the empirical analysis I build on the assumption that changes in firm size and therefore changes in capital are independent of current firm size. In this case, equation (2.14) can be written as

$$I_{it} = \Delta K_{it} + \delta K_{it-1} = g_{\Delta} + \delta K_{it-1} + \omega_{it}, \tag{2.16}$$

where the error term ω_{it} has mean zero and is independent of K_{it}, and $g_{\Delta} = \Delta \bar{K}_{it}$ is the average change in capital.

Since changes in capital are independent from current ones, the equation can be adjusted to

$$I_{it} = g_{\Delta} + \delta K_{it-1} + \omega_{it} = g_{\Delta} + \delta K_{it} + \upsilon_{it} = g_{\Delta} + \delta(\gamma\bar{\tau}_{it})^{\frac{1}{\rho-1}} L^{*}_{it} + \upsilon_{it}, \tag{2.17}$$

where $\upsilon_{it} = -\delta \Delta K_{it} + \omega_{it}$.

In the empirical analysis, I focus on two measures of total investment directly taken from the dataset. The first one uses the total of purchased and self-produced equipment and structures. The second measure additionally includes investments for leased equipment and structures. In both cases, small-sized investments that are not activated on the balance sheet are excluded.

2.5 Estimation

2.5.1 Adjusting Theoretical Predictions to Empirical Outcomes

In the model, predictions are sharp. Firms bunch exactly at the cutoff and there is a firm size interval above the cutoff with no firms. As the empirical counterpart, Figure 2.2 shows the descriptive firm size distribution around the cutoff for total employment excluding vocational trainees in the left panel and for total employment in the right panel. In both panels, there is evidence for excess mass around the firm size cutoff. In the left panel, the density becomes flat at around 200 employees and starts to increase at 220 employees. At 240 employees the density reaches a local maximum before it starts to drop off sharply until 250 employees. For firms above the cutoff, the density follows a typical power law, as is the case for firms with fewer than 200 employees.

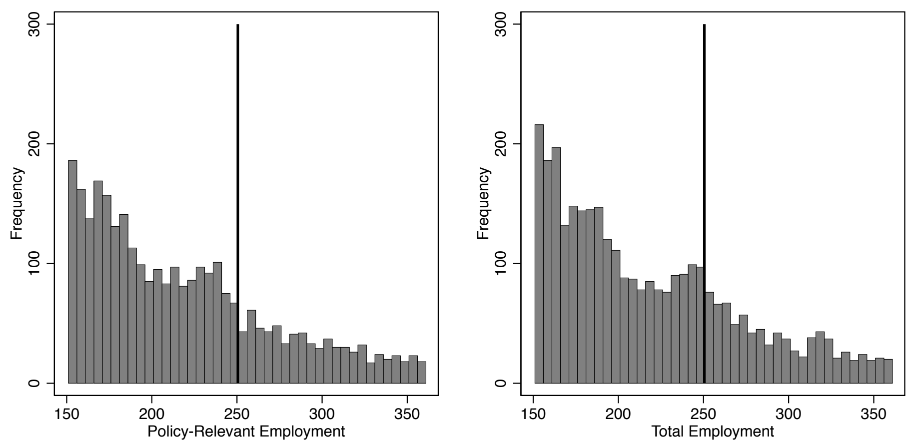


Figure 2.2: Empirical Firm Size Distribution For Total Employment and Policy-Relevant Employment

Note. The firm size distribution is for the pooled sample of manufacturing firms with establishments only in East Germany excluding Berlin over the years 1999 to 2004.

To investigate why the density already drops off before 250 employees, I consider total employment in the right panel. In this case, the density displays similar behavior with an increase in the density before the cutoff but the drop in density only starts at 250 employees. Reasons for such behavior could be that not all firms knew about the exact definition of firm size for the investment tax policy or that they calculated a margin of error.

To summarize, bunching does not occur just at the cutoff, but leads to excess mass for a range of firm sizes around it. Such differences between theory and data have been documented in past research. In my analysis, one likely reason is the measurement error occurring in the firm size variable and I therefore follow Garicano et al. (2016) by incorporating measurement error into the model to address the differences between theory and empirics.

I assume that actual firm size is not perfectly observable but that there is measurement error in the observed data

$$L(\alpha, \epsilon) = L^*(\alpha) e^{\epsilon}, \tag{2.18}$$

where L^* is the equilibrium firm size as defined in equation (2.5), $L(\alpha, \epsilon)$ is the observed firm size in the data and ϵ is unobserved measurement error. I assume that the measurement error is Gaussian noise with mean zero and variance σ^2. This approach captures that the dataset cannot perfectly measure firm size as defined in the model and that the relevant firm size cutoff is slightly different to the information contained in the data.

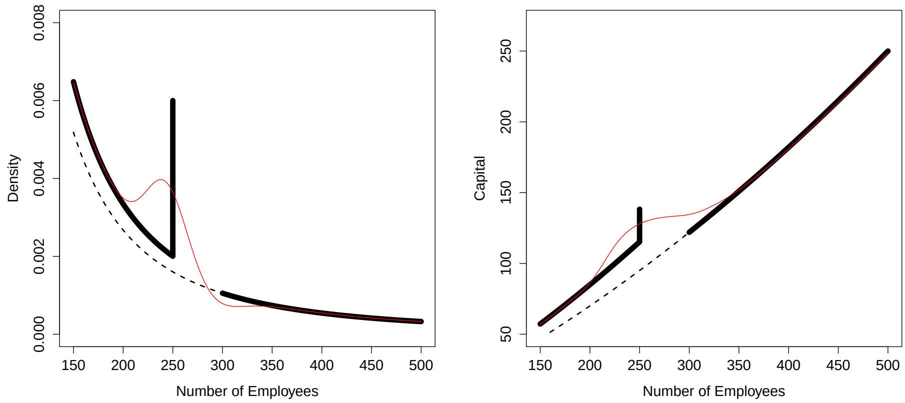


Figure 2.3: Theoretical Firm Size Distribution and Capital-Labor Relation

Note. The graphs show a parametrization for which $L_{min} = 120$, $L_r = 300$, $\beta = 2.3$ and $T = \left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} = 0.8$. The density at the cutoff is capped at 0.006. The actual value is 0.142.

As a result, the observed firm size distribution is

$$\begin{aligned}\chi(L) = \frac{1}{\sigma L}\frac{1}{L_{min}^{1-\beta}}(N^{1-\beta} - \left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} L_r^{1-\beta})\varphi\left(\frac{\ln(L)-\ln(N)}{\sigma}\right) \\ -(1-\beta)\left(\frac{1}{L_{min}}\right)^{1-\beta} L^{-\beta} e^{\frac{\sigma^2}{2}(\beta-1)^2} \\ \cdot\left[\Phi\left(\frac{\ln(L)-\ln(L_{min})}{\sigma}-\sigma(\beta-1)\right)-\Phi\left(\frac{\ln(L)-\ln(N)}{\sigma}-\sigma(\beta-1)\right)\right] \\ -(1-\beta)\left(\frac{1}{L_{min}}\right)^{1-\beta} L^{-\beta} e^{\frac{\sigma^2}{2}(\beta-1)^2}\left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} \\ \cdot\Phi\left(\frac{\ln(L)-\ln(L_r)}{\sigma}-\sigma(\beta-1)\right), \quad (2.19)\end{aligned}$$

where φ is the Gaussian pdf and Φ is the Gaussian cdf.[5] The left panel of Figure 2.3 provides an example distribution. Due to the measurement error, the distribution becomes smooth with observed bunching not only at the cutoff but also around it. There is missing mass to the right of the cutoff, although, it is limited in magnitude. Overall, including measurement error in the model provides predictions in line with the descriptive evidence.

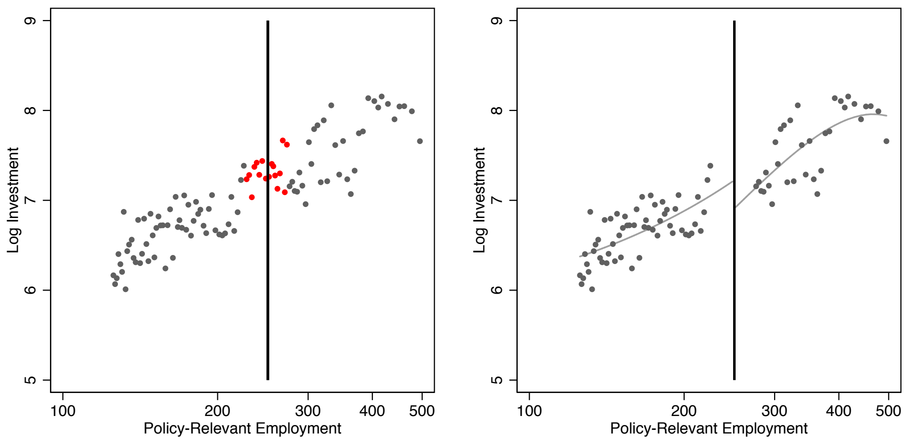


Figure 2.4: Investment Firm Size Relation

Note. The sample consists of all manufacturing firms with establishments only in East Germany excluding Berlin over the years 1999 to 2004. Each data point represents the average value of log total investment of firms with similar firm size. To the left of the cutoff, one data point consists of 50 observations. To the right of the cutoff, one data point consists of 25 observations. The right panel excludes observations with firm size between 225 and 275. Each trend line is from a separate estimation.

To understand the capital-labor relation, Figure 2.4 presents descriptive evidence for the log of total investment including leasing as a function of firm size where each dot represents the average of 50 and 25 observations to the left and right of the cutoff, respectively. The left panel includes observations around the cutoff in the interval of 225 to 275 employees and for these observations investment

[5] Appendix 2.A.2 provides details on the calculation steps.

tends to be higher than for those immediately to the left and right of this interval. This supports the theoretical prediction that bunching firms use relatively more capital for production. There is also some evidence of an upward shift in investment conditional on firm size for firms below the cutoff. For a better visual inspection, the right panel excludes observations in the interval from 225 to 275 employees and includes trend lines. The trend lines confirm the upward shift in investment for firms below the cutoff. The difference is sizable with 32 log points at the cutoff. This is again in line with the theory, that predict a shift towards capital for firms below the cutoff due to larger investment tax credits.

Finally, theory has clear predictions for the slope of the investment-labor relation. In particular, there should be a linear relation between investment and labor which is actually not confirmed in the descriptive evidence with non-linearity evident from the trend lines.

To adjust for the differences between theory and descriptives, I adjust the theoretical framework as follows. First, to capture the possibly of a non-linear relation between capital and labor, I assume the quadratic functional form

$$k(L_i^*) = \begin{cases} a_0 + a_1 L^* + a_2 {L^*}^2 & \text{if } L_{min} \leq L^* < N \\ XN & \text{if } L^* = N \\ b_0 + b_1 L^* + b_2 {L^*}^2 & \text{if } L_r \leq L^*. \end{cases} \tag{2.20}$$

In this definition, I also simplify the relation between capital and firm size at the cutoff by assuming that all bunching firms have the same capital level. This captures the fact, that in the estimation XN measures the average capital level at the cutoff.

Additionally, I introduce measurement error in firm size in the same way as for the empirical firm size distribution.[6] I assume that capital is perfectly observed but that the according firm size is not

[6]In the model, I assume away measurement error in capital. In the estimation, however, measurement error is included through the error term.

and compute the expected capital conditional on firm size, defined as

$$E(K|L) = \int_{\mathbb{R}} K(L^*) f_{L^*|L}(L^*|L) dL^*, \tag{2.21}$$

where $f_{L^*|L}(L^*|L)$ is the distribution of the actual firm size conditional on the observed firm size.

As a result I compute the expected value of capital as[7]

$$E(K|L) = \frac{c^0(K|L) + c^1(K|L) + c^2(K|L) + c^+(K|L)}{\chi(L)}, \tag{2.22}$$

where

$$\begin{aligned}
c^x(K|L=n) = &\\
& - (1-\beta)\frac{a_x}{L_{min}^{1-\beta}} n^{x-\beta} e^{\frac{\sigma^2}{2}(\beta-x-1)^2}\\
\cdot & \left[\Phi\left(\frac{\ln(n)-\ln(L_{min})}{\sigma} - \sigma(\beta-x-1)\right) - \Phi\left(\frac{\ln(n)-\ln(N)}{\sigma} - \sigma(\beta-x-1)\right)\right]\\
& - (1-\beta)\frac{b_x}{L_{min}^{1-\beta}} T n^{x-\beta} e^{\frac{\sigma^2}{2}(\beta-x-1)^2} \Phi\left(\frac{\ln(n)-\ln(L_r)}{\sigma} - \sigma(\beta-x-1)\right),
\end{aligned} \tag{2.23}$$

for $x = 0, 1, 2$ and

$$\begin{aligned}
c^+(K|L=n) &= \frac{d}{dn}\int_{-\infty}^{\ln(n)-\ln(L_r)} \frac{XN}{L_{min}^{1-\beta}}\left[N^{1-\beta} - L_r^{1-\beta}T\right]\frac{1}{\sigma}\varphi\left(\frac{\epsilon}{\sigma}\right) d\epsilon\\
&= \frac{XN}{L_{min}^{1-\beta}}\frac{1}{\sigma n}\left[N^{1-\beta} - L_r^{1-\beta}T\right]\varphi\left(\frac{\ln(n)-\ln(N)}{\sigma}\right).
\end{aligned} \tag{2.24}$$

The right panel of Figure 2.3 provides an example for the capital-labor relation. The measurement error leads to smoothing of the relation at the cutoff, and a decrease in the visibility of the drop of capital above the cutoff. Furthermore, the quadratic functional form leads to a non-linear relation away from the cutoff, evident from the slight curvature.

[7] Appendix 2.A.3 provides additional calculation steps.

2.5.2 Estimation Results

I estimate the firm size density of equation (2.19) and the capital-labor relation of equation (2.22) using a two-step procedure.

As the first step, I use maximum-likelihood for the estimation of the firm size density. The estimation does no rely on parameters of the capital equation and can be estimated independently. I exclude observations with small or very large firm size. I vary the lower bound between 75 to 200 employees and set the upper bound to 10,000 employees. I adjust the firm size density to these truncation values. I also set the coefficient L_{min} to 20, since it is not well identified and is difficult to interpret.

Table 2.1 summarizes the estimation results. Each column presents a separate estimation according to the lower truncation value. The table shows that the lower truncation value matters with the coefficients in column (1) and (2) being particularly distinct from the rest. When taking the results of the first two columns at face value, the coefficient $\hat{T}$ suggests a large upward shift in the density of firms below the cutoff and $\hat{L_r}$ a wide interval of firms bunching at the cutoff. The large measurement error $\hat{\sigma}$, however, masks these large differences. For example, a measurement error of one standard deviation would imply differences between actual and observed firm size of around 50%.

I consider possible misspecification in the firm size distribution a likely factor for these surprising results. The firm distribution relies on the assumption of a power law, that would imply the same underlying parameter β for all firm sizes. In the estimation, however, the truncation value significantly affects $\hat{\beta}$ suggesting that the assumption of a power law over the whole distribution is incorrect. The change is particular strong when comparing column (1) and (2) to the rest of the table.

One way to reduce biases from misspecification in the power law is the selection of smaller firm size intervals over which the assumption of a power law holds true. I therefore focus on the coefficients in

Table 2.1: Empirical Firm Size Distribution – Maximum Likelihood Estimation

	Lower Truncation Value					
	75	100	125	150	175	200
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}$, power law	2.348	2.394	2.612	2.710	2.771	2.829
	(0.021)	(0.028)	(0.029)	(0.034)	(0.041)	(0.049)
$\hat{T}$, shift	0.410	0.451	0.883	0.924	0.944	0.961
distribution	(0.045)	(0.056)	(0.027)	(0.024)	(0.023)	(0.023)
$\hat{Lr}$, marginal	405.139	363.148	272.122	270.075	269.249	268.912
buncher	(96.991)	(102.396)	(6.375)	(5.214)	(4.912)	(4.976)
$\hat{\sigma}$,	0.537	0.523	0.096	0.078	0.071	0.067
measurement error	(0.055)	(0.068)	(0.020)	(0.015)	(0.014)	(0.015)
Observations	9,735	6,715	4,949	3,810	2,982	2,397

Note. Coefficients estimated by maximum likelihood based on the firm size distribution of equation (2.19). Each column includes firms up to a size of 10,000 employees. The lower truncation varies according to the specified value above. Standard errors are in parentheses.

column (3) to (6) that select firm size intervals closer to the cutoff and which are relatively stable. The coefficient of the power law $\hat{\beta}$ for these columns is between 2.612 and 2.829. The coefficient on the shift of the distribution $\hat{T}$ is between 0.883 and 0.961.[8] A value below one signifies an upward shift in the distribution of firms below the cutoff. This means that the scale effect in production is larger than the substitution effect in inputs ($\theta > \rho$). Turning to the estimate for the marginal buncher $\hat{L_r}$, coefficients are stable throughout and range from 268.912 to 272.122. In percentage terms, this means that the change in tax credit rate is large enough to induce a firm size reduction of up to 8% ($\frac{250-272}{272}$). Lastly, the measurement error $\hat{\sigma}$ ranges from 0.067 to 0.096. Thus, an error with the size of one standard deviation leads to differences between actual and observed firm size of 7-10%. Such a deviation seems reasonable given that the

[8] The parameter is defined as $T = \left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}}$.

only available employment measure in the data is the yearly average whereas the policy determined the value according to the beginning of the year.

As the second step, I take these coefficients and include them in the estimation of the capital-labor relation using non-linear least squares. I check for robustness by estimating the equation for the parameters from Table 2.1 column (5) and column (6) separately. As the dependent variable I use both the logarithm of total investment excluding leasing and the logarithm of total investment including leasing. The sample consists of all observations with a firm size between 75 and 1,000.

The upper panel in Table 2.2 summarizes the estimation results. Overall, coefficients are similar in magnitude across specifications. The results suggest that the quadratic functional form has merit with statistical significance of the $\hat{b}_2$ coefficient at the 10% level. To understand the behavior in investment at the cutoff, I calculate the predicted investment at the cutoff under a scenario of elevated tax credits ($A = \hat{a}_0 + \hat{a}_1 250 + \hat{a}_2 250^2$), a scenario of normal tax credits ($B = \hat{b}_0 + \hat{b}_1 250 + \hat{b}_2 250^2$) and for the average of the bunchers ($C = 250X$) and compare the results in the second panel. The log difference of investment between the different tax credit rates at the cutoff is between 35.3 to 36.9 log points. This is a sizable difference, but in terms of elasticity similar to Zwick and Mahon (2017).[9] Turning to the bunching firms, I find that they have on average 31.1 to 35.1 log points higher investments compared to firms receiving the elevated tax credit rate. This is an extreme response considering that it is the average and even the marginal buncher only reduces firm size by 8%.

[9] The result is also in line with Lerche (2018) who uses a differences-in-difference approach.

Table 2.2: Empirical Investment Behavior – Non-linear Least Squares

	Investment with Leasing		Investment w/o Leasing	
	(1)	(2)	(3)	(4)
$\hat{a_0}$, constant below	-62.245	-59.797	-49.194	-47.347
	(101.942)	(100.719)	(96.700)	(95.541)
$\hat{b_0}$, constant above	-2533.087	-2517.435	-2281.454	-2254.684
	(1174.988)	(1145.672)	(1138.807)	(1110.769)
$\hat{a_1}$, linear slope below	3.598	3.545	3.021	2.984
	(1.762)	(1.736)	(1.677)	(1.652)
$\hat{b_1}$, linear slope above	16.365	16.309	14.820	14.721
	(5.173)	(5.073)	(5.019)	(4.923)
$\hat{a_2}$, quadratic slope below	0.009	0.009	0.010	0.010
	(0.007)	(0.007)	(0.007)	(0.007)
$\hat{b_2}$, quadratic slope above	-0.009	-0.009	-0.008	-0.008
	(0.005)	(0.005)	(0.005)	(0.005)
$\hat{X}$, cutoff parameter	7.507	7.672	7.240	7.411
	(1.113)	(1.174)	(1.087)	(1.148)
Observations	8,784	8,784	8,686	8,686
Combination of coefficients				
$\ln A - \ln B$, shift due to capital costs	0.355	0.353	0.369	0.363
	(0.235)	(0.225)	(0.243)	(0.232)
$\ln C - \ln A$, shift due to bunching	0.311	0.329	0.330	0.351
	(0.202)	(0.204)	(0.205)	(0.207)

Note. Coefficients estimated by non-linear least squares based on equation (2.22). The sample includes firms with a size between 75 and 1,000 policy-relevant employees. Columns (1)–(2) use log total investment as dependent variable. Columns (3)–(4) provide a robustness test using log investment excluding investment based on leasing agreements. In columns (1) and (3), firm size parameters are taken from the specification in Table 2.1 with lower truncation value at 175. In columns (2) and (4), firm size parameters are taken from specification in Table 2.1 with lower truncation value at 200. The combinations of coefficients are $A = \hat{a_0} + \hat{a_1}250 + \hat{a_2}250^2$, $B = \hat{b_0} + \hat{b_1}250 + \hat{b_2}250^2$ and $C = 250\hat{X}$. Standard errors are in parentheses.

2.5.3 Recovering Model Parameters

Most of the model parameters are not directly estimated, only the parameters β and σ are. The additional parameters ρ, θ, λ, τ, γ and δ are part of a system of equations.

