Three Essays on Labor Economics

Maria Theresa Koch

Munich 2022

Three Essays on Labor Economics

Inauguraldissertation
zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

vorgelegt von

Maria Theresa Koch

2022

Referent: Prof. Panu Poutvaara, Ph.D.

Korreferent: Prof. Dr. Lutz Bellmann

Promotionsabschlussberatung: 13. Juli 2022

Acknowledgments

Für meine Familie, ohne die alles nichts wäre.

This dissertation would not have been possible without the support of many people. I am grateful to my advisor and coauthor, Panu Poutvaara, for his invaluable guidance and support during the writing of this thesis. This work has benefited greatly from his advice, criticism, and constant encouragement. Furthermore, I would like to thank my second supervisor, Lutz Bellmann, who has always been generous with his time and has supported me in many ways. I would also like to thank Andreas Haufler for serving as the third member of my dissertation committee and for inspiring me together with his team to pursue my Ph.D.

Many great colleagues and friends shaped my time at IAB. I would especially like to thank Hannah Illing, who was a superb coauthor for the second chapter of this dissertation. It was a special pleasure to share all the ups and downs of the dissertation with her. I would also like to thank Gesine Stephan, my mentor in the IAB graduate program ("GradAB"), who always had an open ear for me when needed.

Further, I am very grateful for my time as a "GradAB" and the financial and nonfinancial support that came with it. I am thankful for the interactions with my fellow GradAB Ph.D. students (especially Martin Friedrich, Karolin Hiesinger, Max Kunaschk and Martin Popp), as well as the experiences with all my other colleagues at the IAB (special thanks to the BB department, the BeCovid team, the IL, and my mentoring cohort). I am fortunate for all the experiences and the supportive environment at the IAB. In addition, this work has benefited from comments from participants at various workshops and conferences.

In addition to the scientific experience, I have been able to gain experience in many other areas at IAB, which is very special for a Ph.D. program. These include my internship at the Federal Ministry of Labor and Social Affairs, my time at the Operative Service of the Federal Employment Agency, and my participation in the IAB Mentoring Program. I am also particularly grateful for my time at the IAB Institutsleitung, which was characterized by great teamwork and many valuable experiences. My special thanks go to my Referentinnen colleagues, from whom I was able to learn so much.

This list of acknowledgments would be incomplete without mentioning my family and friends. Thank you to all of them: I was very happy to have you always by my side. A special thanks goes to Käthe, Miriam, and Vroni, on whom I could rely in all situations. But most of all, I am very grateful to my family especially my parents for their continuous support throughout the years.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Innovation is the driving factor behind technological change and economic growth (Romer [1990], Aghion and Howitt [1992]). According to the Romer [1990] model, short-term economic growth is possible by increasing the input of capital and labor. However, due to declining marginal products, this is not sustained growth. Long-term growth depends on innovation and the stock of (human) capital devoted to R&D activities. Both types of growth highlight the importance of a country's workforce, especially for innovation activities: Sufficient and well-trained workers are a prerequisite for a country's economic growth and prosperity. Different from basic theory (Romer [1990]), labor is not exogenous in practice but can be influenced by economic policy and other factors. A high rate of labor market participation in the total workforce is crucial for economic activity. This raises different questions in the labor market context: On the one hand, what can be done to ensure that enough workers in the labor force guarantee short-term growth? And on the other hand, how can we increase the availability of labor in the R&D sector to enable long-term growth?

Especially with regard to the upcoming skills shortage and demographical changes, the topic is highly relevant (Klinger and Fuchs [2019], German Council of Economic Experts (GCEE) [2017], Federal Ministry for Economic Affairs and Climate Action [2022]). To date, there is still an unused potential, especially among specific groups such as migrants, women, and older people (German Council of Economic Experts (GCEE) [2017]). Therefore, the labor market participation of these groups (especially in STEM fields) is an important goal in the upcoming years and decades. The three papers of this dissertation contribute to this goal by examining the labor market participation of individuals at the micro level using different statistical methods. Empirical evidence helps to understand possible improvements in the labor market participation of specific groups and is important for future labor market policies. The first two papers of this thesis focus on the effects of displacement on (migrant) workers and their reintegration into the labor market. The third paper investigates female inventors. For all three papers, I take advantage of the rich administrative employer-employee data provided by the Institute for Employment Research (IAB).

The first and second papers analyze the effects of job displacement due to a mass layoff on individual workers' labor market trajectories. Prior research finds that displacement due to a mass layoff can lead to long-lasting earnings and employment losses for displaced workers (e.g., Jacobson et al. [1993], Couch and Placzek [2010], Davis and von Wachter [2011], Lachowska et al. [2020], Schmieder et al. [2020]), but studies focusing on migrants in this context are scarce. The first paper uses an event study approach to contrast the costs of displacement for native workers compared to migrant workers. The paper provides new insights into the emerging literature focusing on the costs of job loss by worker type (e.g., Meekes and Hassink [2020], Blien et al. [2020]) and investigating these heterogeneities. The second paper identifies the effect of social networks on the re-employment probability of migrants and natives after displacement using an IV estimation strategy. Using newly available geo-coded social security data, I estimate the effect of the employment rate within a displaced worker's neighbor and coworker network on his re-employment probability one year after displacement. This paper contributes to a large body of migration literature arguing that individuals embedded in a network benefit from relevant (social) resources with respect to their labor market outcomes (Battisti et al. [2021], Damm [2014], Edin et al. [2003], Glitz [2017], Granovetter [1977]). The third paper focuses on female inventors. Using a newly available dataset that comprises labor market biographies of inventors in Germany, we investigate female inventors' characteristics and add new evidence to the literature on how to foster innovation by increasing the number of (high-potential but currently underrepresented) inventors (Bell et al. [2019]). In the following, I will present the three papers in more detail:

Paper 1: Who Suffers the Greatest Loss? Costs of Job Displacement for Migrants and Natives.[1]

The first paper investigates the costs of job loss for migrants and natives. Due to increasing immigration to Germany in the past decades, on the one hand, and skill shortages and demographic change, on the other hand, migrants' successful labor market integration is crucial. The research question that the paper aims to answer is how quickly and successful migrants, compared to natives, reintegrate into the labor market after displacement. Using rich administrative register data from Germany from 1996-2017, we compare the labor market outcomes of displaced migrants (individuals with non-German citizenship) and natives. These data include the employment biographies of all employees covered by social security in Germany.

[1]A prior version of this paper was published as discussion paper: H. Illing and T. Koch. Who suffers the greatest loss? Costs of job displacement for migrants and natives. IAB Discussion Paper, 08/2021, 2021, https://www.iab.de/183/section.aspx/Publikation/K210427JCN [last access: 03-15-2022]

Using the variables covered by the extensive data, we contribute to the literature on heterogeneity in the costs of job loss by worker type (see, e.g., Blien et al. [2020] for differences by occupational routine intensity, Helm et al. [2021] for differences by wage group, and Meekes and Hassink [2020] and Illing et al. [2021] for differences by gender).

Following Jacobson et al. [1993], we contrast the labor market outcomes of displaced to nondisplaced workers prior to and after a job loss for migrants and natives, respectively. The idea is that job loss is more likely to be a surprise for workers with high tenure, and therefore, these workers are highly comparable to workers with similar characteristics who keep their job in the same year. This procedure provides us with the cost of job loss, for migrants and natives. However, comparing migrants to natives is still challenging. To make migrants' and natives' costs of job loss more comparable, we reweight migrants to natives based on their observable characteristics (see DiNardo et al. [1996]). This enables us to control for migrants' and natives' different individual characteristics and differential sorting across industries and occupations before displacement.

We find that both migrants and natives face large earnings losses after displacement, when looking at the matching results, but migrants' losses are larger (12,000 EUR vs. 16,000 EUR in the year after losing their job compared to earnings two years earlier). The event study regression model (with worker and year fixed effects) reveals that the decline in migrants' earnings in the year of the layoff is 12 percentage points larger than that of natives. Even five years after displacement, migrants do not fully catch up with natives, which indicates persistent differences. Using the reweighting scheme and making migrants and natives comparable based on individual characteristics, industry, and occupation, the difference in earnings immediately after displacement decreases to 5 percentage points and disappears over time.

Earnings losses may stem from wage and/or employment losses. We find that both wage and employment losses are larger for migrants when looking at the raw results. When taking reweighting into account, the results reveal that while observable characteristics fully explain the gap in wage losses (conditional on finding a job), there is still a gap in employment. Migrants are approximately 5 percentage points less likely to be employed in the year after job loss. This difference decreases to approximately 2 percentage points five years later. The results show a similar pattern for days worked per year: Migrants work approximately 25 fewer days per year in the year after displacement; five years later, the gap is still statistically significant but shrinks to approximately 10 days.

However, what are the possible channels that explain the differences in earnings losses? We investigate this in three steps. First, we explore whether migrants work for different types of establishments after displacement (conditional on employment). We find that migrants indeed work for worse establishments after displacement in terms of wages and establishment fixed effects[2] (Abowd et al. [1999]), and the share of marginally employed workers. However, the earnings gap is much weaker once controlling for observable differences via reweighting. Nevertheless, migrants' sorting into different establishments can only partly explain their higher costs of job loss.

Second, we test whether differences in mobility patterns explain the diverse costs of job loss. As in Huttunen et al. [2018], the results reveal that both migrants and natives expand their regional mobility (change in workplace location, commuting) after a job loss, conditional on employment. We find that migrants are more likely to commute (2 percentage point difference compared to natives) and less likely to change workplaces to a new federal state (a 3 percentage point difference) after displacement. One explanation for migrants' lower mobility is higher constraints, e.g., on the housing market. If migrants face higher difficulties finding new housing, this explains their lower mobility and, in turn, their larger earnings losses.

Third, we investigate the role of local labor market concentration by using three proxies: i) the change in local unemployment rates around the time of displacement, ii) city residency, and iii) the share of same-nationality working-age population in a worker's workplace county. Why should these proxies impact labor market outcomes after a job loss? Previous literature finds that i) migrants' wages assimilate slowly in periods of high unemployment (Bratsberg et al. [2006]), ii) displaced workers face longer unemployment if living in cities (Haller and Heuermann [2019]), and iii) competition among similar workers is particularly high for migrants (e.g., Albert et al. [2021], Beaman [2012]).

To empirically test these hypotheses, we follow Schmieder et al. [2020]. We use a matched difference-in-differences (DID) analysis to measure the relevance of local labor market concentration. First, we find for each displaced worker a simi-

[2]Based on Abowd et al. [1999], a large body of literature finds that persistent wage differentials exist across firms within the same labor market (e.g., Card et al. [2013], Bonhomme et al. [2019], Song et al. [2019]).

lar nondisplaced worker. Second, we build a variable measuring the difference in earnings before and after job loss between each displaced and nondisplaced worker pair. Last, we regress this variable on the three proxies and numerous worker- and establishment-level controls for the displaced workers. We find that displaced workers, irrespective of migration status, face larger earnings losses if local unemployment rates at the time of displacement increase more.

Moreover, earnings losses are higher if one lives in an urban area at the time of displacement, and this negative effect is about twice as high for migrants. In line with Caldwell and Danieli [2021], we also find that migrants working in counties with a higher share of the same-nationality population in the year before displacement have higher costs of job loss. Migrants with similar characteristics in the German labor market, e.g., in terms of recognition of educational degrees, face higher competition for the same types of jobs.

While recent research focuses on the differences in displacement costs among different worker groups (see, e.g., Blien et al. [2020], Illing et al. [2021], Meekes and Hassink [2020]), to date, no study has examined migrants and natives. One unanswered question is whether taste-based or statistical discrimination is the driver behind the migrant-native gap in earnings losses. Previous literature confirms that there is discrimination when new (migrant) workers are hired (e.g., Bertrand and Mullainathan [2004]). With our data at hand, we cannot provide evidence on discrimination in the context of job loss. For example, investigating the role of discrimination, with the help of survey data, is left for future research. However, the results of this paper confirm that migrants and natives face different labor market possibilities after a layoff. Enhancing migrants' labor market participation, e.g., by providing special courses for the recruitment process or facilitating recognition processes, could lead to better labor market outcomes after a displacement for both migrants and natives.

Paper 2: Coworkers, Neighbors, and Job-Search: Migrants' and Natives' Re-employment after a Layoff.

The second paper of this thesis investigates the importance of social networks for people's re-employment probability after a layoff. The paper measures the effect of the employment rate within one's coworker and neighbor networks on the re-employment probability one year after a layoff. A coworker network includes all prior colleagues one to five years prior to displacement, while neighbors include all people living in the same neighborhood in this period. Although previous literature investigates both theoretically and empirically how social networks impact labor market outcomes, no paper so far shows how *different* kinds of networks do this at the individual level. As individuals belonging to different networks simultaneously, this question is highly relevant, yet it is unanswered.

When considering different types of networks, there are various potential reasons why they affect workers differently. Former coworkers belong to the group of work-related contacts, whereas neighbors are private contacts. Therefore, prior coworkers can easily assess workers' skills and are aware of suitable job offers (because they work for firms with open vacancies or look for a job themselves). Neighbors, however, compete less likely for jobs than prior coworkers, who worked in the same industry and selected into the same firms, and neighbors are more aware of geographically close job offers. This paper is the first to investigate how diverse networks affect re-employment probability after a layoff. Prior literature focused on one network at a time, probably because of data limitations.

This paper uses register data from Germany and matches them with new geo-coded residence data. Combining these two high-quality data sources allows us to estimate the effect of different networks on a worker's re-employment probability. For computational reasons, I focus on the universe of workers in Germany's four largest metropolitan areas (Cologne, Frankfurt, Hamburg, Munich)[3] as defined by Kropp and Schwengler [2016] in 1995-2015. The sample comprises all male workers who were displaced due to an establishment closure in 2005-2011. The intuition behind focusing on displaced workers is (i) they lose their jobs involuntarily, due to the establishment closure, irrespective of their abilities and networks, and (ii) the workers are highly comparable. The variation used in this paper comes from the differences in the employment and living biographies of different workers who became unemployed due to an establishment closure. The coworker network includes all workers with whom a person worked in the same establishment in the 5 years prior to layoff. Before identifying the neighbor networks, I first define every displaced worker as being located within a 1 times 1 kilometer square. Neighbor networks are then defined as grids of all workers who lived within 3 times 3 kilometer squares (centered around the 1 times 1 square the displaced worker lives in) in the past five years.

[3]Due to German reunification in 1989 and Berlin's division in the years prior, Berlin is excluded as a special case. Following Kropp and Schwengler [2016], I define metropolitan areas as the cities themselves and the surrounding areas within commuting distance.

To empirically assess the effect of a displaced worker's network employment share on his re-employment probability, I run an IV estimation following Glitz [2017]. In addition to including different control variables in the regression, I further include displacement establishment fixed effects (county fixed effects) for coworker (neighbor) networks. The instrument for the employment share within a network is the share of mass layoffs within the past 5 years within the respective network. The idea behind this IV approach is that conditional on observable characteristics (e.g., industry or county), a displaced worker's re-employment probability is independent of the mass layoffs within this worker's network.

The results reveal that both coworker and neighbor networks positively contribute to a displaced worker's re-employment one year after displacement. However, the paper finds that coworker networks are more relevant for re-employment: For coworker (neighbor) networks, a 10 percentage point increase in the employment rate raises the probability of being employed after one year by 4.9 (0.7) percentage points.

The paper also highlights diverse effects for (i) different workers and (ii) differences in the networks. In a first heterogeneity analysis, I check whether the displaced workers' migration status plays a role. The results show a positive effect of the employment rate within a coworker network for natives (5.0 percentage points) but an insignificant effect for migrants. However, when focusing on only migrants within the displaced worker's network, I find positive effects for both natives (4.7 percentage points) and migrants (5.2 percentage points). Consequently, for migrants, especially migrants within their coworker network, play an important role. The results detect a positive effect on natives, irrespective of their networks' migration status for neighbor networks. The picture is once more different for displaced migrants: While native neighbors' employment share positively impacts their re-employment, there is no significant effect of migrant neighbors. This finding is consistent with papers predicting competition among migrants living in the same neighborhood (Beaman [2012], Albert et al. [2021]).

The second heterogeneity analysis takes, in addition to the migration status, the education status into account. The findings show that the lower the education status is, the more relevant coworkers and neighbor networks are. This is in line with considerations arguing that different networks are of different quality in diffusing open vacancies and assessing workers' skills. Therefore, a neighbor network transferring nonspecialized tasks in the neighborhood may be more beneficial to low- and medium-skilled workers than for experts in their fields, who have to search for jobs outside their neighborhood.

This paper is the first to show how different types of networks affect displaced workers' re-employment after an establishment closure. Additionally, the paper zooms into how the two network types differ for migrants and natives. Furthermore, to compute neighbor networks, the paper uses geo-coded data and defines the displaced workers as the central point within grids of 3 square kilometers in size, irrespective of administrative borders such as counties or municipalities. As the results of this paper show that both network types are relevant for job search, policymakers could use networks more consciously for unemployed individuals, e.g., by utilizing fairs or mentoring programs. Facilitating exchange could result in higher employment rates after a displacement.

In pandemic times, however, personal contact is not always possible. Evaluating the success of virtual formats and groups, e.g., neighborhood groups in Facebook, is an interesting yet unanswered question. Using field experiments and surveys could be an interesting tool to get to know the effects of online acquaintances on re-employment, but this goes beyond the research question of this paper.

Paper 3: Female inventors.
The third paper of this thesis focuses on scarce female inventors, as increasing the number of inventors is crucial for economic growth (Romer [1990]). The paper adds to the economic literature focusing on the factors that determine who becomes an inventor (especially the seminal paper by Bell et al. [2019]). This literature aims to increase the number of inventors, especially among underrepresented groups such as women and minorities. According to Bell et al. [2019], girls are more likely to invent (in a particular field) if they grew up in an area with more female inventors (in that particular field). Additionally, Bell et al. [2019] find that women are underrepresented among the star (most productive) inventors, as are women among all inventors. Missing female role models and networks (exposure effect) are central explanations for these findings. Due to the coronavirus disease (COVID-19) pandemic, Özlem Türeci and Uğur Şahin, who founded BioNTech, are now prominent examples of high potential among currently underrepresented groups. Both researchers are Germans with an immigrant background, and one of them is a woman. These two positive examples show that effectively harnessing the talent of potential in-

ventors by encouraging the "lost Einsteins" to innovate yields significant welfare gains. However, apart from the work of Bell et al. [2019], there is little economic evidence on gender differences among inventors. We contribute to filling this gap by examining the characteristics of female inventors in Germany. Since Germany is among the countries with the highest number of patents, it is an interesting case for investigation.

Women are underrepresented among inventors. This is shown by data from all over the world. Even though the proportion of female inventors has increased in recent decades, gender equality has not yet been achieved. In Germany, the share of female inventors is lower (10 percent in 2015) than in France, Italy, Spain, and the United Kingdom. This article aims to help increase the proportion of female inventors by examining what we know about current female inventors in Germany. The main reason for the gap in the literature is that data on inventors, especially in the labor market context, are scarce. We address this problem by using newly available patent and employment register data on all inventors in Germany who filed a patent at the European patent office between 2000 and 2010. The dataset provides us with full employment biographies of male and female inventors in the German labor market and provides detailed insights.

But when we think about female inventors, we need to start earlier in the employment biography. In many European countries (including Germany), the majority of university graduates are female, except for in STEM subjects (see, e.g., https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Tertiary_education_statistics#Graduates [last access: 12-20-2021]). We know that fewer women with good math grades (an indicator of high potential) become inventors than men (Bell et al. [2019]). Overall, women appear to be choosing not to pursue careers in STEM fields. In this context, the literature refers to the "leaky pipeline," i.e., an ever-decreasing proportion of women reaching each level of higher education (see, e.g., Carrell et al. [2010], Mansour et al. [2021], Buckles [2019]). While the proportion of women in higher education is over half, only 43% of graduates worldwide earn doctorates. Less than one-third of all researchers are women. The proportion of women filing patents as inventors is even lower (11%).

In the first part of the analysis, we descriptively investigate inventors' personal characteristics. A comparison of inventors with a random sample of workers in Germany shows that inventors are on average older, more educated, and earn higher wages.

Female inventors (noninventors) earn less than male inventors (noninventors). The analyses show that women are more likely to work in pharmaceuticals and biotechnology fields than in mechatronics. We also find that women are more likely to work in fields where the total number of patents (in absolute terms) is lower. In the second part, we compare female and male inventors with respect to patent characteristics. We find that female inventors have, on average, a lower number of cited, applied, and rejected patents than their male counterparts. There are also fewer star inventors, i.e., there are fewer highly paid and productive inventors among women than among men (both in absolute numbers and as a share among female/male inventors). Finally, the share of top-coded inventors (above the social security threshold) is lower among women than men.

Furthermore, we run multivariate OLS and logit regressions. In these, we control for additional personal characteristics. These characteristics include (in addition to gender) education, migration status, and a dummy for motherhood among women. The regressions also include numerous fixed effects (firm size, industry, year, and age group). The regression estimates show that female inventors have a higher patent rejection rate and fewer applied and granted patents (all measured at the individual level). Regarding citations between different patent offices, the picture is mixed: For Germany, female inventors have a lower number of citations, while for the European and US Patent Offices, female inventors have a higher number of citations. We have four possible explanations for the mixed picture: (i) Women are more likely to work in teams than alone (Jung and Ejermo [2014], Mauleón and Bordons [2010], Naldi et al. [2004]) and therefore benefit from larger, international networks of the larger teams and are more likely to be cited. (ii) Women who are granted international patents are a highly productive (and selective) group of highly talented women, as international patents are more difficult to obtain and thus they are more frequently cited. (iii) Perhaps women are less discriminated against in more advanced countries with a higher proportion of female inventors. (iv) Different technological fields in patenting could play a role here. Women tend to work more often in biotechnology and related fields that are relatively stronger in many other countries than Germany, while men work in machinery and mechatronics, where Germany is among the leading countries. For this reason, women's patents may be relatively more relevant for research done abroad and men's patents for research done in Germany.

In the last part, we consider mothers among inventors and noninventors. We find that female inventors are less likely to become mothers than a random sample of female employees in Germany ("average" women in the German labor market). For female inventors, the probability of becoming a mother is higher if the inventor is highly productive (as measured by the number of patent grants and citations in Germany). Regarding parental leave, the results show that female inventors are on parental leave longer than "average" women in Germany. Complete exit from the labor market is also more common among female inventors. Female inventors are less likely to return to either part-time or full-time work than a percent random sample of females in Germany, looking at the return to the labor market after the birth of a child. A possible explanation is that working as an inventor requires full-time employment or that jobs in research and development are less family-friendly than other jobs.

This study set out to identify the characteristics of women who decided to become inventors in the past. The paper provides new insights into female inventorship, which is a rather unexplored field due to data limitations. We use the newly available INV-BIO dataset that combines patent data from the German patent office with administrative data from Germany. Germany is an interesting country to investigate since it is among the leading patenting countries regarding the overall number of patents, and its industry includes diverse research fields. It is hoped that this research will contribute to a deeper understanding of female inventors' characteristics. This might be a first step in avoiding lost "Marie Curies" in the future.

The findings of this paper (along with others) can be used to develop targeted interventions aimed at increasing the number of female inventors. Girls who grow up in areas with a high share of female inventors are more likely to become inventors themselves, the so-called exposure effect (Bell et al. [2019]). A steady increase in female inventors and addressing the "leaky pipeline" contribute to higher female participation in the R&D sector. In this context, evaluating the effectiveness of different measures to encourage young women to study STEM subjects is highly relevant. These questions still remain to be answered by future research. In addition, the lower return rate of female inventors to part-time and full-time positions after the birth of a child (compared to "average" women) may indicate that family responsibilities are not easily combinable with a career as an inventor. A reasonable approach to tackle this issue could be to make work in the R&D sector more family-friendly.