The model equations to recover the parameters are

$$\hat{\tau}^{\frac{1}{\hat{\rho}-1}} = \frac{\hat{A}}{\hat{B}} \tag{2.25}$$

$$\hat{\gamma}^{\frac{1}{\hat{\rho}-1}} = \frac{1}{\hat{\delta}}\hat{B} \tag{2.26}$$

$$\left(\frac{\hat{\lambda} + (1-\hat{\lambda})\hat{\gamma}^{\frac{\hat{\rho}}{\hat{\rho}-1}}}{\hat{\lambda} + (1-\hat{\lambda})(\hat{\gamma}\hat{\tau})^{\frac{\hat{\rho}}{\hat{\rho}-1}}}\right)^{\frac{1-\hat{\beta}}{1-\hat{\theta}}\frac{\hat{\rho}-\hat{\theta}}{\hat{\rho}}} = \hat{T} \tag{2.27}$$

$$\hat{L_r}^{1-\hat{\theta}}(\hat{\lambda}+(1-\hat{\lambda})\hat{\gamma}^{\frac{\hat{\rho}}{\hat{\rho}-1}})^{\frac{\hat{\rho}-\hat{\theta}}{\hat{\rho}}}(\hat{\lambda}N^{\hat{\rho}}+(1-\hat{\lambda})K_N^{\hat{\rho}})^{\frac{\hat{\theta}}{\hat{\rho}}}-\hat{\theta}\hat{\lambda}N-(1-\hat{\lambda})\hat{\theta}\hat{\tau}\hat{\gamma}K_N = \hat{L_r}(\hat{\lambda}+(1-\hat{\lambda})\gamma^{\frac{\hat{\rho}}{1-\hat{\rho}}}))(1-\hat{\theta}) \tag{2.28}$$

$$\frac{\int_{L=N}^{\hat{L_r}} K_N(L)(\hat{\beta}-1)L_{min}^{\hat{\beta}-1}L^{\hat{\beta}}dL}{L_{min}^{\hat{\beta}-1}\left[N^{1-\hat{\beta}} - \hat{L_r}^{1-\hat{\beta}}\left[\frac{T(1)}{T(\hat{\tau})}\right]^{\frac{1-\hat{\beta}}{1-\hat{\theta}}\frac{\hat{\rho}-\hat{\theta}}{\hat{\rho}}}\right]} = \hat{X} \tag{2.29}$$

In this setting there are 6 parameters of interest but only 5 equations for identification. The depreciation rate gets introduced as additional term into the estimation since I rely on investment as a proxy for capital. I solve the underidentification problem by setting up an auxiliary regression model that uses information on yearly depreciations available for a subset of firms to get an estimate for the depreciation rate independent of the equations above.

First, yearly depreciations are defined as

$$DA_{it} = \delta K_{it-1} \tag{2.30}$$

where DA_{it} is the amount of depreciation for firm i at time t.

By introducing depreciations into the law of motion in equation (2.13), the intertemporal relation of depreciation is

$$\begin{aligned} DA_{it} &= \delta K_{it-1} = \delta\left((1-\delta)K_{it-2} + I_{it-1}\right) \\ &= (1-\delta)DA_{it-1} + \delta I_{it-1}. \end{aligned} \tag{2.31}$$

Based on this equality, I estimate

$$\Delta DA_{it} = \alpha + \beta_1 DA_{it-1} + \beta_2 I_{it-1} + X_{it}'\beta_3 + \epsilon_{it}, \tag{2.32}$$

where X_{it} is an additional control variable and I assume that all unobserved variables have mean α and a residual term ϵ_{it}. Both β_1 and β_2 can identify the depreciation rate δ. A possible threat to identification stems from the fact, that I do not observe investment perfectly and that other types such as disinvestment are correlated with the regressors. In particular, different investment types are likely correlated with each other. I therefore focus on the coefficient β_1 that is likely to have little influence from unobserved investments. To further reduce concerns of bias, I control for negative investment due to sold machinery.

Table 2.3 summarizes the results. α captures the average influence of unobserved variables. The negative estimate suggests that there are indeed other factors such as negative investment that reduce the capital stock and therefore annual depreciations. The depreciation rate from β_1 is 0.197, meaning that close to 20% of the capital stock depreciates in any given year.

Using the estimated depreciation rate, I solve the system of equations. The system is highly non-linear and does not permit an analytical solution. Standard numerical root-finding algorithms were not capable of finding a sensible interior solution either. I therefore implement a version of the Metropolis-Hastings algorithm that

Table 2.3: Estimation of Depreciation Rate

	Total Investment (1)
α, constant	-138,835.5
	(34,285.87)
β_1, depreciation rate (minus)	-0.197
	(0.004)
β_2, depreciation rate (plus)	0.270
	(0.003)
β_3, control	-767.415
	(77.244)
Observations	29,062

Note. Coefficients estimated using OLS based on equation (2.32). Standard errors in parentheses.

searches the domain space of the parameters by applying directional and random search to find the set of parameters with the lowest sum of square value. I let the algorithm run for 10 million steps and redo the process several times to check the consistency of the results.

For the main analysis I calibrate the return to scale and policy cost parameter. I set θ to 0.85 following Atkeson and Kehoe (2005) and τ to 0.862 based on the tax credit rates of the policy.[10]

Table 2.4 summarizes the results from seven different runs of the optimization algorithm.[11] The sum of squared residuals is similar throughout the runs ranging from 164.15 to 164.55. The residual term is driven by the constraint of equation (2.26) that generates the largest value. SSR is lowest in column (6) for which I recover an elasticity of substitution of 0.083. This value is far lower than one and is evidence that firms do not have Cobb-Douglas production functions. The elasticity of substitution is similar for all other runs

[10] In 1999, the difference in capital costs of firms below and above the cutoff was 0.889 and in 2000–2004, it was 0.857.

[11] For inference, standard errors can be computed using bootstrapping.

Table 2.4: Model Parameters with Parameter Calibration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\hat{\rho}$	-10.893	-9.207	-10.704	-10.551	-9.561	-10.994	-9.596
$\hat{\omega}$	0.084	0.098	0.085	0.087	0.095	0.083	0.094
$\hat{\theta}$ (fixed)	0.850	0.850	0.850	0.850	0.850	0.850	0.850
$\hat{\lambda}$	0.372	0.341	0.527	0.360	0.772	0.462	0.598
$\hat{\tau}$ (fixed)	0.862	0.862	0.862	0.862	0.862	0.862	0.862
$\hat{\gamma}$	4.52e-11	1.35e-09	6.62e-11	9.01e-11	6.62e-10	3.68e-11	6.17e-10
SSR	164.16	164.55	164.22	164.38	164.46	164.15	164.45

Note. Parameters recovered from estimated coefficients using Metropolis-Hastings optimization for 10,000,000 steps. The elasticity of substitution is recovered from $e_{kl} = \frac{1}{1-\rho}$.

with the value ranging from 0.084 to 0.098. These findings are in line with the firm-level evidence in the previous literature that find values significantly below one as well. For example, Moreau (2019) finds a value of 0.12 at the firm level.

I also recover the parameter λ that measures the labor share in production costs. In column (6), the value is 0.462 and in the other columns values range from 0.341 to 0.772. The recovered value fluctuates significantly suggesting that the objective function is relatively flat in terms of λ. Nevertheless, the model does not use information on the labor share and recovers values that are somewhat close to the typical labor share of 66%.

As a robustness check I redo the optimization without calibrating any of the parameters. Table 2.5 summarizes the results. The sum of squared residuals is again similar throughout the runs ranging from 154.127 to 154.941. The SSR is lower since there are two more degrees of freedom for optimization. SSR is lowest in column (5) for which I recover an elasticity of substitution of 0.194. This value is larger than the elasticity recovered using calibrated parameters, but still much lower than 1. The values in all other runs are similar, with the lowest value of 0.098 and the highest value of 0.243.

Table 2.5: Model Parameters without Parameter Calibration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\hat{\rho}$, substitution param.	-9.208	-6.293	-6.598	-5.536	-4.157	-3.113	-7.345
$\hat{e_{kl}}$, substitution elast.	0.098	0.137	0.132	0.153	0.194	0.243	0.120
$\hat{\theta}$, scale parameter	0.492	0.314	0.574	0.378	0.985	0.972	0.042
$\hat{\lambda}$, labor share	0.299	0.699	0.538	0.491	0.279	0.623	0.121
$\hat{\tau}$, tax credit	0.992	0.999	1.000	0.999	0.999	0.999	0.999
$\hat{\gamma}$, normalization term	1.18e-9	4.23e-7	2.27e-7	1.94e-6	3.06e-5	2.51e-4	5.13e-8
SSR	154.569	154.762	154.481	154.767	154.127	154.167	154.941

Note. Parameters recovered from estimated coefficients using Metropolis-Hastings optimization for 10,000,000 steps. The elasticity of substitution is recovered from $e_{kl} = \frac{1}{1-\rho}$.

However the other parameters suggest discrepancies of the model fit. In particular, the recovered tax credit rate $\hat{\tau}$ has a value of 0.999. This would suggest that the policy did not have an effect on capital costs of firms, counter to the descriptive evidence that displays bunching of firms and adjustments in investment behavior at the cutoff. The other parameters are more in line with expectations although fluctuations are sizable.

2.6 Conclusion

The elasticity of substitution between capital and labor is highly relevant for microeconomic and macroeconomic research questions. This paper provides novel evidence for the case of Germany using quasi-experimental variation in a size-dependent investment tax credit policy. The German government provided a reduction in investment costs to manufacturing firms in East Germany for over 20 years through a tax credit program. Since 1994 the tax credit rate was size-dependent and for the period of analysis 1999–2004 the rate was higher for firms up to a simple size cutoff of 250 employees. I set up a theoretical framework by adding size-dependent capital costs to a

Lucas (1978) model to show the possibility of distortions introduced by such a policy design. Such distortions are evident in the descriptive evidence with bunching of firms at the cutoff and adjustments in the investment behavior. I structurally estimate the model using maximum likelihood and non-linear least squares, and recover the model parameters. I find an elasticity of substitution of 0.083 and a range of possible value between 0.083 and 0.098. The estimated elasticity is significantly smaller than one and thus provides evidence that the Cobb-Douglas production function is not representative at the firm-level.

The paper makes various simplifying assumptions to estimate the model. In future research these can be relaxed to explain firm behavior in more detail. The model and estimation uses so far only cross-sectional information and thereby excludes any dynamic effects. However, capital adjustment costs are found to be important for investment decisions (Cooper and Haltiwanger, 2006) and can lead to differences in the estimation of short-run and long-run elasticities (Chirinko et al., 2011). A dynamic investment model of firms can shed light on both the elasticity of substitution and capital adjustment costs and the possible interaction between them. Such an approach can potentially uncover interesting firm behavior at the cutoff, for example the incentives of firms to hold back at the cutoff and invest in anticipation of moving above the cutoff in future time periods. It can be one explanation for the large investments of bunching firms that I find in the estimation. One first step in this direction is the incorporation of firm growth into the currently static estimation of the capital-labor relation.

2.A Appendices

2.A.1 Theoretical Firm Size Distribution

The productivity distribution as per the main text is

$$\phi(\alpha) = c_\alpha \alpha^{-\beta_\alpha} \tag{2.33}$$

where $c_\alpha > 0$ and $\beta_\alpha > 0$.

To derive the firm size distribution, I apply the change of variable formula

$$g(L) = \phi[r^{-1}(L)] \frac{d}{dL} r^{-1}(L), \quad L = r(\alpha) \text{ with r strictly inreasing} \tag{2.34}$$

where $r(\alpha)$ defines the relation between firm size and productivity.

The (preliminary) firm size distribution is

$$\chi(L) = \begin{cases} \frac{c_\alpha}{p}(1-\theta)\left(B\frac{\lambda\theta}{w}\right)^{\frac{\beta-1}{1-\theta}} T(\tau)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} L^{-\beta} & \text{if } L_{min} \leq L < N \\ \frac{1}{p}\int_{\alpha_c}^{\alpha_r} \phi(\alpha) d\alpha = \delta & \text{if } L = N \\ 0 & \text{if } N < L < L_r \\ \frac{c_\alpha}{p}(1-\theta)\left(B\frac{\lambda\theta}{w}\right)^{\frac{\beta-1}{1-\theta}} T(1)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} L^{-\beta} & \text{if } L_r \leq L \end{cases} \tag{2.35}$$

where the parameter p rescales the distribution to conform to the conditions of a probability density function, δ defines the mass of firms bunching at the cutoff and $\beta = \beta_\alpha(1-\theta)+\theta$.

The bunching mass δ is defined on the one hand by the mass of firms with productivity between α_N and α_r, and on the other hand as the residual of the other density terms. Solving the integral, the bunching mass is

$$
\begin{aligned}
&\delta = \int_{\alpha_N}^{\alpha_r} \phi(\alpha) d\alpha = \\
&\int_{\alpha_N^{\frac{1}{1-\theta}}\left(B\frac{\theta\lambda}{w}\right)^{\frac{1}{1-\theta}} T(\tau)^{\frac{\theta-\rho}{\rho(1-\theta)}}}^{\alpha_r^{\frac{1}{1-\theta}}\left(B\frac{\theta\lambda}{w}\right)^{\frac{1}{1-\theta}} T(\tau)^{\frac{\theta-\rho}{\rho(1-\theta)}}} \frac{c_\alpha}{p}(1-\theta)\left(B\frac{\lambda\theta}{w}\right)^{\frac{\beta-1}{1-\theta}} T(\tau)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} L^{-\beta} dL \\
&\qquad = \int_N^{L_r\left[\frac{T(1)}{T(\tau)}\right]^{\frac{\rho-\theta}{\rho(1-\theta)}}} \frac{c_\alpha}{p}(1-\theta)\left(B\frac{\lambda\theta}{w}\right)^{\frac{\beta-1}{1-\theta}} T(\tau)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} L^{-\beta} dL \\
&= \frac{c_\alpha}{p}(1-\theta)\left(B\frac{\lambda\theta}{w}\right)^{\frac{\beta-1}{1-\theta}} \frac{1}{\beta-1} T(\tau)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} \left[N^{1-\beta} - L_r^{1-\beta}\left[\frac{T(1)}{T(\tau)}\right]^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}}\right],
\end{aligned} \tag{2.36}
$$

and solving for the residual term, the bunching mass is

$$
\begin{aligned}
\delta = 1 &- \int_{L_{min}}^{N} \frac{c_\alpha}{p}(1-\theta)\left(B\frac{\lambda\theta}{w}\right)^{\frac{\beta-1}{1-\theta}} T(\tau)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} L^{-\beta} dL \\
&- \int_{L_r}^{\infty} \frac{c_\alpha}{p}(1-\theta)\left(B\frac{\lambda\theta}{w}\right)^{\frac{\beta-1}{1-\theta}} T(1)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} L^{-\beta} dL \\
&= 1 - \frac{c_\alpha}{p}(1-\theta)\left(B\frac{\lambda\theta}{w}\right)^{\frac{\beta-1}{1-\theta}} \frac{1}{\beta-1} \\
\cdot &\left[T(\tau)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} L_{min}^{1-\beta} - T(\tau)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} N^{1-\beta} + T(1)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} L_r^{1-\beta}\right]
\end{aligned} \tag{2.37}
$$

These terms allow to define the firm size distribution explicitly. Setting the terms equal for δ, the equality simplifies to

$$
\frac{c_\alpha}{p}(1-\theta)\left(B\frac{\lambda\theta}{w}\right)^{\frac{\beta-1}{1-\theta}} = (\beta-1) T(\tau)^{\frac{1-\beta}{1-\theta}\frac{\theta-\rho}{\rho}} L_{min}^{\beta-1}. \tag{2.38}
$$

The firm size distribution is then

$$\chi(L^*) = \begin{cases} (\beta-1)L_{min}^{\beta-1}L^{*-\beta} & \text{if } L_{min} \leq L^* < N \\ L_{min}^{\beta-1}\left[N^{1-\beta} - L_r^{1-\beta}\left[\frac{T(1)}{T(\tau)}\right]^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}}\right] & \text{if } L^* = N \\ 0 & \text{if } N < L^* < L_r \\ (\beta-1)L_{min}^{\beta-1}\left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} L^{*-\beta} & \text{if } L_r \leq L^* \end{cases} \tag{2.39}$$

2.A.2 Empirical Firm Size Distribution

To adjust the theoretical model to the empirical predictions I assume differences in observed compared to actual firm size due to measurement error. The relation is defined as

$$L(\alpha, \epsilon) = L^*(\alpha)e^{\epsilon}, \tag{2.40}$$

where L^* is the actual equilibrium firm size , $L(\alpha, \epsilon)$ is the observed firm size and ϵ is Gaussian noise with mean zero and variance σ^2.

To derive the observed empirical firm size distribution, as a first step I consider the conditional cdf

$$P(L < n|\epsilon)$$
$$= \begin{cases} 0 & \text{if } ne^{-\epsilon} \le L_{min} \\ \frac{\beta-1}{L_{min}^{1-\beta}} \int_{L_{min}}^{ne^{-\epsilon}} L^{*-\beta} dL^* & \text{if } L_{min} \le ne^{-\epsilon} < N \\ \frac{\beta-1}{L_{min}^{1-\beta}} \int_{L_{min}}^{N} L^{*-\beta} dL^* + \frac{1}{L_{min}^{1-\beta}} \left[N^{1-\beta} - L_r^{1-\beta} \left[\frac{T(1)}{T(\tau)}\right]^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} \right] & \text{if } N \le ne^{-\epsilon} < L_r \\ \dots + \frac{\beta-1}{L_{min}^{1-\beta}} \left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} \int_{L_r}^{ne^{-\epsilon}} L^{*-\beta} dL^* & \text{if } L_r \le ne^{-\epsilon} \end{cases} \tag{2.41}$$

$$= \begin{cases} 0 & \text{if } \ln(n) - \ln(L_{min}) \le \epsilon \\ 1 - \left(\frac{ne^{-\epsilon}}{L_{min}}\right)^{1-\beta} & \text{if } \ln(n) - \ln(N) < \epsilon \le \ln(n) - \ln(L_{min}) \\ 1 - \left(\frac{L_r}{L_{min}}\right)^{1-\beta} \left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} & \text{if } \ln(n) - \ln(L_r) < \epsilon \le \ln(n) - \ln(N) \\ 1 - \left(\frac{ne^{-\epsilon}}{L_{min}}\right)^{1-\beta} \left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} & \text{if } \epsilon \le \ln(n) - \ln(L_r) \end{cases} \tag{2.42}$$

I can then compute the unconditional cdf by integrating over the distribution of ϵ to obtain

$$\begin{aligned} P(L < n) = \int_{\mathbb{R}} P(L < n|\epsilon) \frac{1}{\sigma} \varphi\left(\frac{\epsilon}{\sigma}\right) d\epsilon \\ = \int_{\ln(n)-\ln(N)}^{\ln(n)-\ln(L_{min})} \left[1 - \left(\frac{ne^{-\epsilon}}{L_{min}}\right)^{1-\beta}\right] \frac{1}{\sigma} \phi\left(\frac{\epsilon}{\sigma}\right) d\epsilon \\ + \int_{\ln(n)-\ln(L_r)}^{\ln(n)-\ln(N)} \left[1 - \left(\frac{L_r}{L_{min}}\right)^{1-\beta} \left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}}\right] \frac{1}{\sigma} \varphi\left(\frac{\epsilon}{\sigma}\right) d\epsilon \\ + \int_{-\infty}^{\ln(n)-\ln(L_r)} \left[1 - \left(\frac{ne^{-\epsilon}}{L_{min}}\right)^{1-\beta} \left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} e^{-\epsilon(1-\beta)}\right] \frac{1}{\sigma} \varphi\left(\frac{\epsilon}{\sigma}\right) d\epsilon \end{aligned}$$

$$
\begin{aligned}
&= \Phi\left(\frac{\ln(n)-\ln(L_{min})}{\sigma}\right) - \left(\frac{L_r}{L_{min}}\right)^{1-\beta}\left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} \\
&\qquad \cdot\left[\Phi\left(\frac{\ln(n)-\ln(N)}{\sigma}\right) - \Phi\left(\frac{\ln(n)-\ln(L_r)}{\sigma}\right)\right] \\
&- \left(\frac{n}{L_{min}}\right)^{1-\beta} e^{\frac{\sigma^2}{2}(\beta-1)^2}\left[\Phi\left(\frac{\ln(n)-\ln(L_{min})}{\sigma} - \sigma(\beta-1)\right) - \Phi\left(\frac{\ln(n)-\ln(N)}{\sigma} - \sigma(\beta-1)\right)\right] \\
&\quad - \left(\frac{n}{L_{min}}\right)^{1-\beta} e^{\frac{\sigma^2}{2}(\beta-1)^2}\left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}}\Phi\left(\frac{\ln(n)-\ln(L_r)}{\sigma} - \sigma(\beta-1)\right),
\end{aligned}
\tag{2.43}
$$

where φ is the Gaussian pdf, Φ is the Gaussian cdf, and I use that $\frac{\partial}{\partial\epsilon}e^{\frac{\sigma^2}{2}(\beta-1)^2}\Phi(\frac{\epsilon}{\sigma} - \sigma(\beta-1)) = \frac{1}{\sigma}\varphi(\frac{\epsilon}{\sigma})e^{\epsilon(\beta-1)}$.

As a final step I take the derivate with respect to n to compute the corresponding pdf as

$$
\begin{aligned}
\chi(n) = \frac{1}{\sigma n}\frac{1}{L_{min}^{1-\beta}}&\left(N^{1-\beta} - \left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}} L_r^{1-\beta}\right)\varphi\left(\frac{\ln(n)-\ln(N)}{\sigma}\right) \\
&\qquad - (1-\beta)\left(\frac{1}{L_{min}}\right)^{1-\beta} n^{-\beta}e^{\frac{\sigma^2}{2}(\beta-1)^2} \\
&\cdot\left[\Phi\left(\frac{\ln(n)-\ln(L_{min})}{\sigma} - \sigma(\beta-1)\right) - \Phi\left(\frac{\ln(n)-\ln(N)}{\sigma} - \sigma(\beta-1)\right)\right] \\
-(1-\beta)\left(\frac{1}{L_{min}}\right)^{1-\beta} & n^{-\beta}e^{\frac{\sigma^2}{2}(\beta-1)^2}\left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}}\Phi\left(\frac{\ln(n)-\ln(L_r)}{\sigma} - \sigma(\beta-1)\right)
\end{aligned}
\tag{2.44}
$$

2.A.3 Empirical Capital-Labor Relation

$$
\begin{aligned}
E(K|L) = \int_{\mathbb{R}} K(L^*)f_{L^*|L}(L^*|L)dL^* &= \int_{\mathbb{R}} K(L^*)\frac{f(L^*,L)}{f_L(L)}dL^* \\
&= \frac{\int_{\mathbb{R}} K(L^*)f(L|L^*)f_{L^*}(L^*)dL^*}{f_L(L)},
\end{aligned}
\tag{2.45}
$$

where $f_{L^*|L}(L^*|L)$ is the distribution of the actual firm size conditional on the observed firm size, $f(L^*,L)$ is the joint distribution of actual and observed firm size.