Concluding remarks

The questions I raised in the first paragraph of this thesis "What can be done to ensure that enough workers in the labor force guarantee short-term growth?" and "How can we increase the availability of labor in the R&D sector to enable long-term growth?" are too complex to answer in a single dissertation. More appropriate would be the image of a mosaic to which three small stones have now been added. However, to obtain a comprehensive picture, much more research is needed to support politicians in their decision-making process.

Particularly against the backdrop of demographic change and the shortage of skilled workers, securing a sufficient labor supply is a key issue for politics and future research. Demographic aging will noticeably reduce the German labor force over the next 15 years (Klinger and Fuchs [2019]). Fuchs et al. [2017] conclude that Germany would need an annual net immigration of more than 400,000 from now on to maintain the number of potentially employable people at the current level. Thus, in the future, controlled immigration and making full use of the German labor market potential will play a key role in dampening the effects of demographic change. The rising labor force participation rates of women, older people, and recognized asylum seekers currently offset the declining labor force. Nevertheless, the labor force potential has not yet been exhausted, especially among women, migrants, and older people there is still an unused labor market potential. In the future, more use should be made of the workforce potential to mitigate the shortage of skilled workers caused by demographic change.

How can future research contribute to this process? First, forecasts to calculate the labor force potential are essential to address this change. These forecasts should also consider different aspects, such as the regional and industrial heterogeneity within Germany. In this context, it must also be borne in mind that the probable decline in the labor force potential does not necessarily lead to a shortage of skilled workers, as macroeconomic adjustment reactions can be expected (Weber [2016]). For instance, better training of employees and technological change could lead to the necessity of fewer employees. The high unemployment rate of low-skilled workers shows that part of the available labor potential is hardly used. With qualification requirements tending to rise in the future, higher investment in education could help mitigate the consequences of a declining labor force potential for the economy and the labor market in the longer term. Further, particular attention must be paid to the role of digitization. On the one hand, if productivity, e.g., as a result of increasing digitization and per capita national income rise, this could cushion the consequences of demographic change and counteract welfare losses. On the other hand, productivity increases in an aging society are not easy but require intensive efforts to further educate workers.

However, in addition to forecasting, it remains essential to (re)integrate the unemployed and part-time employed into the labor market as well as possible. Therefore, further research on this is needed, as shown in the first and second paper of this dissertation. The use of previously unexplored data sources, such as geo-coded data or telephone data, is an interesting new tool to study different aspects relevant for labor market integration, e.g., the influence of networks and social interaction. However, new forms of work (such as increasing working time flexibility) are also an interesting field for further discoveries. Especially considering the COVID-19 pandemic, home office options have changed dramatically and will probably never be the same as before the pandemic. It is important to have evidence-based results on how the trend toward home offices will affect the workforce in the long term, e.g., whether older or female workers will be willing to increase their working hours because remote work is now possible. However, artificial intelligence and new digital possibilities will also lead to serious changes in the labor market, and research into the consequences of digital and environmental change is crucial to future economic growth. These research areas are of central importance for securing Germany's economic prosperity in the long term and countering the upcoming hurdles due to the shortage of labor.

However, what guidance does current research give to policymakers to counter future labor shortages? Working conditions should be adapted and made more flexible so that currently unemployed people can take up work and part-time employees can increase their working hours. Examples would be (i) the further expansion of all-day childcare, (ii) a more flexible management of working time (e.g., the German Council of Economic Experts (GCEE) [2017] recommends replacing the maximum working time per day with a maximum working time per week), and (iii) the expansion of home office options. Relatively low investments to improve the reconciliation of work and family life can significantly impact women's labor force participation. This is also in line with our findings in paper three of this dissertation. According to Krebs and Scheffel [2016], annual investments in kindergartens and schools of 10 billion Euros could increase the labor supply by nearly 500,000 full-time equivalent jobs over five years.

The German Council of Economic Experts further advocates labor market-oriented immigration of skilled workers to Germany (German Council of Economic Experts (GCEE) [2017]). According to Miguélez and Moreno [2014], the keys to attracting skilled immigrants are job opportunities and wage levels. The changes in the immigration regulations of the past decades (EU blue card; simplified labor market access for persons from third countries, Skilled Migration Act) facilitated access to the German labor market (Organization for Economic Cooperation and Development (OECD) [2013], Expert Council on Integration and Migration [2015]). However, according to the German Council of Economic Experts and the German expert council on Integration and Migration, there is still room for improvement (German Council of Economic Experts (GCEE) [2017], Expert Council on Integration and Migration [2021]). For instance, the access to the labor market for well-integrated and working but rejected asylum seekers ("Spurwechsel") should be further facilitated. Another tool could be to enable visas for job search in Germany (especially in shortage sectors). Additionally, people from third countries who are willing to, should be enabled to start a vocational training in Germany. Further, access to the labor market should be eased and current hurdles (e.g., in the educational recognition process) overcome (Brücker et al. [2021]). Another argument that also paper two of this thesis points towards is the importance of social networks when thinking of labor market integration. The expert council on Integration and Migration highlights the importance of networks and role models, especially when thinking of the low rates of people with migration background working in the public sector (Expert Council on Integration and Migration [2021]). According to these experts, various studies show that people with a migration background are sometimes discriminated against in hiring procedures. It is therefore questionable to what extent diversity is actually consistently valued in the labor market (Expert Council on Integration and Migration [2021]). In this spirit, a welcoming culture in Germany is key for the success of all these interventions.

Chapter 2

Who Suffers the Greatest Loss? Costs of Job Displacement for Migrants and Natives.

Joint work with Hannah Illing.

2.1 Introduction

To fill the vacancies left by demographic change, firms in many countries rely on steady inflows of migrant workers.[1] For example, most of the increase in the US (65%) and the EU (92%) workforce in 2005-2015 was due to immigration (Rouzet et al. [2019]). Yet for destination countries to fully benefit from immigration, successful long-term labor market integration of migrants is crucial. In particular, destination countries need to ensure that migrants remain attached to the labor market, even after an economic shock such as job displacement.

Given how important immigration is for ageing labor markets, there is surprisingly little causal evidence on how well migrants integrate into destination countries' labor markets in the long run. While there exists descriptive evidence on the educational and labor market performance of first and second generation migrants (e.g., Dustmann et al. [2012], Algan et al. [2010]), there is no comprehensive micro-level study investigating how migrants, compared to natives, react to a labor market shock such as losing their job.

In an ideal world in which migrants were fully integrated in destination countries' labor markets, migrants' and natives' response to economic shocks should not differ. Yet, previous research points to migrants being a particularly vulnerable group in their destination country's labor market: Migrants are at high risk of losing their job during recessions (e.g., Fairlie et al. [2020], Montenovo et al. [2020]), and are less likely to be hired if unemployed (Forsythe and Wu [2021]). They moreover have lower average wages (e.g., Battisti et al. [2021], Borjas [1995]) and worse networks than natives (e.g., Gërxhani and Kosyakova [2020], Glitz [2017]). At the same time, their employment recovery after the first wave of the COVID-19 Pandemic was quicker than that of natives (Borjas and Cassidy [2021]). Thus, it is ex ante unclear whether and how labor market shocks may affect migrants and natives differently.

In this paper, we investigate this question by analyzing how job displacement disrupts workers' careers, and how migrant workers' post-layoff trajectories differ from natives'.[2] For this purpose, we use rich administrative employer-employee data from

[1] Many OECD countries are facing labor supply shortages as a result of demographic change. Societies are ageing fast: The share of individuals aged 65 or older in OECD countries has increased from less than 9% in 1960 to more than 17% in 2017, and it is projected to rise to 27% by 2050 (Organization for Economic Cooperation and Development (OECD) [2019]).

[2] Throughout this paper, we define migrants as individuals with non-German citizenship.

Germany provided by the Institute for Employment Research (IAB), which span more than 20 years (1996-2017). These data comprise the universe of employees covered by social security, and they are directly filed by employers, making them both representative and highly reliable. We use the rich set of individual characteristics recorded in the data to follow the growing literature on heterogeneity in the costs of job loss by worker type (see, e.g., Helm et al. [2021] for differences by wage group, Blien et al. [2020] for differences by occupational routine intensity, and Illing et al. [2021] and Meekes and Hassink [2020] for differences by gender).

Our empirical strategy builds on the seminal paper by Jacobson et al. [1993], who compare the labor market outcomes (e.g., earnings, log daily wages, and employment) of displaced to nondisplaced workers before and after job loss. After identifying displacement events, we use propensity score matching on a range of individual and establishment characteristics to find a suitable match for each displaced worker. We match within migration status, and estimate event study regressions separately for migrant and native matched worker pairs.[3]

Exploiting layoff events to compare the labor market trajectories of migrant and native workers comes with the advantage that we can keep workers' reasons for leaving an establishment constant. Workers may exit employment contracts for all sorts of reasons. For example, they may quit an establishment to switch to a position with higher pay, or establishments may fire workers if they are bad matches, and this may differ by migration status. Focusing on involuntary, unexpected job displacement thus helps us to better compare migrants' and natives' post-shock labor market trajectories. To ensure that job displacement is unexpected, we follow the job displacement literature (e.g., Schmieder et al. [2020], Davis and von Wachter [2011], Jacobson et al. [1993]) and focus on workers with at least three years of tenure in the German labor market. We thus compare natives to a sample of migrant workers with relatively high labor market experience in Germany.[4]

The key challenge of our study is to make migrants comparable to natives. Migrants in our sample are, on average, less educated (11.2 vs. 12.3 years), younger (37.9 vs.

[3]While we estimate our main results for a sample of men, we show in the Online Appendix, Section C, that women's labor market trajectories follow a very similar pattern.

[4]In a robustness check, we vary this restriction to one and two years of tenure, respectively, and this does not change our main results. In the Online Appendix Figure 2.8 we moreover show that the time migrant workers have spent in the German labor market upon job displacement matters surprisingly little for the magnitude of their earnings losses.

39.4 years), and earn lower daily wages (89.2 EUR vs. 102.3 EUR) in the year before displacement. These differences in observable characteristics most likely influence how migrants respond to job displacement compared to native workers.

To isolate the role of observable characteristics in explaining the gap, we use a reweighting scheme first proposed by DiNardo et al. [1996] and first applied to the context of job loss by Illing et al. [2021]. For this purpose, we estimate a probit regression where the dependent variable is a dummy for native displaced worker on a set of different individual characteristics, 1-digit industry dummies, and 1-digit occupation dummies, all measured before displacement. Migrant workers then receive a weight that is based on their propensity scores. We always report both the raw coefficients (without reweighting) and the adjusted coefficients net of observable characteristics (with reweighting).

Descriptively, we find that both migrants and natives face large average earnings losses after displacement, with substantially larger losses for migrants (16,000 EUR vs. 12,000 EUR in the year after losing their job, compared to earnings two years earlier). The results from our event study regression model, where we control for worker and year fixed effects, show that the raw decline in migrants' relative earnings in the year of the layoff compared to two years ealier is 12 percentage points larger than that of natives.[5] The long-term trend moreover shows that migrants do not catch up with natives even five years after displacement. Once we reweight the distribution of migrants' characteristics to natives' based on individual characteristics, 1-digit industry dummies, and 1-digit occupations, the earnings gap immediately after displacement shrinks to 5 percentage points and closes over time (no differences from year 4 after displacement).

We then decompose earnings into log daily wages (conditional on employment) and employment. We find that while observable characteristics fully explain the gap in wage losses (conditional on finding a job), the gap in employment losses persists even after reweighting. Differences in observable characteristics can thus explain why migrants earn lower wages after job displacement, but they cannot explain

[5] If we include spells with zero earnings (and thus account for workers in unemployment or workers temporarily unobserved due to, e.g., self-employment), this difference increases to 80 log points, meaning that the effects on migrants are 1.8 times the effects on natives. We construct a panel where we keep workers in the sample if they disappear from the social security data in a given year and appear again in a future year. If they fully disappear from the data, we only include them up to the last year they are observed in the data.

why migrants are less likely to become re-employed. Migrants are approximately 5 percentage points less likely to be employed in the year after losing their job, and this gap shrinks to approximately 2 percentage points five years after job loss. We observe a similar pattern for days worked per year: Migrants work approximately 25 fewer days per year than natives in the year after displacement; five years later, the difference is still statistically significant but reduced to approximately 10 days. Overall, migrants are more likely to sort into part-time rather than full-time employment after job displacement.

Next, we analyze whether migrants' and natives' mobility patterns (conditional on finding a new job) differ. In line with Huttunen et al. [2018], we find that both migrants and natives expand their regional mobility - both in terms of changing workplace location and commuting - after job loss. Our results suggest that migrants are slightly more likely to commute after job loss (a 2 percentage points difference compared to natives) and slightly less likely to move workplaces to a new federal state (a 3 percentage points difference). Migrants may thus face higher relocation constraints than natives (e.g., because of housing market tightness), and their lower propensity to relocate partly explains their larger earnings losses.[6] Overall, this is suggestive evidence for migrants searching for jobs more locally.[7]

The broad comparison of migrants and natives may hide substantial heterogeneity within groups. We thus document the adjusted migrant-native gap in costs of job displacement within educational groups. Our results show that workers with vocational training (corresponding to approximately 12 years of education) or without vocational training (10 years of education) drive this gap, while there is no gap for high-skilled workers (with university degrees). While for native workers, earnings losses do not vary substantially across education groups, the least-educated migrant workers face by far the largest losses. For example, earnings losses of migrants without vocational training are about 1.34 times the size of high-skilled migrants' losses, and migrants with vocational training lose about 1.2 times as much.

We moreover show that even within the group of migrant workers, there is substan-

[6]Online Appendix Table 2.9 shows that commuting and relocation patterns explain about 7% of the migrant-native earnings gap.

[7]We moreover show that migrants are about 4 percentage points more likely than natives to completely disappear from the social security data after job displacement, suggesting that they may return to their home countries. Conditional on not dropping out, they are also 5 percentage points less likely than natives to search for jobs formally, pointing to differences in their job search behavior which may negatively affect their job finding rate.

tial heterogeneity in costs of job displacement depending on migrant workers' origin.[8] Migrants from Turkey, Asia and the Middle East, and Africa face the largest earnings, wage, and employment losses. They also sort into establishments with lower wage premia, which explains 20-30% of their higher wage losses. Compared to the other origin groups and conditional on re-employment, these migrants are least likely to relocate to a different federal state (NUTS-1) after job displacement, and have a relatively high propensity to switch 1-digit occupations, thus potentially losing human capital. In contrast, migrants from the former USSR fare slightly better than natives: they have both a higher re-employment probability and switch to establishments with higher wage premia.

Last, we explore the importance of local labor market concentration, proxied by three measures: i) the change in local unemployment rates around time of displacement, ii) city residency, and iii) the share of same-nationality working age population in a worker's workplace county. We expect that these three proxies are relevant because prior literature has shown that i) migrants' wage assimilation is particularly slow in periods of high unemployment (Bratsberg et al. [2006]), ii) displaced workers' unemployment duration is particularly high if they live in cities (Haller and Heuermann [2019]), and iii) within-network competition may harm migrants' labor market integration (e.g., Albert et al. [2021], Beaman [2012]).

Our results suggest that displaced workers, irrespective of nationality, face greater earnings losses if local unemployment rates at the time of displacement increase more. Moreover, earnings losses are greater if displaced workers live in a city at the time of displacement, and this effect is approximately double the size for migrants. In addition, migrants working in counties with a higher share of the same-nationality population in the year before displacement face substantially higher costs of job loss. These findings suggest, in line with Caldwell and Danieli [2021], that a greater concentration of similar workers at the time of displacement is a crucial factor driving displaced workers' earnings losses. Migrants in particular seem to compete with workers of the same origin for the same types of jobs.

This paper contributes to the literature investigating migrants' labor market integration (e.g., Dustmann et al. [2012], Algan et al. [2010]) by exploiting job displacement as an exogenous shock to both migrants' and natives' labor market trajectories.

[8] A number of studies have documented that the degree of migrants' labor market integration varies across origin groups, possibly driven by cultural distance (see, e.g., Bedaso [2021], Lundborg [2013], Edin et al. [2003]).

While much of the literature investigating migrants' labor market success is descriptive, our empirical approach of combining event study regressions à la Jacobson et al. [1993] with worker-level reweighting allows us to estimate causal effects of migrant-native earnings, wage, and employment gaps. We moreover add a detailed analysis of the potential mechanisms for migrants' slower recovery, showing to what extent establishment sorting, geographic mobility, and local labor market concentration contribute to the migrant-native earnings gap.

We moreover contribute to the literature investigating how sensitive migrants are to adverse economic shocks. A recent paper by Borjas and Cassidy [2021] finds that migrants particularly suffered from displacement during the early phase of the Covid-19 pandemic, partly because they are less likely to work in jobs that can be performed remotely. In the same spirit, other studies have shown that migrants' entry wages during recessions are lower than natives' (see, e.g., Speer [2016], Kondo [2015], Kahn [2010]) and that migrants' or people of color's unemployment rate is particularly sensitive to business cycle conditions and local unemployment rates (e.g., Hoynes et al. [2012], Bratsberg et al. [2006], Altonji and Blank [1999]). The main difference from our study is that whereas most of these papers analyze aggregate outcomes, we follow individual workers' careers before and after job loss. The high-quality administrative employer-employee data from Germany allow us to show how each worker's earnings, wage, and employment trajectory evolved up to five years before and after job loss. We can thus directly compare how involuntary displacement affects migrant workers relative to native workers at the individual level.

Last, we contribute to the literature on the individual costs of job loss by adding evidence on migrant workers. Many studies have documented large and persistent earnings losses for displaced workers (see, e.g., Schmieder et al. [2020], von Wachter et al. [2011], Couch and Placzek [2010], Jacobson et al. [1993]) but without differentiating between specific groups. While there is an emerging literature on the costs of job loss by worker type (e.g., Helm et al. [2021], Illing et al. [2021], Blien et al. [2020], Meekes and Hassink [2020]), no study to date focuses on migrant workers.[9]

The remainder of the paper proceeds as follows. Section 2.2 provides an overview

[9]One exception is the study by Hardoy and Schøne [2014] who focus on displacement effects from plant closings over the business cycle in Norway. In comparison to this paper, they, however, study only two outcomes: employment probability and employment duration. Moreover, they are not clear on the extent to which differences in observable characteristics drive differences in employment outcomes between migrants and natives.

of our data and sample construction. Section 2.3 describes our empirical strategy, including the event study regression model and the reweighting technique. Section 2.4 reports our main results. In Section 2.5, we explore heterogeneities by educational attainment, origin group, and local labor market concentration. Section 2.6 presents our robustness checks, and Section 2.7 concludes the paper.

2.2 Data and Sample Construction

In this section, we first describe the linked employer-employee data that we use for our analysis. Second, we discuss how we define layoff events and construct our baseline sample. Third, we explain our propensity score matching algorithm, which we use to find a unique control worker for each displaced worker.

2.2.1 German Administrative Data

For our empirical analysis, we use high-quality social security data provided by the IAB. Our primary data source is a random 12.5 percent sample of the universe of workers subject to social security contributions in 1996-2017, which stems from the *Integrated Employment Biographies (IEB), version 14*.[10] This data includes information on both workers' employment and unemployment spells with daily precision. We thus observe a detailed set of labor market characteristics for each worker, including wage, employment status, and the exact number of days worked. Moreover, the data contain an extensive set of individual characteristics, such as nationality, age, education, industry, occupation, and workplace at the municipality level (local administrative units (LAU)).

We use a unique establishment identifier to combine our worker-level sample with establishment data from the *Establishment History Panel (BHP)*, from which we obtain variables such as establishment size, average establishment wage, number of migrant workers in the establishment, and number of marginally employed workers in the establishment.

[10]These data stem from administrative sources and are therefore highly reliable. Note, however, that these data do not include the self-employed, civil servants, or the informal sector. One limitation of our study is therefore that we cannot observe whether more migrants than natives sort into self-employment or into the informal sector after displacement.

Based on the code provided by Dauth and Eppelsheimer [2020], we then construct a worker-level panel as of June 30 each year. If workers leave the data and do not return until the end of our observation period 2017, they drop out of our sample upon exit.[11] If workers only temporarily leave the data, we assign them zero earnings, zero employment, and missing wages for the missing spells. To ensure the validity of the data, we further conduct two imputation procedures. First, we correct implausible education entries following Fitzenberger et al. [2006]. Second, we impute wages censored at the contribution assessment ceiling in Germany following Gartner [2005] and Dustmann et al. [2009].

2.2.2 Layoff Events

Next, we use the universe of workers in Germany to identify mass layoff events in 2001-2011. To ensure that our results are comparable with state-of-the-art studies from the US and other countries, we follow Hethey-Maier and Schmieder [2013] in their identification of mass layoffs in the German data. In our definition, a layoff occurs between June 30 in t=-1 and June 30 in t=0 if an establishment (i) completely closes down, or (ii) reduces its workforce by at least 30 percent. To identify genuine mass layoffs, we follow the standard literature on job displacement (e.g., Schmieder et al. [2020], Davis and von Wachter [2011], Jacobson et al. [1993]) and restrict our sample to establishments with a minimum of 50 employees in the year before the layoff and without major employment fluctuations in the years before. This definition follows common approaches in the US literature and thus ensures the comparability of our study.

One threat to the identification of mass layoffs in administrative data are mergers, takeovers, spinoffs, and id changes. To eliminate such events from our data and thus avoid measurement error, we construct a matrix of worker flows between establishments by year following Hethey-Maier and Schmieder [2013]. If more than 30

[11] We drop these workers because they could potentially include migrants who have moved abroad (e.g., returned to their native country) or selected into self-employment or informal sector employment. If more migrants than natives left social security employment for these reasons, we would otherwise overestimate our results. Note that in the data, we observe both workers' employment and unemployment spells; workers who drop out of the data are not registered with the employment agency at all, and thus do not receive unemployment benefits in Germany. Appendix Figure 2.9 shows that migrants are about 4 percentage points more likely to disappear from the social security records after displacement (mean outcome for natives: 5%).

percent of displaced workers move to the same successor establishment, we exclude this establishment from our sample.

2.2.3 Baseline Sample

In the next step, we apply a number of baseline restrictions to the random sample of workers. Following Schmieder et al. [2020], we only consider workers subject to the following baseline restrictions in a given baseline year: male workers with at least 3 years of tenure who are full-time employed in an establishment with at least 50 employees and aged 25-50.[12]

These baseline restrictions allow us to compare our results to prior literature from the US. However, they come at the expense of the representativity of our sample. For example, Illing et al. [2021] show that the costs of job loss differ substantially between men and women, and throughout this paper, we focus on men.[13] Moreover, given that we focus on workers with at least three years of tenure upon displacement in our baseline analysis, we focus on relatively well-integrated migrants. In the robustness section, we relax the tenure restriction to one and two years of tenure. The results reveal that the migrant-native gap in costs of job displacement is comparable across different tenure definitions.

In line with our layoff definition from Section 2.2.2, we define a worker in our sample as displaced between June 30 in t=-1 and June 30 in t=0 if (i) the establishment lays off at least 30 percent of its workforce between t=-1 and t=0 and (ii) the worker leaves the establishment between t=-1 and t=0 and is not employed in the displacement establishment in the following ten years. Workers in our sample are displaced in 2001-2011. Restricting our observation period to 1996-2017 thus ensures that we can follow workers for at least five years prior to and five years after displacement, as long as they are registered in the social-security data during this period.

---

[12] The focus on workers with high tenure and full-time employment ensures that if workers switch jobs, they likely do so involuntarily. For high-tenure workers, job-to-job mobility in Germany is very low, as German law offers employees a high level of protection. We moreover focus on prime-age workers to ensure that workers have already fully entered the labor market and do not yet have access to partial retirement programs. Furthermore, the outside options after job loss may differ by sex (e.g., fertility for women), which is why we exclude women from our baseline sample. All of these restrictions enable a clearer interpretation of our results.

[13] Appendix Section C replicates our main results for a sample of women.