I use these insides to implement the following computation similar to the one for the empirical firm size distribution. I first calculate the cumulative capital of firms that have less than a specific observed firm size conditioning on the measurement error in firm size. I then average over all types of measurement error to get to the unconditional cumulative capital of firms below the observed firm size. The increase of capital is driven by the amount of capital used by a specific firm and by the number of firms with the same observed firm size. To get to the expected capital use for a firm, I differentiate with respect to firm size and divide by firm density. As formula this is

$$E(K|L=n) = \frac{\frac{d}{dn}\int_{\mathbb{R}} C(K|L \leq n, \epsilon)\frac{1}{\sigma}\varphi\left(\frac{\epsilon}{\sigma}\right)}{f_L(L)}, \tag{2.46}$$

where $C(K|L \leq n, \epsilon)$ is the cumulative capital function conditional on the upper limit n and the measurement error ϵ. The denominator is the observed firm size distribution $\chi(L)$ and therefore defined as in equation (2.44).

Specifically the cumulative capital function is

$$C(K|L \leq n, \epsilon) \tag{2.47}$$

$$= \begin{cases} 0 & \text{if } ne^{-\epsilon} \leq L_{min} \\ k(L^*)\frac{\beta-1}{L_{min}^{1-\beta}}\int_{L_{min}}^{ne^{-\epsilon}} L^{*-\beta}dL^* & \text{if } L_{min} \leq ne^{-\epsilon} < N \\ k(L^*)\frac{\beta-1}{L_{min}^{1-\beta}}\int_{L_{min}}^{ne^{-\epsilon}} L^{*-\beta}dL^* + \frac{XN}{L_{min}^{1-\beta}}\left[N^{1-\beta} - L_r^{1-\beta}T\right] & \text{if } N \leq ne^{-\epsilon} < L_r \\ \cdots + k(L^*)\frac{\beta-1}{L_{min}^{1-\beta}}\int_{L_r}^{ne^{-\epsilon}} L^{*-\beta}dL^* & \text{if } L_r \leq ne^{-\epsilon} \end{cases} \tag{2.48}$$

To allow for flexibility in the functional form of capital and subsuming the interval of capital at the cutoff into an average, I assume that

$$k(L_i^*) = \begin{cases} a0 + a1L^* & \text{if } L_{min} \leq L^* < N \\ XN & \text{if } L^* = N \\ b0 + b1L^* & \text{if } L_r \leq L^*, \end{cases} \tag{2.49}$$

Since the terms are additive, I consider them separately as L^{*x} where x is the polynomial order. I compute

$$C^x(K|L \le n, \epsilon) \tag{2.50}$$

$$= \begin{cases} 0 & \text{if } \ln(n) - \ln(L_{min}) \le \epsilon \\ \frac{\beta-1}{\beta-x-1} \frac{a_x}{L_{min}^{1-\beta}} \left[L_{min}^{1+x-\beta} - (ne^{-\epsilon})^{1+x-\beta}\right] & \text{if } \ln(n)-\ln(N) < \epsilon \le \ln(n) - \ln(L_{min}) \\ \frac{\beta-1}{\beta-x-1} \frac{a_x}{L_{min}^{1-\beta}} \left[L_{min}^{1+x-\beta} - N^{1+x-\beta}\right] & \text{if } \ln(n)-\ln(L_r) < \epsilon \le \ln(n)-\ln(N) \\ \dots + \frac{\beta-1}{\beta-x-1} \frac{b_x}{L_{min}^{1-\beta}} T \left[L_r^{1+x-\beta} - (ne^{-\epsilon})^{1+x-\beta}\right] & \text{if } \epsilon \le \ln(n)-\ln(L_r), \end{cases} \tag{2.51}$$

where $T = \left(\frac{T(1)}{T(\tau)}\right)^{\frac{1-\beta}{1-\theta}\frac{\rho-\theta}{\rho}}$.

The unconditional cumulative capital is then

$$\begin{aligned} C^x(K|L < n) &= \int_{\mathbb{R}} C^x(K|L < n, \epsilon) \frac{1}{\sigma} \varphi\left(\frac{\epsilon}{\sigma}\right) d\epsilon \\ &= \int_{\ln(n)-\ln(N)}^{\ln(n)-\ln(L_{min})} \frac{\beta-1}{\beta-x-1} \frac{a_x}{L_{min}^{1-\beta}} \left[L_{min}^{1+x-\beta} - (ne^{-\epsilon})^{1+x-\beta}\right] \frac{1}{\sigma} \phi\left(\frac{\epsilon}{\sigma}\right) d\epsilon \\ &\quad + \int_{\ln(n)-\ln(L_r)}^{\ln(n)-\ln(N)} \frac{\beta-1}{\beta-x-1} \frac{a_x}{L_{min}^{1-\beta}} \left[L_{min}^{1+x-\beta} - N^{1+x-\beta}\right] \frac{1}{\sigma} \varphi\left(\frac{\epsilon}{\sigma}\right) d\epsilon \\ &\quad + \int_{-\infty}^{\ln(n)-\ln(L_r)} \dots + \frac{\beta-1}{\beta-x-1} \frac{b_x}{L_{min}^{1-\beta}} T \left[L_r^{1+x-\beta} - (ne^{-\epsilon})^{1+x-\beta}\right] \frac{1}{\sigma} \varphi\left(\frac{\epsilon}{\sigma}\right) d\epsilon \\ &= \frac{\beta-1}{\beta-x-1} \frac{a_x}{L_{min}^{1-\beta}} L_{min}^{1+x-\beta} \Phi\left(\frac{\ln(n)-\ln(L_{min})}{\sigma}\right) \\ &\quad - \frac{\beta-1}{\beta-x-1} \frac{a_x}{L_{min}^{1-\beta}} n^{1+x-\beta} e^{\frac{\sigma^2}{2}(\beta-x-1)^2} \\ &\quad \cdot \left[\Phi\left(\frac{\ln(n)-\ln(L_{min})}{\sigma} - \sigma(\beta-x-1)\right) - \Phi\left(\frac{\ln(n)-\ln(N)}{\sigma} - \sigma(\beta-x-1)\right)\right] \\ &\quad - \frac{\beta-1}{\beta-x-1} \frac{a_x}{L_{min}^{1-\beta}} N^{1+x-\beta} \Phi\left(\frac{\ln(n)-\ln(L_r)}{\sigma}\right) + \frac{\beta-1}{\beta-x-1} \frac{b_x}{L_{min}^{1-\beta}} T L_r^{1+x-\beta} \Phi\left(\frac{\ln(n)-\ln(L_r)}{\sigma}\right) \\ &\quad - \frac{\beta-1}{\beta-x-1} \frac{b_x}{L_{min}^{1-\beta}} T n^{1+x-\beta} e^{\frac{\sigma^2}{2}(\beta-x-1)^2} \Phi\left(\frac{\ln(n)-\ln(L_r)}{\sigma} - \sigma(\beta-x-1)\right), \end{aligned} \tag{2.52}$$

where I use $e^{\frac{\sigma^2}{2}(\beta-p-1)^2}\Phi(\frac{\epsilon}{\sigma}-\sigma(\beta-p-1)) = \frac{1}{\sigma}\varphi(\frac{\epsilon}{\sigma})e^{\epsilon(\beta-p-1)}$ for any p.

By taking the derivative I get to the slope of cumulative capital

$$
\begin{aligned}
c^x(K|L=n) = \frac{dC^x(K|L<n)}{dn} = & \\
& -(1-\beta)\frac{a_x}{L_{min}^{1-\beta}}n^{x-\beta}e^{\frac{\sigma^2}{2}(\beta-x-1)^2} \\
\cdot \left[\Phi\left(\frac{\ln(n)-\ln(L_{min})}{\sigma}-\sigma(\beta-x-1)\right)-\Phi\left(\frac{\ln(n)-\ln(N)}{\sigma}-\sigma(\beta-x-1)\right)\right] & \\
-(1-\beta)\frac{b_x}{L_{min}^{1-\beta}}Tn^{x-\beta}e^{\frac{\sigma^2}{2}(\beta-x-1)^2}\Phi\left(\frac{\ln(n)-\ln(L_r)}{\sigma}-\sigma(\beta-x-1)\right) & \quad (2.53)
\end{aligned}
$$

As a final step I consider firms with actual firm size at the cutoff that I have excluded so far from the computations:

$$
\begin{aligned}
c^+(K|L=n) = \frac{C^+(K|L<n)}{dn} &= \frac{d}{dn}\int_{-\infty}^{\ln(n)-\ln(L_r)} \frac{XN}{L_{min}^{1-\beta}}\left[N^{1-\beta}-L_r^{1-\beta}T\right]\frac{1}{\sigma}\varphi\left(\frac{\epsilon}{\sigma}\right)d\epsilon \\
&= \frac{XN}{L_{min}^{1-\beta}}\frac{1}{\sigma n}\left[N^{1-\beta}-L_r^{1-\beta}T\right]\varphi\left(\frac{\ln(n)-\ln(N)}{\sigma}\right) \quad (2.54)
\end{aligned}
$$

By putting all pieces together, the expected capital conditional on observed firm size is

$$
E(K|L) = \frac{c^0(K|L)+c^1(K|L)+c^+(K|L)}{\chi(L)} \quad (2.55)
$$

2.A.4 Marginal Buncher

The estimation establishes a value for the firm size of the marginal buncher. The value is driven by the underlying model parameters. In particular, the marginal buncher depends on the profit at the cutoff.

At the cutoff, profits for the marginal buncher are

$$
\max_{K_i} p_N\alpha_r f(N,K_i) - wN - r\tau K_i = p_r\alpha_r f(L_r,K_r) - wL_r - rK_r, \quad (2.56)
$$

equal to equation (2.7) where K_i is unknown.

For the maximization problem I substitute the productivity α_r according to the optimal labor formula.

$$\max_{K_i} L_r^{1-\theta} \left(\frac{\theta\lambda}{w}\right)^{-1} \left(\lambda + (1-\lambda)\left(\frac{1}{\gamma}\right)^{\frac{\rho}{1-\rho}}\right)^{\frac{\rho-\theta}{\rho}} f(N, K_i)^{\theta} - wN - r\tau K_i, \quad (2.57)$$

This leads to the first order condition that implicitly determines the level of capital at the cutoff for the marginal buncher K_{N_r} as

$$FOC : L_r^{1-\theta}(\gamma\tau)^{-1} \left(\lambda + (1-\lambda)\left(\frac{1}{\gamma}\right)^{\frac{\rho}{1-\rho}}\right)^{\frac{\rho-\theta}{\rho}} = \left(\lambda N^{\rho} + (1-\lambda)K_{N_r}^{\rho}\right)^{\frac{\rho-\theta}{\rho}} K_{N_r}^{1-\rho} \quad (2.58)$$

I simplify the profit at L_r by substituting optimal capital and the productivity term to compute profits in terms of firm size as

$$\pi(L_r) = \frac{L_r w}{\theta\lambda} \left(\lambda + (1-\lambda)\left(\frac{1}{\gamma}\right)^{\frac{\rho}{1-\rho}}\right)(1-\theta) \quad (2.59)$$

Expanding both profit equations by a constant $\frac{1}{1-\lambda}$ to get rid of wage w and capital rental rate r the final system of equations is

$$\begin{aligned}
&\frac{1}{1-\lambda} L_r^{1-\theta} \left(\lambda + (1-\lambda)\left(\frac{1}{\gamma}\right)^{\frac{\rho}{1-\rho}}\right)^{\frac{\rho-\theta}{\rho}} \left(\lambda N^{\rho} + (1-\lambda)K_{N_r}^{\rho}\right)^{\frac{\theta}{\rho}} - \theta\frac{\lambda}{1-\lambda}N - \theta\tau\gamma K_{N_r} \\
&\qquad = \frac{1}{1-\lambda} L_r \left(\lambda + (1-\lambda)\left(\frac{1}{\gamma}\right)^{\frac{\rho}{1-\rho}}\right)(1-\theta) \\
&FOC : L_r^{1-\theta}(\gamma\tau)^{-1} \left(\lambda + (1-\lambda)\left(\frac{1}{\gamma}\right)^{\frac{\rho}{1-\rho}}\right)^{\frac{\rho-\theta}{\rho}} = \left(\lambda N^{\rho} + (1-\lambda)K_{N_r}^{\rho}\right)^{\frac{\rho-\theta}{\rho}} K_{N_r}^{1-\rho}
\end{aligned} \quad (2.60)$$

Chapter 3

OCCUPATIONAL RECOGNITION AND IMMIGRANT LABOR MARKET OUTCOMES

with Herbert Brücker, Albrecht Glitz and Agnese Romiti

3.1 Introduction

It is a well documented fact in most developed economies that immigrants perform significantly worse in the labor market than their native counterparts (see, e.g., Dustmann and Frattini, 2013). In many cases, the main reason appears to be a lack of human capital, which pushes immigrants into low paying and precarious jobs and prohibits them from moving into more desirable segments of the labor market. However, even when immigrants accumulated valuable skills in their countries of origin prior to migration, the transferability of these skills to the host country economy is often problematic, partly because of insufficient language skills (Chiswick and Miller, 2003), partly because of the limited signaling function of foreign qualifications which makes it difficult for native employers to assess immigrants' occupational skills.[1] In addition, legal restrictions often prohibit immigrants from working in certain occupations (Sweetman et al., 2015). Kleiner (2017), for instance, reports that the share of the US workforce holding an occupational license increased from less than 5 percent in the 1950s to about 25 percent in 2015. Koumenta and Pagliero (2016) document a similarly important role of occupational regulation in the EU, where the share of the workforce with a license reached 22 percent in 2015, with Denmark ranking lowest (14 percent) and Germany ranking highest (33 percent).

While occupational regulation is meant to ensure a minimum quality standard within a profession (e.g. Leland, 1979, Bryson and Kleiner, 2010), its prevalence is likely to have a particularly detrimental effect on the labor market outcomes of immigrants. Without formal recognition of their foreign qualifications, immigrants would often not be able to work in licensed occupations nor would they be able to credibly signal their occupational skills to native employers, who are all too often unfamiliar with the skill content of foreign qualifications. This may lead to an underutilization of immigrants' skills as suggested by the widespread occupational downgrading immigrants experience in many labor markets after arrival (see, for example, Friedberg, 2001, for Israel, Mattoo et al., 2008, for the US, and Dustmann et al., 2013, for the UK). Facilitating the recognition of foreign qualifications might be a way to overcome this inefficiency and fundamentally improve the economic integration of immigrants in their host countries.

In this paper, we estimate the impact of occupational recognition on immigrants' labor market outcomes. To obtain recognition for their foreign credentials, immigrants in Germany are required to go through a formal process, at the end of which, if successful,

[1] One manifestation of the low transferability of human capital are the remarkably low returns to foreign education and experience observed in many destination countries (see Dustmann and Glitz, 2011, for a comprehensive overview of this literature).

the responsible authorities certify the equivalence between the immigrants' foreign qualification and its German counterpart. From a labor market perspective, occupational recognition affects labor market outcomes through two main mechanisms. First, a successful recognition gives the immigrants access to segments of the labor market that they could previously not enter. These regulated segments tend to be characterized by high wages, both because of high returns to skills and because of monopoly rents from occupational licensing (see e.g. Stigler, 1971, Kleiner and Krueger, 2010, 2013, or Gittleman et al., 2018).[2] Second, occupational recognition reduces uncertainty about the skills of immigrant workers, which allows employers both in the regulated and unregulated segment of the labor market to better screen in the hiring process, leading to higher quality matches between workers and firms (Arrow, 1973, Spence, 1973). Both mechanisms thus suggest a positive impact of occupational recognition on immigrants' employment outcomes and wages.

Identifying the causal impact of occupational recognition is not straightforward due to self-selection on the part of the immigrants. Presumably, those immigrants who obtain occupational recognition would also perform comparatively well in the labor market if they had not received it, even conditional on other observable characteristics. This is because having obtained recognition reflects a specific set of skills that is likely to be generally valued in the labor mar-

[2]For evidence on the positive association between occupational licensing and wages in specific professions in the US, see Pagliero (2011) for lawyers, Timmons and Thornton (2008) for radiologic technologists, Timmons and Thornton (2010) for barbers, Thornton and Timmons (2013) for massage therapists, and Angrist and Guryan (2008) for teachers' certification. The positive wage effects, however, do not necessarily lead to a higher quality of the offered services as shown, for example, by Angrist and Guryan (2008) who find increases of 3-5 percent in the wages of teachers with state-mandated teacher testing in the US but no increase in the quality of teaching. Kleiner and Kudrle (2000) and Kleiner et al. (2014) come to similar conclusions for the dentistry and medical doctor professions, respectively, where more stringent licensing requirements lead to higher prices but no improvement in quality.

ket, both in the regulated and unregulated segment. In addition, immigrants who decide to go through the costly application process are likely to differ from those who do not in terms of unobservable characteristics such as ambition and motivation, factors that on their own would be associated with better labor market outcomes. We deal with these issues by exploiting a novel German data set that links detailed survey information on the exact timing of the application process for recognition with comprehensive social security data on the respondents' entire work histories in Germany. Taking advantage of the longitudinal dimension of our data, we estimate both static and dynamic difference-in-differences specifications, comparing the labor market outcomes of immigrants who obtain full recognition to those of immigrants who either never apply or have not yet received full recognition themselves. While the estimates from the static models allow us to assess the average effects of occupational recognition on labor market outcomes in our sample, the estimates from the dynamic specifications provide information on the precise evolution of the employment and wage effects over time.

Our empirical findings show substantial positive effects of occupational recognition on employment and wages. On average, immigrants in our sample who obtained full recognition in the past are 16.5 percentage points more likely to be employed and earn 15.1 percent higher wages than comparable immigrants who have either not applied or not yet received recognition themselves. We show that these employment effects are primarily driven by successful immigrants moving into occupations that were previously not accessible because of licensing restrictions. These movements into regulated occupations occur both out of non-employment and by workers moving from unregulated to regulated occupations.

Turning to the dynamic processes underlying these average effects, our estimates show that the probability of being employed relative to the control group increases rapidly with the receipt of occupational recognition, reaching 17.1 percentage points within the first twelve months. In subsequent years, the employment gap continues to widen, though at a lower pace, reaching a value of 24.5 percentage points three years after recognition. The wage gains from occupational recognition take a little longer to materialize but increase steadily after obtaining recognition, reaching 19.8 percent after three years. There is no evidence of any significant anticipation effects, neither in the employment nor in the wage regressions. The relative shift into the regulated segment of the labor market starts directly after recognition, primarily through movements out of non-employment. Movements from unregulated to regulated occupations, in contrast, only start intensifying with some delay.

Studying the heterogeneity of these effects across different subgroups of immigrants, our findings suggest that occupational recognition is beneficial for all groups considered. The effects on employment, wages, and access to regulated occupations are positive for all education levels and particularly large for individuals holding a foreign doctoral degree. When looking at the type of occupation for which individuals apply for recognition, our estimates are largest for the group of physicians, dentists, veterinarians and pharmacists for whom recognition is mandatory to practice their profession. However, occupational recognition improves the employment and wage outcomes also for those groups of workers who do not have mandatory recognition requirements, indicating that the certification of the quality of training received in the home country has an independent value in the German labor market.

While our administrative data do not allow us to analyze directly the quality of immigrants' work in regulated occupations vis-a-vis that of their native counterparts, we estimate standard earnings assimilation profiles in which we allow the speed of convergence to change with the recognition of immigrants' foreign qualifications. We show that earnings growth relative to natives accelerates after obtaining recognition, and that the earnings of immigrants who receive full recognition eventually fully converge to those of comparable natives, which could be interpreted as evidence for a similar quality in the services provided by immigrants and natives.

Our paper relates to the literature on the economic assimilation of immigrants (see, e.g., Borjas, 1995, or Lubotsky, 2007) in that it studies a specific mechanism through which immigrants may be held back in the host country's labor market. In comparison to this extensive literature, the evidence regarding the impact of occupational recognition on immigrant labor market outcomes is scarce.[3] Kugler and Sauer (2005) address this research question by exploiting the fact that Soviet trained physicians who immigrated to Israel in the early 1990s were exogenously assigned to different re-training tracks that differentially affected the probability of eventually obtaining a medical license. Their instrumental variable estimates show substantial monetary returns from obtaining a medical license of the order of 200 percent of monthly earnings within 3 to 4 years after arrival in Israel. Gomez et al. (2015) study the effect of occupational licensing on immigrant labor market outcomes in Canada, using annual data from the Survey of Labour and Income Dynamics (SLID). Controlling for time-invariant unobserved heterogeneity, their estimates show that immigrants receive a 20 log points earnings premium for working in a licensed occupation but are also 20 percent less likely to work in such an occupation than natives with similar observable characteristics. In similar regressions based on the Longitudinal Survey of Immigrants to Australia (LSIA), Tani (2018) finds that immigrants working in licensed occupations earn around 15 log points higher wages than comparable immigrants working in unlicensed occupations. Focussing more specifically on the role of occupational recognition on labor market outcomes, Chapman and Iredale (1993) find that immigrant men who unsuccessfully apply for recognition in Australia earn 15 to 30 percent lower wages than their successful counterparts, while Tani (2015) provides some evidence that the official assessment of immigrants' foreign educational degrees after arrival in Australia is associated with significantly higher wage rates.

[3]For an overview of the literature on occupational regulation and its interplay with the recognition of foreign qualifications, see Sweetman et al. (2015).

While the qualitative results of these studies are similar to some of ours, there are a number of important differences. First, rather than approaching the question of how occupational recognition affects immigrant labor market outcomes indirectly by studying the effects of working in a licensed occupation on wages, we analyze this question directly by focussing on the actual occupational recognition process. Since access to licensed occupations is only one channel through which occupational recognition can improve immigrants' labor market outcomes, our analysis thus provides a more comprehensive assessment of this important labor market institution. Second, apart from wages, we also consider employment and occupational mobility as distinct outcomes in our empirical analysis. Third, we analyze the effects of occupational recognition for a broader set of qualifications, including both post-secondary education and vocational training. Finally, we exploit unique information about the precise timing of the recognition process to estimate dynamic effects at monthly frequency, allowing us to identify both short- and long-run effects and to argue more convincingly for a causal relationship between occupational recognition and immigrants' labor market outcomes.

The paper is structured as follows. The next section describes the institutional setting in which the occupational recognition process takes place in Germany. Section 3.3 presents the empirical model and identification strategy. Section 3.4 describes our data set and provides some key summary statistics. Section 3.5 presents the main results together with a number of robustness checks and further supportive analysis. Section 3.6 links our findings to the earnings assimilation process of immigrants in Germany. Section 3.7 concludes the paper.

3.2 Institutional Setting

For an immigrant about to enter the German labor market, the distinction between regulated and unregulated occupations is of central importance. As many other European countries, Germany has a long tradition of regulated occupations dating back to medieval times. The entry and practice of regulated occupations is thereby governed by legal or administrative provisions that require proof of specific professional qualifications. Only individuals who have the required qualifications or, in the case of immigrants, obtained formal recognition of their foreign qualifications, are entitled to work in regulated occupations and use the corresponding professional job titles.[4] As of 2018, the regulated segment of the German labor market comprises 419 occupations (Bundesagentur für Arbeit, 2018), of which 29 percent are professions in the health sector (e.g. physicians, psychotherapists, pharmacists, nurses, physiotherapists), 27 percent professions in the technical sector (e.g. architects, engineers, physicists), 17 percent professions in the public sector (e.g. civil servants, policemen, firemen), 12 percent professions in the educational sector (e.g. teachers, educators, social workers), 7 percent professions in the transport sector (e.g. pilots), and 2 percent legal professions (e.g. lawyers, judges, attorneys).[5]

The authorities in charge of the recognition process for regulated occupations in Germany are very heterogeneous, depending on the particular occupation pursued. In the important health sector, the recognition of the degrees of physicians, dentists, pharmacists and nurses is regulated by governmental health authorities at the state

[4]In practice, occupational regulation can take many different forms with the literature mainly distinguishing between *registration*, *certification* and *licensure*. While there are no uniform definitions of these types of regulation, only licensure is generally viewed as being exclusionary in that it restricts access to certain occupations (see e.g. Kleiner and Krueger, 2013 or Sweetman et al., 2015). In distinguishing between regulated and unregulated occupations, we follow the German terminology which uses the terms *regulated occupation* and *licensed occupation* synonymously. For more details about the recognition process and the legal background in Germany, see https://www.anerkennung-in-deutschland.de.

[5]About three-quarters of the regulated occupations in Germany require an academic degree, sometimes in conjunction with further training. The remaining quarter of occupations require vocational training degrees or an occupational training in the public sector.

(*Länder*) level, in case of specialists (*Fachärzte*) additionally by the respective chambers. The entry to most occupations in the education sector, in turn, is regulated by educational authorities at the state level, and the entry to most regulated technical occupations by either governmental authorities or chambers, also at the state level. In contrast, in some selected occupations, for instance in the transport sector, the responsible authorities operate at the national level while for some occupations relevant for local authorities, the municipalities themselves are in charge of the recognition process.

In contrast to regulated occupations, formal recognition is not a precondition for the practice of unregulated occupations. Immigrants may work in these occupations without a license and thus without obtaining recognition for their foreign qualifications. For most unregulated occupations, however, immigrants can voluntarily apply for an assessment of their foreign qualifications. In case of a successful evaluation, the notice received at the end of this process serves as an official and legally secure document confirming the equivalence of the foreign qualification with the relevant German reference qualification. Examples of unregulated occupations where this type of certification is possible are so-called training occupations (e.g. office management clerks, mechanics or electricians) and advanced training occupations (e.g. master craftsman qualifications, certified advisors, certified senior clerks, specialist commercial clerks or business economists).[6] The most important authorities for the certification process of unregulated occupations are the chambers of industry and commerce (*Industrie- und Handelskammern*) and the chambers of crafts (*Handwerkskammern*). While the chambers of industry and commerce have set up a central authority at the national level responsible for the recognition of foreign qualifications, the chambers of crafts are organized at the state level.

[6] All training occupations, i.e. occupations for which training takes place within the dual system, are unregulated in Germany. In contrast, recognition is compulsory in order to work as a self-employed in some craft trades that require a license.

In order to apply for recognition, immigrants are not required to hold German citizenship or be in the possession of a residence permit for Germany. There is also no need to be living in Germany at the time of application, allowing immigrants to initiate the process while still being located abroad. Applications for occupational recognition need to be accompanied by extensive documentation: proof of identity, tabular summary of the training courses completed including previous occupational activity if relevant, proof of vocational qualification, proof of relevant occupational experience, evidence of other qualifications (e.g. continuing vocational training courses), a declaration of having not previously submitted an application, and evidence of the intention to work in Germany (which does not apply to nationals of the EU/EEA/Switzerland and persons residing in the EU/EEA/Switzerland). All documents must be submitted in German, with the relevant translations made by publicly authorized or certified interpreters or translators. Applications are subject to an administrative fee ranging between 100 and 600 euros depending on the occupation and the federal state in which the application is submitted. The costs of fees and other expenses, for instance for translations and certifications of documents, must be borne by the applicants themselves.[7] Since 2005, a proof of language proficiency can be made an additional requirement for the recognition of foreign credentials, as for example in the case of physicians.

These administrative features of the application process suggest that the bureaucratic hurdles to obtain occupational recognition in Germany are not negligible. According to our survey data, among

[7]In some circumstances, and on an individual case basis, these fees may be paid by other administrative entities. For example, prior to submitting an application, unemployed applicants or applicants registered as job seekers can seek clarification from their local employment offices or job centres whether they will cover the costs of the procedure. The labor administration authorities only provide such support if they consider the recognition of a foreign training qualification necessary for the holder to be integrated into the labor market. In these cases, adaptation measures such as continuing training courses or examination preparations may also be funded.

those immigrants who hold a foreign certificate and could therefore, in principle, apply for occupational recognition, only 35.8 percent end up doing so. The main reasons put forward for not applying are that a recognition is not considered important by the respondent (38.1 percent), that an application would have no chance of succeeding (12.9 percent), that the respondent does not know how to apply (6.6 percent), that the procedure is too bureaucratic or time-consuming (6.6 percent) and that important documentation is missing (4.6 percent). Monetary costs, in contrast, seem to constitute only a minor obstacle to applying (2.8 percent).

At the end of the recognition process, there are three possible outcomes: denial, partial recognition and full recognition.[8] In the case of partial recognition, which is a possible outcome only in the context of unregulated occupations, the assessment notification issued by the responsible authorities includes a detailed description of the existing qualifications as well as the knowledge that is still missing relative to the German reference qualification. The notification also provides concrete suggestions for training or apprenticeship measures which, if completed successfully, can then lead to a new application. A decision of full recognition, in turn, certifies the equivalence of the foreign qualification with the relevant German reference qualification and gives the worker full access to the relevant occupation and job title.

During most of our sample period, the recognition of European professional and vocational qualifications was regulated at the European level.[9] In contrast, for immigrants from third countries outside the EU, the EAA and Switzerland, there was no common official procedure regulating the recognition of foreign qualifications. In the absence of a legal basis, decisions on the equivalence between foreign

[8]For more details about the potential outcomes, see https://www.bq-portal.de/de

[9]The relevant legislation was the EU Directive 2005/36/EC on the recognition of professional qualifications, which came into force on 20 October 2005 and was introduced in Germany in 2007.

and German qualifications for this group of immigrants were more idiosyncratic, with the applicant's country of origin often playing a decisive role for the outcome of the application. This unsatisfactory situation largely motivated the introduction of the Federal Recognition Act (*Anerkennungsgesetz*) in April 2012 whose aim was to simplify, standardize and accelerate the procedure for the recognition of foreign qualifications governed by federal law, and open up such procedures to groups not covered by previous legislation.[10] However, 80 percent of immigrants in our sample applied for recognition before April 2012, so that our estimates largely reflect observations under the old legislative regime.

3.3 Empirical Framework

In the administrative component of our data set, we are able to continuously track immigrants after their arrival in Germany. We also know from the survey component if and when they receive occupational recognition. We exploit this information to compare the labor market outcomes of individuals after successful recognition with those of individuals who have either not yet received recognition or never applied for it. To facilitate the interpretation of our results, and because of limited sample sizes, we only consider full recognitions as successful and exclude individuals with partial or denied recognition.

[10] An additional shortcoming before the introduction of the Recognition Act was the absence of a binding time frame for processing the applications which lead to sometimes unnecessarily lengthy procedures. With the introduction of the Recognition Act, the maximum duration for the recognition process was mandated, with the responsible authorities now having to make a decision within 3 months of receipt of the applicant's full documentation (with a single extension possible in difficult cases). This acceleration of the recognition process is already noticeable in our sample, where the average duration between application and final decision was 5.5 months before the introduction of the Recognition Act (with a standard deviation of 13.4 months) and 3.8 months afterwards (with a standard deviation of 3.6 months).

Adopting a standard difference-in-differences approach, we start with the following fixed effects regression to obtain an overall estimate of the impact of recognition:

$$\mathrm{y}_{it} = \beta \mathrm{CertRecog}_{it} + X'_{it}\gamma + \lambda_t + \lambda_p + \lambda_i + \varepsilon_{it}. \tag{3.1}$$

The variable y_{it} denotes a specific labor market outcome of individual i at time t. In particular, we examine the impact of occupational recognition on an immigrant's employment, wages, and an index tracking the degree of regulation of the observed occupation (which we discuss in more detail in the next section). The first two outcomes provide general insights into the effects of occupational recognition on immigrants' labor market performance and are particularly important when viewed in the context of the rather poor employment and wage outcomes of immigrants, documented in much of the migration literature (for Germany, see, for example, Algan et al., 2010). The latter outcome is more specific to our setup and provides insights into the mechanism through which occupational recognition affects labor market outcomes. In particular, it sheds light on the central question whether occupational recognition indeed allows immigrants to move into regulated occupations. By running the regressions first without conditioning on immigrants' employment status, assigning a level of zero regulation to non-employment, and then conditional on employment, we are able to assess whether the movements into regulated occupations occur primarily out of non-employment or through gradual job changes from unregulated to regulated occupations.

The main regressor of interest, $\mathrm{CertRecog}_{it}$, is a dummy variable taking the value one if individual i has a foreign qualification that was recognized before or in time period t. For individuals who never apply, this value is zero for all time periods. We are interested in identifying β, the causal effect of occupational recognition on labor market outcomes. For this, we require that, in the absence of recognition, the outcomes of individuals who receive full recognition would have evolved in the same way as those of individuals who have either not yet applied or who never apply during our observation window.

Below we explain how we assess the validity of this crucial identification assumption based on observable differences in the pre-trends between treatment and control group. To control for general changes in labor market conditions, for example due to seasonal variation or business cycle fluctuations, we include time (month $\times$ year) fixed effects (λ_t) in our estimation of equation (3.1). We also add a full set of months since migration fixed effects (λ_p) which capture the dynamic evolution of immigrants' labor market outcomes as a result of their ongoing integration into the host country's economy. To account for time-invariant observable and unobservable heterogeneity, we further include a full set of individual fixed effects (λ_i). Their inclusion accounts for much of the personal characteristics associated with better labor market outcomes and the selection into the occupational recognition process, such as country of origin, gender, the level of education before migration, and time-invariant ability and motivation. In addition to the comprehensive set of fixed effects, we also control for a quadratic term in age[11] in the spirit of Mincerian wage equations and a proxy for German language proficiency (X_{it}) to capture further heterogeneity in the labor market trajectories of immigrants.[12] We cluster standard errors at the individual level as suggested for difference-in-differences estimations by Bertrand et al. (2004), thus allowing the error terms to be heteroscedastic and arbitrarily correlated over time for a given individual.

To evaluate the sensitivity of our results to changes in behavior after applying for recognition, we also include an indicator variable that switches on during the time period between initial application and final recognition in an alternative specification. It is possible that after submitting their application, individuals wait for the outcome of

[11]Since we include both individual and time fixed effects, the linear age effect is not separately identified.

[12]The survey provides information on self-reported language proficiency at two points in time, before migration and at the time of the interview. Linearly interpolating between the two data points, we construct proxies for language proficiency at monthly intervals.

the recognition process and, if unemployed, search less intensively for a new job or, if employed, stop working altogether or put less effort into their on-going jobs (and thus earn lower wages). On the other hand, being in the process of applying for occupational recognition may already serve as a positive signal in the labor market, improving applicants' labor market outcomes. By including the application dummy, we ensure that our estimate of β, which measures the change in the outcome variable after recognition relative to the control group, are not confounded by this type of anticipatory behavior.

While specification (3.1) provides a useful summary measure of the average impact of occupational recognition on employment, wages and the degree of regulation in immigrants' occupations, it conceals valuable information about the dynamic process through which the effects of recognition evolve over time. As an extension, we therefore introduce individual dummy variables for the months around the date of recognition as additional regressors, allowing us to distinguish between short- and long-term labor market effects in an event study type setup. More specifically, we use the regression model:

$$
\begin{aligned}
\mathrm{y}_{it} = &\sum_{q=-24}^{-1} \delta_{t-q}\mathrm{CertRecogMth}_{i,t-q} + \delta_{t+25}\mathrm{CertRecog}_{i,t+25} \\
&+ \sum_{q=1}^{60} \delta_{t-q}\mathrm{CertRecogMth}_{i,t-q} + \delta_{t-61}\mathrm{CertRecog}_{i,t-61} \\
&\qquad + X_{it}'\gamma + \lambda_t + \lambda_p + \lambda_i + \varepsilon_{it},
\end{aligned}
\tag{3.2}
$$

where the dummy variables $\mathrm{CertRecogMth}_{i,t-q}$, which equal one if individual i's qualification was recognized in period $t-q$, now capture the effect of occupational recognition in specific months around the recognition date. We create these dummy variables starting 24 months before the recognition date and ending 60 months thereafter. All dummy variables are equal to one only in the relevant time period and zero otherwise. For example, $\mathrm{CertRecogMth}_{i,t-10}$ is equal to one when the successful recognition was ten months before period t, so that the corresponding estimate δ_{t-10} measures the effect of recognition ten months after it was obtained. $\text{CertRecog}_{i,t-61}$ is a dummy variable for individuals having a foreign qualification that was recognized before or in period $t-61$. Thus, δ_{t-61} picks up the long-run average effect of recognition on labor market outcomes during all months more than five years after the recognition date. Similarly, $\text{CertRecog}_{i,t+25}$ is a dummy variable for all periods at least 25 months before an individual's recognition date. By definition, non-applicants get assigned zero for all these dummy variables. Importantly, equation (3.2) does not include a separate dummy variable for the time period when recognition was actually obtained ($q = 0$), so that the estimated dynamic effects of recognition are measured relative to this baseline period.[13] Just as for the static analysis, it is possible to control for the timing of the application by including a dummy for the application period as an additional regressor.

The main concern regarding our difference-in-differences approach is that unobserved time-varying factors related to both labor market outcomes and the recognition process might confound our estimation results. The inclusion of separate dummy variables for the months prior to recognition allows us to directly assess the relevance of this type of endogeneity as it would typically manifest itself through a violation of the parallel trends assumption. For instance, if some positive labor market shock (e.g. landing a new job) incentivizes an immigrant to apply for recognition (maybe because that would allow the worker to further advance in the new job), diverging trends in labor market outcomes relative to the control group should already materialize before the official recognition is received. Conversely, if in anticipation of a positive recognition outcome, applicants hold back in the labor market even before submitting their application, a deterioration in their labor market trajectories relative to non-applicants should show up in the pre-recognition period. The observation of

[13]Any level differences in outcomes between treatment and control group in the time period when recognition was obtained are absorbed by the individual fixed effects λ_i, so that the effect of recognition in this baseline period is essentially normalized to zero.

insignificant estimates close to zero in all months prior to the actual recognition date and significant effects moving away from zero soon after would lend support to a causal interpretation of our findings.

While the relatively small sample size of treated individuals with full recognition in our data prevents us from following alternative approaches for the estimation of dynamic treatment effects (see e.g. Fredriksson and Johansson, 2008, Crépon et al., 2009, or Vikström, 2017), we also use a pooled version of the synthetic control method developed by Abadie et al. (2010) to further check the robustness of our findings. In this approach, each immigrant who receives full recognition is matched to an appropriate control group of immigrants who never applied for recognition but whose labor market outcomes in the period prior to application are similar to those of the treated immigrant. Appendix 3.A provides more details on the implementation of this alternative procedure and documents the corresponding findings, which largely corroborate our main regression-based results.

3.4 Data

The basis of our empirical analysis are the first three waves of a novel longitudinal survey of people with migration background in Germany, the IAB-SOEP Migration Sample (Brücker et al., 2014). This survey, jointly conducted by the Institute for Employment Research (IAB) and the German Socio-Economic Panel (SOEP), was initiated in 2013 and designed to oversample recent immigrants who arrived in Germany after 1994.[14] The initial sample comprised around 5,000 first-

[14]The sampling of anchor persons proceeded as follows. In a first step, the IEB records were restricted to individuals who first appeared in the data after 1994. Individuals with a migration background were then identified based on their foreign, i.e. non-German, citizenship or their participation in measures of the Federal Employment Agency specifically designed for persons with a migration background (e.g. language classes). A short screening interview was then conducted with each cooperating anchor person after which around 30 percent of all households were screened out because anchor persons turned out not to be

and second-generation immigrants who were then interviewed on an annual basis, with a refreshment sample added every year to deal with sample attrition. The most innovative feature of this data set is its linkage with the German administrative data of the IEB (the so-called *Integrierte Erwerbsbiografie*), which comprise full employment histories of the universe of workers covered by the social security system in Germany during the period 1975 to 2014.[15] For data protection reasons, respondents to the survey component of the IAB-SOEP Migration Sample were asked to give their prior consent to the record linkage by signing a corresponding statement. The overall approval rate was about 50 percent, giving rise to a linked sample of 2,606 individuals: 1,992 from the first wave, 48 from the second wave, and 566 from the third wave. Out of this sample, we only consider first generation immigrants in our analysis and further exclude those individuals with missing information on the variables of interest.

The linked IAB-SOEP Migration Sample is particularly suited for our analysis for two reasons. First, the survey component contains detailed information on occupational qualifications obtained both before migration and after arrival in Germany. Importantly, this includes a full module devoted to the recognition process of foreign qualifications, with information about the month and year when the application process was initiated and the month and year when a final decision (denial, partial recognition, full recognition) was obtained.[16] Second, the social security component of the data allows us to observe an immigrant's entire work history after arrival in Germany. Linking the information about the precise timing of the recognition process to the spell structure of the administrative data, we can ob-

part of the target population. In more than half of the cases, screen-out was due to immigration before 1995 and in about one-third of the cases to not having a migration background. Note that other interviewed household members might have arrived in Germany before 1995.

[15]Civil servants, self-employed and military personnel are thus excluded from the IEB.

[16]There are also few cases where the status is pending and the individual still waiting for the result of the application. We exclude those cases from our analysis.

serve each individual's labor market outcomes before, during, and after the application process at monthly intervals.

We construct all our monthly outcome variables from the administrative spell data of the IEB. Employment is measured as the share of days during which an individual is in contractual employment in a given month (thus varying between 0 and 1).[17] Wages in the IEB are measured as log gross daily wages which we average across all full-time spells in a given month and translate into hourly wages by dividing by 8.[18] As indicated before, we also use an index tracking the degree of regulation in an immigrant's current occupation. The use of an index is necessary because even though each 8-digit occupation in the German system can be unambiguously classified as either regulated (licensed) or unregulated, occupations in the IEB data are not recorded at such fine level of disaggregation. We therefore employ the mapping constructed by Vicari (2014) in which, based on information from the full IEB-registry for the year 2012, each 3-digit occupation is assigned an index that represents the share of 8-digit subcategories within that occupation that requires a formal recognition of foreign qualifications in order to be accessible for immigrants. Weighting each 8-digit occupation by its relative size among the working population, the index ranges from zero (no subcategories requiring recognition) to one (all subcategories requiring recognition). We use this continuous index as a proxy for working in a regulated occupation.[19]

[17]The administrative data refer only to formal employment so that we cannot observe movements from informal to formal employment.

[18]Wages in the administrative data are right-censored at the social security contribution ceiling. This does not constitute a major issue in the context of this study since immigrants in Germany tend to earn wages well below the censoring limit.