2.2.4 Propensity Score Matching

We cannot simply compare displaced to nondisplaced workers in our sample, since they may differ in individual characteristics, which could bias our regression coefficients. We thus follow the job loss literature, in particular Schmieder et al. [2020], and apply propensity score matching to assign each displaced worker a suitable nondisplaced control worker match. We consider only displaced workers and potential controls who satisfy our baseline restrictions in a given baseline year. We then estimate a probit regression, where the outcome variable is a dummy for being displaced. We include the following explanatory variables: establishment size in t=-1, log wage in t=-3 and t=-4, years of education in t=-1, tenure in t=-1, and age in t=-1. We only allow exact matches within cells of baseline year, 1-digit industries, and migration status. This means that we only match displaced migrants to nondisplaced migrants, and displaced natives to nondisplaced natives. We assign each worker a control worker with the closest propensity score (without replacement).[14]

This matching algorithm leaves us with a highly comparable control group of nondisplaced workers for migrants and natives. Table 2.1 presents summary statistics on the individual characteristics of displaced compared to nondisplaced workers in the year before displacement. While Columns (1) and (2) show migrant workers' characteristics, Columns (3) and (4) report native workers' characteristics.

Panel A of Table 2.1 shows that the matched workers exhibit very similar predisplacement means in individual characteristics such as years of education and tenure. In contrast, displaced workers' wages, earnings, and days worked are lower than those of matched controls. The main reason for this is our definition of displacement: As workers are displaced between June 30 in t=-1 and t=0, the average wages of the displaced worker sample are already lower in t=-1 by construction. Another explanation are anticipation effects (Ashenfelter [1978]). Note that for this reason, we match on log wages in t=-3 and t=-4 for our propensity score matching algorithm. As Figures 2.1 and 2.7 show, both levels and trends in earnings are, however, remarkably similar for displaced workers and nondisplaced workers in all periods leading up to t=-1.

Panel B of Table 2.1 focuses on regional characteristics. It shows that the majority of displaced and nondisplaced workers in our sample live in cities in West Germany.

[14]See Section 2.6 for a number of robustness checks with respect to the matching specification.

The change in municipality unemployment rates between t=-1 and t=0 is substantially larger for displaced workers, suggesting that some of the layoffs affect local labor markets.

Panel C of Table 2.1 shows that matched workers work for establishments that are similar in terms of worker composition. One difference is that displaced workers tend to work in slightly larger establishments. Approximately one third of workers are displaced from a complete establishment closure (100 percent layoff rate). We moreover see that a tiny fraction of nondisplaced workers are laid off in complete closures. This is because we do not impose any restrictions with respect to employment on this control group after pseudo-treatment following Schmieder et al. [2020]. Some of the control workers are thus also laid-off in future years.

When comparing displaced migrants to natives, a few differences stand out: Migrants have substantially lower wages and consequently lower yearly earnings (30,000 EUR vs. 35,000 EUR). They report fewer years in formal education (11.2 vs. 12.3 years of education). The vast majority of migrants live in cities (80 percent), compared to only 57 percent of natives. Migrants also work in different types of establishments: These are, on average, smaller, have a substantially higher average share of migrant workers (25 percent vs. 7.4 percent) and a lower share of high-skilled workers (7.9 percent vs. 12 percent).

Online Appendix Tables 2.6 and 2.7 report the pre-displacement distributions of migrants and natives (and their respective matched control group) across industries and occupations. In groups of migration status and due to our exact matching within industry cells, the distribution of displaced and nondisplaced workers across industries is the same. However, there are differences between migrants and natives; e.g., migrants are more likely to work in food manufacturing, in the hospitality sector, and in the production goods sector. Natives, in turn, are more likely to work in education, the non-profit sector, and public administration. With respect to occupations, migrants are more likely to work in occupations with simple, manual tasks. Natives more often work in high-skilled occupations such as engineering, qualified services, and qualified administrative tasks.

This shows that directly comparing migrant to native workers is a challenge. For example, if we found that migrants' earnings losses after job displacement are greater, then this could simply be due to the fact that they on average work in occupations with fewer vacancies. While this raw gap is interesting per se, our goal is to understand whether a migrant-native gap remains even net of observable characteristics. For our regression analysis, we will therefore reweight migrants to natives with respect to individual characteristics, industries, and occupations, using the reweighting scheme first proposed by DiNardo et al. [1996] and first applied in the context of job displacement by Illing et al. [2021].

2.3 Estimating Costs of Job Displacement by Migration Status

2.3.1 Event Study Regressions

When analyzing the effects of job loss on migrants' and natives' labor market outcomes, we follow the seminal study by Jacobson et al. [1993] and apply an event study regression model with worker and time fixed effects. Specifically, we estimate the following regression specification separately for migrants and natives:

$$y_{itc} = \sum_{j=-5, j\neq -3}^{j=5} \alpha_j \times I(t = c+1+j) \times \text{Disp}_i + \sum_{j=-5, j\neq -3}^{j=5} \beta_j \times I(t = c+1+j) + \pi_t + \gamma_i + X_{it}\beta + \varepsilon_{itc} \tag{2.1}$$

where the dependent variable y_{itc} denotes average labor market outcomes (e.g., log yearly earnings, log daily wages, employment, number of days worked) of individual i, belonging to cohort c in year t.[15] Disp_i is a dummy indicating whether a worker is displaced, which is interacted with dummies $I(t = c + 1 + j)$ for years t=-5 to t=5 since job loss. We omit period t=-3 as the reference category, as it should not be affected by anticipation effects (Ashenfelter [1978]). The coefficients of interest are α_j, which present the change in labor market outcomes of displaced workers relative to the trends of the nondisplaced control group. Following Schmieder et al. [2020], we include dummies for the year since displacement in the regression equation. In addition, π_t adds year fixed effects, γ_i captures individual fixed effects, and X_{it} is a vector of time-varying age polynomials. We cluster standard errors at the worker level.

[15]For all workers laid off in year t=0, the baseline year is t=-1, which is also their cohort, c.

2.3.2 Making Migrants and Natives Comparable: Reweighting

When comparing migrants' and natives' labor market trajectories after job displacement, it is crucial to control for differences in individual characteristics and sorting into different industries and occupations. To account for these differences, we follow Illing et al. [2021] and complement our event study regression with a reweighting scheme first proposed by DiNardo et al. [1996]. Thus, we reweight migrants to native workers in terms of observable characteristics before job loss. Migrants who are more similar to natives on characteristics such as years of education and tenure receive a higher weight. The intuition is that after reweighting migrants to natives, we can attribute the differences in their outcomes after job loss to how they respond to displacement or to the difficulties they face, rather than to their characteristics.

Econometrically, we approach this as follows: First, we estimate a simple probit regression model:

$$Pr(native_i = 1|X_i) = \varphi\left[X_i'\beta\right] \tag{2.2}$$

where the dependent variable is a dummy that takes value 1 for all native workers. We regress this dummy on a matrix X_i of individual and establishment characteristics. The individual characteristics are log wage (t=-3, t=-4), age (t=-1), years of education (t=-1), tenure (t=-1), and a dummy for city residency (t=-1). In addition, we control for establishment size (t=-1), 1-digit industry (t=-1) and 1-digit occupations (t=-1) following the definition of Blossfeld [1987].[16]

For each displaced migrant worker, we then use the estimated propensity score $\hat{ps}$ to assign them an individual weight $\frac{\hat{ps}}{1-\hat{ps}}$. Following Illing et al. [2021], we compute these weights only for displaced migrants and then ensure that the weights are constant within matched worker pairs. In a robustness check in Section 2.6, we show that our results do not change if we reweight natives to migrants, instead.

Table 2.2 presents summary statistics of displaced workers in our sample in t=-1. Column (1) shows the characteristics of a random 2-percent sample of migrants in

[16] Table 2.8 in the Online Appendix shows the explanatory power of these reweighting variables for the migrant-native earnings gap. The following characteristics turn out to be particularly important: Age and education, city residency, occupations, and industries.

Germany, which we compare to our baseline sample of migrants (Column (2)) and migrants after reweighting (Column (3)). Migrants in our sample have substantially higher tenure, wages, and earnings than the random sample of migrants. A similar pattern holds for a random sample of native workers (Column (4)) compared to native workers in our sample (Column (5)). This reflects our baseline restrictions, which ensure that we focus on a sample of high-tenured workers with strong attachment to the labor market. Note, however, that our results are robust to relaxing the tenure restriction to one or two years, respectively (see Table 2.5).

Comparing our reweighted sample of migrants (Column (3)) to baseline native workers (Column (5)) shows that they are very similar in terms of characteristics. After reweighting, hardly any differences between migrants and natives remain in terms of characteristics such as years of education, age, earnings, and establishment types. Any migrant-native gaps remaining after reweighting are thus likely due either to their different responses to economic shocks, or differences in labor demand.

2.4 The Cost of Job Displacement for Migrants and Natives

Section 2.4 presents our main results. As a benchmark without controls, we first present descriptive statistics (Section 2.4.1). We proceed with the results of both the event study regression model on its own and combined with the reweighting scheme (Section 2.4.2).

2.4.1 The Evolution of Migrant and Native Earnings without Controls

We first present descriptive statistics on how average yearly earnings develop before and after job displacement by migration status. Panel A of Figure 2.1 shows how earnings (without controls) evolve differently for displaced (green line) and nondisplaced (blue line) natives in the five years before and after job loss. While trends and levels in pretreatment earnings are very similar between displaced workers and matched controls, displaced workers' earnings start decreasing from t=-1 onwards. Between t=-2 and t=0, displaced workers' earnings decrease from approximately

EUR 37,000 to EUR 25,000. While they recover slightly in the years following job loss, they do not catch up with average earnings in the control group even five years after displacement. Panel B of Figure 2.1 shows earnings losses for migrant workers. Displaced migrants' average earnings are already lower than natives' pre displacement, and they lose more, both in absolute and relative terms: Their earnings drop from roughly EUR 33,000 in t=-2 to EUR 17,000 in t=0.[17] Thus, while natives lose about 32% of their original earnings, migrants, at 48%, have much higher relative losses. This earnings difference between nondisplaced and displaced migrants decreases but persists for up to five years.

Figure 2.1 moreover shows that for the control groups of nondisplaced workers, log earnings slightly fall from t=1 onwards. Recall that in the year before (pseudo-) displacement, both displaced and nondisplaced workers have to be employed with three years of tenure. This ensures that both groups display relatively stable employment careers before job loss. Starting with period t=0, we, however, allow nondisplaced workers to leave social security records for reasons such as unemployment, self-employment, or parental leave; their average earnings thus mechanically decrease. This does not present a threat to the validity of our analysis, as we think of our control group as a random sample of workers, which we do not want to artificially restrict to being employed. Given the decreasing trends in the control group, even if there were biases, we would under- rather than overestimate our effects.

2.4.2 Estimating Costs of Job Displacement by Migration Status

Event Study Coefficients with and without Reweighting Figure 2.2 presents the results from our event study regression model, where the solid green line shows the trajectory for natives, the dashed blue line shows the trajectory for migrants, and the dashed light blue line shows the trajectory for the reweighted sample of migrants where we control for differences in observable characteristics by migration status. Panel (a) presents our results for yearly earnings relative to earnings in

[17]Note that migrants who have contributed to social security in Germany for at least 12 months within the past 30 months are entitled to receive unemployment benefits according to the same rules as German workers. Due to our baseline restrictions, all displaced workers in our sample have at least three years of experience in the German labor market upon displacement, and should thus be eligible for these benefits.

t=-2.[18] The results underscore the descriptive results from Figure 2.1. Relative (unweighted) earnings decline significantly both for native (38 log points) and migrant (51 log points) displaced workers between t=-1 and t=0. Neither displaced migrants nor displaced natives have fully recovered 5 years after job loss. From t=0 onwards, migrants' average recovery rate is faster, yet even five years after job loss, their earnings are lower than natives'. Overall, our findings of large and persistent earnings losses after job loss are in line - in terms of magnitude and pattern - with existing studies from the US and Germany (e.g., Schmieder et al. [2020], Jacobson et al. [1993]).

The light blue line shows that reweighting migrants to natives and thus essentially controlling for differences in individual and establishment characteristics roughly halves the raw migrant-native gap in earnings losses. Nevertheless, a significant gap remains, showing that differences in observable characteristics cannot fully explain the differences.

To better understand what drives these differences, we next decompose earnings losses into wage and employment losses. Panel (b) of Figure 2.2 shows that migrants face substantially higher raw wage losses than natives (30 vs. 18 log points in t=0), but observable characteristics can almost fully explain this difference. In contrast, as Panels (c) and (d) show, observable characteristics cannot explain the migrant-native gap in employment.

Panel (c) shows that migrants and natives are both less likely to have a job in social-security employment in the years following their job displacement. Here, the outcome variable is a dummy for being employed at least once in a given year (this includes full-time, part-time, and marginal employment). Migrants' employment decreases substantially and more than natives' (20 vs. 13 percentage points), and observable characteristics cannot explain the difference. Even five years after displacement, migrants have not fully caught up with natives.[19]

[18] This measure helps us to include observations with 0 earnings and is more easily interpretable than log(earnings+1). Similar measures have been used in recent job displacement papers such as Illing et al. [2021] and Blien et al. [2020]. Note that we exclude outlier pairs where the relative earnings measure exceeds 100 at least once during the observation period. This affects less than 0.7% of our sample. See section 2.7 for other earnings measures.

[19] Our findings are in line with recent work on the Dutch labor market by Meekes and Hassink [2020]. While the focus of their paper is gender differences in job flexibility outcomes after job loss, they also show in their online appendix that relative to individuals born in the Netherlands, the foreign born (non-natives) are less likely to become re-employed (10 percentage points) after job loss. They find no differential effect on hourly wages.

Panel (d) presents a very similar pattern with respect to days worked per year. Again, the reduction is larger for migrants (approximately 150 days) than natives (approximately 110 days). The gap hardly closes if we reweight migrants to natives based on observable characteristics (light blue line). Migrants never fully catch up with natives, even though the gap substantially shrinks from t=3 onwards; after five years, neither group has fully recovered from displacement in terms of days worked. Finally, Panels (e) and (f) show that migrants are more likely to take up part-time rather than full-time employment after layoff. This offers an additional explanation for migrants' higher earnings losses and suggests that they sort into worse employment contracts.

Overall, Figure 2.2 provides two key takeaways. First, if migrants find a new job after displacement, their wage losses are slightly higher, but observable characteristics can explain this gap.[20] Second, migrants experience greater difficulty finding a new (full-time) job than natives in the first place. Neither individual characteristics nor differential sorting across industries and occupations can explain this employment gap.

Job Search Behavior One obvious question is whether differences in job search behavior can explain part of this employment gap. Figure 2.9 in the Online Appendix offers suggestive evidence that migrants do react differently to job displacement in terms of how, and possibly where, they look for jobs. For our baseline sample and displaced workers only, it shows how reweighted migrants' job search outcomes differ from natives'.

First, the figure shows that about 5% of native workers completely disappear from our sample after layoff[21], and that this share is about twice the size for migrants. This suggests that migrants do react differently to displacement: For example, they may be more likely to return to their country of origin, become self-employed, or move to the informal sector.

Second, conditional on not dropping out of the data, we see that migrants are 5

[20]In the Online Appendix, section 2.7, we discuss the role of differential sorting into specific types of establishments after job displacement (in terms of establishment wage premia, share of marginally employed workers, and share of migrant workers).

[21]This means that they drop out of the social-security records between t=1 to t=5 and do not return. Note that in our baseline analysis, these workers drop out of our sample in the year they are last observed, such that they do not bias our employment estimates.

percentage points less likely to ever report a job seeker spell (mean for natives: 64%). This is suggestive evidence for migrants searching for jobs in more informal ways, which could negatively affect their job search success.

We moreover explore 4 job search outcomes measured at the time of the displacement, conditional on workers ever reporting a job seeker spell. These show that migrants are somewhat more likely to look for "any contract" as opposed to a "permanent contract", and that there is no search difference with respect to the type of job (full-time vs. part-time).

2.4.3 The Role of Geographic Mobility

We now turn to discussing one characteristic that could explain differences in the success of finding a new job: geographic mobility. Within-country geographic mobility is an important tool to adjust regional labor market imbalances and, hence, raise local labor market efficiency (Blanchard and Katz [1992]). Displaced workers who relocate permanently may be rewarded with higher job search success, yet this depends on the reason for moving (Huttunen et al. [2018]). Nudges for displaced workers to relocate are particularly high if they work in highly concentrated labor markets with fewer outside options (Haller and Heuermann [2019]). While previous literature has shown that migrants tend to be more geographically mobile than natives (e.g., Borjas [2001], Cadena and Kovak [2016]), this pattern may reverse in regions with tight housing markets (Clark and Drever [2000]).

In this section, we make use of the geographic information recorded in the IAB data. We know the municipality, county, and federal state in which a worker lives and works.[22] It is important to keep in mind that we only observe this information for *employed* natives and migrants. To draw conclusions on all displaced workers, we have to assume that employed workers' mobility patterns reflect mobility patterns in the overall population of migrants and natives.

Panel (a) of Figure 2.3 reports event study coefficients for workplace changes as the

[22] Germany exhibits widespread federalism. Therefore, there exist different administrative units (according to size): (i) federal states and city states (NUTS 1), (ii) administrative districts (NUTS 2), (iii) counties and cities (NUTS 3), and (iv) municipalities (LAU). In 2010, there were a total of 11,993 municipalities and 401 counties in Germany. According to data provided by the *German Federal Statistical Office*, on average, a municipality had 4,954 inhabitants, and a county had 186,596 inhabitants in 2010.

outcome variable. Specifically, we create a dummy variable indicating whether the workplace municipality differs from the workplace municipality in t=-1. In line with our expectation, displaced workers' likelihood of moving workplaces substantially increases following job displacement. In t=0, displaced natives were approximately 58 percent more likely to change workplaces than nondisplaced controls. For migrants, this number is slightly lower (approximately 50 percent). Once we control for observable characteristics, hardly any differences between migrants and natives remain.[23]

Panel (b) of Figure 2.3 shows that mobility across federal states follows a similar pattern. Approximately 19 percent of displaced natives changed their workplace to a different federal state from t=-1 to t=0. In contrast, only 11 percent of migrants moved to a different federal state after displacement. After reweighting migrants to natives, this difference reduces to 3 percentage points but remains statistically significant.

Finally, Panel (c) of Figure 2.3 shows how commuting patterns evolve after displacement, where commuting is defined as working and living in different municipalities. It shows that following displacement, the likelihood of commuting increases substantially. Slightly more migrants (6 percent) than natives (4 percent) start commuting following displacement.[24]

Overall, our results on geographic mobility suggest that migrants face higher mobility constraints (e.g., due to tight housing markets or because migrants are particularly dependent on local networks) in terms of relocating permanently after displacement.[25] While they seem to compensate for this by commuting slightly more, this may not be enough to catch up in terms of job search success. However, we only observe mobility outcomes for *employed* workers. Another explanation of the pattern that we observe is thus that migrants receive fewer job offers at greater geographic distances.

[23]This result is robust to adapting the mobility definition to include only workplace moves over a distance of more than 50 km.

[24]This result is robust to defining commuting at the county rather than municipality level.

[25]Table 2.9 shows that changes in commuting behavior or relocation across federal states explain about 7% of the migrant-native gap in earnings.

2.5 Heterogeneity in Costs of Job Displacement

The broad comparisons of migrant and native workers may hide substantial heterogeneity across different groups of workers. In this section, we shed more light on three factors by which costs of job displacement may differ: Educational attainment, migrants' origin group, and local labor market conditions.

For our empirical approach, we follow Schmieder et al. [2020] and estimate a difference-in-differences (DID) type of regression model, where we proceed in two steps. In the first step, within each matched worker pair, we construct an individual-level measure of earnings losses (and other outcomes), which we call the DID outcome. For this purpose, we compute the mean difference in earnings (and other outcomes) before and after job loss within each displaced and nondisplaced worker match:

$$\Delta y_{DID,ic} = \Delta y_{DP,ic} - \Delta y_{NDP,ic} \tag{2.3}$$

where $\Delta y_{DP,ic}$ reports the difference in average earnings for displaced worker i in cohort c before (t=-5 to t=-2) vs. after (t=0 to t=3) job loss. $\Delta y_{NDP,ic}$ reports the corresponding measure for the matched nondisplaced worker. $\Delta y_{DID,ic}$ then indicates the extent to which these differences in means vary within matched worker pairs, the "individual treatment effect" from job loss.

In the second step, we estimate OLS regression models for displaced workers only, where we regress the outcome variable $\Delta y_{DID,ic}$ on dummies for educational attainment, origin group, or proxies for local labor market concentration. In addition, we always control for a vector X_{ic} with individual characteristics, 1-digit industries, and 1-digit occupations measured in the year before displacement. We cluster standard errors at the baseline establishment or baseline county level.

2.5.1 Educational Attainment

Previous research has shown that migrants' labor market integration varies substantially with their educational attainment (Battisti et al. [2021], Brücker et al. [2021]). Educational attainment may also play a role in costs of job displacement, yet the mechanism is ex ante unclear. On the one hand, earnings losses from job displacement may be relatively low for high-skilled workers if there is a high demand for their types of skills, and their job covers many non-routine tasks (Blien et al. [2020]) . On the other hand, high-skilled workers may possess very specific human capital, which they cannot easily transfer across positions.

For skilled migrants, it may be particularly difficult to transfer their skills across different types of jobs, given that they face difficulties in getting these skills acknowledged (Brücker et al. [2021]). Yet demand for skilled labor may counteract this effect: A Syrian physician displaced from his job may be re-employed much more quickly than a Syrian worker working on the assembly line.

To investigate the role of educational attainment, we regress our measure for the individual treatment effect from job displacement, as defined in Equation 2.3, on 3 dummies for educational attainment[26] and interactions of these dummies with migration status. We then plot marginal effects from this regression in Figure 2.4.

Focusing on Figure 2.4, Panel (a), offers 3 takeaways: First, workers with a university degree suffer the smallest earnings losses from job displacement, regardless of their migration status. Second, while migrant workers with a university degree suffer somewhat larger earnings losses than native workers with a university degree, this difference is not statistically significant. Third, within the groups of workers with/without vocational training, migrant workers always suffer substantially larger earnings losses than natives. The gap is by far the largest for workers without vocational training: Earnings losses for migrant workers in this group are about 1.5 times the size of native workers' losses (60 vs. 40 log points).

Interestingly, the overall pattern differs by migration status. For native workers, earnings losses from job displacement do not vary much across education groups. This could reflect the counteracting mechanisms discussed above, where high-skilled workers have the highest losses in terms of their "employer-employee match premium" but face less competition when searching for new jobs. Yet for migrants, workers with a university degree have the smallest earnings losses, follwed by migrants with vocational training, and migrants without vocational training losing most.

Panels (b)-(d) of Figure 2.4 show that this is a pattern which is consistent for the probability to be employed (Panel (b)), log wages (Panel (c)), and establishment

[26]The dummies take three values: "No vocational training", "vocational training", and "university degree" (the latter includes universities of applied sciences).

wage premia (Panel (d)). Less educated migrants are less likely to become re-employed, earn lower wages, and sort into worse establishments after job displacement. Within the group of high-skilled workers, there is no difference by migration status.

While both differences in labor demand and labor supply by education and migration group can possibly explain this pattern, we provide suggestive evidence that labor supply may play a role. Online Appendix Figure 2.11 shows that conditional on employment, medium- and low-skilled migrants are less likely than natives in the same education group to relocate to a different federal state (approximately 3-4 percentage points). They are also more likely to switch 1-digit occupations, thus potentially losing more in terms of human capital.

2.5.2 Differences in Costs of Job Displacement by Origin Group

The degree of labor market integration can vary substantially by origin group (see, e.g., Algan et al. [2010]). In this section, we therefore report our measures for migrants' costs of job displacement for 9 different origin groups as defined by Battisti et al. [2021][27]. For this purpose, we regress the individual difference-in-differences outcome as defined in Equation 2.3 on the 9 origin groups dummies, where native worker is the omitted category. Figure 2.5 reports the results, showing that indeed, there is substantial heterogeneity in costs of job displacement by origin group.

Three groups stand out in particular: Migrants from "Turkey", from "Asia and the Middle East", and from "Africa" have by far the largest earnings, wage, and employment losses from job displacement. Conditional on employment, they also sort into establishments with lower wage premia (Panel (d)). A potential explanation for these origin-group-differences may be cultural distance, but both labor supply and labor demand may be potentially important.

Costs of job displacement for most other groups are either slightly larger or do not differ from natives'. The only group which is doing slightly better than natives are migrants from the former USSR: They are both more likely to become re-employed

[27]See Online Appendix table 2.16 for a more detailed overview on the definition of these origin groups.

(Panel (a)) and sort into better establishments (Panel (d)). This is in line with previous research showing that, e.g., mathematicians from the former USSR were particularly productive (Borjas and Doran [2012]).

Additional outcomes, which we show in the Online Appendix Figure 2.12, suggest that migrants from these different origin groups search for jobs in different ways. For instance, immigrants from Turkey are the only group with a substantially lower likelihood to commute than native workers, whereas migrants from the former USSR are more likely to commute. Turkish migrants are also the only group for which the share of migrant coworkers increases (compared to natives) after job displacement, suggesting that ethnic networks may be particularly important for their job search. Moreover, migrants from the three origin groups with the largest earnings losses are all substantially less likely to relocate to a different state after displacement. This may be due to local job search, or barries to mobility such as financial constraints or discrimination on the housing market.

2.5.3 Location at Time of Displacement

We explore one final type of heterogeneity: the location at time of displacement. Does the concentration of the local environment at the time of job displacement matter? We believe that concentration matters in two ways. First, if displaced workers live and work in labor markets with a high concentration of similar workers, then finding a new job will be particularly challenging for them (e.g., Caldwell and Danieli [2021], Haller and Heuermann [2019]), and this may hold in particular for migrants (Bratsberg et al. [2006]). On the one hand, prospective employers may find it difficult to judge migrants' skill portfolio, especially if they did not receive their qualifications in Germany (Brücker et al. [2021]). They may thus perceive asymmetric information to be a more severe issue when hiring migrants and prefer to hire native workers instead.

On the other hand, establishments may display taste-based or statistical discrimination against migrants. If labor supply is very elastic and employers can choose between a migrant and native candidate, they may thus opt for the native worker. Second, migrants may compete for jobs among each other. While previous studies have shown that migrants benefit from better social networks (e.g., Edin et al. [2003], Munshi [2003]), migrants may also suffer from within-network competition (e.g., Albert et al. [2021], Beaman [2012], Calvo-Armengol and Jackson [2004]). Migrants living in counties with a particularly high share of same-nationality population may compete for a limited number of jobs.[28]

We thus regress the individual difference-in-differences measure of earnings losses (and other outcomes) from Equation 2.3 on three proxies for local labor market concentration. We start with including UR_{ic}, which measures the percentage change in the unemployment rate in the workplace municipality between t=-1 and t=0 for displaced worker i in cohort c:

$$\Delta y_{DID,ic} = \alpha Mig_i + \beta_1 UR_{ic} + \beta_2 UR_{ic} * Mig_i + \phi X_{ic} + \varepsilon_{ic} \tag{2.4}$$

The intuition behind this is as follows: If the local unemployment rate increases as workers lose their job, it will be more difficult for them to find a new job. Note that to some extent, this measure also controls for the size of the mass layoff event: If a big employer, e.g., a car manufacturer, closes down its establishment, this will affect the local labor market more severely than if a small service firm goes bankrupt.

Next, we add $City_{ic}$ to the difference-in-differences regression, which is a dummy indicating whether a worker lives in a city at t=-1:[29]

$$\Delta y_{DID,ic} = \alpha Mig_i + \gamma_1 City_{ic} + \gamma_2 City_{ic} * Mig_i + \phi X_{ic} + \varepsilon_{ic} \tag{2.5}$$

The expected effect of living in a city is ambiguous: On the one hand, we expect cities to offer more opportunities, on the other hand, there may be more competition for some types of jobs (Haller and Heuermann [2019]).

Last, we include a dummy for $EthnicShare_{ic}$, which reports the share of the working age population of a worker's nationality by the total working age population in his workplace county at t=-1:[30]

[28]We assume that migrants with the same nationality are similar in terms of characteristics, e.g., because of similar education systems in their countries of origin, and therefore substitutes.

[29]To define cities, we use a municipality classification proposed by the *German Federal Institute for Research on Building, Urban Affairs and Spatial Development* (BBSR), which is based on population size and administrative function (see German Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR) [2021]). [last access: 05-30-2021].

[30]We use data on working age population by nationality and county from the *German Federal Statistical Office (Destatis)*. For more detail, see Section 2.7 in the appendix.

$$\Delta y_{DID,ic} = \alpha Mig_i + \delta_1 EthnicShare_{ic} + \delta_2 EthnicShare_{ic} * Mig_i + \phi X_{ic} + \varepsilon_{ic} \quad (2.6)$$

$EthnicShare_{ic}$ can be a proxy for the extent to which migrant workers compete for jobs among each other. Note however, that it could also be a proxy for ethnic enclaves, where the empirical literature is inconclusive on whether they contribute to or harm migrant workers' labor market integration (e.g., Chiswick and Miller [2005], Edin et al. [2003]). With the data at hand, we cannot distinguish between the substitution vs. ethnic enclave mechanisms.

The Impact of Local Labor Market Concentration on Earnings Table 2.3 reports the results from Equations 2.4-2.6, where we consecutively include controls. The outcome variable is log(earnings). In all regressions, we control for a set of individual and establishment characteristics.[31] The average loss of earnings is 36 log points for native workers (see the mean of the dependent variable), and for migrants, this loss increases by additional 19 log points (Column (1)). This confirms our results from Section 2.4.2 that even after controlling for observable characteristics, migrants face larger earnings losses.

We then add our first proxy for local labor market concentration and local unemployment rate changes in Column (2). The coefficient implies that a 1% increase in the municipality unemployment rate from t=-1 to t=0 increases earnings losses - regardless of migration status - by 11%. This supports our hypothesis that higher local unemployment rates reduce workers' outside options and thus increase displaced workers' earnings losses. The coefficient on the interaction of local unemployment rate changes and the migrant dummy is negative but estimated very imprecisely. In Column (3), we include city residency as another proxy for concentration. The coefficients confirm the negative correlation between living in a city at the time of displacement and earnings losses, as documented by Haller and Heuermann [2019]. Earnings losses of displaced workers who live in cities at the time of displacement are 5.6% larger. This effect is approximately twice the size for migrant workers.

[31] These are age, age squared, years of education, tenure, experience, full-time employment, log firm size (all measured in t=-1), log wage in t=-3, 1-digit industries (in t=-1), and occupations according to Blossfeld [1987] (in t=-1).

Column (4) provides suggestive evidence for a negative relationship between migrants' earnings and the concentration of similar workers, proxied by the working-age population of a worker's nationality as a share of the total working-age population in his workplace county in the year before job displacement. This is in line with literature highlighting a competition effect between similar migrants (Beaman [2012], Albert et al. [2021]). Note that if we simultaneously control for all concentration proxies (Column (6)), the interaction of the migrant dummy with city residency shrinks and loses its significance. This suggests that a large part of the city effect for migrants can be explained by a higher share of the same-nationality working-age population in cities.

To complete the picture, we estimate versions of Equation 2.4 for additional outcome variables. Panel A of Table 2.4 reports coefficients on the migrant dummy for regressions with individual, industry, and occupation controls. The table confirms the overall pattern from the event study regression model: Migrants' employment (Columns (1) and (2)) and wage (Column (3)) losses after displacement are substantially larger than natives'. This can partly be explained by migrants selecting into establishments with lower wage premia (Column (5)) and a higher share of marginally employed workers (Column (7)).

We next add our concentration proxy controls in Panel B. The respective coefficients broadly confirm the pattern that we already observed in Table 2.4: A larger increase in the local unemployment rate change from t=-1 to t=0 is associated with greater losses in terms of employment, irrespective of migrant status. Workers living in cities at the time of displacement face larger employment and wage losses; for migrants, this "city penalty" on wage losses is particularly high. Migrants living in counties with a higher share of the same-nationality population face particularly large wage and employment losses.

We do not want to interpret the magnitude of the coefficient on the interaction between migrants and shares of the same nationality since the effect may vary substantially depending on a migrant's position in the share distribution. To show this, we regress the individual DID term, $\Delta y_{DID,ic}$, for log earnings on 18 categories for the share of same-nationality working age population in t=-1. We plot the respective marginal effects in Figure 2.6, where the x-axis reports the 18 categories. While earnings losses for natives (Panel (a), solid green line) are constant and do not vary substantially by the percentage share of same-nationality working age population, there is a clear pattern for migrants (Panel (a), dashed blue line): Earnings losses are particularly high for migrants working in counties with the highest share of same-nationality working age population (8-10%). This pattern is driven by larger log wage losses (Panel (b)) and larger employment losses, both on the extensive and intensive margins (Panels (c) and (d)).

2.6 Robustness

In the following, we perform a variety of robustness checks to show that our results are robust to varying our baseline restrictions, matching specifications, and to the direction of reweighting. Table 2.5 reports the results, where Column (1) shows the baseline coefficients. We report four outcome variables: Earnings relative to t=-2 (Panel A), log wages (Panel B), employment (Panel C), and days worked in full-time employment (Panel D).

Baseline Tenure Restriction As discussed before, our baseline analysis focuses on a sample of workers who are highly attached to the labor market (3 years of tenure). This could bias the migrant-native gap if high-tenure migrants are particularly well-integrated into the German labor market, and their re-employment probability is thus higher than that of other migrants. In this case, we would underestimate the gap. In Columns (2) and (3) of Table 2.5, we therefore relax the tenure restriction to 1 and 2 years, respectively. The table shows that for workers with 1 year of tenure upon displacement, the migrant-native earnings gap is essentially the same, while it is even slightly smaller for workers with 2 years of tenure (Panel A). While the migrant-native gap in employment does not substantially differ across tenure restrictions (Panels C and D), the difference in wages is substantially larger for workers with 1 year of tenure, and smaller for workers with 2 years of tenure (Panel B). Overall, the changing baseline tenure restrictions does not substantially impact the migrant-native gap in costs of job displacement.[32]

[32] Appendix Figure 2.8 shows that there is surprisingly little variation in costs of job loss by the number of years a worker has been recorded in the administrative (admin) employer-employee data at the time of job displacement. Losses for workers with lower tenure in the admin data are somewhat lower, perhaps because these are younger workers who are more flexible. Regardless of their time in the admin data, migrants always have greater earnings losses than natives, even though this gap somewhat closes for workers who have been in the admin data for more than 27 years (yet note that this is partly due to lower observations and thus more noisy estimates).

Propensity Score Matching Next, we explore whether our results are sensitive to the matching algorithm. For example, one might be worried that we match on log wages in t=-3 and t=-4, and thus essentially on trends in one of our outcome variables. We therefore run a propensity score matching analysis where we match displaced to nondisplaced workers on education, tenure, and establishment size only. Column (4) of Table 2.5 shows that this somewhat increases the migrant-native earnings gap (from 6.4% to 10%). This is mainly due to a larger difference in employment, with migrants losing about 33 days per year more in full-time employment (compared to 25 days in the baseline sample). Moreover, one might be worried about anticipation effects or other trends in the years before job displacement. We therefore conduct a matching analysis where we specifically only match on characteristics recorded in t=-4, and thus well before job displacement. Column (5) shows that this somewhat decreases the migrant-native earnings gap (6.4% to 5%) but it remains statistically significant. In addition, recent work has shown that the effects of mass layoffs spill over to local labor markets (Gathmann et al. [2020]). We therefore conduct a matching exercise where we match exactly on counties (NUTS-3 regions) instead of 1-digit industries in t=-1. Column (6) of Table 2.5 shows that our main results do not change.

Time Window and Reweighting We moreover check how sensitive our results are to the time window we selected. For this purpose, we run alternative specifications where we define the individual difference-in-differences estimate as the difference in the respective outcome from t=-10 to t=-2 vs. t=0 to t=10 (as opposed to t=-5 to t=-2 vs. t=0 to t=3 in the baseline regressions). Column (7) of Table 2.5 shows that if we consider this larger time window, the migrant-native earnings gap decreases (6.4% vs. 4.7%). In this specification, there is no migrant-native wage gap, and the employment gap between migrants and natives is smaller, yet still statistically significant.

We moreover change the direction of reweighting and reweight natives to migrants (Column (8)). For this purpose, we use the same reweighting algorithm as described in Section 2.4.2. The only difference is that instead of a dummy for native workers as an outcome variable in our probit regression, we now regress a dummy for *migrant* workers on a set of predisplacement individual characteristics, 1-digit industries, and occupations as defined by Blossfeld [1987]. This hardly changes our results.[33]

Complete Closures, Displacing Establishments Thus far, our sample includes both workers displaced from complete establishment closures and from layoffs where only part of the workforce is laid off. Workers displaced in mass layoffs may be different from workers laid off during complete establishment closures (Gibbons and Katz [1991]): If establishments decide whom to lay off, they will be more likely to first fire workers of low ability, without family obligations, or bad matches. Being laid off could thus be a negative signal to future employers.[34] Imagine, for example, that some migrants have worse language skills than their coworkers, and are thus more likely to be laid off during a mass layoff. These differences in language skills (which we do not observe and thus cannot control for) moreover negatively affect their re-employment probability. If displaced migrants in our sample constituted a negative selection in terms of language skills compared to the average migrant population, then we would overestimate migrants' costs of job displacement.

To solve this, we restrict the sample to workers laid off in a complete establishment closure only, where we assume that neither migrants nor natives will constitute a negative selection. As Online Appendix Table 2.15 shows, the individual difference-in-differences results are very comparable to our baseline regressions.

In a second robustness check, we moreover add fixed effects for the establishment from which workers are displaced to our regression model. We do this because workers may sort into specific establishments prior to displacement. For example, in Table 2.1 we showed that migrants have a greater propensity to work in smaller establishments with higher shares of migrant workers and lower shares of high-skilled workers. Again, our individual difference-in-differences results are remarkably stable (Online Appendix Table 2.14).

Financial Crisis, East Germany Migrants particularly suffer during recessions (e.g., Borjas and Cassidy [2021], Fairlie et al. [2020], Montenovo et al. [2020], Freeman et al. [1973]), so the financial crisis, which is included in our period of analysis,

[33]For more results, see Online Appendix Table 2.13.

[34]Gibbons and Katz [1991] show that workers displaced from mass layoffs have larger wage losses and higher unemployment durations than workers laid off in complete closures.

may bias our results in the direction of particularly large earnings losses for migrants. We thus estimate Equation 2.1 only for baseline years up to 2007, ensuring that none of the workers in our analysis sample lose their jobs during the financial crisis. Reassuringly, Table 2.10 in the Online Appendix shows that our results are robust to excluding the financial crisis years. Migrants displaced in 2001-2007 face substantially larger earnings losses (Columns (1) and (2)), wage losses (Columns (3) and (4)), employment losses (Columns (5) and (6)), and losses in yearly days worked (Columns (7) and (8)) than native workers.[35]

Similarly, we run a robustness check where we exclude workers displaced from East German establishments from our sample. Our observation period starts only six years after German reunification and covers a time when East Germany underwent major economic transitions. This could lead to different displacement effects for workers in East Germany, and for migrants in East Germany, reintegration into the labor market could be particularly difficult. Reassuringly, our results are robust to estimating our regression based on a sample of workers displaced in West Germany only (Online Appendix Table 2.12).

2.7 Conclusion

In this paper, we investigate differences in the costs of job displacement for migrants compared to native workers. We show that migrants face larger costs of job loss than natives: The raw earnings gap between displaced migrants and natives is 12 percentage points in the year of displacement, and it reduces to 5 percentage points if we control for pre-displacement differences in observables.

Decomposing earnings into wages and employment shows that in terms of the raw difference, migrants have both substantially larger wage and employment losses. Once we control for individual characteristics, 1-digit industries, and 1-digit occupations, the migrant-native wage gap closes while the employment gap persists both in the short run (5 ppt) and in the long run (2ppt). We moreover find that while

[35] Since our post-job-loss period spans five years, restricting the observation period to 2007, the year before the financial crisis, could not suffice - the crisis could also have reduced job search success in t=1 up to t=5. We therefore run an additional robustness check, where we only include matched worker pairs with baseline years up to 2003 in our sample (see Table 2.11 in the Online Appendix). The resulting patterns are very similar to our main results: Migrants face larger earnings and employment losses.

migrants are slightly more likely to commute (2 percentage points), they are less likely to permanently relocate to a new federal state (3 percentage points) following job loss, pointing to relocation constraints.

With respect to heterogeneity, we show that costs of job displacement are particularly high for the least educated migrant workers, while there is no migrant-native gap for high-skilled workers. In addition, while some origin groups have particularly large earnings losses (e.g., migrants from Turkey, Asia and the Middle East, and Africa), other groups do not differ substantially from native workers. Migrants from the former USSR even fare slightly better.

Finally, we show that local labor market concentration upon displacement is an important contributor to displaced workers' costs of job loss. Displaced workers living in municipalities with a higher increase in local unemployment rates or in cities face greater losses. One important factor driving the migrant-native gap in earnings losses is competition by same-nationality workers: The higher the share of the working-age population of the same nationality in their workplace county, the larger migrants' earnings losses are.

Policymakers interested in improving migrants' labor market outcomes should pay attention to our finding that migrants face substantial difficulties in job search after displacement. When searching for a job, migrants may therefore need a different type of training than natives (e.g., language courses or training targeted at learning how the job application process in their destination country works). Such programs should target low-skilled migrants, in particular. For authorities, it may be worthwhile to invest in different types of trainings for unemployed individuals, depending on their migration status.

Table 2.1: Worker Characteristics of Displaced Workers and Matches in t=-1

	(1) Nondisplaced Migrants	(2) Displaced Migrants	(3) Nondisplaced Natives	(4) Displaced Natives
Panel A: Individual Characteristics				
Years of Education	11.2	11.2	12.3	12.3
	[1.68]	[1.61]	[1.76]	[1.77]
Age	37.9	37.9	39.4	39.4
	[6.83]	[6.68]	[6.82]	[6.71]
Tenure	6.37	6.38	6.19	6.20
	[2.60]	[2.56]	[2.46]	[2.43]
Real Daily Wage (EUR)	91.3	89.2	104.1	102.3
	[30.1]	[30.8]	[36.1]	[36.7]
Total Yearly Earnings	33644.5	30194.9	38028.3	35477.8
	[11159.3]	[11844.1]	[13486.1]	[14189.6]
Days Worked in Year	362.7	335.5	362.8	344.2
	[15.1]	[53.9]	[14.1]	[45.6]
Panel B: Regional Characteristics				
Lives in City	0.77	0.80	0.55	0.57
	[0.42]	[0.40]	[0.50]	[0.50]
Lives in East Germany	0.031	0.041	0.22	0.25
	[0.17]	[0.20]	[0.42]	[0.43]
Local UR Change (between $t = -1$ and $t = 0$)	0.014	0.027	0.019	0.035
	[0.14]	[0.14]	[0.13]	[0.14]
Panel C: Establishment Characteristics				
Establishment Size	277.3	291.1	328.9	347.2
	[532.0]	[490.4]	[723.2]	[636.8]
Share Migrant Workers	0.22	0.25	0.064	0.074
	[0.19]	[0.19]	[0.085]	[0.095]
Share High-Skilled Workers	0.079	0.079	0.12	0.12
	[0.12]	[0.12]	[0.16]	[0.16]
Share Marginally Employed Workers	0.078	0.059	0.054	0.041
	[0.15]	[0.13]	[0.11]	[0.095]
Displaced from Complete Closure	0.00011	0.32	0.000077	0.32
	[0.011]	[0.47]	[0.0088]	[0.47]
Number of Observations	17605	17605	129701	129701

Notes: Characteristics of displaced and matched, nondisplaced workers in the year prior to the displacement year. Workers satisfy the following baseline restrictions: Aged 24 to 50, working full time in predisplacement year, at least 3 years of tenure, and establishment has at least 50 employees. Nondisplaced workers are matched to displaced workers using propensity score matching within year and industry cells. The nondisplaced sample of workers is a random sample of workers (one per displaced worker) who satisfy the same baseline restrictions. UR is an abbreviation for unemployment rate. Standard deviations in brackets. Source: IEB and BHP.

Table 2.2: Comparing Displaced Workers in t=-1 to a Random Sample of Workers

	(1) Random 2-percent Sample Migrants	(2) Baseline Displacement Sample Migrants	(3) Reweighted Displacement Sample Migrants	(4) Random 2-percent Sample Natives	(5) Baseline Displacement Sample Natives
Panel A: Individual Characteristics					
Years of Education	11.2	11.2	12.1	12.0	12.3
	[2.05]	[1.61]	[2.20]	[1.94]	[1.77]
Age	37.8	37.9	39.8	40.4	39.4
	[12.5]	[6.68]	[6.71]	[13.3]	[6.71]
Tenure	2.37	6.38	6.05	2.93	6.20
	[2.07]	[2.56]	[2.34]	[2.17]	[2.43]
Real Daily Wage	57.5	89.2	105.2	68.7	102.3
	[48.8]	[30.8]	[37.5]	[53.0]	[36.7]
Total Yearly Earnings	13620.3	30194.9	35928.0	20661.7	35477.8
	[16493.5]	[11844.1]	[14285.7]	[18855.8]	[14189.6]
Days per year working	214.8	335.5	338.7	281.9	344.2
	[158.6]	[53.9]	[51.1]	[135.1]	[45.6]
Panel B: Regional Characteristics					
Lives in City	0.64	0.80	0.58	0.44	0.57
	[0.48]	[0.40]	[0.49]	[0.50]	[0.50]
Lives in East Germany	0.063	0.041	0.060	0.19	0.25
	[0.24]	[0.20]	[0.24]	[0.39]	[0.43]
Panel C: Establishment Characteristics					
Size of establishment	1000.3	291.1	334.0	782.1	347.2
	[3922.8]	[490.4]	[640.6]	[3473.1]	[636.8]
Share Migrant Workers	0.30	0.24	0.18	0.053	0.075
	[0.27]	[0.19]	[0.18]	[0.086]	[0.095]
Share High-Skilled Workers	0.099	0.079	0.13	0.13	0.12
	[0.16]	[0.12]	[0.18]	[0.17]	[0.16]
Share Marginally Employed Workers	0.21	0.059	0.049	0.17	0.041
	[0.28]	[0.13]	[0.11]	[0.26]	[0.095]
Displaced from Complete Closure	.	0.32	0.32	.	0.32
		[0.47]	[0.47]		[0.47]
Number of Observations	574167	17605	17605	5882551	129701

Notes: This table summarizes characteristics of different samples of (displaced) migrants and natives. Columns (1) and (4) show characteristics of a random 2-percent sample of all workers subject to social security in Germany 2000-2010. Columns (2) and (5) represent all displaced workers in our baseline sample. We measure characteristics in the year prior displacement (t=-1). Column (3) reports migrants in our sample reweighted to natives. Standard deviations in brackets. Source: IEB, BBSR.