[19]Note that if the distribution of immigrants with full recognition across 8-digit subcategories were the same as that of the existing working population, the interpretation of our parameter of interest β would be the same whether we use our continuous regulation index on the 3-digit level as the dependent variable or a binary measure on the 8-digit level for whether or not a specific occupation

Table 3.1: Regulated and Unregulated Occupations

	Index of Regulation (1)	Fraction of Working Pop. % (2)	Mean Wage (3)	Annual Rate of Wage Growth % (4)	First 3 Years Rate of Wage Growth % (5)
Panel A. First 10 Occupations with High Degree of Regulation					
Occupations in human medicine and dentistry	1.000	0.544	16.443	4.655	32.508
Occ. in veterinary medicine & non-medical animal health practitioners	1.000	0.020	12.335	4.808	22.104
Teachers in schools of general education	0.991	0.351	13.228	3.143	13.089
Occ. in police and criminal investigation, jurisdiction and penal institutions	0.875	0.038	9.270	2.586	10.782
Occupations in nursing, emergency medical services and obstetrics	0.760	2.223	9.458	3.523	19.827
Occupations in technical research and development	0.753	1.752	14.015	2.795	11.179
Occupations in construction scheduling and supervision, and architecture	0.708	0.816	13.907	2.786	13.534
Occupations in geriatric care	0.628	0.102	7.034	7.272	18.007
Occ. in education & social work, and pedagogic specialists in social care work	0.445	2.151	9.454	3.365	17.045
Ship's officers and masters	0.442	0.072	11.827	2.620	12.450
First 10 occupations (unweighted average)	0.760	0.807	11.697	3.755	17.052
Panel B. Last 10 Occupations with Low Degree of Regulation					
Sales occupations in retail trade (without product specialisation)	0.000	4.262	6.844	3.689	16.922
Driver of vehicles in road traffic	0.000	3.497	8.849	1.613	9.523
Occupations in metalworking	0.000	3.083	9.684	2.714	13.842
Trading occupations	0.000	1.581	11.048	3.747	16.606
Gastronomy occupations	0.000	1.230	5.443	3.552	13.457
Drivers & operators of construction & transportation vehicles, and equipment	0.000	0.793	9.896	1.779	6.780
Occupations in housekeeping and consumer counselling	0.000	0.653	5.977	2.637	11.331
Occupations in technical media design	0.000	0.419	10.684	3.150	17.591
Occupations in advertising and marketing	0.000	0.339	11.779	4.327	14.162
Occupations in hotels	0.000	0.272	7.074	4.007	14.243
Last 10 occupations (unweighted average)	0.000	1.613	8.728	3.121	13.446

Note: Data source: IEB data. Panel A refers to the first 10 occupations with the highest value of the regulation index. Panel B refers to the last 10 occupations with the lowest value of the regulation index. The index is provided by Vicari (2014) and is weighted according to the working population in each occupation in the full IEB registry in 2012. All descriptive values are computed using a 2 percent sample of the full (including immigrants and natives) IEB registry and refer to the years 1975-2014. Wages refer to the average real gross hourly wage considering all full-time spells. To mitigate the effect of outliers, we exclude the top and bottom 0.1 percentiles of the wage distribution. The rate of annual wage growth (column 4) refers to the within occupation relative difference in wages across two consecutive years. The first 3-year rate of wage growth (column 5) refers to the within occupation wage difference between the first and third year in a given occupation, relative to the first year wage.

To provide some examples, Table 3.1 reports the ten 3-digit occupations with the highest (Panel A) and lowest (Panel B) share of regulated 8-digit occupations.[20] Apart from the value of the regulation index, we report the fraction of the working population employed in each of these occupation, the average hourly wage in the occupation, the annual rate of wage growth and the rate of wage growth over the first three years in an occupation. The descriptive evidence shows that average wages in the ten occupations with the highest degree of regulation are significantly higher than average wages in the ten occupations with the lowest degree of regulation, 11.70 vs. 8.73 euros per hour. In addition, occupations with a higher degree of regulation are also characterized by faster wage growth. For example, those working in the ten most regulated occupations have an average annual (first 3-year) wage growth of 3.76 (17.05) percent compared to 3.12 (13.45) percent for those working in the ten least regulated occupations. These positive associations between wage levels and wage growth on the one hand and the degree of occupational regulation on the other hand is also more generally detectable in the data. For example, regressing occupation-specific log hourly wages and annual wage growth rates on the regulation index yields positive and highly significant coefficients of 0.425 (0.001) and 0.373 (0.020), respectively.

As mentioned above, we restrict our sample to foreign-born individuals who either eventually receive full recognition or never apply for recognition during our observation window.[21] Out of this group, we select all individuals who migrated to Germany aged 18 or older and who remained in Germany thereafter. We further only consider observations for prime working age individuals aged between 25 and 59 and exclude individuals with a known incapacity for work. Fi-

is regulated. In both cases, β would reflect the increase in the probability of working in a regulated occupation.

[20] The reported order of occupations is obtained after sorting by the index value and the fraction of the working population.

[21] The samples of immigrants whose application was denied (33) or who obtained only partial recognition (45) are too small to study separately in a meaningful way.

nally, we condition on having requested recognition before 2015 to be able to observe post-recognition outcomes in the administrative data (which end in 2014). Our final estimation sample consists of 1,218 individuals, of which 140 receive full recognition and 1,078 never apply for recognition, either because they do not have a foreign certificate with which to apply (568) or because they have one but choose not to apply (510).

Table 3.2 shows a number of descriptive statistics for our estimation sample which comprises individuals who receive full recognition (column 1) and individuals who did not apply for recognition (column 4). For completeness, we also report descriptive statistics for those in the survey who only received partial recognition (column 2) or were denied recognition (column 3). Focusing first on the full recognition sample, we see that 42.9 percent of the immigrants are men, aged 41.8 years on average in their last observable spell in our data. The schooling level of these immigrants is relatively high with 11.0 years of education (not counting tertiary education). The table also provides information about the typical migration and recognition process. On average, immigrants entered Germany when they were 31.3 years old. After that, they take on average about 8 months before making an official recognition request. One of the reasons for this delay could be the demanding recognition process which is one of the most important reasons reported by those deciding not to apply (12.9 percent), together with the lack of knowledge about how to apply (6.6 percent) and the bureaucratic and time-consuming nature of the process (6.6 percent). After on average 5.2 months, successful immigrants get to know the result of their application. However, as indicated by the large standard deviation of 12.1 months, there is significant variation in the waiting times.

Table 3.2 also provides information about each group's labor market outcomes, both during the first year after arrival in Germany and across all available time periods. In general, there are significant improvements in the employment rate between the first year and subsequent periods, particularly for those who applied for recogni-

Table 3.2: Descriptive Statistics by Recognition Outcome

	Full Recognition	Partial Recognition	Denied Recognition	Non-Applicant
Panel A. Immigrants				
Male %	42.9	48.5	33.3	46.6
	(49.7)	(50.8)	(47.7)	(49.9)
Yrs. Schooling	11.0	10.1	10.0	10.4
	(1.7)	(2.0)	(1.4)	(2.1)
Age Last Spell	41.8	43.1	44.9	41.2
	(9.6)	(8.5)	(8.5)	(9.6)
Age at first Migration	31.3	29.5	32.8	31.3
	(7.4)	(7.2)	(8.5)	(8.9)
Age at Request of Recognition	32.1	32.3	35.4	
	(7.5)	(9.8)	(9.2)	
Time Request to Result (Month)	5.2	12.2	4.1	
	(12.1)	(23.2)	(6.8)	
West %	9.3	0.0	2.2	12.2
	(29.1)	(0.0)	(14.9)	(32.7)
East Europe %	12.9	12.1	4.4	16.3
	(33.6)	(33.1)	(20.8)	(37.0)
South East Europe %	25.7	15.2	8.9	22.4
	(43.9)	(36.4)	(28.8)	(41.7)
USSR %	35.7	57.6	68.9	28.6
	(48.1)	(50.2)	(46.8)	(45.2)
Others %	16.4	15.2	15.6	20.6
	(37.2)	(36.4)	(36.7)	(40.5)
Panel B. Observations - First Year In Germany				
Employed %	29.7	13.0	7.6	31.2
	(45.7)	(33.7)	(26.5)	(46.3)
Index Regulation %	10.6	3.7	1.1	2.4
	(27.6)	(16.0)	(9.7)	(10.3)
Real Hourly Wage	12.7	7.5	5.2	9.0
	(5.2)	(2.4)	(2.7)	(5.4)
Panel C. Observations - Average Over Time				
Employed %	66.1	52.9	53.5	58.3
	(47.3)	(49.9)	(49.9)	(49.3)
Index Regulation %	13.7	11.9	7.7	3.8
	(27.4)	(25.4)	(22.0)	(12.4)
Real Hourly Wage	10.7	8.8	7.5	8.7
	(5.1)	(4.2)	(3.5)	(4.3)
Individuals	140	33	45	1,078

Note. Data source: IAB-SOEP Migration Sample linked to IEB data. Statistics depicted are means with standard deviations in parentheses. Statistics are based on individuals in upper panel and on monthly observations in the lower two panels. Employed % compares time periods of employment to times of employment and non-employment. Because information on regulated occupations is not available at the level of the single occupation, but only at the aggregate level of the regulation index provided by the IAB, each occupation has a degree of regulation corresponding to the regulation index ranging between 0 and 1. The table reports the average regulation index for the respective groups in the sample. Real hourly wages are constructed from daily wage information using only full-time spells and assuming that full-time employment is 8 hours per day.

tion. Average hourly wages for the full recognition and non-applicant group, in contrast, do not increase over time which is most likely due to strong positive selection into employment in the first year after arrival. When comparing across immigrant groups, there is substantial heterogeneity. Immigrants who obtain full recognition perform better in terms of wages relative to all other groups and in terms of initial employment relative to the two other applicant groups. They also tend to be younger when making their request than those immigrants whose application is eventually denied. Across all groups, the largest group in terms of country of origin are immigrants from the former USSR, mostly ethnic Germans, followed by immigrants from South East Europe. Given the heterogeneity in observable characteristics between the different immigrant groups, we analyze the robustness of our main results by replicating the analysis on the restricted sample of immigrants who eventually all received full recognition, thus only exploiting the differential timing of their recognition process for identification.

Unfortunately, until the third wave, the IAB-SOEP Migration Sample did not ask respondents explicitly for which specific occupation or field of study they requested recognition. If that information were available, we could separately study the labor market effects for regulated and unregulated occupations, which would allow us to distinguish the pure signalling effect of occupational recognition from the effect arising due to better access to certain occupations.[22] What we do observe in all three waves of the data, however, is the general type of certificate for which recognition is being requested, with the highest fraction applying for the recognition of a college/university degree (57.0 percent), followed by a vocational training (36.0 per-

[22]Table 3.A1 in the appendix reports the occupational distribution for the 38 respondents in the refreshment sample of the third wave who received full recognition of their qualifications. In line with official aggregate figures, most of these occupations are indeed regulated (71.1 percent) and require comparatively high skill levels, such as nurses and doctors (23.7 percent of recognitions), engineers (13.2 percent), veterinaries (10.5 percent) and teachers (7.9 percent).

cent), a doctoral degree (4.5 percent) and some other education (2.5 percent). In the first two waves, we also observe the type of authority to which immigrants applied for recognition, which can be used as a proxy for seeking recognition of a regulated or unregulated occupation (see Section 3.5.3).

3.5 Main Results

In this section, we first present estimates of the average impact of recognition on employment, wages and the regulation index and check the robustness of these findings to different sample definitions. We then graphically show the results from our dynamic specification, followed by an analysis of heterogeneous treatment effects in terms of immigrants' characteristics and key features of their recognition process.

3.5.1 Static Effects

In Panel A of Table 3.3, we report the static results from our baseline specification (3.1). In Panel B, we add a dummy that turns on during the application period as an additional control variable to deal with any potential anticipatory behavior on the part of the applicants. The estimate in column (1) of Panel A shows that obtaining full occupational recognition increases the share of days in employment per month by 16.0 percentage points, suggesting that occupational recognition helps immigrants find and maintain employment. In the specification including the dummy for the application period (Panel B), the effect of receiving full recognition increases slightly to 16.5 percentage points. The point estimate for having applied, in turn, is close to zero and statistically not significant, indicating that applying in itself neither serves as a positive signal in the labor market nor does it reduce employment outcomes, for example because of a lower job search intensity in anticipation of the final result of the

Table 3.3: Occupational Recognition and Average Labor Market Outcomes

	Employment (1)	Log Wages (Full-time) (2)	Regulation Index (3)	Regulation Index (Employed) (4)
Panel A				
Received full recognition	0.160***	0.157*	0.150***	0.114**
	(0.050)	(0.080)	(0.033)	(0.056)
Panel B				
Application period	0.024	-0.053	0.009	0.065
	(0.067)	(0.105)	(0.035)	(0.065)
Received full recognition	0.165***	0.141	0.152***	0.129*
	(0.052)	(0.103)	(0.035)	(0.068)
Individuals	1,218	830	1,218	1,081
with recognition	140	114	140	132
without recognition	1,078	716	1,078	949
Observations	136,306	50,971	129,471	74,003

Note. Data source: IAB-SOEP Migration Sample linked to IEB data. Panel A reports the estimates based on specification (3.1), Panel B adds a dummy variable for the application period as discussed in the text. The dependent variable is the share of days in employment per month in column (1), log real hourly wages for full-time employees averaged over all spells in a given month in column (2), the index of occupational regulation, assigning a value of zero to the non-employed, in column (3), and the index of occupational regulation in column (4). Additional controls are individual fixed effects, time fixed effects, time since migration fixed effects, age squared, and German language proficiency.
Standard errors in parentheses are clustered at the individual level: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

application.[23] In most of the following discussion of our findings, we nonetheless focus on the specification with an included dummy for the application period. Column (2) shows the results of occupational recognition for log wages. Full recognition increases wages by 17.0

[23] An observationally equivalent explanation would be that both effects exist but that they compensate each other.

percent (15.7 log points) according to Panel A and 15.1 percent (14.1 log points) according to Panel B, suggesting that recognition enables immigrants to more effectively utilize their skills in the host country's labor market. Note, however, that the coefficient in Panel B is not significant at conventional levels.

Column (3) shows that after recognition, immigrants move increasingly into more regulated jobs, with the regulation index of their occupations increasing by around 15 percentage points on average. Since, for this estimation, we keep non-employed immigrants in the sample and set their regulation indices equal to zero, some of the positive effect is likely driven by the significant movement from non-employment to employment shown in column (1). However, given a mean regulation index of 0.066 for employed immigrants without recognition (0.130 for the 90th percentile), the estimated coefficient is large, suggesting that part of the increase is also driven by movements from unregulated to regulated occupations. To investigate this possibility, we study the effect of occupational recognition on the regulation index conditional on being employed in column (4). For the subset of employed workers, full occupational recognition leads to a move into occupations that are on average 12.9 percentage points more likely to be regulated. The similarity between the results in the last two columns suggests that movements into more regulated occupations happen to a similar extent from non-employment and unregulated jobs.

Table 3.A3 in the appendix provides robustness checks with respect to our sample selection procedure by introducing additional restrictions one at a time. Column (4) restates the baseline results of Table 3.3 with our preferred and most restrictive sample. In column (1), we impose only the restriction of having migrated after the age of 18. Compared to our baseline results the effects are smaller, notably for the wage outcome. In column (2), we then exclude individuals who have an incapacity for work. The estimated effects of full recognition on employment, wages and the regulation index all increase somewhat, with the largest impact being on the employment

outcome where the estimate increases from 0.149 to 0.172. In column (3), we impose the additional restriction of only including observations for individuals of prime working age (age 25-59). This leads to a lowering of the employment effect towards our baseline estimate but otherwise only minor changes. Finally, we exclude individuals that migrated to Germany more than once in column (4) which leads to an increase in the estimate for log real wages. We exclude these individuals in our preferred specification since we do not know their labor market outcomes during their time outside of Germany. Overall, the particular sample selection rules do not seem to have a large impact on the magnitude of our main estimates.

Table 3.A4 in the appendix shows how our estimates of the impact of occupational recognition on the different labor market outcomes vary with the set of control variables included in the specification. After controlling for time since migration and individual fixed effects, the further inclusion of time fixed effects, the quadratic age profile and the German proficiency control has little impact on our point estimates.

3.5.2 Dynamic Effects

We now turn our attention to the results from the dynamic specification given in equation (3.2). In all reported estimations, we include a dummy for the application period and use the same sample restrictions as for the static main results in Table 3.3. For better readability, we plot the estimates of the period-specific effects δ_{t-q} graphically together with their corresponding 90% and 95% confidence intervals. Figure 3.1 displays the effects of occupational recognition on employment (upper left panel), log real wages (upper right panel), the regulation index including the non-employed (lower left panel) and the regulation index conditional on employment (lower right panel) in the 24 months before and 60 months after recognition.

In the months after receiving full recognition, the difference in the share of days per month in employment increases rapidly relative to

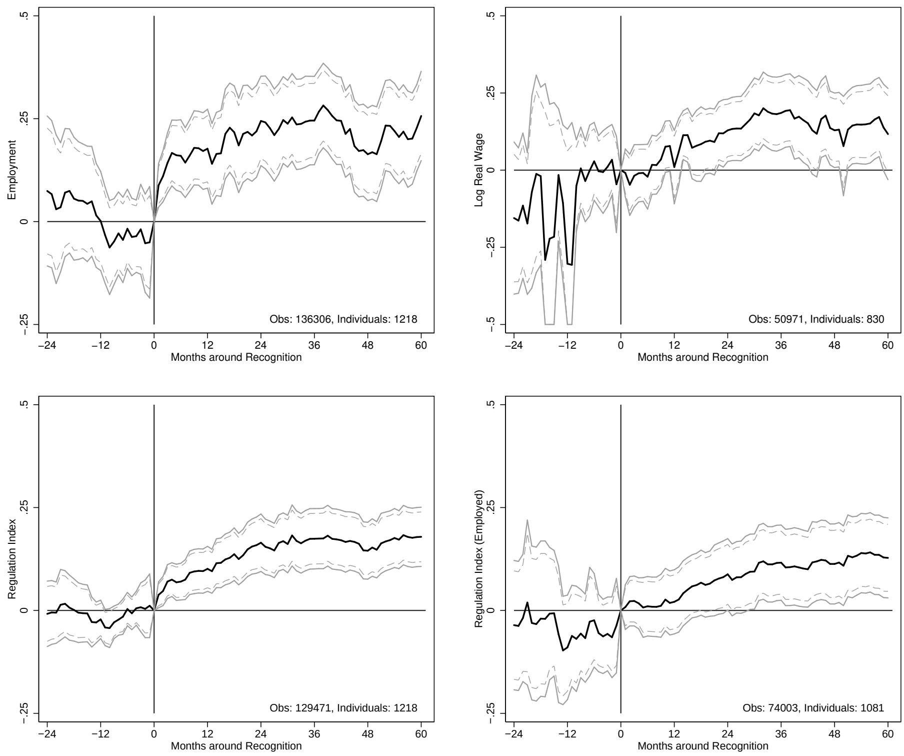


Figure 3.1: Dynamic Effects of Occupational Recognition

Note. Data source: IAB-SOEP Migration Sample linked to IEB data. The figures report the coefficients of the period dummies obtained from estimating regressions of specification (3.2) including a dummy for the application period. The dependent variable is the share of days in employment per month (upper left panel), log real wages for full-time employees (upper right panel), the index of occupational regulation, assigning a value of zero to the non-employed (lower left panel) and the index of occupational regulation (lower right panel). Additional controls are: the long-run average effect after recognition ($\text{CertRecog}_{i,t-61}$), the long-run average effect before recognition ($\text{CertRecog}_{i,t+25}$) an indicator for the application period, individual fixed effects, time fixed effects, time since migration fixed effects, age squared, and German language proficiency. 90% and 95% confidence intervals displayed using clustered standard errors at the individual level. Values of the confidence interval in the wage graph are cut at -0.5 for presentation purposes.

the control group, reaching 17.1 percentage points after 12 months. After that, the employment gap continues to grow albeit at a slower rate, reaching a value of 24.5 percentage points three years after recognition and stabilizing thereafter at slightly lower levels. This pattern suggests that occupational recognition increases the labor market opportunities of immigrants relatively quickly following the positive decision, and that their employability stays higher even in the long run, most likely due to their wider access to jobs. Reassuringly, there is no discernible difference in employment rates between those immigrants who obtain recognition within the following 24 months and those who do not, as indicated by the small and insignificant parameter estimates prior to the recognition date.

The corresponding dynamic pattern for log wages (upper right panel), shows an increase of the relative wage differential over time without any immediate jump. After receiving recognition, there is an increase in hourly wages that reaches 8.1 percent (7.8 log points) after one year and 19.8 percent (18.1 log points) after three years. From then onwards, the wage differential relative to those without occupational recognition levels off and coefficients fluctuate around a difference of around 16 percent. The reason for the delayed onset of significant wage gains from occupational recognition could be due to the fact that it takes time for immigrants to locate jobs in the higher paying and now accessible regulated segment of the labor market. It could also be that employers' remain initially skeptical regarding the equivalence between foreign and native credentials, and that this skepticism is only overcome with time. While somewhat more noisy due to the smaller sample size of employed immigrants, there is once again no evidence of a significant wage gap in the months prior to recognition, especially in the immediately preceeding year, lending credibility to the claim that the subsequent positive wage effects are indeed causally related to the occupational recognition.

The dynamic results with respect to the occupational regulation index in the lower panel of the figure provide further insights into the ways immigrants gain employment after recognition by entering increasingly more regulated occupations. When including non-employed individuals in the estimation (lower left panel), there is a rapid increase in the regulation index starting immediately after recognition by 9.6 percentage points after 12 months. Subsequently, the occupations chosen by immigrants with successful occupational recognition continue to have a higher regulation index compared to those of immigrants' without recognition, with the gap increasing to 17.4 percentage points after three years. This delay until all occupational adjustments after recognition materialize is likely due to difficulties of locating a suitable job in the regulated market segment for some migrants.

When considering the effect of recognition on the regulation index conditional on employment (lower right panel), the pattern is slightly different. In this case, we do not observe differentials in the regulation of occupations between immigrants with and without occupational recognition until about 12 months after recognition, mirroring the corresponding pattern for log wages. Only after this initial time period, the relative movements into more regulated occupations become significant, evident by a steady increase in our sequence of estimates. After three years, the relative increase in the probability of working in a regulated occupation amounts to 11.5 percentage points and remains more or less constant over the remaining time period. Taken together, these two dynamic regressions show that a successful recognition is helpful in securing employment in regulated occupations. Initially, these employment gains are mostly due to non-employed workers finding jobs in the regulated segment but after some delay, there is also a shift among employed workers into more regulated occupations. These observations are in line with the suggested mechanism underlying the slow wage growth. Securing a regulated occupation does not directly imply higher wages. But the continuous employment in these occupations, which tend to be the jobs with higher wages and faster wage growth, generate the observed long-term wage effects.

The evidence presented in this section suggests that immigrants who have not yet applied for recognition and those who never apply can serve as a reasonable control group in our difference-in-differences setting. As a robustness check, we redo the analysis but restrict the sample to only those immigrants who eventually all get full recognition. By focussing a priori on this group of immigrants, we reduce observable and unobservable heterogeneity in the sample, and identify the parameters of interest exclusively from the differential timing of the recognition processes across individuals (compare e.g. Arai and Thoursie, 2009, for a similar approach). As Figure 3.A4 and Table 3.A2 in the appendix show, our main results are robust to this alternative identification strategy, with average employment and wage effects slightly higher and movements into regulated occupations slightly lower. Similarly, the results from the pooled synthetic control method reported in Appendix 3.A confirm that occupational recognition has positive effects on immigrants' employment, hourly wages and probability of working in a regulated occupation.

3.5.3 Heterogeneous Effects

Our results so far speak to the overall static and dynamic effects of occupational recognition on immigrants' labor market outcomes. In this section, we study the heterogeneity of these effects across a number of different dimensions. Because of our relatively small sample size, several of the estimates in this section suffer from low precision, making it hard to draw strong conclusions. Table 3.4 presents results where we allow the treatment effect to vary by the type of foreign certificate for which immigrants applied for recognition. As mentioned before, we do not observe the exact certified occupation or field of study of a successful applicant, but we do observe the broad educational category for which recognition is requested, allowing us to distinguish four groups: vocational training, college/university degree, doctoral degree and any other education.[24]

[24] The reported education levels in the survey are, in decreasing order: 1. doctoral degree, 2. university education, 3. college education, 4. vocational school, 5. apprenticeship, 6. practical training, 7. other education, 8. missing. We aggregate groups 2 and 3 into the group "college/university degree" and groups 4-6 into the group "vocational training". Since every immigrant with full recognition

Table 3.4: Static Effects by Type of Recognized Certificate

	Employment (1)	Log Wages (Full-time) (2)	Regulation Index (3)	Regulation Index (Employed) (4)
Application period	0.026	-0.041	0.005	0.062
	(0.065)	(0.104)	(0.035)	(0.066)
Full recognition of vocational training	0.269***	0.061	0.164**	0.189
	(0.055)	(0.174)	(0.065)	(0.122)
college/university degree	0.101	0.161	0.132***	0.135**
	(0.070)	(0.122)	(0.039)	(0.065)
doctoral degree	0.456***	0.305	0.431**	-0.261
	(0.095)	(0.261)	(0.178)	(0.601)
other education	-0.148	0.261***	-0.026***	-0.008
	(0.130)	(0.028)	(0.009)	(0.015)
Individuals	1,218	830	1,218	1,081
with recognition	140	114	140	132
without recognition	1,078	716	1,078	949
Observations	136,306	50,971	129,471	74,003

Note. Data source: IAB-SOEP Migration Sample linked to IEB data. The estimates are based on specification (3.1) including a dummy for the application period and separate treatment dummies for individuals with vocational training, university/college degree, doctoral degree and other education as their highest level of foreign training for which they requested recognition. The dependent variable is the share of days in employment per month in column (1), log real wages for full-time employees in column (2), the index of occupational regulation, assigning a value of zero to non-employed in column (3), and the index of occupational regulation in column (4). Additional controls are individual fixed effects, time fixed effects, time since migration fixed effects, age squared, and German language proficiency. For individuals with several foreign certificates, the highest in terms of educational value is chosen.
Standard errors in parentheses are clustered at the individual level: * p<0.10, ** p<0.05, *** p<0.01.

The empirical results suggest that the recognition process is important for most types of qualifications. Except for the category of *other education*, all coefficients for employment and wage regressions are positive though in several case not statistically significant. The

provided valid information about his or her education level, there is no "missing" category in Table 3.4.

group with a doctoral degree benefits the most with an employment increase of 45.6 percentage points and a wage effect of 35.7 percent (30.5 log points), followed by the group with vocational training with an employment effect of 26.9 percentage points and an insignificant wage effect of 6.3 percent (6.1 log points). The movement into regulated occupations is similar for the groups with vocational training and college/university degrees. For the group with doctoral studies, the movement from non-employment into regulated occupations is particularly important. Conditional on being employed, the coefficient is actually negative, although not significant, suggesting that these immigrants remain unemployed until they get a position in their desired regulated occupation.

A complementary analysis considers heterogeneous effects across the different types of authorities to which immigrants apply for recognition. To which specific institutions immigrants must apply depends on the particular occupation or field of study for which they seek recognition. Different authorities are associated with more or less regulated occupations, allowing us to use the information on the recognizing authority as a proxy for recognition of a regulated versus unregulated occupation. We distinguish between five broad groups: the Chamber of Crafts, the Chamber of Industry and Commerce, and the Office for the Recognition of Foreign University Degrees, all of which are dealing primarily with unregulated occupations, and the Chambers of Physicians, Dentists, Veterinarians and Pharmacists, and Other Institutions, which are dealing primarily with regulated occupations.[25]

As shown in Table 3.5, for trained physicians, dentists, veterinarians, and pharmacists, the benefits from obtaining a recognition are substantial, with an employment effect of 50.6 percentage points and a wage effect of 235.0 percent (120.9 log points). While this wage

[25] To identify the particular authority responsible for the recognition of specific occupations we use information from https://www.anerkennung-in-deutschland.de/. The exact assignment can deviate within occupation, since rules vary by region and the information provided only reflects the current situation.

effect appears large, it is comparable to the findings by Kugler and Sauer (2005) who find a return to a medical license for immigrants in Israel between 180 and 340 percent. There is also a large impact on the probability of working in a regulated occupation, with an increase of the regulation index by 28.4 percentage points, conditional on employment. This reflects the fact that physicians, dentists, veterinarians, and pharmacists are all licensed occupations and hence a formal recognition indispensable for working in these occupations.

Immigrants who obtain recognition from the Office for the Recognition of Foreign University Degrees also experience substantial wage gains of 54.3 percent (43.4 log points) but the employment responses are relatively small, reflecting the fact that most of the relevant occupations are unregulated and thus already accessible prior to obtaining recognition. The same is true for the Chamber of Crafts and the Chamber of Industry and Commerce, where the effect on the probability of working in a regulated occupations, conditional on employment, is once again not as important. Interestingly, for these two cases there are, however, still sizeable employment and wage effects: the wage effect for recognitions from the Chamber of Crafts is 38.6 percent (32.7 log points) and the employment effect for recognitions from the Chamber of Industry and Commerce is 25.8 percentage points.[26] Together with the positive wage effects estimated for recognitions from the Office for the Recognition of Foreign University Degrees, these results suggest that even for unregulated occupations a formal recognition in Germany has significant positive effects on subsequent labor market outcomes, possibly due to its role in signalling immigrants' skills to potential employers.

An important finding in the literature on immigrant assimilation is that the transferability of immigrants' skills depends on the closeness between the education system of the origin country and the host

[26]Different labor market institutions, such as unionization and other occupation-specific regulations, might explain why recognitions from the Chamber of Crafts primarily affect wages while recognitions from the Chamber of Industry and Commerce mostly affect the employment margin.

Table 3.5: Static Effects by Type of Recognizing Authority

	Employment (1)	Log Wage (Full-time) (2)	Regulation Index (3)	Regulation Index (Employed) (4)
Application period	0.023	-0.043	0.005	0.063
	(0.074)	(0.092)	(0.032)	(0.062)
Full recognition from				
Chamber of Crafts	0.096	0.327***	0.061	0.004
	(0.120)	(0.065)	(0.058)	(0.026)
Chamber of Industrye and Commerce	0.258***	-0.005	0.067	0.193
	(0.065)	(0.251)	(0.065)	(0.181)
Office Recognition University Degree	0.109	0.434***	0.102	0.072
	(0.150)	(0.106)	(0.065)	(0.280)
Chambers of Physicians, etc.	0.506***	1.209***	0.440***	0.284***
	(0.038)	(0.115)	(0.169)	(0.064)
Other Institutions	0.114	0.028	0.197***	0.195**
	(0.083)	(0.151)	(0.062)	(0.093)
Individuals	833	600	833	750
with recognition	99	82	99	93
without recognition	734	518	734	657
Observations	122,905	46,484	116,316	66,996

Note. Data source: IAB-SOEP Migration Sample linked to IEB data. The estimates are based on specification (3.1) including a dummy for the application period and separate treatment dummies for recognition through the Chamber of Crafts, Chamber of Industry and Commerce, Office for the Recognition of Foreign University Degrees, Chambers of Physicians, Dentists, Veterinarians and Pharmacists, and Other Institutions. The dependent variable is the share of days in employment per month in column (1), log real wages for full-time employees in column (2), the index of occupational regulation, assigning a value of zero to the non-employed in column (3), and the index of occupational regulation in column (4). Additional controls are individual fixed effects, time fixed effects, time since migration fixed effects, age squared, and German language proficiency. Standard errors in parentheses are clustered at the individual level: * p<0.10, ** p<0.05, *** p<0.01.

country. A natural question in this context is whether the effect of occupational recognition also varies with the characteristics of the immigrants' home countries. Using GDP per capita as a proxy for the closeness between the home country and Germany, the estimates in Table 3.6 show that the effect of recognition is quite homogeneous

Table 3.6: Static Effects by GDP in Country of Origin

	Employment (1)	Log Wage (Full-time) (2)	Regulation Index (3)	Regulation Index (Employed) (4)
Application period	0.031	-0.053	0.007	0.068
	(0.068)	(0.100)	(0.035)	(0.072)
Received full recognition	0.177***	0.143	0.149***	0.138*
	(0.052)	(0.100)	(0.035)	(0.077)
Received full recognition	-0.000	-0.016**	0.001	0.000
× GDP/capita 2015 (demeaned)	(0.004)	(0.007)	(0.004)	(0.003)
Mean GDP/capita	5.49	5.80	5.56	5.76
Individuals	1,140	780	1,140	1,014
with recognition	133	107	133	125
without recognition	1,007	673	1,007	889
Observations	124,982	46,925	118,439	68,362

Note. Data source: IAB-SOEP Migration Sample linked to IEB data. Estimates based on specification (3.1) including a dummy for the application period and an interaction term with demeaned GDP per capita. The dependent variable is the share of days in employment per month in column (1), log real wages for full-time employees in column (2), the index of occupational regulation, assigning a value of zero to the non-employed in column (3), and the index of occupational regulation in column (4). Additional controls are individual fixed effects, time fixed effects, time since migration fixed effects, age squared, and German language proficiency. The mean GDP/capita is the average among included individuals weighted by their number of observations (in \$1,000). GDP information is taken from World Bank database. Standard errors in parentheses are clustered at the individual level: * p<0.10, ** p<0.05, *** p<0.01.

across home countries. By demeaning the interaction variable, the coefficient on the main recognition dummy is close to the average effect we estimate in our baseline specification. The coefficients of the interaction terms, in turn, are very close to zero, with the only exception being the impact on wages, where an increase of GDP per capita by \$1,000 leads to a 1.6 percent smaller increase in wages. This is not surprising since immigrants from richer countries are likely to earn higher wages in the German labor market to start with due to the better quality and transferability of their home country specific human capital, so that they have less to gain from obtaining occupational recognition than immigrants from poorer countries.

3.6 Implications for Immigrant Earnings Assimilation

Our results so far have shown significant positive long-run effects of occupational recognition on immigrants' employment and wage outcomes. In this section, we put these gains into perspective by relating them to standard earnings assimilation profiles of immigrants in Germany. For this purpose, we merge a 1 percent random sample of native German workers in the IEB to our IAB-SOEP Migration Sample and jointly estimate the following immigrant and native earnings equations:

$$\begin{aligned}
\text{Immigrants:} \quad & \log w_{it} = && \phi_m' X_{it} + \alpha_m \cdot \text{age}_{it} + \beta \cdot \text{ysm}_{it} \\
& && + \gamma \cdot \text{ysr}_{it} + \delta C_i + \theta_m \pi_t + \varepsilon_{it} \qquad (3.3) \\
\text{Natives:} \quad & \log w_{it} = && \phi_n' X_{it} + \alpha_n \cdot \text{age}_{it} + \theta_n \pi_t + \varepsilon_{it},
\end{aligned}$$

where w_{it} are total monthly earnings of individual i at time t, X_{it} is a vector of socioeconomic characteristics (educational attainment[27], gender, federal state of residence), age_{it} represents a quartic function of the individual's age, ysm_{it} represents a quartic function of the number of years since migration, ysr_{it} represents a quartic function of the number of years passed since the result of the recognition process was obtained (set to zero for all immigrants who never applied for recognition), C_i is a vector of dummy variables indicating an immigrant's arrival cohort (1970-1994, 1995-2005, 2005-2013), and π_t is a vector of year fixed effects. Since aging, cohort and period effects are perfectly collinear, we impose the standard assumption that period effects are the same for immigrants and natives ($\theta_m = \theta_n$) as suggested by Borjas (1995). We estimate this model using all available monthly native and immigrant observations, clustering our standard errors at the individual level. The immigrants in the sample belong to four distinct groups: immigrants who never applied for recognition,

[27] We use the imputed education variable obtained by applying the IP1 algorithm developed by Fitzenberger et al. (2005).

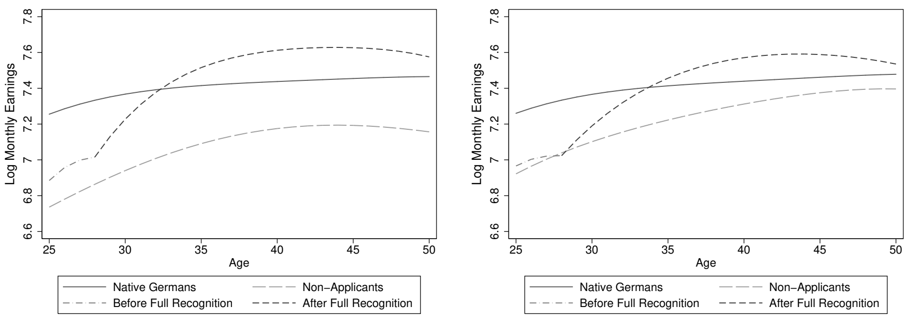


Figure 3.2: Effect of Recognition on Immigrant Assimilation Profiles

Note: The displayed simulations of earnings profiles in the left and right panel are based on parameter estimates reported in columns (2) and (4) of Table 3.A5, respectively. Immigrants are assumed to enter Germany at the age of 25, with the comparison being relative to natives of the same age. We compute each profile for the mean values of all socioeconomic characteristics in the sample, thus accounting for observable differences in educational attainment, gender, federal state of residence and time period between the different immigrant groups and natives. The intercepts of the different immigrant groups reflect their weighted mean cohort effects. The left panel shows the predicted earnings profiles without controlling for occupations, the right panel the profiles after controlling for 3-digit occupations in the IEB data.

immigrants who applied but were denied recognition, immigrants who applied and gained partial recognition, and immigrants who applied and gained full recognition. We drop immigrants who applied for recognition but whose decision is pending at the time of the survey from the sample. In the estimation, we allow the age, years since migration and years since recognition profiles to vary between each of the four immigrant groups.

Rather than presenting the full regression results, which can be found in Table 3.A5 in the appendix, we use the estimates from the two-equation regression model in (3.3) to predict native and immigrant earnings profiles (compare column (2) of Table 3.A5). We simulate earnings profiles for immigrants who enter Germany at the age of 25 and compare them to the corresponding earnings profile of natives of the same age. We compute each profile for the mean values of all socioeconomic characteristics in the sample, thus netting out the effects arising from observable differences in educational attainment, gender, federal state of residence and time period between the different immigrant groups and natives. The intercepts of the four immigrant groups reflect the weighted means of their cohort effects.[28] For clarity, the left panel of Figure 3.2 only depicts the predicted log earnings profiles of native Germans, immigrant non-applicants, and immigrants who eventually receive full recognition, suppressing the corresponding profiles for immigrants whose application is denied and immigrants who only receive partial recognition, which together make up only a small fraction of the overall sample.

Immigrants who never apply for recognition (who make up 81.5 percent of the immigrant sample) initially face an earnings gap relative to native Germans of 40.5 percent (51.9 log points) which steadily declines over time, levelling off at around 22.5 percent (25.5 log points) after 15 years of residence in Germany. The earnings of immigrants who eventually obtain full recognition (11.6 percent of the immigrant sample) grow initially at a similar rate but start from a more advantageous position, with an earnings gap upon arrival of only 31.0 percent (37.1 log points). After obtaining full recognition, which for these simulations we assume to occur after three years of residence in Germany (the mean duration between arrival and recognition in the assimilation sample), the speed of convergence of these immigrants' earnings increases substantially (dashed line), leading to a catch-up and eventual overtaking of native earnings after about 8 years, with a maximum positive earnings advantage of around 19.8 percent (18.0 log points) observed after 17 years in the country. However, due to the small sample size, we lack precision in the estimates for the immigrant group with full recognition, so that from 5 years since migration onwards, their earnings gap relative to natives is no longer statistically significant. These findings suggest that occupa-

[28]Similarly to Bratsberg et al. (2006), we allow the returns to education and gender to vary between natives and immigrants, but not between different immigrant groups. We further assume that the region effects are the same for immigrants and natives.

tional recognition has a significant effect on the speed of immigrants' economic assimilation in Germany.

Part of the reason for why immigrants who obtain full recognition may outperform the average native in the left panel of Figure 3.2 is their greater likelihood of working in high-paying occupations, for example in the health sector. Controlling for educational attainment partly accounts for such heterogeneity but even within the group of say university-educated workers, immigrants with occupational recognition are likely to be working in more attractive occupations. In the right panel of Figure 3.2, we depict predicted assimilation profiles from an extended specification in which we control for a full set of 3-digit occupation dummies (compare column (4) of Table 3.A5). Since part of the growth in immigrants' earnings over time is due to their climbing of the occupational ladder, one would generally not want to control for occupation in these types of assimilation regressions. Including occupation fixed effects, however, improves the comparability of natives and immigrants in our sample and, importantly, reveals information about the relative earnings of immigrants and natives within the same occupations, which could be interpreted as a proxy for the quality of the services provided by immigrants relative to natives in the same types of jobs.

As the right panel of Figure 3.2 shows, holding the occupational distribution constant across groups, reduces somewhat the earnings gaps of the different immigrants groups relative to natives. The initial gaps for non-applicants and immigrants who eventually obtain full recognition are now almost identical, 28.7 percent (33.8 log points) and 25.4 percent (29.4 log points) respectively. As before, we do observe an acceleration of the speed of assimilation at the time recognition is obtained and an eventual overtaking of native earnings after 10 years, with the maximum gap amounting to a statistically not significant 14.9 percent (13.9 log points) 17 years after arrival. The good relative performance of immigrants who obtain full recognition for their foreign qualifications is therefore not just due to their advantageous distribution across occupations relative to the representative sample of natives that serves as the comparison group. Rather, it appears that even conditional on occupation, these immigrants perform at least at the same level as their native counterparts, mitigating concerns that occupational recognition leads to a dilution of occupational standards and suggesting that the formal recognition process in Germany does a reasonable job in ensuring the equivalence of foreign qualifications with their native counterparts.

3.7 Conclusion

In this paper, we analyze how the formal recognition of immigrants' foreign qualifications affects their subsequent labor market outcomes. For our analysis, we exploit a novel linked survey-social security data set which, besides including comprehensive information about workers' entire work histories, explicitly asks participants, if applicable, about the timing of their recognition process in Germany. This allows us to assess in detail how occupational recognition affects immigrant labor market outcomes, both from a static and dynamic point view. Comparing the labor market outcomes of immigrants who obtain full recognition to those of immigrants who either never apply or have not yet received full recognition themselves, the evidence from our dynamic difference-in-differences specification suggests large and long-lasting positive effects of occupational recognition on immigrants' labor market outcomes, with a 24.5 percentage point higher employment rate and a 19.8 percent higher hourly wage three years after obtaining recognition. We further document that occupational recognition indeed induces workers to enter regulated occupations, both directly out of non-employment and, with some delay, through horizontal movements of employed workers from unregulated into regulated occupations.

Further heterogeneity analysis suggests that formal recognition is not only beneficial with respect to regulated occupations but also when it comes to occupations that are freely accessible even in the absence of recognition. This important finding suggests that, besides granting access to regulated occupations, the certification of foreign qualifications also plays a signalling role in the German labor market, eliminating uncertainty about an immigrant worker's occupational skills. The signalling value of formal recognition appears to be particularly large for immigrants from less developed countries, who, at least in terms of wages, benefit significantly more from the recognition of their qualifications. This could be due to the higher initial degree of uncertainty in the German labor market regarding these immigrants' qualifications, which means there is more to gain from a formal certification of these qualifications' equivalence with their native counterparts.

We conclude by showing that occupational recognition leads to a significant acceleration of immigrants' earnings growth relative to natives. Recognizing immigrants' foreign credentials may thus be an effective way of tapping into their human capital and fostering their integration into the host country's economy. More generally, our results suggest that part of the substantial employment and wage gaps between natives and immigrants around the world may be due to the lack of formal recognition of the latter's occupational qualifications. The large positive wage effects and the eventual full convergence to native earnings indicate that, at least in Germany, foreign credentials, once declared equivalent to native ones, are indeed valued in the labor market, mitigating fears of a watering-down of occupational standards.

3.A Appendices

3.A.1 Synthetic Control Method

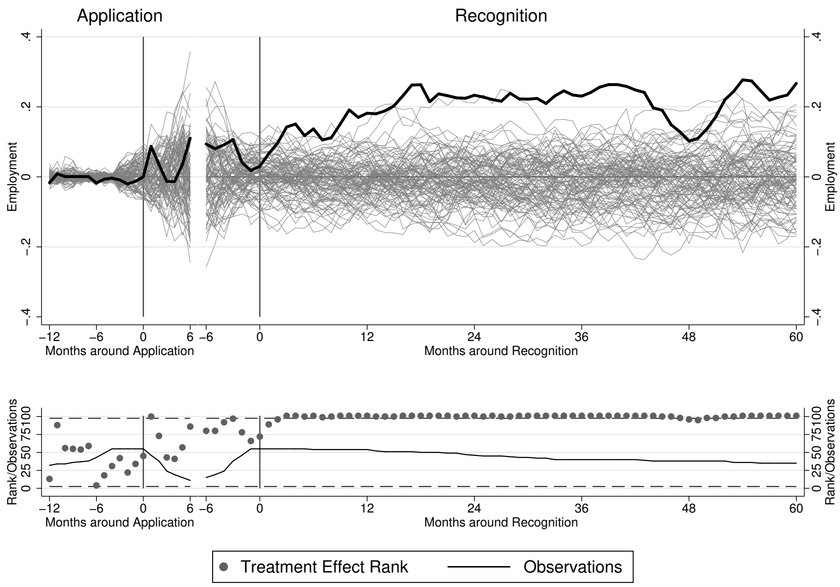


Figure 3.A1: Dynamic Employment Effects of Occupational Recognition

Note: Data source: IAB-SOEP Migration Sample linked to IEB data. The displayed estimates along the thick black lines are the average differentials in employment in each pre- and post-treatment period between all treated units and their synthetic control groups. The thin gray lines depict 100 placebo estimations, in which we iteratively apply the synthetic control method to randomly picked non-treated immigrants in each treated immigrant's donor pool.