Table 2.3: Explaining Earnings Losses by Local Labor Market Concentration

	(1) Δ Log (Earnings)	(2) Δ Log (Earnings)	(3) Δ Log (Earnings)	(4) Δ Log (Earnings)	(5) Δ Log (Earnings)	(6) Δ Log (Earnings)
Migrant	-0.19**	-0.19**	-0.13**	-0.20	-0.12**	-0.25*
	(0.016)	(0.017)	(0.022)	(0.13)	(0.023)	(0.12)
Local UR Change		-0.11**			-0.12**	-0.12**
		(0.041)			(0.041)	(0.042)
Migrant*UR Change		-0.090			-0.14	-0.15
		(0.12)			(0.12)	(0.12)
City Residency			-0.056**		-0.058**	-0.064**
			(0.012)		(0.012)	(0.011)
Migrant*City Residency			-0.058*		-0.063*	-0.040
			(0.027)		(0.027)	(0.028)
Share Same Nationality				-0.053		-0.17
				(0.14)		(0.12)
Migrant*Share Same Nationality				-3.03**		-2.96**
				(0.82)		(0.78)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	127653	126524	126924	123542	125834	122635
R^2	0.049	0.050	0.050	0.050	0.050	0.052
Mean Dep. Var (Native)	-0.36	-0.36	-0.36	-0.36	-0.36	-0.36

Notes: This table shows to what extent local labor market conditions contribute to the migrant-native earnings gap after job displacement. The outcome variable is based on the individual difference-in-differences estimate which measures the change in log earnings before (t=-5 to t=-2) vs. after (t=0 to t=3) job loss for displaced vs. non-displaced workers within matched worker pairs. We successively add controls for local unemployment rate (UR) changes reported at the municipality (LAU) level (Column (2)), city residency (Column (3)), and the share of co-ethnic working age population in a county (NUTS 3) (Column (4)), all measured in t=-1. Columns (5) and (6) show the coefficients when (all) controls are included simultaneously. All columns control for baseline characteristics (age, age squared, years of education, tenure, experience, full-time work, log establishment size, 1-digit industries, 1-digit occupations (all in t=-1), and log wage (in t=-3)). The regression sample includes displaced workers, only. ** and * refer to statistical significance at the 1 and 5 percent level for standard errors clustered at the baseline county level. Workers in our sample are displaced in 2001-2011, and they are observed from 1996 to 2017. Source: IEB, BBSR, Destatis.

Table 2.4: Explaining Costs of Job Displacement by Local Labor Market Concentration

	(1) Δ Employed	(2) Δ Days Worked	(3) Δ Log Daily Wage	(4) Δ Commutes	(5) Δ AKM Effect	(6) Δ Share Migrants	(7) Δ Share Marginally Employed
Panel A: Controlling for Individual Characteristics, Industry, and Occupation							
Migrant	-0.040**	-21.1**	-0.11**	-0.0070	-0.030**	0.0014	0.028**
	(0.0044)	(2.10)	(0.011)	(0.0096)	(0.0066)	(0.0035)	(0.0031)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	133338	133338	121866	121676	94866	119631	119291
R^2	0.020	0.034	0.047	0.018	0.093	0.007	0.021
Mean Dep. Var (Native)	-0.094	-58.7	-0.17	0.027	-0.072	-0.0098	0.034
Panel B: Adding Controls for Local Unemployment Rate Change, City Residency, and Share of Coethnic Neighbors							
Migrant	0.012	-26.5	-0.15*	-0.016	0.0096	0.17**	-0.027
	(0.031)	(16.2)	(0.070)	(0.069)	(0.046)	(0.016)	(0.016)
Local UR Change	-0.014	-15.8*	-0.020	0.017	-0.039	0.0071	0.0069
	(0.011)	(6.21)	(0.023)	(0.017)	(0.044)	(0.0065)	(0.0070)
Migrant*UR Change	-0.0100	-6.46	-0.087	-0.064	0.053	-0.011	0.0089
	(0.027)	(14.8)	(0.076)	(0.053)	(0.044)	(0.032)	(0.025)
City Residency	-0.018**	-9.72**	-0.022**	0.059**	0.0024	0.00026	0.0037**
	(0.0035)	(1.61)	(0.0062)	(0.0099)	(0.0063)	(0.0012)	(0.0014)
Migrant*City Residency	0.0047	3.63	-0.073**	-0.024	-0.033**	-0.00053	0.018**
	(0.0066)	(3.10)	(0.019)	(0.019)	(0.0079)	(0.0062)	(0.0061)
Share Same Nationality	0.045	-12.1	-0.13	-0.0063	0.013	0.18**	-0.039*
	(0.033)	(17.1)	(0.073)	(0.073)	(0.050)	(0.017)	(0.016)
Migrant*Share Same Nationality	-0.76**	-329.3**	-2.06**	0.66	-0.70	0.13	0.38*
	(0.18)	(86.2)	(0.57)	(0.47)	(0.44)	(0.20)	(0.15)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	128092	128092	117075	116885	91178	115078	114745
R^2	0.021	0.035	0.049	0.021	0.095	0.015	0.022
Mean Dep. Var (Native)	-0.094	-58.7	-0.17	0.027	-0.072	-0.0098	0.034

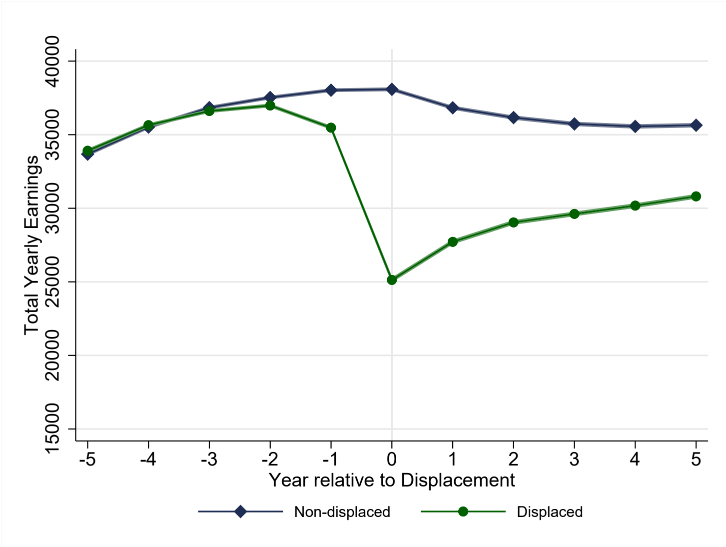
Notes: This table shows to what extent local labor market conditions contribute to migrant-native gaps in labor market outcomes after job displacement. All outcome variables are based on individual difference-in-differences estimates which measure the change in the outcome (e.g. log dailys wages) before (t=-5 to t=-2) vs. after (t=0 to t=3) job loss for displaced vs. non-displaced workers within matched worker pairs. Panel A reports coefficients when controlling for baseline characteristics only (age, age squared, years of education, tenure, experience, full-time work, log establishment size, 1-digit industries, 1-digit occupations (all in t=-1), and log wage (in t=-3)). Panel B reports coefficients when adding controls for local unemployment rate (UR) changes reported at the municipality (LAU) level, city residency, and the share of coethnic working age population in a county (NUTS 3), all measured in t=-1. The AKM effect is a proxy for wage differentials across firms, based on Abowd et al. [1999]. The regression sample includes displaced workers, only. ** and * refer to statistical significance at the 1 and 5 percent level for standard errors clustered at the baseline county level. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. Source: IEB, BBSR, Destatis.

Table 2.5: Robustness Checks: Matching and Reweighting

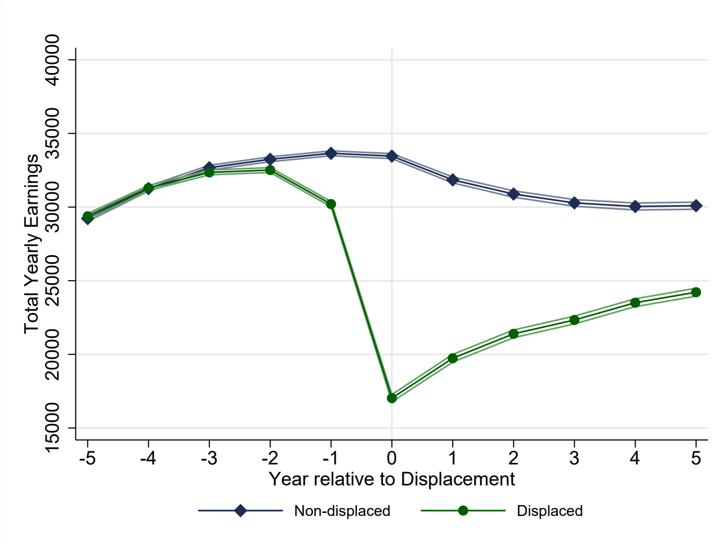
	(1) Baseline	(2) 1 Year Tenure	(3) 2 Years Tenure	(4) Matching w/o Wages	(5) Matching in t=-4	(6) County Matching	(7) Time Window	(8) Reweight. Nat. to Mig.
Panel A: Earnings Rel. to Year -2								
Migrant	-0.064	-0.064	-0.056	-0.10	-0.050	-0.070	-0.047	-0.068
	(0.011)**	(0.013)**	(0.011)**	(0.034)**	(0.013)**	(0.013)**	(0.010)**	(0.0094)**
Observations	132270	176391	153192	203795	145143	127281	130664	131802
R^2	0.003	0.017	0.002	0.001	0.001	0.003	0.001	0.003
Mean Dep. Var Men	-.225	-.238	-.226	-.266	-.259	-.232	-.162	-.226
	(.002)	(.002)	(.003)	(.004)	(.003)	-.232	(.002)	(.002)
Panel B: Log Wages								
Migrant	-0.041	-0.064	-0.016	-0.040	-0.041	-0.033	-0.0058	-0.10
	(0.013)**	(0.011)**	(0.013)	(0.013)**	(0.013)**	(0.016)*	(0.015)	(0.013)**
Observations	120766	161864	140244	182831	132425	116336	124353	120431
R^2	0.001	0.024	0.000	0.001	0.001	0.001	0.000	0.005
Mean Dep. Var Men	-.176	-.176	-.177	-.174	-.189	-.181	-.159	-.177
	(.002)	(.002)	(.003)	(.002)	(.002)	-.181	(.002)	(.002)
Panel C: Employment								
Migrant	-0.036	-0.030	-0.036	-0.038	-0.033	-0.042	-0.023	-0.034
	(0.0070)**	(0.0051)**	(0.0064)**	(0.0043)**	(0.0064)**	(0.0072)**	(0.0057)**	(0.0041)**
Observations	132270	176391	153192	203795	145143	127281	130664	131802
R^2	0.004	0.030	0.004	0.004	0.003	0.005	0.002	0.003
Mean Dep. Var Men	-.094	-.089	-.092	-.091	-.095	-.093	-.067	-.094
	(.001)	(.001)	(.001)	(.001)	(.001)	-.093	(.001)	(.001)
Panel D: Days Worked Full-time								
Migrant	-24.6	-25.0	-24.0	-32.9	-21.0	-24.8	-18.9	-26.5
	(3.07)**	(2.40)**	(3.08)**	(2.19)**	(2.85)**	(3.28)**	(2.93)**	(2.38)**
Observations	132270	176391	153192	203795	145143	127281	130664	131802
R^2	0.007	0.033	0.007	0.010	0.005	0.007	0.004	0.008
Mean Dep. Var Men	-67.5	-65.54	-66.39	-64.155	-69.2	-68.417	-45.206	-67.654
	(.415)	(.375)	(.536)	(.376)	(.419)	-68.417	(.429)	(.414)

Notes: Each column in this table represents a different robustness check of a weighted difference-in-differences regression. All outcome variables are based on the individual difference-in-differences estimate which measures differences in the outcome before (t=-5 to t=-2) vs. after (t=0 to t=3) job loss for displaced vs. non-displaced workers. Column (1) reports the baseline coefficients. Column (2) and (3) report results when relaxing the baseline tenure restriction to 1 and 2 years, respectively. Columns (4) and (5) report results of a matching specification where we do not match on trends in wages and where we match on characteristics in t=-4, only. Column (6) reports results when we match exactly on counties (NUTS 3 regions) instead of 1-digit industries in t=-1. Column (7) reports results for a longer time window (10 years pre vs. 10 years post displacement), and Column (8) reports results when we reweight natives to migrants. We cluster standard errors at the county level at time of displacement (constant within matched worker pairs). * and ** correspond to 5 and 1 percent signficance levels, respectively. Workers in our sample are displaced in 2001-2011, and they are observed from 1996 to 2017. Source: IEB.

Figure 2.1: Native and Migrant Workers' Earnings - No Controls



(A) Total Yearly Earnings in EUR - Natives



(B) Total Yearly Earnings in EUR - Migrants

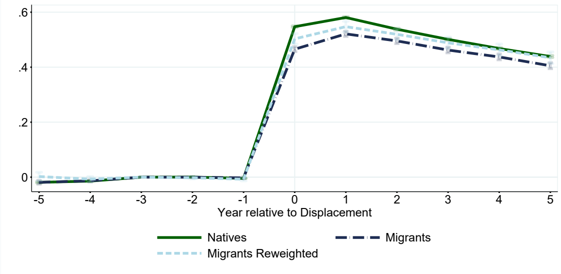
Notes: This figure plots raw earnings for displaced compared to nondisplaced workers, respectively for natives (Panel A) and migrants (Panel B). The blue diamonds show earnings trajectories for nondisplaced workers, and the green circles show earnings trajectories for workers displaced between $t = -1$ and $t = 0$. Displaced workers are matched to nondisplaced workers using propensity score matching. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. In $t = -1$, we observe 35,210 migrants and 259,402 natives. Source: IEB.

Figure 2.2: Labor Market Outcomes by Migration Status

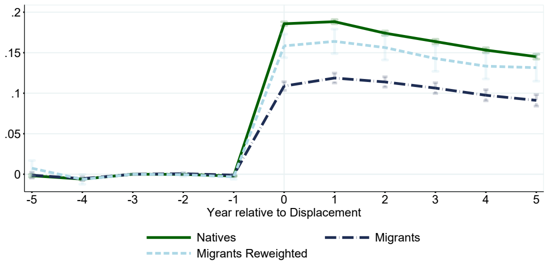
Notes: This figure plots event study regression coefficients on the differential evolution of the following outcomes for displaced vs. nondisplaced workers: earnings relative to t=-2 (Panel A), log wages (Panel B), employment probability (Panel C), days worked per year (Panel D), days worked in part-time employment per year (Panel E), and days worked in full-time employment per year (Panel F). The solid green line reports the results for our sample of native workers, the dashed blue line reports the results for our sample of migrant workers, and the light blue line reports the results for our sample of reweighted migrant workers. Reweighting characteristics are log wage (t=-3, t=-4), age (t=-1), years of education (t=-1), tenure (t=-1), city residency (t=-1), establishment size (t=-1), 1-digit industry (t=-1), and 1-digit occupation (t=-1). Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the individual level. Regressions control for year fixed effects, year since displacement fixed effects, age polynomials, and worker fixed effects. We omit t=-3 as the reference category. Displaced workers are matched to nondisplaced workers using propensity score matching. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. Source: IEB.

Figure 2.3: Geographic Mobility by Migration Status

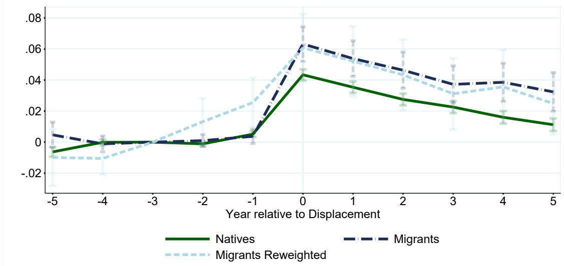
(A) Changed Workplace Municipality since t=-1



(B) Changed Workplace State since t=-1

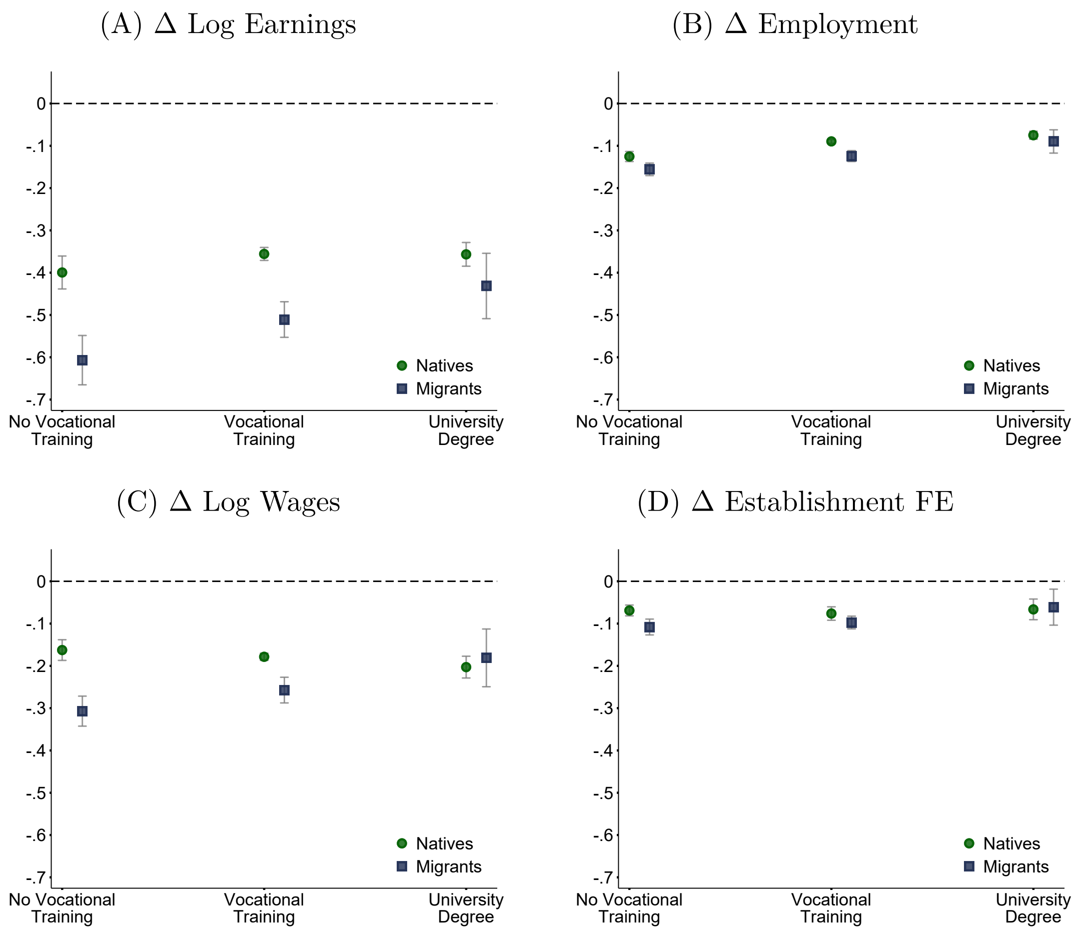


(C) Commutes



Notes: This figure plots event study regression coefficients on the differential evolution of the following mobility outcomes for displaced vs. nondisplaced workers: the propensity to change workplace to a different municipality (LAU) from t=-1 (Panel A), the propensity to change workplace to a different federal state (NUTS 1) from t=-1 (Panel B), and the propensity to commute (Panel C). We define commuting as working and living in different municipalities. The solid green line reports the results for our sample of native workers, the dashed blue line reports the results for our sample of migrant workers, and the light blue line reports the results for our sample of reweighted migrant workers. Reweighting characteristics are log wage (t=-3, t=-4), age (t=-1), years of education (t=-1), tenure (t=-1), city resident (t=-1), establishment size (t=-1), 1-digit industry (t=-1), and 1-digit occupation (t=-1). Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the individual level. Regressions control for year fixed effects, year since displacement fixed effects, age polynomials, and worker fixed effects. We omit t=-3 as the reference category. Displaced workers are matched to nondisplaced workers using propensity score matching. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. Source: IEB.

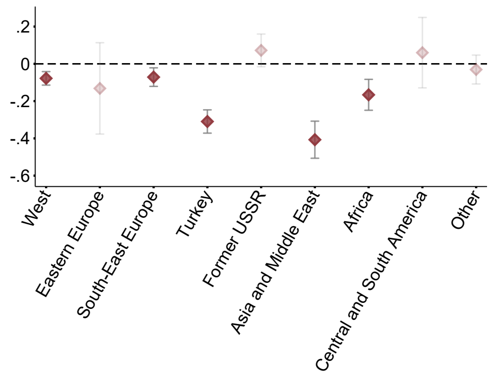
Figure 2.4: Costs of Job Displacement by Education Level and Migration Status



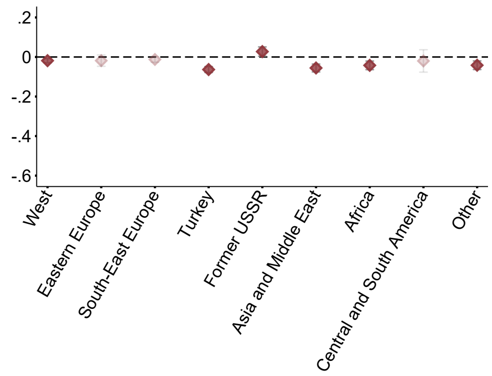
Notes: This figure shows how costs of job displacement differ by education level and migration status. We regress workers' individual difference-in-differences outcomes on dummies for 3 educational groups and an interaction of these educational groups with migration status. We then plot marginal effects at the respective educational groups. Each difference-in-differences outcome measures differences in the outcome before (t=-5 to t=-2) vs. after (t=0 to t=3) job loss for displaced vs. nondisplaced workers. Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the displacement establishment level. All regressions control for individual characteristics (age, age squared, years of education, tenure, experience, full-time work, log wage in t=-3, and log establishment size), 1-digit industries, and occupations according to Blossfeld [1987] in the year before displacement. Source: IEB.

Figure 2.5: Costs of Job Displacement by Origin Group

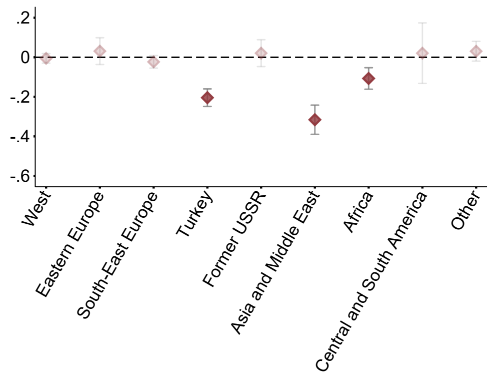
(A) Δ Log Earnings



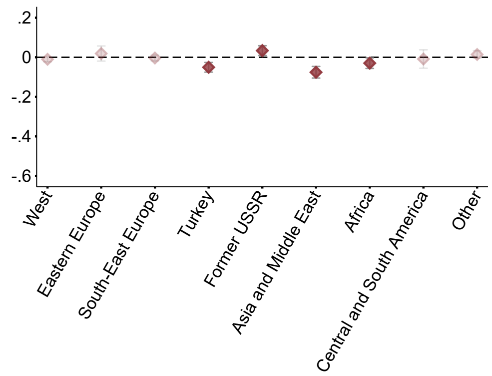
(B) Δ Employment



(C) Δ Log Wages

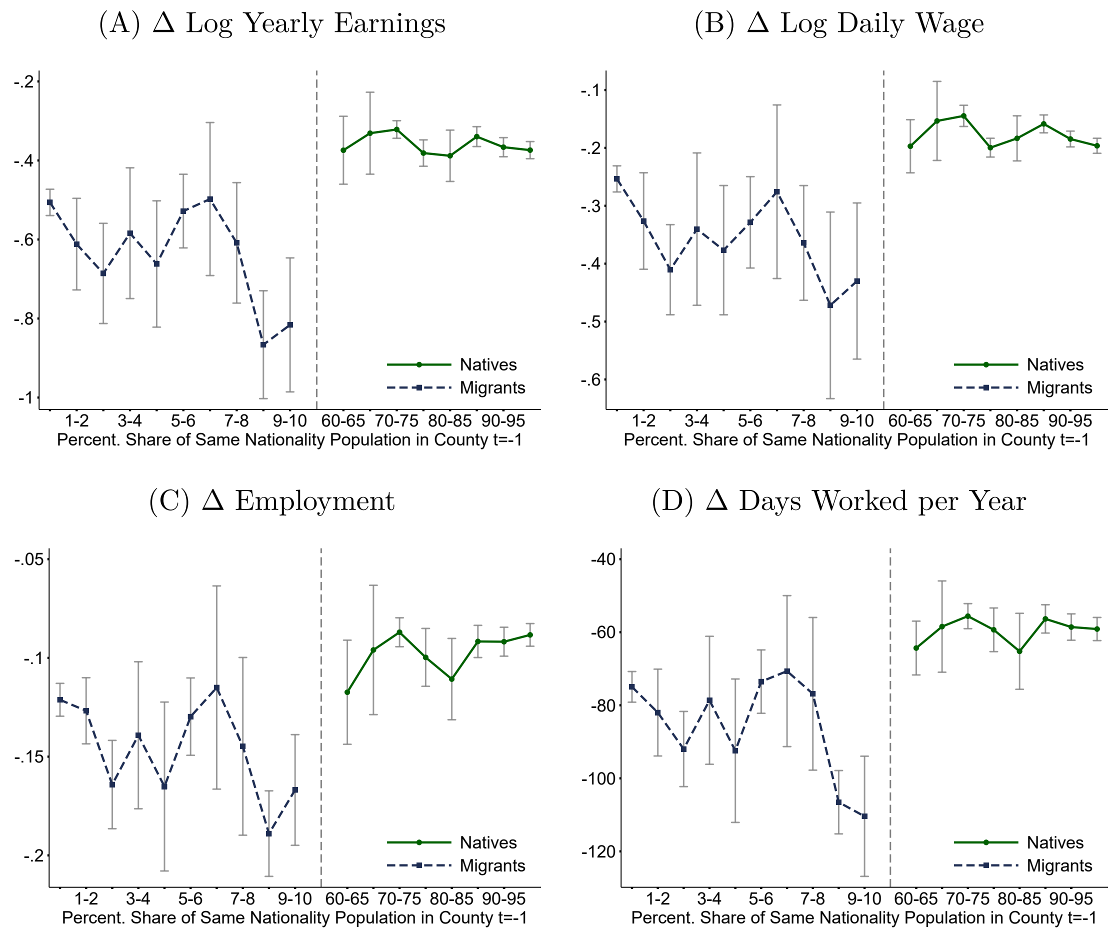


(D) Δ Establishment FE



Notes: This figure shows how costs of job displacement differ by origin group. Each panel plots coefficients from a separate OLS regression where we regress workers' individual difference-in-differences outcomes on dummies for the 9 origin groups, with "German origin" as omitted category. Each difference-in-differences outcome measures differences in the outcome before (t=-5 to t=-2) vs. after (t=0 to t=3) job loss for displaced vs. nondisplaced workers. Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the displacement establishment level. All regressions control for individual characteristics (age, age squared, years of education, tenure, experience, full-time work, log wage in t=-3, and log establishment size), 1-digit industries, and occupations according to Blossfeld [1987] in the year before displacement. Source: IEB.

Figure 2.6: Costs of Job Displacement by Share of Same-Nationality Working Age Population in t=-1



Notes: This figure shows how costs of job displacement differ by the share of the same-nationality working-age population in a worker's workplace county in t=-1. This share ranges from 0 to 10% for migrants and from 60 to 100% for natives. Panel (a) reports coefficients for log earnings, Panel (b) reports coefficients for log wages, Panel (c) reports coefficients for employment probability, and Panel (d) reports coefficients for number of days worked per year. We regress workers' individual difference-in-differences outcomes on the categories of same-nationality share reported on the x-axis. The solid green line reports the results for our sample of native workers, and the dashed blue line reports the results for our sample of migrant workers. Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the displacement county level. Our regression controls for individual characteristics (age, age squared, years of education, tenure, experience, full-time work, log wage in t=-3, and log establishment size), 1-digit industries, and occupations according to Blossfeld [1987] in the year before displacement. Source: IEB and Destatis.

Online Appendix

Appendix A: Population Data and Additional Results

Population Data

In order to analyze the role of local ethnic shares, we use the dataset *Population and Employment, Foreign Population, Results of the Central Register of Foreigners, Destatis, 2019*. It is based on official records from the German foreigners' registration office and thus highly reliable.

This dataset reports the population in Germany on December 31. It thus contains the exact population of a given nationality by age and county. We have access to this data for each year in the period 1998-2017. To construct our ethnic share measure, we restrict the data to the working age population, i.e. individuals aged 15-65. We then divide nationalites into groups of origin according to Battisti et al. [2021] (see also Table 2.16). To give one example: Rather than analyzing the share of Polish citizens by itself, we group them into a cluster of Central European countries (Polish, Czech, Hungarian, Slovakian, and Slovenian citizens). The idea is that on the one hand, individuals from these countries have a similar educational background and are thus likely substitutes for eachother. On the other hand, these countries are culturally closely related, and Central European citizens may thus form ethnic clusters.

In a last step, we divide the number of each nationality group in a given county by the overall working age population in that county on December 31:

$$\frac{P_{oct}}{P_{oct} + N_{ct}}$$

where P_{oct} is the number of working age citizens from a given nationality group o, in county c, and at time t. N_{ct} is the number of working age natives in county c and at time t.

Figure 2.15 shows how the share of same-nationality working age population is distributed among displaced workers. Not surprisingly, it takes much higher values for Germans (60-100%) than migrants (0-10%). Even though the share is skewed towards 0 for migrant workers, there is substantial variation in the ethnic share: About 16% of displaced migrants live in counties with an ethnic share of 5% or more, and one third of displaced migrant workers live in counties with an ethnic share of at least 3%.

Note that the population data comes with a drawback: For the majority of foreigners' registration offices, the jurisdictions coincide with German counties. However, in the federal states of Saarland, Hesse, and Brandenburg, a county-specific assignment of data is not always possible. Therefore, it is not possible to determine the percentage of the working-age population of a certain nationality for all German counties over the whole period. For instance, in the year 2017, 10 out of 401 German counties could not be merged (Kassel city and the county of Kassel, all six counties of Saarland, Cottbus, and the county of Spree-Neiße). This is only a minor issue for our analysis, as the vast majority of counties - especially the five largest metropolitan areas: Berlin, Cologne, Frankfurt, Hamburg, and Munich - are included in the sample.

Alternative Earnings Measures - Raw Gap

Panel (a) of Figure 2.7 presents the event study coefficients for yearly earnings losses relative to earnings in t=-2 which we have already discussed in Section 2.3 in the paper.

Note that in Panels (b) and (c) of Figure 2.7, we report log(earnings) and log(earnings+1) as alternative outcome variables. While including workers with zero earnings in Panel (c) substantially increases the size of our coefficients, the overall pattern holds: Both migrant and native displaced workers face large earnings losses, with a substantial gap between migrant and native displaced workers. We observe the same pattern in Panel (d), which shows total yearly earnings (in EUR).

The Role of Establishment Characteristics

If displacement had similar effects on migrant and native workers, then we would expect them to, on average, sort into similar establishments. Yet as we have shown, migrant and native workers differ in observable characteristics, and workers with particular characteristics may sort into specific types of establishments. In an additional analysis, we therefore estimate our main regression equation with establishment-specific outcome variables, including a specification where we reweight migrant to native workers in terms of individual characteristics, industries, and occupations.

The solid green and dashed dark blue lines in Panels (a)-(b) of Figure 2.11 show that both displaced migrants and natives sort into "worse" establishments after displacement. These establishments have lower wage premia (Panel (a)), and have higher shares of marginally employed workers (Panel (b)). Looking at the raw migrant-native gap only, our results suggest that migrants sort into substantially worse establishments, with an even larger reduction in wage premia, and a higher share of marginally employed coworkers. However, once we control for observable characteristics, these differences largely disappear (dashed light blue lines in Panels (a) and (b)), suggesting that observables can largely explain the differential sorting of migrants and natives. Note that while the migrant-native gap with respect to establishment fixed effects (Panel (a)) slightly closes after reweighting, the difference remains statistically significant, suggesting that a greater share in migrants' wage losses can be explained by a loss in establishment wage premia.

Panel (c) of Figure 2.11 shows that after job displacement, both migrants and natives sort into establishments with a lower share of migrant workers compared to control workers.[36] Initially, this share is particularly low for migrants but they catch up with natives as time passes. For the re-weighted sample, the difference in establishments' migrant share disappears starting from the second year after displacement. Given that previous literature has identified ethnic networks as an important source of information in migrants' job search, this pattern is somewhat surprising.

The Correlation of Wage Residuals with Costs of Job Displacement

So far, we have documented a number of observable characteristics and how they relate to costs of job displacement. One obvious question is whether a worker's earn-

[36] Note that for the share of migrant workers in an establishment, we compute the "leave-one-out mean", as otherwise the share mechanically increases if displaced migrants start working at a new establishment.

ings losses after job displacement also relate to their unobservable characteristics, such as ability.

As an additional analysis, we thus follow Borjas et al. [2019] and compute residuals from a Mincerian wage regression to show whether they correlate with workers' earnings, wage, and employment losses. Our regression sample includes displaced workers in the baseline year (t=-1). We regress wages on a number of observable characteristics: 7 education group dummies[37], potential experience, experience squared, tenure, a dummy for full-time employment on June 30, age dummies, year dummies, and log firmsize. The outcome variable is log wages.

We think of the residuals from this regression in two ways: First, they may reflect employer-employee-specific match quality in the job that workers are displaced from. Workers with particularly high wage residuals may thus also face greater costs of job displacement, because they "have more to lose". Second, the unobserved component of the Mincerian wage regressions may reflect how a given worker is selected in terms of their unobservable characteristics. In that case, we may expect that workers with higher residuals find a new job quicker, and have lower earnings losses. We plot the cumulative distribution of the wage residuals in Figure 2.13, which shows that migrants have overall lower match quality, or constitute a negative selection relative to natives.

The binscatter plots in Figure 2.14, which correlate deciles of wage residuals (x-axis) with the individual difference-in-differences outcome as described in Equation 2.3, provide evidence for both of these potential mechanisms: For earnings and wages, there is a weak U-shape pattern, suggesting that both workers in the highest and lowest deciles lose relatively little. Overall, this pattern is stronger for migrant workers, perhaps reflecting the fact that migrant workers in the upper decile constitute a particularly positive selection.

[37]These include the following: No training, middle school (Volks-, Haupt-, Realschule) without vocational training, middle school with vocational training, secondary school (Abitur) without vocational trianing, secondary school with vocational training, university of applied sciences degree, university degree.

Appendix B: Tables and Figures

Table 2.6: Workers' Distribution Across Industries in t=-1

	(1) Nondisplaced Migrants	(2) Displaced Migrants	(3) Nondisplaced Natives	(4) Displaced Natives
Agriculture	0.00023	0.00023	0.00084	0.00084
	[0.015]	[0.015]	[0.029]	[0.029]
Mining, Energy	0.034	0.034	0.023	0.023
	[0.18]	[0.18]	[0.15]	[0.15]
Food Manufacturing	0.064	0.064	0.037	0.037
	[0.24]	[0.24]	[0.19]	[0.19]
Consumption Goods	0.10	0.10	0.070	0.070
	[0.30]	[0.30]	[0.25]	[0.25]
Production Goods	0.12	0.12	0.084	0.084
	[0.33]	[0.33]	[0.28]	[0.28]
Investment Goods	0.16	0.16	0.15	0.15
	[0.37]	[0.37]	[0.36]	[0.36]
Construction	0.039	0.039	0.086	0.086
	[0.19]	[0.19]	[0.28]	[0.28]
Retail	0.11	0.11	0.13	0.13
	[0.32]	[0.32]	[0.34]	[0.34]
Traffic, Telecommunication	0.075	0.075	0.069	0.069
	[0.26]	[0.26]	[0.25]	[0.25]
Credit, Insurance	0.0043	0.0043	0.015	0.015
	[0.066]	[0.066]	[0.12]	[0.12]
Restaurants	0.021	0.021	0.0052	0.0052
	[0.14]	[0.14]	[0.072]	[0.072]
Education	0.0022	0.0022	0.020	0.020
	[0.046]	[0.046]	[0.14]	[0.14]
Health	0.0051	0.0051	0.012	0.012
	[0.071]	[0.071]	[0.11]	[0.11]
Commercial Services	0.23	0.23	0.24	0.24
	[0.42]	[0.42]	[0.43]	[0.43]
Other Services	0.022	0.022	0.028	0.028
	[0.15]	[0.15]	[0.16]	[0.16]
Non-Profit	0.0092	0.0092	0.013	0.013
	[0.095]	[0.095]	[0.11]	[0.11]
Public Administration	0.0022	0.0022	0.018	0.018
	[0.047]	[0.047]	[0.13]	[0.13]
Number of Observations	17605	17605	129701	129701

Notes: Distribution across industries of displaced and nondisplaced workers in the year prior to the displacement year. Workers satisfy the following baseline restrictions: Aged 24 to 50, working full-time in predisplacement year, at least 3 years of tenure, and establishment has at least 50 employees. Nondisplaced workers are matched to displaced workers using propensity score matching within year and industry cells. The nondisplaced sample of workers is a random sample of workers (one per displaced worker) who satisfy the same baseline restrictions. Standard deviations in brackets. Source: IEB.

Table 2.7: Workers' Distribution Across Occupations in t=-1

	(1) Nondisplaced Migrants	(2) Displaced Migrants	(3) Nondisplaced Natives	(4) Displaced Natives
Agriculture, gardening, work with animals	0.0066	0.0043	0.0072	0.0041
	[0.081]	[0.065]	[0.085]	[0.064]
Simple, manual tasks	0.42	0.46	0.22	0.24
	[0.49]	[0.50]	[0.41]	[0.42]
Qualified, manual tasks	0.18	0.17	0.24	0.26
	[0.38]	[0.38]	[0.43]	[0.44]
Technician	0.025	0.029	0.073	0.075
	[0.16]	[0.17]	[0.26]	[0.26]
Engineer	0.017	0.015	0.043	0.038
	[0.13]	[0.12]	[0.20]	[0.19]
Simple services	0.23	0.20	0.14	0.12
	[0.42]	[0.40]	[0.35]	[0.33]
Qualified services	0.013	0.012	0.019	0.017
	[0.11]	[0.11]	[0.14]	[0.13]
Semi-professions	0.0050	0.0047	0.016	0.015
	[0.071]	[0.069]	[0.13]	[0.12]
Professions	0.0039	0.0041	0.0084	0.011
	[0.062]	[0.064]	[0.091]	[0.10]
Simple commercial and admin. tasks	0.023	0.021	0.039	0.035
	[0.15]	[0.14]	[0.19]	[0.18]
Qualified commercial and admin. tasks	0.065	0.061	0.16	0.16
	[0.25]	[0.24]	[0.37]	[0.36]
Manager	0.010	0.012	0.029	0.029
	[0.10]	[0.11]	[0.17]	[0.17]
Not classified	0.0026	0.0025	0.0032	0.0030
	[0.050]	[0.050]	[0.057]	[0.055]
Number of Observations	17605	17605	129701	129701

Notes: Distribution across occupations according to Blossfeld [1987] of displaced and nondisplaced workers in the year prior to the displacement year. Workers satisfy the following baseline restrictions: Aged 24 to 50, working full-time in predisplacement year, at least 3 years of tenure, and establishment has at least 50 employees. Nondisplaced workers are matched to displaced workers using propensity score matching within year and industry cells. The nondisplaced sample of workers is a random sample of workers (one per displaced worker) who satisfy the same baseline restrictions. Standard deviations in brackets. Source: IEB.

Table 2.8: The Importance of Reweighting Variables in Explaining the Migrant-Native Gap in Earnings Rel. To t=-2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Migrant	-0.093**	-0.087**	-0.087**	-0.087**	-0.084**	-0.084**	-0.081**	-0.077**
	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.0099)	(0.0095)	(0.0091)
Age in t-1		-0.0015**	-0.0014**	-0.0017**	-0.0017**	-0.0017**	-0.0016**	-0.0020**
		(0.00033)	(0.00032)	(0.00030)	(0.00030)	(0.00029)	(0.00028)	(0.00026)
Education in t-1		0.0075**	0.0074**	0.0044**	0.0045**	0.0046**	0.0042**	0.0030*
		(0.0013)	(0.0013)	(0.0014)	(0.0014)	(0.0013)	(0.0012)	(0.0012)
Tenure in t-1			-0.00074	-0.0013	-0.0013	-0.0011	-0.00098	0.00016
			(0.0013)	(0.0013)	(0.0013)	(0.0012)	(0.0012)	(0.0010)
Log wage in t-3				0.010	0.0093	0.0055	0.0048	0.018**
				(0.0066)	(0.0066)	(0.0065)	(0.0067)	(0.0067)
Log wage in t-4				0.014*	0.014*	0.013*	0.013*	0.019**
				(0.0059)	(0.0059)	(0.0058)	(0.0059)	(0.0059)
City Resident in t-1					-0.0077*	-0.0098**	-0.011**	-0.015**
					(0.0038)	(0.0032)	(0.0033)	(0.0029)
Log(Firmsize) in t-1						0.016**	0.016**	0.0061
						(0.0052)	(0.0054)	(0.0045)
Observations	266136	266136	266136	266136	264576	264576	264576	264576
Occupation Controls	No	No	No	No	No	No	Yes	Yes
Industry Controls	No	No	No	No	No	No	No	Yes

Notes: Each column in each panel returns the coefficients from the OLS regression. The dependent variable is earnings relative to earnings in t=-2. The displacement occurred between t=-1 and t=0. Controls correspond to our reweighting variables (all measured in t=-1 if not stated otherwise): Age, education (in years), tenure (in years), log(wage) (in t=-3 and t=-4), a dummy for city residency, log(firmsize), 1-digit occupations, and 1-digit industries. We cluster standard errors at the displacement establishment level (constant within matched worker pairs). * and ** correspond to 5 and 1 percent significance levels, respectively. Source: IEB.

Table 2.9: The Power of Geographic Mobility in Explaining the Migrant-Native Gap in Earnings Rel. To t=-2

	(1)	(2)	(3)	(4)
Migrant	-0.081**	-0.080**	-0.074**	-0.075**
	(0.0088)	(0.0088)	(0.0087)	(0.0087)
Commutes after Displ.		0.18**		0.17**
		(0.0047)		(0.0046)
Moves State after Displ.			0.20**	0.16**
			(0.0061)	(0.0056)
Observations	132226	132226	132226	132226
R^2	0.036	0.058	0.049	0.067
Mean Dep. Var (Native)	-0.23	-0.23	-0.23	-0.23

Notes: This table shows to what extent geographic mobility can explain the migrant-native gap in earnings relative to t=-2. In each regression, we control for the following variables (measured in t=-1 if not stated otherwise): Age, age squared, education (in years), tenure (in years), experience (in years), a dummy for working full-time, log(wages) (in t=-3 and t=-4), log(firmsize), 1-digit industry dummies, and 1-digit occupation dummies. We cluster standard errors at the displacement establishment level (constant within matched worker pairs). ** and * refer to statistical significance at the 1 and 5 percent level, respectively. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017.

Table 2.10: Restricting the Sample to Baseline Years up to 2007 (Pre Financial Crisis)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log (Earnings)		Log Wage		Employment		Days Worked	
	Natives	Migrants	Natives	Migrants	Natives	Migrants	Natives	Migrants
Year (Disp) t-5	0.015**	0.036**	0.0053**	0.0014	0.0024**	-0.0012	2.13**	6.02**
	(0.0018)	(0.010)	(0.0016)	(0.010)	(0.00052)	(0.0033)	(0.31)	(1.91)
Year (Disp) t-4	0.014**	0.015*	0.00039	0.011	-0.000023	-0.00045**	1.46**	2.95*
	(0.0011)	(0.0072)	(0.0014)	(0.0081)	(0.000024)	(0.00016)	(0.17)	(1.22)
Year (Disp) t-2	-0.011**	-0.014**	-0.016**	0.0089	0.0014**	0.00046	0.11	-0.65
	(0.00091)	(0.0039)	(0.0013)	(0.0097)	(0.00028)	(0.00087)	(0.16)	(0.60)
Year (Disp) t-1	-0.085**	-0.11**	-0.020**	-0.0063	-0.0000051	0.00092**	-18.8**	-24.1**
	(0.0011)	(0.0062)	(0.0015)	(0.0094)	(0.000036)	(0.00033)	(0.17)	(1.01)
Year (Disp) t	-0.58**	-0.70**	-0.22**	-0.24**	-0.14**	-0.18**	-114.1**	-135.2**
	(0.0032)	(0.019)	(0.0026)	(0.016)	(0.0011)	(0.0068)	(0.45)	(2.86)
Year (Disp) t+1	-0.36**	-0.50**	-0.19**	-0.23**	-0.12**	-0.17**	-68.9**	-92.3**
	(0.0032)	(0.021)	(0.0024)	(0.018)	(0.0011)	(0.0078)	(0.47)	(3.16)
Year (Disp) t+2	-0.27**	-0.37**	-0.17**	-0.21**	-0.094**	-0.14**	-47.4**	-67.6**
	(0.0032)	(0.020)	(0.0025)	(0.017)	(0.0011)	(0.0077)	(0.48)	(3.09)
Year (Disp) t+3	-0.23**	-0.30**	-0.16**	-0.17**	-0.078**	-0.11**	-37.5**	-53.8**
	(0.0032)	(0.019)	(0.0026)	(0.016)	(0.0012)	(0.0076)	(0.49)	(3.12)
Year (Disp) t+4	-0.19**	-0.24**	-0.15**	-0.12**	-0.067**	-0.092**	-30.8**	-42.4**
	(0.0032)	(0.021)	(0.0027)	(0.022)	(0.0012)	(0.0075)	(0.49)	(3.05)
Year (Disp) t+5	-0.17**	-0.17**	-0.14**	-0.10**	-0.059**	-0.089**	-26.1**	-36.1**
	(0.0032)	(0.019)	(0.0027)	(0.017)	(0.0012)	(0.0083)	(0.48)	(3.31)
Observations	2215070	265244	2144405	254099	2311627	282494	2311627	282494
R^2	0.104	0.115	0.050	0.049	0.072	0.104	0.153	0.190
Mean of dep. var	10.4	10.2	4.62	4.42	0.96	0.94	332.5	321.3

Notes: The table returns coefficients α_j from regression equation (1). The displacement occurred between t=-1 and t=0. The sample is restricted to pre financial crisis baseline years, e.g., all years up to 2007. Year $t = -3$ is omitted as the baseline category. The outcome variables are log (earnings) (columns 1 and 2), log wage (columns 3 and 4), employment (columns 5 and 6), and days worked (columns 7 and 8). In all columns, we control for year since displacement, year, and age polynomials. Standard errors are clustered at the individual level. Migrants are reweighted to natives using individual characteristics, industries, and occupations. ** and * refer to statistical significance at the 1 and 5 percent level, respectively. Source: IEB.