As a robustness check for our dynamic estimation, we apply a pooled version of the synthetic control method proposed by Abadie et al. (2010). In contrast to our main approach, each immigrant who receives recognition (the treatment) is here matched to a set of other immigrants who never applied for recognition but whose labor market outcomes in the period prior to application are similar to those of

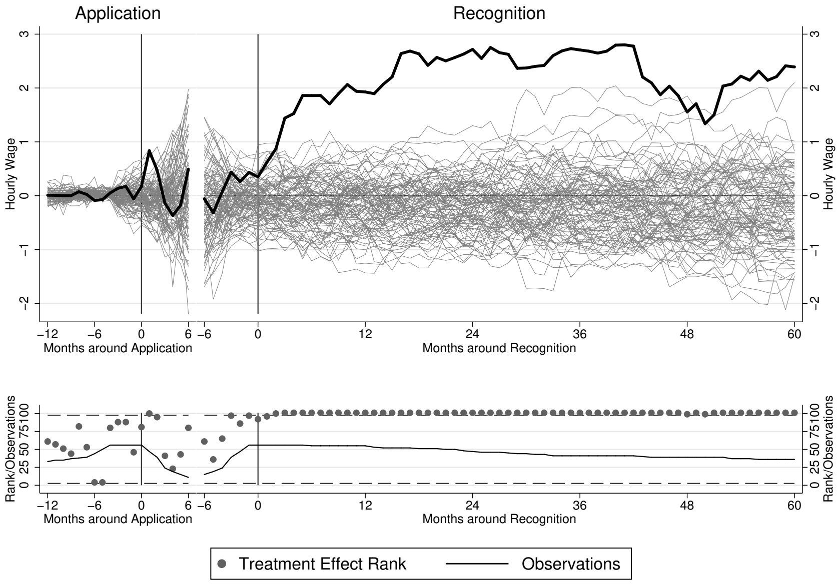


Figure 3.A2: Dynamic Wage Effects of Occupational Recognition

Note: Data source: IAB-SOEP Migration Sample linked to IEB data. The displayed estimates along the thick black lines are the average differentials in hourly wages in each pre- and post-treatment period between all treated units and their synthetic control groups, including zeros for non-employed individuals. The thin gray lines depict 100 placebo estimations, in which we iteratively apply the synthetic control method to randomly picked non-treated immigrants in each treated immigrant's donor pool.

the treated immigrant. We obtain a synthetic control group for each treated immigrant and then average the dynamic treatment effects in each pre- and post-treatment month across all treated individuals in the sample in those months. Note that we match directly on the corresponding outcome variables in the year prior to application, excluding the last three months to test for anticipation effects.

The thick black lines in Figures 3.A1 and 3.A2 show the resulting dynamic impacts of occupational recognition on employment and hourly wages between 12 months before the application period and 60 months after recognition. We look at hourly wages rather than log hourly wages since otherwise it would be difficult to find potential control individuals with positive wages in precisely the same months as the treated individuals. This implies that part of the estimated impacts on hourly wages are driven by individuals finding employment and starting to earn non-zero wages. Overall, the dynamic patterns are similar to those obtained from our regression-based approach, with substantial and relatively quick increases in both employment and hourly wages in the months immediately after recognition, continuing divergence at a slower pace for a couple of years, and a flattening out of the two profiles thereafter.

To assess the statistical significance of the dynamic effects from the synthetic control group method, we perform 100 placebo estimations in which, for each iteration, we randomly pick for each treated immigrant an untreated immigrant from his or her donor pool, assign the same hypothetical application and recognition dates as for the treated immigrant, find a suitable synthetic control group for this placebo immigrant, and then aggregate all dynamic impact estimates across all placebo immigrants. As illustrated by the thin gray lines in Figures 3.A1 and 3.A2, the estimated effects of actual occupational recognition are large relative to the distribution of dynamic placebo effects, suggesting that they pick up real employment and wage effects. Contrary to the regression-based results reported in Table 3.3, we find some indication for a significant positive effect of applying itself on the probability of being employed although this effect only extends to the first month after submitting the application.

To facilitate the assessment of the statistical significance of the estimated treatment effects in each period, we depict their rank among the distribution of placebo effects (gray dots) and the underlying number of treated individuals (black line) for each period in a separate plot underneath the main graphs. Note that the sample size of treated individuals used in these estimations is substantially smaller than in our main approach since we need to condition on observing individuals for at least one period prior to their application and for at least one period between their application and their recognition

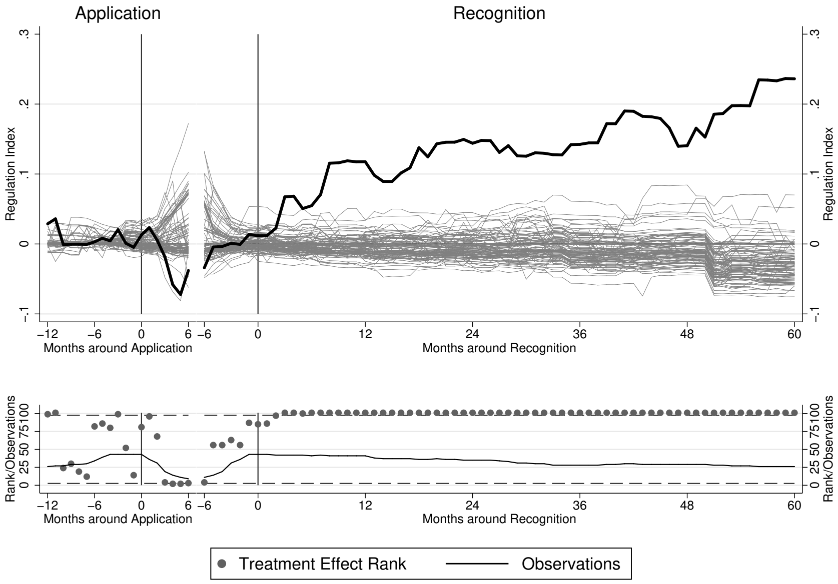


Figure 3.A3: Dynamic Effects of Occupational Recognition on the Degree of Regulation

Note: Data source: IAB-SOEP Migration Sample linked to IEB data. The displayed estimates along the thick black lines are the average differentials in the regulation index in each pre- and post-treatment period between all treated units and their synthetic control groups, including zeros for non-employed individuals. The thin gray lines depict 100 placebo estimations, in which we iteratively apply the synthetic control method to randomly picked non-treated immigrants in each treated immigrant's donor pool.

date. Individuals who apply in the month they are first observed in the IEB data or individuals who obtain the result of their application in the same month in which they apply are thus excluded from the estimation sample.

Figure 3.A3 displays the corresponding dynamic effects for the average occupational regulation index, where the index is set to zero for non-employed individuals as in our main approach without conditioning on employment. The latter is not feasible under the synthetic control approach as it would require finding suitable control

individuals with exactly the same monthly employment histories as the treated individuals. Similar to the pattern documented in the bottom left panel of Figure 3.1, there is a swift increase in the regulation index after obtaining full recognition which continues more or less uninterruptedly throughout the entire post-recognition period, amounting to a value of almost 0.25 after five years.

Overall, while not entirely comparable in terms of the outcome variables considered, we view the evidence from the synthetic control method as supportive of the main findings from our regression-based difference-in-differences approach, indicating significant and quantitatively large effects of occupational recognition on immigrants' labor market outcomes.

3.A.2 Further Tables and Figures

Table 3.A1: Distribution of Occupations for Requested Recognition

Occupation	%
Doctor	13.16
Engineer	13.16
Nurse	10.53
Veterinary	10.53
Teacher	7.89
Civil Servant (executive officer)	2.63
Pharmacist	2.63
Midwife	2.63
Shop Assistant	2.63
Physioterapist	2.63
Correspondent in foreign language	2.63
Agrotechnical Assistant (state approved)	2.63
IT-Assistant (state approved)	2.63
Vocational College in Electronics (state approved)	2.63
Business Economist	2.63
Biologic Laboratory Technician (state approved)	2.63
Marketing Specialist	2.63
Cook	2.63
Food Inspector	2.63
Financial advisor	2.63
Manufacturer	2.63
Reseacher	2.63
Total	100.00

Note. Data source: IAB-SOEP Migration Sample, third wave. The table refers to the distribution of occupations for which recognition was requested. Only individuals obtaining full recognition are considered.

Table 3.A2: Static Effects of Occupational Recognition - Excluding Non-Applicants

	Employment	Log Wage (Full-time)	Regulation Index	Regulation Index (Employed)
	(1)	(2)	(3)	(4)
Panel A				
Received full recognition	0.186***	0.154*	0.137***	0.105
	(0.062)	(0.081)	(0.037)	(0.064)
Panel B				
Application period	0.020	0.031	-0.007	0.048
	(0.067)	(0.097)	(0.040)	(0.071)
Received full recognition	0.191***	0.163	0.136***	0.116
	(0.065)	(0.105)	(0.041)	(0.077)
Individuals	140	114	140	132
Observations	17,170	8,563	16,405	10,581

Note. Data source: IAB-SOEP Migration Sample linked to IEB data. Panel A reports the estimates based on specification (3.1), Panel B adds a dummy variable for the application period as discussed in the text. The dependent variable is the share of days in employment per month in column (1), log real hourly wages for full-time employees averaged over all spells in a given month in column (2), the index of occupational regulation, assigning a value of zero to the non-employed, in column (3), and the index of occupational regulation in column (4). Additional controls are individual fixed effects, time fixed effects, time since migration fixed effects, age squared, and German language proficiency. The sample comprises only immigrants who eventually receive full recognition, and who migrated to Germany at the age of at least 18, stayed in Germany after arrival and do not have any reported incapacity for work. Observations are only included when migrant's age is at least 25 and less than 60.

Standard errors in parentheses are clustered at the individual level: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 3.A3: Impact of Different Sample Selection Procedures

	Migration after 18yr (1)	(1) + w/o incapacity (2)	(2) + working age (3)	(3) + stay in Germany (4)
Employment				
Application period	-0.008	0.023	0.051	0.024
	(0.058)	(0.055)	(0.060)	(0.067)
Received full recognition	0.149***	0.172***	0.162***	0.165***
	(0.055)	(0.049)	(0.055)	(0.052)
Individuals	1,470	1,412	1,346	1,218
with recognition	166	159	158	140
without recognition	1,304	1,253	1,188	1,078
Observations	189,027	176,994	155,566	136,306
Log Real Wage				
Application period	-0.139	-0.117	-0.113	-0.053
	(0.097)	(0.098)	(0.102)	(0.105)
Received full recognition	0.070	0.089	0.086	0.141
	(0.096)	(0.103)	(0.106)	(0.103)
Individuals	1,019	976	924	830
with recognition	135	129	128	114
without recognition	884	847	796	716
Observations	62,982	59,280	55,765	50,971
Regulation Index				
Application period	0.020	0.025	0.028	0.009
	(0.029)	(0.030)	(0.033)	(0.035)
Received full recognition	0.144***	0.153***	0.157***	0.152***
	(0.032)	(0.033)	(0.037)	(0.035)
Individuals	1,470	1,412	1,346	1,218
with recognition	166	159	158	140
without recognition	1,304	1,253	1,188	1,078
Observations	181,088	169,313	148,378	129,471
Regulation Index (Employed)				
Application period	0.041	0.045	0.040	0.065
	(0.051)	(0.053)	(0.057)	(0.065)
Received full recognition	0.118**	0.123**	0.123**	0.129*
	(0.054)	(0.056)	(0.059)	(0.068)
Individuals	1,316	1,268	1,198	1,081
with recognition	159	152	150	132
without recognition	1,157	1,116	1,048	949
Observations	92,140	87,004	80,782	74,003

Note. Data source: IAB-SOEP Migration Sample linked to IEB data. All estimations based on specification (3.1) including a dummy for the application period. The dependent variable is the share of days in employment per month in Panel A, log real hourly wages for full-time employees averaged over all spells in a given month in Panel B, the index of occupational regulation, assigning a value of zero to the non-employed, in Panel C, and the index of occupational regulation in Panel D. Additional controls are individual fixed effects, time fixed effects, time since migration fixed effects, age squared, and German language proficiency. The sample comprises immigrants who either receive full recognition or never apply. Additional selection rules are described in the heading.
Standard errors in parentheses are clustered at the individual level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 3.A4: Impact of Control Variables

	(1)	(2)	(3)	(4)	(5)
Employment					
Application period	-0.237***	-0.108*	0.032	0.027	0.024
	(0.066)	(0.065)	(0.066)	(0.066)	(0.067)
Received full recognition	0.153***	0.127***	0.170***	0.166***	0.165***
	(0.034)	(0.035)	(0.051)	(0.051)	(0.052)
Individuals	1,218	1,218	1,218	1,218	1,218
with recognition	140	140	140	140	140
without recognition	1,078	1,078	1,078	1,078	1,078
Observations	136,306	136,306	136,306	136,306	136,306
Log Real Wages (Full-time)					
Application period	-0.061	-0.026	-0.034	-0.047	-0.053
	(0.167)	(0.161)	(0.101)	(0.105)	(0.105)
Received full recognition	0.242***	0.242***	0.136	0.148	0.141
	(0.056)	(0.055)	(0.101)	(0.104)	(0.103)
Individuals	830	830	830	830	830
with recognition	114	114	114	114	114
without recognition	716	716	716	716	716
Observations	50,971	50,971	50,971	50,971	50,971
Regulation Index					
Application period	0.019	0.023	0.013	0.011	0.009
	(0.018)	(0.019)	(0.035)	(0.035)	(0.035)
Received full recognition	0.117***	0.117***	0.155***	0.153***	0.152***
	(0.022)	(0.021)	(0.035)	(0.035)	(0.035)
Individuals	1,218	1,218	1,218	1,218	1,218
with recognition	140	140	140	140	140
without recognition	1,078	1,078	1,078	1,078	1,078
Observations	129,471	129,471	129,471	129,471	129,471
Regulation Index (Employed)					
Application Period	0.096*	0.082	0.068	0.066	0.065
	(0.054)	(0.053)	(0.066)	(0.066)	(0.065)
Received full recognition	0.149***	0.150***	0.131*	0.130*	0.129*
	(0.030)	(0.029)	(0.068)	(0.068)	(0.068)
Individuals	1,081	1,081	1,081	1,081	1,081
with recognition	132	132	132	132	132
without recognition	949	949	949	949	949
Observations	74,003	74,003	74,003	74,003	74,003
Time since migration fixed effects		Yes	Yes	Yes	Yes
Individual fixed effects			Yes	Yes	Yes
Time fixed effects				Yes	Yes
Controls					Yes

Note. Data source: IAB-SOEP Migration Sample linked to IEB data. All estimations based on specification (3.1) including a dummy for the application period. The dependent variable is the share of days in employment per month in Panel A, log real hourly wages for full-time employees averaged over all spells in a given month in Panel B, the index of occupational regulation, assigning a value of zero to the non-employed, in Panel C, and the index of occupational regulation in Panel D. Sample selection is according to the results in Table 3.3. Additional controls are specified for each column in the table. The category *Controls* includes age squared and German language proficiency.
Standard errors in parentheses are clustered at the individual level. * p<0.10, ** p<0.05, *** p<0.01.

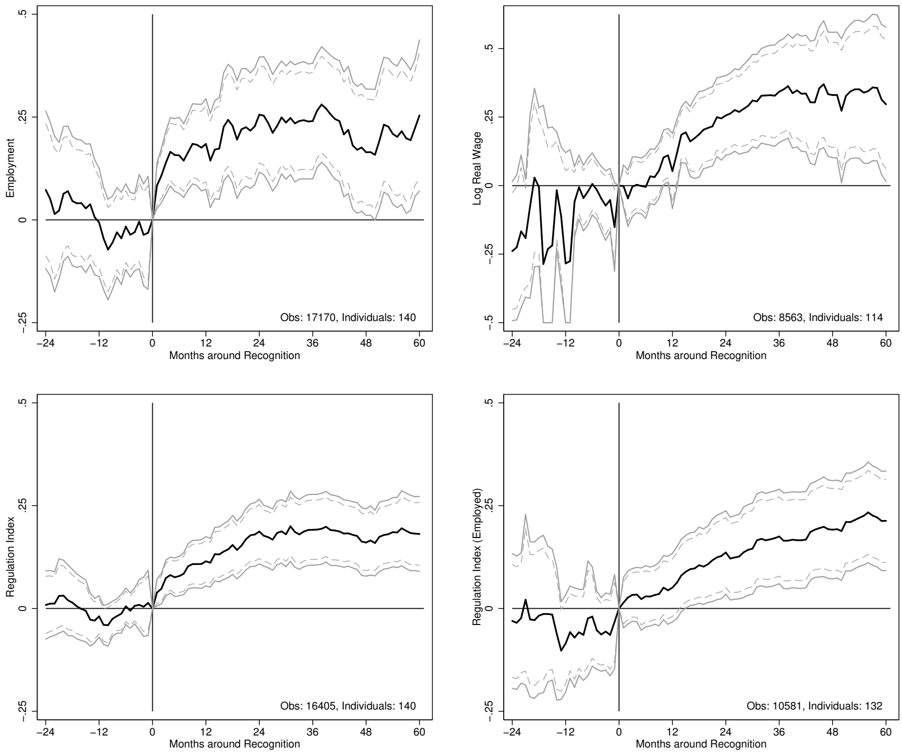


Figure 3.A4: Dynamic Effects of Occupational Recognition - Excluding Non-Applicants

Note. Data source: IAB-SOEP Migration Sample linked to IEB data. The figures report the coefficients of the period dummies obtained from estimating regressions of specification (3.2) including a dummy for the application period. The dependent variable is the share of days in employment per month (upper left panel), log real wages for full-time employees (upper right panel), the index of occupational regulation, assigning a value of zero to the non-employed (lower left panel) and the index of occupational regulation (lower right panel). Additional controls are: the long-run average effect after recognition ($\text{CertRecog}_{i,t-61}$), the long-run average effect before recognition ($\text{CertRecog}_{i,t+25}$) an indicator for the application period, individual fixed effects, time fixed effects, time since migration fixed effects, age squared, and German language proficiency. The sample only comprises immigrants who eventually receive full recognition, and who migrated to Germany at the age of at least 18, stayed in Germany after arrival and do not have any reported incapacity for work. Observations are only included when migrant's age is at least 25 and less than 60. 90% and 95% confidence intervals displayed using clustered standard errors at the individual level. Values of the confidence interval in the wage graph are cut at -0.5 for presentation purposes.