Table 2.11: Restricting the Sample to Baseline Years up to 2003 (Pre Financial Crisis)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log (Earnings)		Log Wage		Employment		Days Worked	
	Natives	Migrants	Natives	Migrants	Natives	Migrants	Natives	Migrants
Year (Disp) t-5	0.018**	0.042**	0.0042*	-0.0027	0.0029**	0.0031	2.66**	8.44**
	(0.0023)	(0.014)	(0.0019)	(0.012)	(0.00066)	(0.0040)	(0.39)	(2.47)
Year (Disp) t-4	0.014**	0.022*	-0.00071	0.0099	-0.000048	-0.00042*	1.55**	4.43**
	(0.0015)	(0.0100)	(0.0016)	(0.010)	(0.000030)	(0.00019)	(0.22)	(1.67)
Year (Disp) t-2	-0.0084**	-0.017**	-0.014**	0.0077	0.0022**	0.00067	0.51*	-0.91
	(0.0012)	(0.0051)	(0.0017)	(0.012)	(0.00041)	(0.0013)	(0.21)	(0.84)
Year (Disp) t-1	-0.081**	-0.11**	-0.024**	-0.016	-0.0000073	0.0013*	-17.6**	-23.0**
	(0.0013)	(0.0083)	(0.0018)	(0.012)	(0.000040)	(0.00052)	(0.21)	(1.29)
Year (Disp) t	-0.59**	-0.69**	-0.20**	-0.20**	-0.15**	-0.18**	-119.9**	-138.2**
	(0.0041)	(0.025)	(0.0031)	(0.019)	(0.0013)	(0.0093)	(0.56)	(3.89)
Year (Disp) t+1	-0.37**	-0.51**	-0.18**	-0.22**	-0.13**	-0.17**	-72.6**	-96.1**
	(0.0041)	(0.029)	(0.0030)	(0.022)	(0.0014)	(0.011)	(0.60)	(4.32)
Year (Disp) t+2	-0.27**	-0.38**	-0.16**	-0.20**	-0.10**	-0.15**	-50.0**	-73.1**
	(0.0041)	(0.027)	(0.0031)	(0.020)	(0.0015)	(0.011)	(0.62)	(4.26)
Year (Disp) t+3	-0.23**	-0.30**	-0.15**	-0.15**	-0.083**	-0.12**	-39.5**	-56.6**
	(0.0041)	(0.026)	(0.0033)	(0.020)	(0.0015)	(0.010)	(0.62)	(4.26)
Year (Disp) t+4	-0.19**	-0.25**	-0.14**	-0.11**	-0.071**	-0.095**	-32.0**	-45.3**
	(0.0041)	(0.028)	(0.0033)	(0.031)	(0.0015)	(0.010)	(0.62)	(4.09)
Year (Disp) t+5	-0.17**	-0.16**	-0.13**	-0.11**	-0.063**	-0.097**	-27.0**	-38.2**
	(0.0040)	(0.025)	(0.0034)	(0.022)	(0.0015)	(0.011)	(0.62)	(4.51)
Observations	1469255	150594	1418583	143951	1540502	161467	1540502	161467
R^2	0.103	0.112	0.049	0.046	0.077	0.109	0.162	0.198
Mean of dep. var	10.3	10.2	4.62	4.45	0.95	0.93	329.7	317.8

Notes: The table returns coefficients α_j from regression equation (1). The displacement occurred between t=-1 and t=0. The sample is restricted to the pre financial crisis baseline years, e.g., all years up to 2003. Year $t = -3$ is omitted as the baseline cateogry. The outcome variables are log (earnings) (columns 1 and 2), log wage (columns 3 and 4), employment (columns 5 and 6), and days worked (columns 7 and 8). In all columns, we control for year since displacement, year, and age polynomials. Standard errors are clustered at the individual level. Migrants are reweighted to natives using individual characteristics, industries, and occupations. ** and * refer to statistical significance at the 1 and 5 percent level, respectively. Source: IEB.

Table 2.12: Restricting the Sample to Workplace in West Germany at Time of Displacement

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log (Earnings)		Log Wage		Employment		Days Worked	
	Natives	Migrants	Natives	Migrants	Natives	Migrants	Natives	Migrants
Year (Disp) t-5	0.025**	0.0081	0.014**	-0.056	0.0037**	-0.0043	3.40**	11.3
	(0.0030)	(0.037)	(0.0022)	(0.034)	(0.00081)	(0.019)	(0.49)	(8.19)
Year (Disp) t-4	0.018**	0.059	0.0039*	0.017	-0.00012**	-0.00030	2.11**	12.0*
	(0.0018)	(0.034)	(0.0018)	(0.025)	(0.000043)	(0.00079)	(0.27)	(5.18)
Year (Disp) t-2	-0.015**	-0.018	-0.020**	-0.0017	0.00068	0.0033	-0.35	1.55
	(0.0014)	(0.017)	(0.0017)	(0.022)	(0.00040)	(0.0022)	(0.24)	(1.76)
Year (Disp) t-1	-0.091**	-0.10**	-0.026**	-0.031	0.000026	0.00026	-20.0**	-22.3**
	(0.0016)	(0.019)	(0.0019)	(0.021)	(0.000064)	(0.0015)	(0.27)	(2.96)
Year (Disp) t	-0.61**	-0.79**	-0.20**	-0.20**	-0.14**	-0.20**	-120.4**	-148.8**
	(0.0049)	(0.092)	(0.0037)	(0.052)	(0.0016)	(0.023)	(0.66)	(12.3)
Year (Disp) t+1	-0.35**	-0.56**	-0.17**	-0.28**	-0.12**	-0.17**	-70.2**	-91.9**
	(0.0049)	(0.095)	(0.0034)	(0.058)	(0.0017)	(0.031)	(0.71)	(13.5)
Year (Disp) t+2	-0.26**	-0.41**	-0.15**	-0.27**	-0.093**	-0.15**	-48.0**	-66.9**
	(0.0048)	(0.069)	(0.0035)	(0.052)	(0.0017)	(0.023)	(0.72)	(10.5)
Year (Disp) t+3	-0.22**	-0.42**	-0.14**	-0.20**	-0.078**	-0.13**	-38.3**	-67.4**
	(0.0049)	(0.083)	(0.0036)	(0.049)	(0.0018)	(0.023)	(0.73)	(11.4)
Year (Disp) t+4	-0.19**	-0.28**	-0.13**	-0.15**	-0.069**	-0.15**	-32.0**	-58.4**
	(0.0048)	(0.062)	(0.0037)	(0.050)	(0.0018)	(0.028)	(0.73)	(11.6)
Year (Disp) t+5	-0.17**	-0.21**	-0.12**	-0.15**	-0.061**	-0.13**	-27.8**	-47.6**
	(0.0047)	(0.061)	(0.0038)	(0.046)	(0.0018)	(0.025)	(0.73)	(11.0)
Observations	1021363	23973	983096	22811	1068103	25903	1068103	25903
R^2	0.110	0.135	0.054	0.061	0.074	0.118	0.164	0.209
Mean of dep. var	10.2	10.1	4.43	4.33	0.96	0.93	328.7	313.8

Notes: The table returns coefficients α_j from regression equation (1). The displacement occurred between t=-1 and t=0. The sample is restricted to workers employed in West Germany at the time of displacement. Year $t = -3$ is omitted as the baseline category. The outcome variables are log (earnings) (columns 1 and 2), log wage (columns 3 and 4), employment (columns 5 and 6), and days worked (columns 7 and 8). In all columns, we control for year since displacement, year, and age polynomials. Standard errors are clustered at the individual level. Migrants are reweighted to natives using individual characteristics, industries, and occupations. ** and * refer to statistical significance at the 1 and 5 percent level, respectively. Source: IEB.

Table 2.13: Reweighting Natives to Migrants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log (Earnings)		Log Wage		Employment		Days Worked	
	Natives	Migrants	Natives	Migrants	Natives	Migrants	Natives	Migrants
Year (Disp) t-5	0.0066*	0.027**	0.017**	0.013**	-0.00033	0.00012	-1.27*	1.92*
	(0.0033)	(0.0054)	(0.0025)	(0.0039)	(0.00088)	(0.0014)	(0.52)	(0.86)
Year (Disp) t-4	0.015**	0.014**	0.0084**	0.0036	0.00031**	0.000017	1.13**	1.79**
	(0.0021)	(0.0029)	(0.0022)	(0.0030)	(0.000045)	(0.000076)	(0.28)	(0.45)
Year (Disp) t-2	-0.018**	-0.016**	-0.022**	-0.017**	0.0016**	0.00032	-0.011	-0.24
	(0.0017)	(0.0023)	(0.0018)	(0.0026)	(0.00041)	(0.00057)	(0.24)	(0.35)
Year (Disp) t-1	-0.10**	-0.12**	-0.028**	-0.025**	-0.00072**	-0.000047	-21.8**	-27.0**
	(0.0020)	(0.0028)	(0.0020)	(0.0029)	(0.000072)	(0.00014)	(0.27)	(0.47)
Year (Disp) t	-0.66**	-0.91**	-0.26**	-0.43**	-0.14**	-0.19**	-120.4**	-149.4**
	(0.0050)	(0.0099)	(0.0040)	(0.0090)	(0.0015)	(0.0030)	(0.63)	(1.16)
Year (Disp) t+1	-0.43**	-0.62**	-0.23**	-0.33**	-0.12**	-0.18**	-74.2**	-96.6**
	(0.0048)	(0.010)	(0.0036)	(0.0078)	(0.0016)	(0.0033)	(0.66)	(1.30)
Year (Disp) t+2	-0.33**	-0.47**	-0.21**	-0.29**	-0.096**	-0.13**	-52.5**	-67.6**
	(0.0048)	(0.010)	(0.0036)	(0.0077)	(0.0016)	(0.0032)	(0.68)	(1.32)
Year (Disp) t+3	-0.29**	-0.39**	-0.19**	-0.26**	-0.082**	-0.11**	-42.5**	-53.5**
	(0.0048)	(0.010)	(0.0037)	(0.0078)	(0.0016)	(0.0033)	(0.68)	(1.34)
Year (Disp) t+4	-0.25**	-0.31**	-0.18**	-0.22**	-0.070**	-0.086**	-35.3**	-40.7**
	(0.0048)	(0.0099)	(0.0037)	(0.0078)	(0.0016)	(0.0033)	(0.67)	(1.34)
Year (Disp) t+5	-0.23**	-0.27**	-0.17**	-0.20**	-0.063**	-0.075**	-30.2**	-34.8**
	(0.0049)	(0.0100)	(0.0039)	(0.0079)	(0.0016)	(0.0032)	(0.66)	(1.33)
Observations	2589001	355810	2507729	341462	2696370	376467	2696370	376467
R^2	0.120	0.147	0.069	0.078	0.069	0.103	0.154	0.203
Mean of dep. var	10.3	10.2	4.60	4.40	0.96	0.95	334.0	323.5

Notes: The table returns coefficients α_j from regression equation (1). The displacement occurred between t=-1 and t=0. Year $t = -3$ is omitted as the baseline category. The outcome variables are log (earnings) (columns 1 and 2), log wage (columns 3 and 4), employment (columns 5 and 6), and days worked (columns 7 and 8). In all columns, we control for year since displacement, year, and age polynomials. Standard errors are clustered at the individual level. Natives are reweighted to migrants using individual characteristics, industries, and occupations. ** and * refer to statistical significance at the 1 and 5 percent level, respectively. Source: IEB.

Table 2.14: Explaining Costs of Job Loss by Local Labor Market Concentration and Controlling for Displacing Establishment

	(1) Δ Log (Earnings)	(2) Δ Employed	(3) Δ Days Worked	(4) Δ Log Wage	(5) Δ Commutes	(6) Δ AKM Effect	(7) Δ Share Migrants	(8) Δ Share Marg. Employed
Panel A: Controlling for Individual Characteristics, Industry, and Occupation								
Migrant	-0.20**	-0.040**	-22.1**	-0.12**	-0.00098	-0.032**	0.016**	0.024**
	(0.016)	(0.0049)	(2.20)	(0.012)	(0.0084)	(0.0050)	(0.0029)	(0.0031)
Observations	127653	133338	133338	121866	121676	94866	119631	119291
R^2	0.063	0.031	0.048	0.057	0.036	0.189	0.043	0.034
Mean of dep. var	-0.39	-0.099	-61.6	-0.19	0.028	-0.077	-0.010	0.038
Panel B: Adding Controls for Local Unemployment Rate Change, City Residency and Share of Coethnic Neighbors								
Migrant	0.33**	0.090*	46.0**	0.13	0.18**	0.044	0.048*	-0.028
	(0.11)	(0.036)	(16.2)	(0.078)	(0.059)	(0.033)	(0.020)	(0.018)
Local UR Change	-0.14**	-0.017	-18.2**	-0.053*	0.037*	-0.014	0.0064	0.0076
	(0.044)	(0.011)	(6.68)	(0.022)	(0.015)	(0.015)	(0.0058)	(0.0066)
Migrant*UR Change	-0.15	-0.0062	-5.27	-0.086	-0.055	0.035	-0.024	0.014
	(0.11)	(0.025)	(13.6)	(0.069)	(0.047)	(0.039)	(0.026)	(0.023)
City Residency	-0.058**	-0.015**	-8.81**	-0.017**	0.053**	-0.0020	0.0023*	0.0035**
	(0.0077)	(0.0024)	(1.06)	(0.0053)	(0.010)	(0.0023)	(0.00093)	(0.0013)
Migrant*City Residency	-0.049	0.0033	2.44	-0.077**	-0.034	-0.029**	0.00041	0.018**
	(0.028)	(0.0068)	(3.26)	(0.020)	(0.019)	(0.0072)	(0.0059)	(0.0059)
Share Same Nationality	0.49**	0.14**	70.6**	0.19*	0.18**	0.052	0.040	-0.039*
	(0.12)	(0.039)	(17.9)	(0.084)	(0.065)	(0.036)	(0.022)	(0.020)
Migrant*Share Same Nationality	-3.68**	-0.83**	-420.3**	-2.45**	-0.12	-0.95**	0.32	0.38*
	(0.81)	(0.17)	(93.8)	(0.59)	(0.38)	(0.32)	(0.21)	(0.15)
Observations	122635	128092	128092	117075	116885	91177	115078	114745
R^2	0.065	0.033	0.049	0.059	0.039	0.193	0.043	0.035
Mean of dep. var	-0.39	-0.099	-61.6	-0.19	0.028	-0.077	-0.010	0.038

Notes: This table shows to what extent local labor market conditions contribute to migrant-native gaps in labor market outcomes after job displacement. All regressions control for displacing establishment fixed effects. All outcome variables are based on individual difference-in-differences estimates which measure the change in the outcome (e.g. log dailys wages) before (t=-5 to t=-2) vs. after (t=0 to t=3) job loss for displaced vs. non-displaced workers within matched worker pairs. Panel A reports coefficients when controlling for baseline characteristics only (age, age squared, years of education, tenure, experience, full-time work, log establishment size, 1-digit industries, 1-digit occupations (all in t=-1), and log wage (in t=-3)). Panel B reports coefficients when adding controls for local unemployment rate (UR) changes reported at the municipality (LAU) level, city residency, and the share of coethnic working age population in a county (NUTS 3), all measured in t=-1. The AKM effect is a proxy for wage differentials across firms, based on Abowd et al. [1999]). The regression sample includes displaced workers, only. ** and * refer to statistical significance at the 1 and 5 percent level for standard errors clustered at the baseline county level. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. Source: IEB, BBSR, Destatis.

Table 2.15: Explaining Costs of Job Loss by Local Labor Market Concentration - Only Complete Closures

	(1) Δ Log (Earnings)	(2) Δ Employed	(3) Δ Days Worked	(4) Δ Log Wage	(5) Δ Commutes	(6) Δ AKM Effect	(7) Δ Share Migrants	(8) Δ Share Marg. Employed
Panel A: Controlling for Individual Characteristics, Industry, and Occupation								
Migrant	-0.18**	-0.044**	-19.3**	-0.11**	-0.0053	-0.027	-0.0038	0.029**
	(0.024)	(0.0067)	(3.10)	(0.015)	(0.013)	(0.015)	(0.0066)	(0.0053)
Observations	40851	42824	42824	39568	39252	26903	39135	38760
R^2	0.051	0.025	0.037	0.046	0.018	0.175	0.010	0.029
Mean of dep. var	-0.36	-0.092	-56.1	-0.20	0.038	-0.11	-0.011	0.038
Panel B: Adding Controls for Local Unemployment Rate Change, City Residency and Share of Coethnic Neighbors								
Migrant	-0.35*	-0.011	-44.4*	-0.18	-0.031	-0.0095	0.19**	0.023
	(0.15)	(0.041)	(18.1)	(0.12)	(0.14)	(0.12)	(0.039)	(0.024)
Local UR Change	-0.19**	-0.043*	-27.0**	0.0015	0.048	-0.11	0.025	0.011
	(0.066)	(0.021)	(9.86)	(0.048)	(0.049)	(0.091)	(0.022)	(0.014)
Migrant*UR Change	-0.13	-0.033	-10.6	-0.076	0.030	-0.0049	-0.0017	-0.026
	(0.25)	(0.048)	(26.5)	(0.14)	(0.093)	(0.10)	(0.067)	(0.043)
City Residency	-0.059**	-0.022**	-9.72**	-0.030**	0.053**	0.0026	-0.0034	0.0028
	(0.012)	(0.0035)	(1.58)	(0.0084)	(0.016)	(0.0094)	(0.0025)	(0.0023)
Migrant*City Residency	-0.059	0.0029	2.67	-0.096**	-0.012	-0.045*	-0.0042	0.018
	(0.041)	(0.011)	(4.54)	(0.028)	(0.031)	(0.018)	(0.0100)	(0.010)
Share Same Nationality	-0.31*	0.012	-37.1	-0.19	-0.019	-0.040	0.21**	0.016
	(0.15)	(0.042)	(19.1)	(0.12)	(0.15)	(0.13)	(0.041)	(0.026)
Migrant*Share Same Nationality	-4.02**	-1.18**	-529.8**	-2.15*	0.51	-1.63	0.10	0.57*
	(1.49)	(0.35)	(146.5)	(1.07)	(0.47)	(1.30)	(0.32)	(0.25)
Observations	39365	41257	41257	38120	37807	25812	37710	37344
R^2	0.054	0.027	0.039	0.048	0.020	0.185	0.021	0.031
Mean of dep. var	-0.36	-0.092	-56.1	-0.20	0.038	-0.11	-0.011	0.038

Notes: This table shows to what extent local labor market conditions contribute to migrant-native gaps in labor market outcomes after job displacement. Regression sample restricted to workers displaced from a complete establishment closure. All outcome variables are based on individual difference-in-differences estimates which measure the change in the outcome (e.g. log dailys wages) before (t=-5 to t=-2) vs. after (t=0 to t=3) job loss for displaced vs. non-displaced workers within matched worker pairs. Panel A reports coefficients when controlling for baseline characteristics only (age, age squared, years of education, tenure, experience, full-time work, log establishment size, 1-digit industries, 1-digit occupations (all in t=-1), and log wage (in t=-3)). Panel B reports coefficients when adding controls for local unemployment rate (UR) changes reported at the municipality (LAU) level, city residency, and the share of coethnic working age population in a county (NUTS 3), all measured in t=-1. The AKM effect is a proxy for wage differentials across firms, based on Abowd et al. [1999]. The regression sample includes displaced workers, only. ** and * refer to statistical significance at the 1 and 5 percent level for standard errors clustered at the baseline county level. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. Source: IEB, BBSR, Destatis.

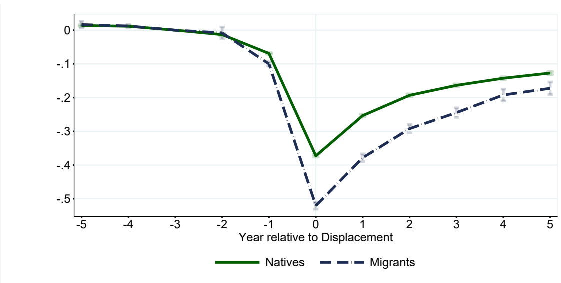
Table 2.16: Overview of Origin Groups as in Battisti et al. (2021)

	(1) Group name	(2) Countries	
1	Germany	Germany	
2	Western incl. Western European Countries	Australia	New Zealand
		Austria	Norway
		Canada	Portugal
		Denmark	Samoa
		Finland	Spain
		France	Sweden
		Greece	Switzerland
		Italy	United Kingdom
		Ireland	USA
		Netherlands	
3	Eastern Europe	Czech Republic	Slovakia
		Hungary	Slovenia
		Poland	
4	South-Eastern Europe	Albania	Former Jugoslavia
		Bosnia and Herzegovina	Northmazedonia
		Bulgaria	Mazedonia
		Kosovo	Romania
		Croatia	Serbia
5	Turkey	Turkey	
6	Former USSR	Armenia	Lithuania
		Azerbaijan	Moldova
		Belarus	Russian Federation
		Estonia	Tajikistan
		Georgia	Turkmenistan
		Kazakhstan	Ukraine
		Kyrgyzstan	Uzbekistan
		Latvia	
7	Asia and Middle East		
8	Africa		
9	Central and South America		
10	Other		

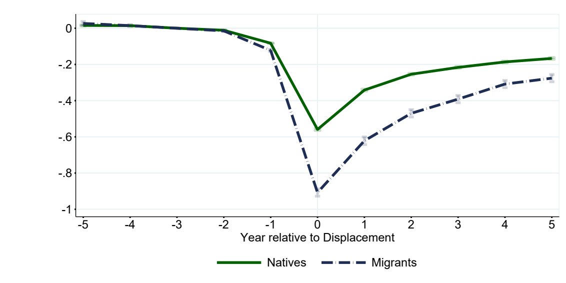
Notes: This table shows how we assign migrants to origin groups following Battisti et al. [2021]. We use these origin groups for our heterogeneity analysis in figure 2.5 and figure 2.12. The category "Other" contains origin countries which rarely appear in our data. These include, amongst other islands, the Fiji Islands, the Marshall Islands, and Andorra.

Figure 2.7: 4 Measures of Earnings by Migration Status

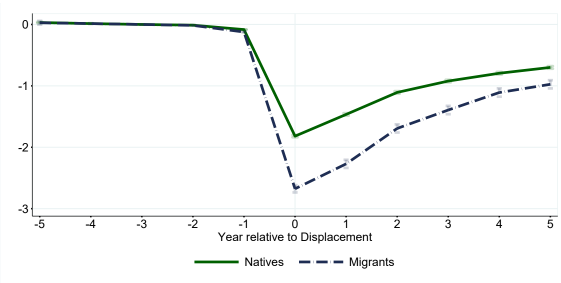
(A) Yearly Earnings Rel. To t=-2



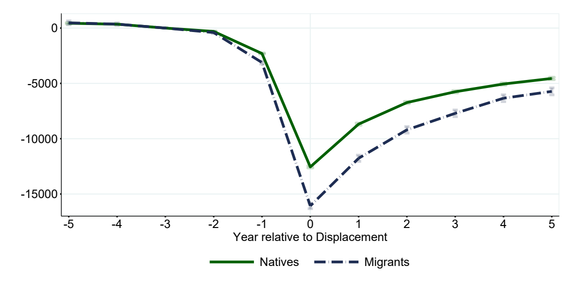
(B) Yearly Log Earnings



(C) Yearly Log(Earnings+1)



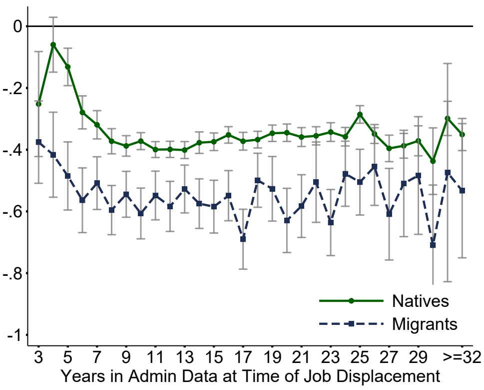
(D) Total Yearly Earnings in EUR



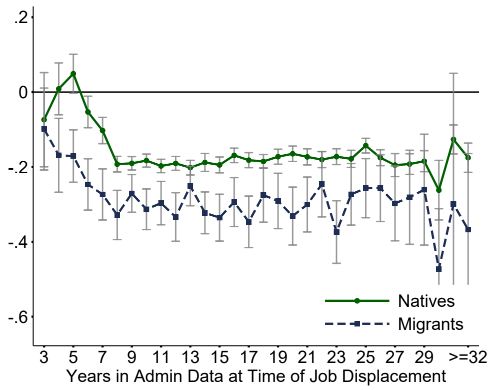
Notes: This figure plots event study regression coefficients on the differential earnings evolution for displaced vs. nondisplaced workers for 4 different outcomes: earnings relative to t=-2 (Panel A), yearly log earnings (Panel B), yearly log(earnings+1) (Panel C), and yearly earnings in EUR (Panel D). The solid green line reports the results for our sample of native workers, and the dashed blue line reports the results for our sample of migrant workers. Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the individual level. Regressions control for year fixed effects, year since displacement fixed effects, age polynomials, and worker fixed effects. We omit t=-3 as the reference category. Displaced workers are matched to nondisplaced workers using propensity score matching. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. Source: IEB.