Table 3.A5: Assimilation Regressions

	No Occupation Controls				With Occupation Controls			
	(1)		(2)		(3		(4)	
Never	1.184	(6.285)	1.138	(6.285)	-1.366	(5.923)	-1.379	(5.925)
Denied	24.739	(30.449)	33.459	(35.290)	39.808	(31.468)	56.424	(34.324)
Partial	7.880	(29.353)	17.998	(28.714)	36.511	(25.968)	42.049	(26.109)
Full	-18.379	(14.747)	-12.493	(14.999)	-12.638	(12.603)	-9.005	(12.717)
Never × YSM	0.004	(0.028)	0.002	(0.028)	0.004	(0.025)	0.004	(0.025)
Denied × YSM	-0.045	(0.197)	0.071	(0.198)	-0.074	(0.239)	0.152	(0.248)
Partial × YSM	0.235	(0.151)	0.190	(0.172)	0.341*	(0.172)	0.328	(0.175)
Full × YSM	-0.003	(0.058)	-0.031	(0.058)	-0.037	(0.051)	-0.061	(0.051)
Never × $YSM^2/10$	0.014	(0.040)	0.017	(0.040)	0.007	(0.036)	0.007	(0.036)
Denied × $YSM^2/10$	0.230	(0.353)	0.011	(0.365)	0.268	(0.424)	-0.090	(0.439)
Partial × $YSM^2/10$	-0.330	(0.245)	-0.278	(0.254)	-0.473	(0.256)	-0.436	(0.252)
Full × $YSM^2/10$	0.044	(0.075)	-0.073	(0.074)	0.071	(0.068)	-0.029	(0.070)
Never × $YSM^3/100$	-0.005	(0.021)	-0.007	(0.021)	-0.003	(0.019)	-0.003	(0.019)
Denied × $YSM^3/100$	-0.203	(0.235)	-0.066	(0.243)	-0.230	(0.277)	-0.015	(0.285)
Partial × $YSM^3/100$	0.208	(0.161)	0.168	(0.147)	0.268	(0.156)	0.228	(0.149)
Full × $YSM^3/100$	-0.022	(0.035)	0.088*	(0.041)	-0.034	(0.033)	0.052	(0.038)
Never × $YSM^4/1000$	0.000	(0.003)	0.001	(0.003)	0.000	(0.003)	0.000	(0.003)
Denied × $YSM^4/1000$	0.050	(0.050)	0.023	(0.052)	0.057	(0.058)	0.015	(0.060)
Partial × $YSM^4/1000$	-0.043	(0.035)	-0.036	(0.030)	-0.051	(0.032)	-0.041	(0.031)
Full × $YSM^4/1000$	0.003	(0.005)	-0.018*	(0.007)	0.005	(0.005)	-0.011	(0.006)
Age	0.412***	(0.011)	0.412***	(0.011)	0.388***	(0.011)	0.388***	(0.011)
$Age^2/10$	-0.144***	(0.004)	-0.144***	(0.004)	-0.137***	(0.004)	-0.137***	(0.004)
$Age^3/1000$	0.226***	(0.007)	0.226***	(0.007)	0.216***	(0.006)	0.216***	(0.006)
$Age^4/100000$	-0.134***	(0.004)	-0.134***	(0.004)	-0.128***	(0.004)	-0.128***	(0.004)
Never × Age	-0.244	(0.647)	-0.240	(0.647)	0.081	(0.605)	0.081	(0.605)
Denied × Age	-2.318	(3.165)	-3.249	(3.647)	-3.940	(3.239)	-5.704	(3.545)
Partial × Age	-1.130	(2.956)	-2.189	(2.869)	-4.078	(2.621)	-4.668	(2.632)
Full × Age	1.722	(1.510)	1.082	(1.533)	1.120	(1.288)	0.719	(1.297)
Never × $Age^2/10$	0.120	(0.243)	0.119	(0.243)	-0.026	(0.226)	-0.026	(0.226)
Denied × $Age^2/10$	0.739	(1.205)	1.093	(1.371)	1.384	(1.224)	2.044	(1.337)
Partial × $Age^2/10$	0.514	(1.090)	0.924	(1.049)	1.615	(0.966)	1.846	(0.970)
Full × $Age^2/10$	-0.583	(0.565)	-0.329	(0.573)	-0.358	(0.482)	-0.198	(0.485)
Never × $Age^3/1000$	-0.243	(0.399)	-0.241	(0.399)	0.038	(0.367)	0.038	(0.367)
Denied × $Age^3/1000$	-0.968	(1.992)	-1.547	(2.235)	-2.081	(2.015)	-3.142	(2.190)
Partial × $Age^3/1000$	-0.975	(1.749)	-1.666	(1.660)	-2.768	(1.549)	-3.162*	(1.550)
Full × $Age^3/1000$	0.839	(0.919)	0.403	(0.930)	0.483	(0.785)	0.205	(0.788)
Never × $Age^4/100000$	0.171	(0.240)	0.170	(0.240)	-0.022	(0.219)	-0.022	(0.219)
Denied × $Age^4/100000$	0.428	(1.208)	0.773	(1.335)	1.131	(1.219)	1.754	(1.315)
Partial× $Age^4/100000$	0.647	(1.032)	1.077	(0.963)	1.727	(0.912)	1.974*	(0.908)
Full × $Age^4/100000$	-0.438	(0.548)	-0.166	(0.554)	-0.233	(0.470)	-0.060	(0.472)
Medium Edu	0.262***	(0.002)	0.262***	(0.002)	0.119***	(0.002)	0.119***	(0.002)
High Edu	0.673***	(0.003)	0.673***	(0.003)	0.347***	(0.003)	0.347***	(0.003)
Immigrant × Medium Edu	-0.055	(0.054)	-0.053	(0.053)	0.035	(0.048)	0.036	(0.048)
Immigrant × High Edu	0.190**	(0.072)	0.184*	(0.073)	0.218***	(0.060)	0.216***	(0.060)
Female	-0.515***	(0.002)	-0.515***	(0.002)	-0.441***	(0.002)	-0.441***	(0.002)
Immigrant × Female	-0.111*	(0.048)	-0.100*	(0.049)	-0.014	(0.041)	-0.007	(0.041)
Cohort 1970-1994	0.123	(0.087)	0.145	(0.087)	0.103	(0.075)	0.105	(0.075)
Cohort 1995-2004	0.050	(0.057)	0.054	(0.057)	0.059	(0.053)	0.061	(0.053)
Denied × YSR			-0.017	(0.044)			-0.050	(0.048)
Partial × YSR			0.027	(0.026)			-0.001	(0.023)
Full × YSR			0.117***	(0.026)			0.097***	(0.023)
Denied × $YSR^2/10$			0.023	(0.026)			0.031	(0.023)
Partial × $YSR^2/10$			-0.018	(0.048)			0.009	(0.045)
Full × $YSR^2/10$			-0.010	(0.012)			0.002	(0.009)
Denied × $YSR^3/100$			0.018	(0.031)			0.018	(0.029)
Partial × $YSR^3/100$			0.005	(0.013)			0.008	(0.010)
Full × $YSR^3/100$			-0.057***	(0.014)			-0.046***	(0.012)
Denied × $YSR^4/1000$			-0.012	(0.012)			-0.012	(0.011)
Partial × $YSR^4/1000$			0.008	(0.014)			-0.002	(0.012)
Full × $YSR^4/1000$			0.014***	(0.004)			0.011***	(0.003)
Constant	2.227***	(0.110)	2.227***	(0.110)	1.859***	(0.192)	1.859***	(0.192)
R-squared	0.27		0.27		0.39		0.39	
Observations	88,283,926		88,283,926		86,444,335		86,444,335	

Note: The dependent variable are log monthly earnings, conditional on working at least one day in a given month. The omitted categories are males, low educational attainment, immigrant cohort 2005-2014, period 1975. Standard errors are clustered at the individual level. The sample comprises monthly observations of 571,581 individuals in columns (1) and (2) and 569,104 individuals in columns (3) and (4).

Bibliography

Aaronson, D. and French, E. (2007). Product market evidence on the employment effects of the minimum wage. *Journal of Labor Economics*, 25(1):167–200.

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.

Abel, A. B. (1980). Empirical investment equations: An integrative framework. *Carnegie-Rochester Conference Series on Public Policy*, 12:39 – 91.

Abel, A. B. and Blanchard, O. J. (1986). The present value of profits and cyclical movements in investment. *Econometrica*, 54(2):249–273.

Abel, A. B. and Eberly, J. C. (1994). A unified model of investment under uncertainty. *The American Economic Review*, 84(5):1369–1384.

Acemoglu, D. (2010). When does labor scarcity encourage innovation? *Journal of Political Economy*, 118(6):1037–1078.

Acemoglu, D. and Autor, D. (2011). Chapter 12 - skills, tasks and technologies: Implications for employment and earnings. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4 of *Handbook of Labor Economics*, pages 1043 – 1171. Elsevier.

Akerman, A., Gaarder, I., and Mogstad, M. (2015). The Skill Complementarity of Broadband Internet. *The Quarterly Journal of Economics*, 130(4):1781–1824.

Algan, Y., Dustmann, C., Glitz, A., and Manning, A. (2010). The economic situation of first- and second-generation immigrants in france, germany, and the uk. *The Economic Journal*, 120(542):F4–F30.

Angrist, J. D. and Guryan, J. (2008). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review*, 27(5):483 – 503.

Arai, M. and Thoursie, P. S. (2009). Renouncing personal names: An empirical examination of surname change and earnings. *Journal of Labor Economics*, 27(1):127–147.

Arrow, K. J. (1973). Higher education as a filter. *Journal of Public Economics*, 2(3):193–216.

Atkeson, A. and Kehoe, P. (2005). Modeling and measuring organization capital. *Journal of Political Economy*, 113(5):1026–1053.

Axtell, R. L. (2001). Zipf distribution of U.S. firm sizes. *Science*, 293(5536):1818–1820.

Becker, S. O., Egger, P. H., and von Ehrlich, M. (2010). Going NUTS: The effect of EU structural funds on regional performance. *Journal of Public Economics*, 94(9 - 10):578 – 590.

Becker, S. O., Egger, P. H., and von Ehrlich, M. (2013). Absorptive capacity and the growth and investment effects of regional transfers: A regression discontinuity design with heterogeneous treatment effects. *American Economic Journal: Economic Policy*, 5(4):29–77.

Benzarti, Y. and Harju, J. (2019). Are taxes turning humans into machines? Using payroll tax variation to estimate the capital-labor elasticity of substitution. Working paper.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Blanchard, O. and Perotti, R. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *The Quarterly Journal of Economics*, 117(4):1329–1368.

Bond, S. and Reenen, J. V. (2007). Chapter 65 - Microeconometric models of investment and employment. *Handbook of Econometrics*, 6:4417 – 4498.

Borjas, G. J. (1995). Assimilation and changes in cohort quality revisited: What happened to immigrant earnings in the 1980s? *Journal of Labor Economics*, 13(2):pp. 201–245.

Bratsberg, B., Barth, E., and Raaum, O. (2006). Local unemployment and the relative wages of immigrants: Evidence from the current population surveys. *The Review of Economics and Statistics*, 88(2):243–263.

Bronzini, R. and de Blasio, G. (2006). Evaluating the impact of investment incentives: The case of Italy's law 488/1992. *Journal of Urban Economics*, 60(2):327 – 349.

Bryson, A. and Kleiner, M. M. (2010). The regulation of occupations. *British Journal of Industrial Relations*, 48(4):670–675.

Brücker, H., Kroh, M., Bartsch, S., Göbel, J., Kühne, S., Lindau, E., Trübswetter, P., Tucci, I., and Schupp, J. (2014). The new IAB-SOEP migration sample: An introduction into the methodology and the contents. *SOEP Survey Papers 216, Series C.*

Bundesagentur für Arbeit (2018). Berufenet: Reglementierte Berufe.

Busso, M., Gregory, J., and Kline, P. (2013). Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review*, 103(2):897–947.

Cantore, C., León-Ledesma, M., McAdam, P., and Willman, A. (2014). Shocking stuff: Technology, hours, and factor substitution. *Journal of the European Economic Association*, 12(1):108–128.

Cerqua, A. and Pellegrini, G. (2014). Do subsidies to private capital boost firms' growth? A multiple regression discontinuity design approach. *Journal of Public Economics*, 109:114 – 126.

Chapman, B. J. and Iredale, R. R. (1993). Immigrant qualifications: Recognition and relative wage outcomes. *International Migration Review*, 27(2):pp. 359–387.

Chirinko, R. S. (2000). Investment tax credits. Working Paper Series 243, CESifo.

Chirinko, R. S. (2002). Corporate taxation, capital formation,and the substitution elasticity between labor and capital. *National Tax Journal*, 55(2):339 – 355.

Chirinko, R. S., Fazzari, S. M., and Meyer, A. P. (2011). A new approach to estimating production function parameters: The elusive capital—labor substitution elasticity. *Journal of Business & Economic Statistics*, 29(4):587–594.

Chirinko, R. S. and Wilson, D. J. (2008). State investment tax incentives: A zero-sum game? *Journal of Public Economics*, 92(12):2362 – 2384.

Chiswick, B. R. and Miller, P. W. (2003). The complementarity of language and other human capital: Immigrant earnings in Canada. *Economics of Education Review*, 22(5):469 – 480.

Collard-Wexler, A. and De Loecker, J. (2016). Production function estimation with measurement error in inputs. Working Paper 22437, National Bureau of Economic Research.

Combes, P.-P. and Gobillon, L. (2015). Chapter 5 - The empirics of agglomeration economies. In Duranton, G., Henderson, J. V., and Strange, W. C., editors, *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics*, pages 247 – 348. Elsevier.

Cooper, R. W. and Haltiwanger, J. C. (2006). On the nature of capital adjustment costs. *The Review of Economic Studies*, 73(3):611–633.

Criscuolo, C., Martin, R., Overman, H. G., and Reenen, J. V. (2016). The causal effects of an industrial policy. Discussion Paper 1113, Centre for Economic Performance.

Crépon, B., Ferracci, M., Jolivet, G., and van den Berg, G. J. (2009). Active labor market policy effects in a dynamic setting. *Journal of the European Economic Association*, 7(2-3):595–605.

Cummins, J. G., Hassett, K. A., Hubbard, R. G., Hall, R. E., and Caballero, R. J. (1994). A reconsideration of investment behavior using tax reforms as natural experiments. *Brookings Papers on Economic Activity*, 1994(2):1–74.

Dettmann, E., Brachert, M., and Titze, M. (2016). Identifying the effects of place-based policies: Causal evidence from germany. Working Paper 5901, CESifo.

Diamond, P., McFadden, D., and Rodriguez, M. (1978). Measurement of the elasticity of factor substitution and bias of technical change. In Fuss, M. and McFadden, D., editors, *Production Economics: A Dual Approach to Theory and Applications*, volume 2, chapter 5. McMaster University Archive for the History of Economic Thought.

Dornbusch, R., Fischer, S., and Samuelson, P. A. (1980). Heckscher-Ohlin trade theory with a continuum of goods. *The Quarterly Journal of Economics*, 95(2):203–224.

Dustmann, C. and Frattini, T. (2013). Immigration: The European experience. In Card, D. and Raphael, S., editors, *Immigration, Poverty, and Socioeconomic Inequality*, chapter 13. Russell Sage Foundation.

Dustmann, C., Frattini, T., and Preston, I. P. (2013). The effect of immigration along the distribution of wages. *The Review of Economic Studies*, 80(1):145–173.

Dustmann, C. and Glitz, A. (2011). Chapter 4 - Migration and education. In Eric A. Hanushek, S. M. and Woessmann, L., editors, *Handbook of The Economics of Education*, volume 4 of *Handbook of the Economics of Education*, pages 327 – 439. Elsevier.

Dustmann, C. and Glitz, A. (2015). How do industries and firms respond to changes in local labor supply? *Journal of Labor Economics*, 33(3):711–750.

Edgerton, J. (2010). Investment incentives and corporate tax asymmetries. *Journal of Public Economics*, 94(11):936 – 952.

Etzel, T. and Siegloch, S. (2018). Efficiency and equity effects of place-based policies: Evidence from capital subsidies in east germany. Working paper.

Fitzenberger, B., Osikominu, A., and Völter, R. (2005). Imputation rules to improve the education variable in the IAB employment subsample. *FDZ-Methodenreport*, 03/2005.

Fredriksson, P. and Johansson, P. (2008). Dynamic treatment assignment: The consequences for evaluations using observational data. *Journal of Business & Economic Statistics*, 26(4):435–445.

Friedberg, R. M. (2001). The impact of mass migration on the Israeli labor market. *Quarterly Journal of Economics*, 116(4):1373–1408.

Garicano, L., Lelarge, C., and Van Reenen, J. (2016). Firm size distortions and the productivity distribution: Evidence from France. *American Economic Review*, 106(11):3439–79.

Gathmann, C., Helm, I., and Schönberg, U. (2018). Spillover effects of mass layoffs. *Journal of the European Economic Association*, forthcoming.

Gittleman, M., Klee, M. A., and Kleiner, M. M. (2018). Analyzing the labor market outcomes of occupational licensing. *Industrial Relations*, 57(1):57–100.

Glaeser, E. L. and Gottlieb, J. D. (2008). The economics of place-making policies. *Brookings Papers on Economic Activity*, 39(1 (Spring):155–253.

Gomez, R., Gunderson, M., Huang, X., and Zhang, T. (2015). Do immigrants gain or lose by occupational licensing? *Canadian Public Policy*, 41(Supplement 1):S80–S97.

Gourio, F. and Roys, N. (2014). Size-dependent regulations, firm size distribution, and reallocation. *Quantitative Economics*, 5(2):377–416.

Greenstone, M., Hornbeck, R., and Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 118(3):536–598.

Guner, N., Ventura, G., and Xu, Y. (2008). Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, 11(4):721 – 744.

Hall, R. E. and Jorgenson, D. W. (1967). Tax policy and investment behavior. *American Economic Review*, 57(3):391–414.

Hamermesh, D. S. (1993). *Labor Demand.* Princeton University Press Princton, N.J.

Hassett, K. A. and Hubbard, R. G. (2002). Chapter 20 - Tax policy and business investment. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume 3 of *Handbook of Public Economics*, pages 1293 – 1343. Elsevier.

Hayashi, F. (1982). Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica*, 50(1):213–224.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, 64(4):605–654.

Hethey-Maier, T. and Schmieder, J. F. (2013). Does the use of worker flows improve the analysis of establishment turnover? Evidence from German administrative data. *Schmollers Jahrbuch*, 133(4):477–510.

House, C. L. and Shapiro, M. D. (2008). Temporary investment tax incentives: Theory with evidence from bonus depreciation. *American Economic Review*, 98(3):737–68.

Karabarbounis, L. and Neiman, B. (2013). The global decline of the labor share. *The Quarterly Journal of Economics*, 129(1):61–103.

Kleiner, M. M. (2017). The influence of occupational licensing and regulation. *IZA World of Labor*, 392.

Kleiner, M. M. and Krueger, A. B. (2010). The prevalence and effects of occupational licensing. *British Journal of Industrial Relations*, 48(4):676–687.

Kleiner, M. M. and Krueger, A. B. (2013). Analyzing the extent and influence of occupational licensing on the labor market. *Journal of Labor Economics*, 31(2):S173–202.

Kleiner, M. M. and Kudrle, R. T. (2000). Does regulation affect economic outcomes? The case of dentistry. *The Journal of Law and Economics*, 43(2):547–582.

Kleiner, M. M., Marier, A., Park, K. W., and Wing, C. (2014). Relaxing occupational licensing requirements: Analyzing wages and prices for a medical service. NBER working paper no. 19906.

Kline, P. and Moretti, E. (2014a). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority. *The Quarterly Journal of Economics*, 129(1):275.

Kline, P. and Moretti, E. (2014b). People, places, and public policy: Some simple welfare economics of local economic development programs. *Annual Review of Economics*, 6(1):629–662.

Klump, R. and Grandville, O. D. L. (2000). Economic growth and the elasticity of substitution: Two theorems and some suggestions. *The American Economic Review*, 90(1):282–291.

Koh, D. and Santaeulàlia-Llopis, R. (2017). Countercyclical elasticity of substitution. Working Paper 946, Barcelona GSE.

Koumenta, M. and Pagliero, M. (2016). Measuring prevalence and labour market impacts of occupational regulation in the EU. Technical report.

Krusell, P., Ohanian, L. E., Ríos-Rull, J.-V., and Violante, G. L. (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica*, 68(5):1029–1053.

Kugler, A. D. and Sauer, R. M. (2005). Doctors without borders? Relicensing requirements and negative selection in the market for physicians. *Journal of Labor Economics*, 23(3):437–466.

Leland, H. E. (1979). Quacks, lemons, and licensing: A theory of minimum quality standards. *Journal of Political Economy*, 87(6):1328–1346.

León-Ledesma, M. A., McAdam, P., and Willman, A. (2010). Identifying the elasticity of substitution with biased technical change. *American Economic Review*, 100(4):1330–57.

Lerche, A. (2018). Investment tax credits and the response of firms. Working paper.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2):317–341.

Lubotsky, D. (2007). Chutes or ladders? A longitudinal analysis of immigrant earnings. *Journal of Political Economy*, 115(5):pp. 820–867.

Lucas, R. E. (1978). On the size distribution of business firms. *The Bell Journal of Economics*, 9(2):508–523.

Maffini, G., Xing, J., and Devereux, M. P. (2016). The impact of investment incentives: Evidence from UK corporation tax returns. Working Paper 16/01, Oxford University Centre for Business Taxation.

Mattoo, A., Neagu, I. C., and Ozden, C. (2008). Brain waste? Educated immigrants in the US labor market. *Journal of Development Economics*, 87(2):255–269.

Mertens, K. and Ravn, M. O. (2013). The dynamic effects of personal and corporate income tax changes in the United States. *American Economic Review*, 103(4):1212–47.

Moreau, F. (2019). Inferring capital-labor substitution from firm-level distortions. Working paper.

Moretti, E. (2010). Local multipliers. *American Economic Review*, 100(2):373–77.

Moretti, E. (2011). Chapter 14 - Local labor markets. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4 of *Handbook of Labor Economics*, pages 1237 – 1313. Elsevier.

Moretti, E. and Thulin, P. (2013). Local multipliers and human capital in the United States and Sweden. *Industrial and Corporate Change*, 22(1):339–362.

Neumark, D. and Simpson, H. (2015). Chapter 18 - Place-based policies. In Duranton, G., Henderson, J. V., and Strange, W. C., editors, *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics*, pages 1197 – 1287. Elsevier.

Oberfield, E. and Raval, D. (2014). Micro data and macro technology. Working Paper 20452, National Bureau of Economic Research.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297.

Pagliero, M. (2011). What is the objective of professional licensing? Evidence from the US market for lawyers. *International Journal of Industrial Organization*, 29(4):473 – 483.

Raval, D. R. (2019). The micro elasticity of substitution and non-neutral technology. *The RAND Journal of Economics*, 50(1):147–167.

Romer, C. D. and Romer, D. H. (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review*, 100(3):763–801.

Rowthorn, R. (1999). Unemployment, wage bargaining and capital-labour substitution. *Cambridge Journal of Economics*, 23(4):413–425.

Sauer, S. and Strobel, T. (2015). Technologieintensität von Investitionen nach Branchen. Report 15, CESifo Institute.

Spence, M. (1973). Job market signalling. *Quarterly Journal of Economics*, 87(3):355–374.

Stigler, G. (1971). The theory of economic regulation. *Bell Journal of Economics*, 2(1):3–21.

Summers, L. H. (1981). Taxation and corporate investment: A q-theory approach. *Brookings Papers on Economic Activity*, (1):67–140.

Sweetman, A., McDonald, J. T., and Hawthorne, L. (2015). Occupational regulation and foreign qualification recognition: An overview. *Canadian Public Policy*, 41(S1):1–13.

Tani, M. (2015). Local signals and the returns to foreign education. Technical report, IZA Discussion Paper No 9597.

Tani, M. (2018). Selective immigration, occupational licensing, and labour market outcomes of foreign-trained migrants. Technical report, IZA Discussion Paper No 11370.

Thornton, R. J. and Timmons, E. J. (2013). Licensing one of the world's oldest professions: Massage. *The Journal of Law and Economics*, 56(2):371–388.

Timmons, E. J. and Thornton, R. J. (2008). The effects of licensing on the wages of radiologic technologists. *Journal of Labor Research*, 29(4):333–346.

Timmons, E. J. and Thornton, R. J. (2010). The licensing of barbers in the USA. *British Journal of Industrial Relations*, 48(4):740–757.

Vicari, B. (2014). Degree of standardised certification of occupations: An indicator for measuring institutional characteristics of occupations. *FDZ-Methodenreport*, 4.

Vikström, J. (2017). Dynamic treatment assignment and evaluation of active labor market policies. *Labour Economics*, 49:42–54.

Zwick, E. and Mahon, J. (2017). Tax policy and heterogeneous investment behavior. *American Economic Review*, 107(1):217–48.