Figure 2.8: Costs of Job Loss by Years in Administrative Data

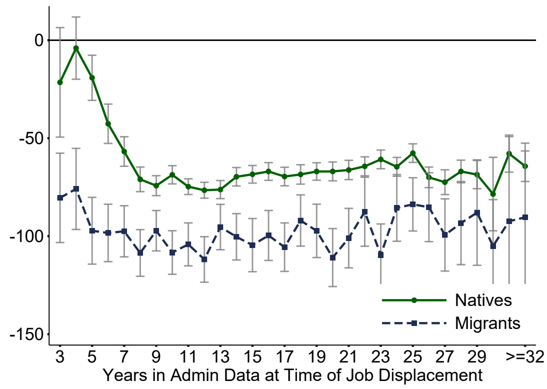
(A) Log Yearly Earnings



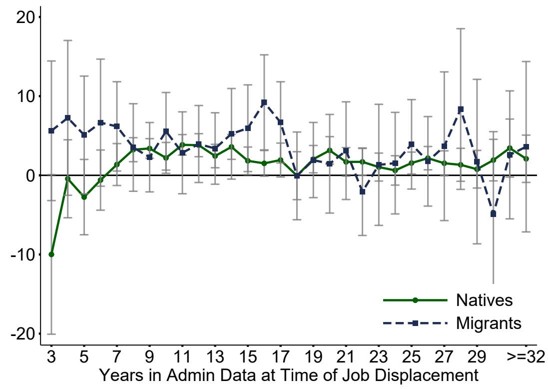
(B) Log Daily Wage



(C) Days Worked in Full-time Employment



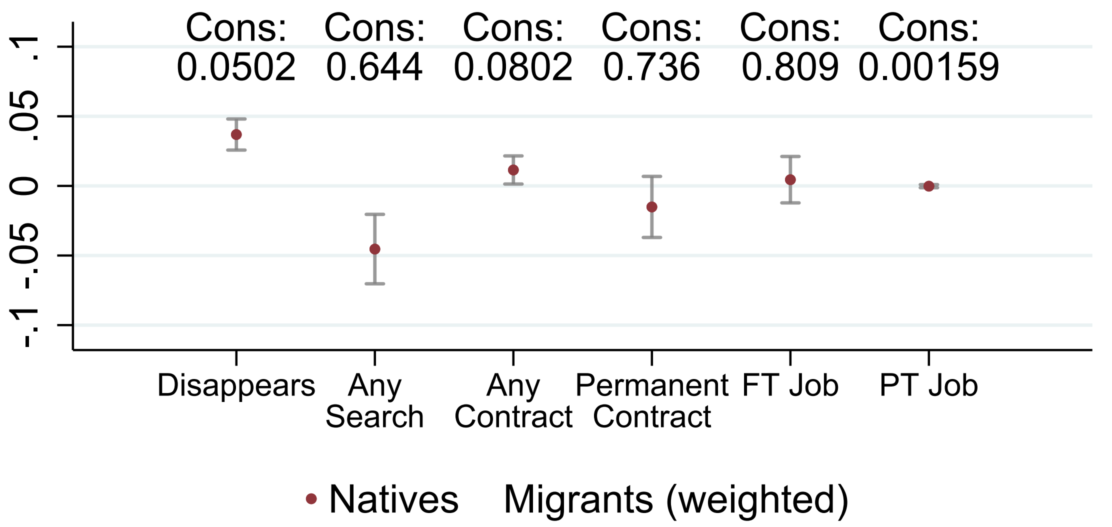
(D) Days Worked in Part-time Employment



Notes: This figure shows how costs of job loss differ by the time (in years) a workers has been registered in the German administrative data at t=-1. Panel A reports log(earnings) per year, Panel B reports log(wage), Panel C reports the number of full-time days worked per year, and Panel D reports the number of part-time days worked per year. We regress workers' individual difference-in-differences outcomes on dummies for years in admin data (x-axis), as well as individual, industry, and occupation controls. The solid green line reports the results for our sample of native workers, and the dashed blue line reports the results for our sample of migrant workers. Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the displacement establishment level. Our regression controls for individual characteristics (age, age squared, years of education, tenure, experience, full-time work, log wage in t=-3, and log firm size), 1-digit industries, and occupations according to Blossfeld [1987] in the year before displacement. Source: IEB and Destatis.

Figure 2.9: Differences in Job Search Behavior

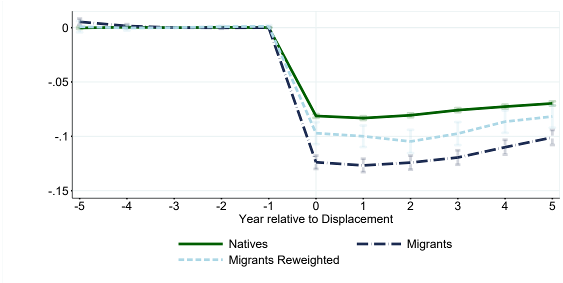
(A) Different Search Outcomes



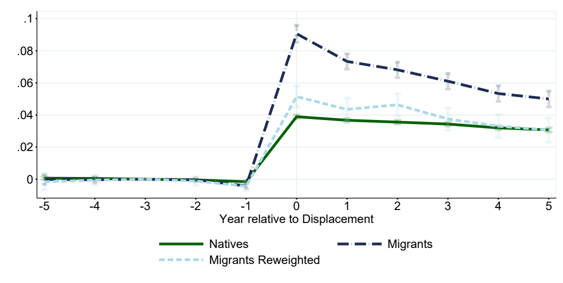
Notes: This figure shows to what extent job search behavior differs between migrants and natives. Each red dot reports the coefficient on the migrant dummy for different outcomes (listed on the x-axis) in a regression where we reweight migrants to natives and cluster standard errors at the displacement establishment level. The sample is based on our main regression sample but consists of displaced workers, only. The constant reports the mean value of a given outcome for displaced natives. The first outcome, "disappears", is a dummy which equals 1 if a worker drops out of the social-security records and does not return up to year t=5 after displacement. The second outcome, "any search", is a dummy which equals 1 if a worker ever has a job seeker spell (ASU), conditional on not dropping out after displacement. Outcomes 3-6 are conditional on ever having a job seeker spell, and measured at the time of layoff: The third outcome, "any contract", is a dummy equal to 1 if workers report that they are searching for any employment contract. The fourth outcome, "permanent contract", is a dummy equal to 1 if workers report that they are searching for permanent contracts only. Finally "FT Job" equals 1 if workers are searching for full-time jobs, whereas "PT Job" equals 1 if workers are searching for part-time jobs. Source: IEB.

Figure 2.10: Types of Establishments by Migration Status

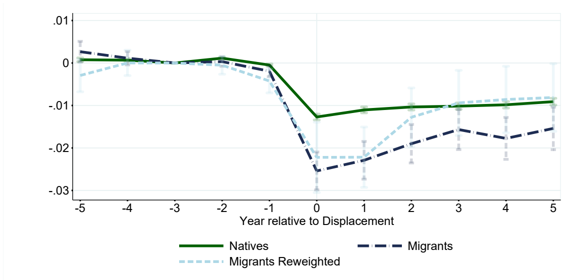
(A) AKM Establishment Effect



(B) Share of Marginally Employed Workers



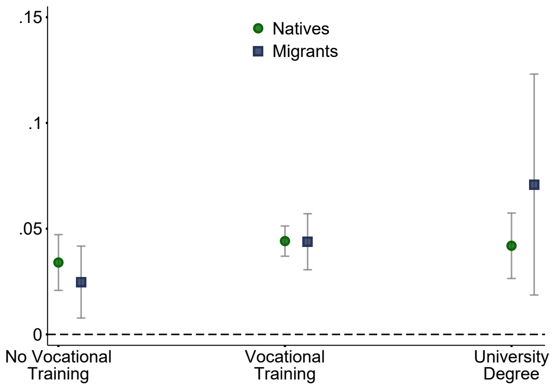
(C) Share of Migrant Workers



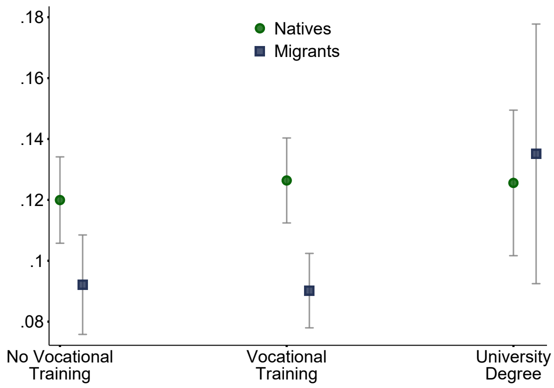
Notes: This figure plots event study regression coefficients on the differential evolution of the following outcomes for displaced vs. nondisplaced workers: AKM-style establishment fixed effects (Panel A) (the AKM effect is a proxy for wage differentials across firms, based on Abowd et al. [1999]), the share of marginally employed workers in an establishment (Panel B), and the share of migrant workers in an establishment (Panel C, leave-one-out mean). The solid green line reports the results for our sample of native workers, the dashed blue line reports the results for our sample of migrant workers, and the light blue line reports the results for our sample of reweighted migrant workers. Reweighting characteristics are log wage (t=-3, t=-4), age (t=-1), years of education (t=-1), tenure (t=-1), city residency (t=-1), establishment size (t=-1), 1-digit industry (t=-1), and 1-digit occupation (t=-1). Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the individual level. Regressions control for year fixed effects, year since displacement fixed effects, age polynomials, and worker fixed effects. We omit t=-3 as the reference category. Displaced workers are matched to nondisplaced workers using propensity score matching. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. Source: IEB, BHP.

Figure 2.11: Additional Outcomes by Education Level and Migration Status

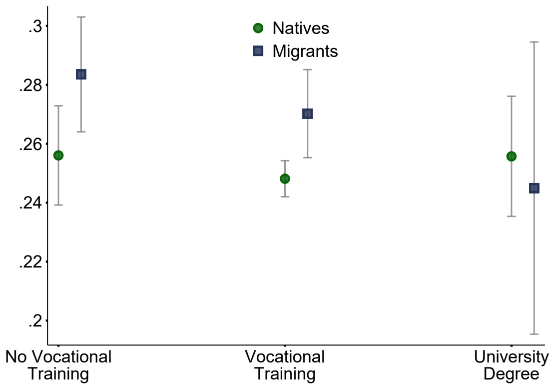
(A) Δ Commuting



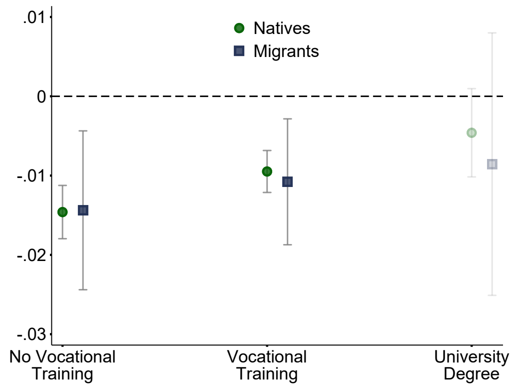
(B) Δ Relocating to Different State



(C) Δ Changed 1-Digit Occupation

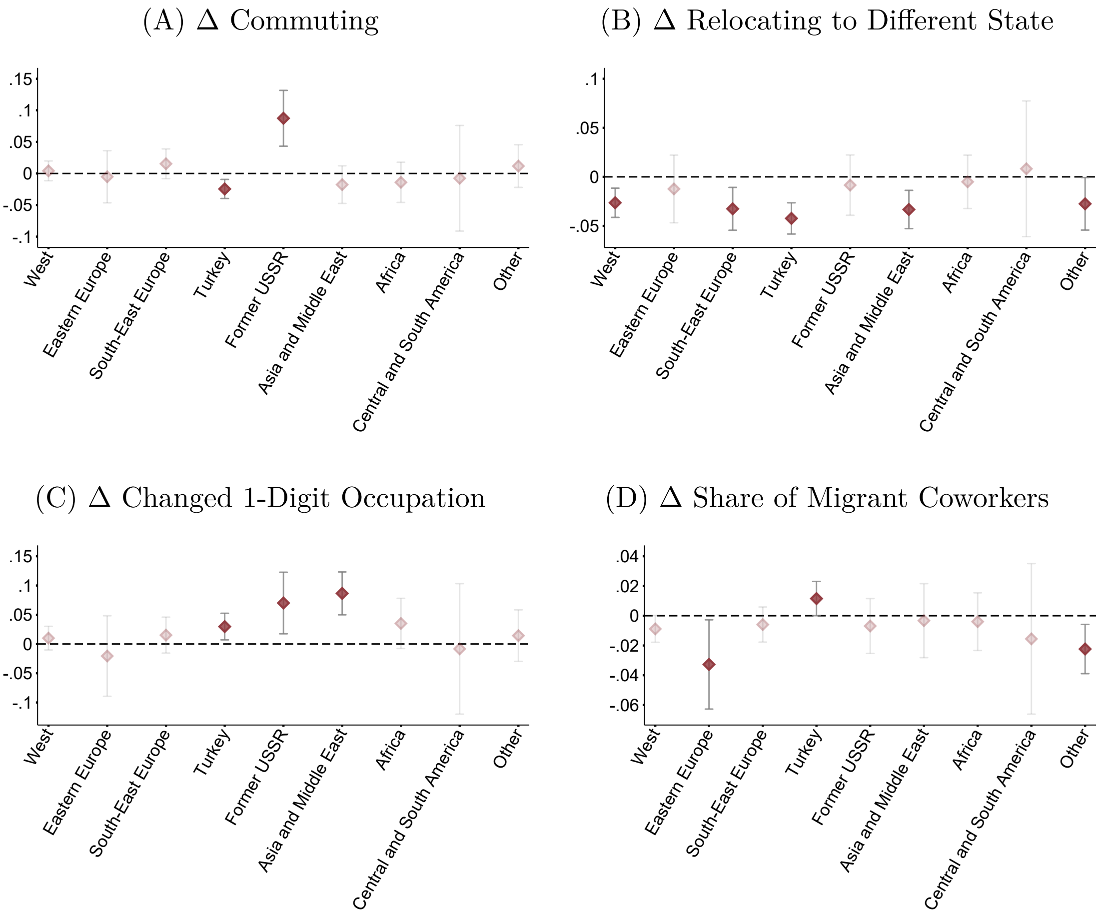


(D) Δ Share of Migrant Coworkers



Notes: This figure shows how costs of job displacement differ by education group and migration status. Each panel plots coefficients from a separate OLS regression where we regress workers' individual difference-in-differences outcomes on dummies for the 9 origin groups, with "German origin" as omitted category. We define commuting as working and living in different municipalities (LAU). "Relocating to different state" (NUTS-1) is defined relative to the baseline year (t=-1). Each difference-in-differences outcome measures differences in the outcome before (t=-5 to t=-2) vs. after (t=0 to t=3) job loss for displaced vs. nondisplaced workers. Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the displacement establishment level. All regressions control for individual characteristics (age, age squared, years of education, tenure, experience, full-time work, log wage in t=-3, and log establishment size), 1-digit industries, and occupations according to Blossfeld [1987] in the year before displacement. Source: IEB.

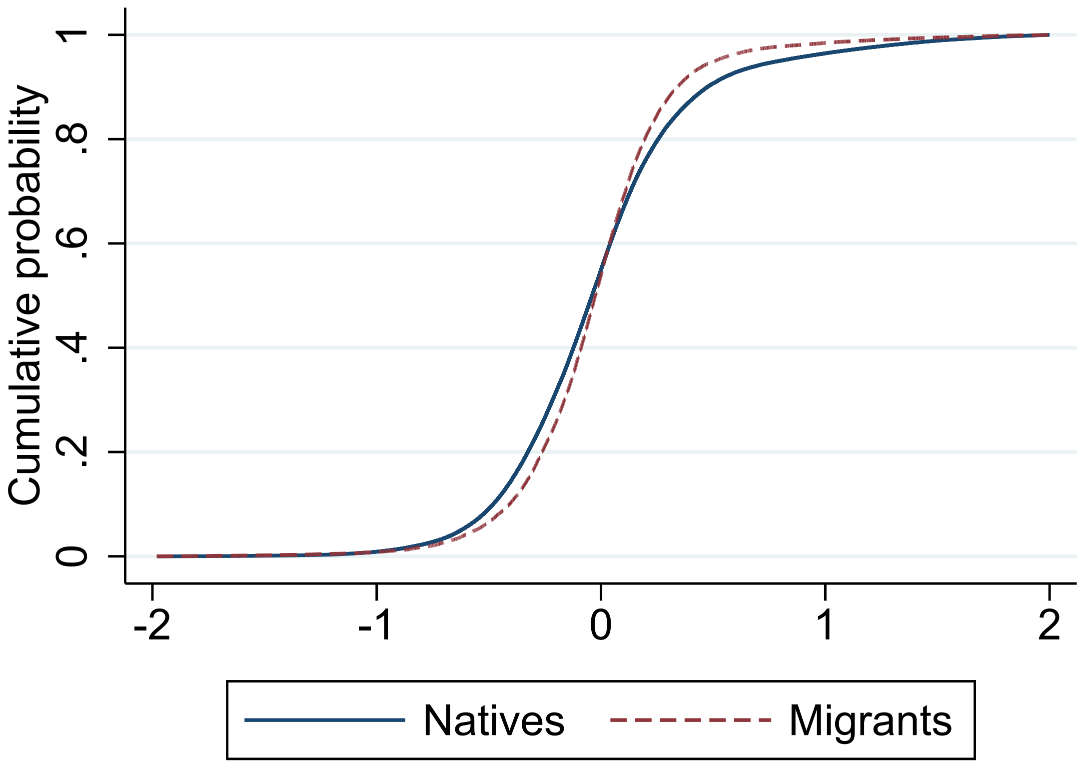
Figure 2.12: Additional Outcomes by Origin Group



Notes: This figure shows how costs of job displacement differ by origin group. Each panel plots coefficients from a separate OLS regression where we regress workers' individual difference-in-differences outcomes on dummies for the 9 origin groups, with "German origin" as omitted category. We define commuting as working and living in different municipalities (LAU). "Relocating to different state" (NUTS-1) is defined relative to the baseline year (t=-1). Each difference-in-differences outcome measures differences in the outcome before (t=-5 to t=-2) vs. after (t=0 to t=3) job loss for displaced vs. nondisplaced workers. Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the displacement establishment level. All regressions control for individual characteristics (age, age squared, years of education, tenure, experience, full-time work, log wage in t=-3, and log establishment size), 1-digit industries, and occupations according to Blossfeld [1987] in the year before displacement. Source: IEB.

Figure 2.13: Distribution Function of the Wage Residuals from a Mincerian Wage Regression by Migration Status

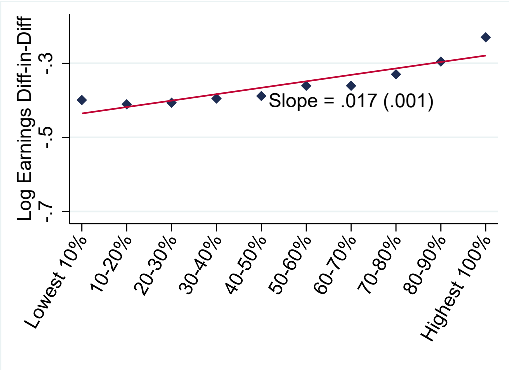
(A) [Distribution Function of Wage Residuals by Migration Status



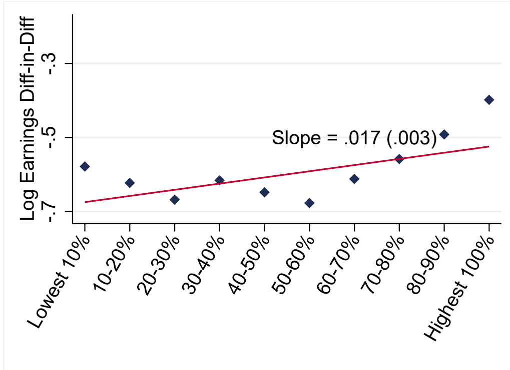
Notes: This figure shows the distribution function of the wage residuals from a Mincerian wage regression for migrants (dashed red line) and natives (solid blue line). Our regression sample includes displaced workers in the baseline year (t=-1). We regress wages on a number of observable characteristics: 7 education group dummiesThese include the following: No training, middle school (Volks-, Haupt-, Realschule) without vocational training, middle school with vocational training, secondary school (Abitur) without vocational trianing, secondary school with vocational training, university of applied sciences degree, university degree, potential experience, experience squared, tenure, a dummy for full-time employment on June 30, age dummies, year dummies, and log firmsize. The outcome variable is log wages. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. In $t = -1$, we observe 17,605 displaced migrants and 129,701 displaced natives. Source: IEB.

Figure 2.14: Binscatter Plots: Wage Residuals vs. Costs of Job Displacement

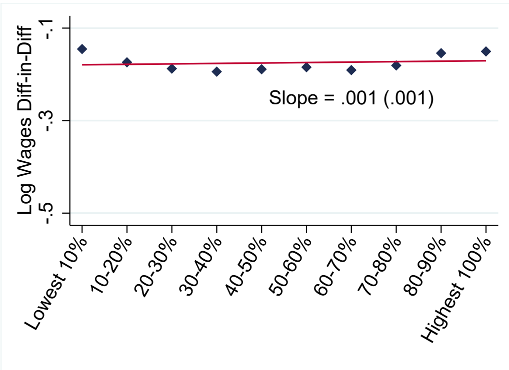
(A) Δ Log Earnings vs. Wage Residuals - Natives



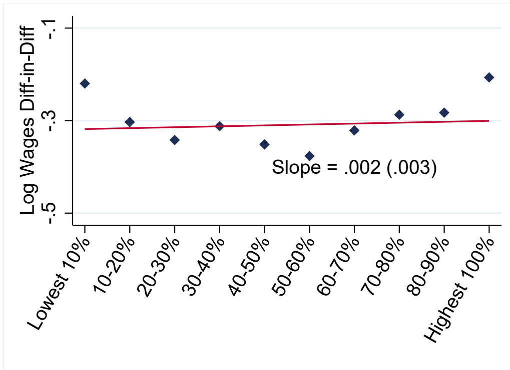
(B) Δ Log Earnings vs. Wage Residuals - Migrants



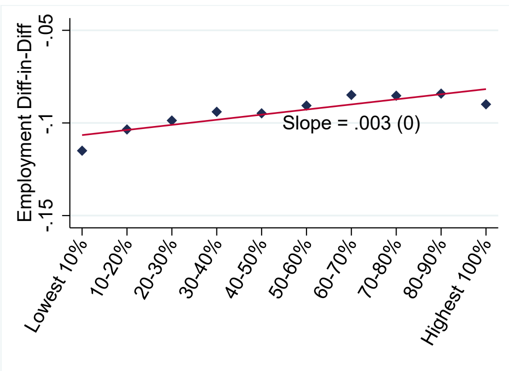
(C) Δ Log Wages vs. Wage Residuals - Natives



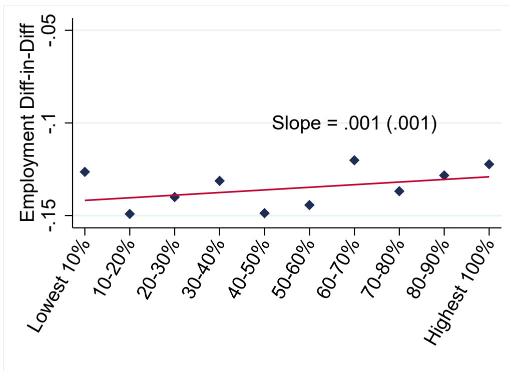
(D) Δ Log Wages vs. Wage Residuals - Migrants



(E) Δ Employment vs. Wage Residuals - Natives



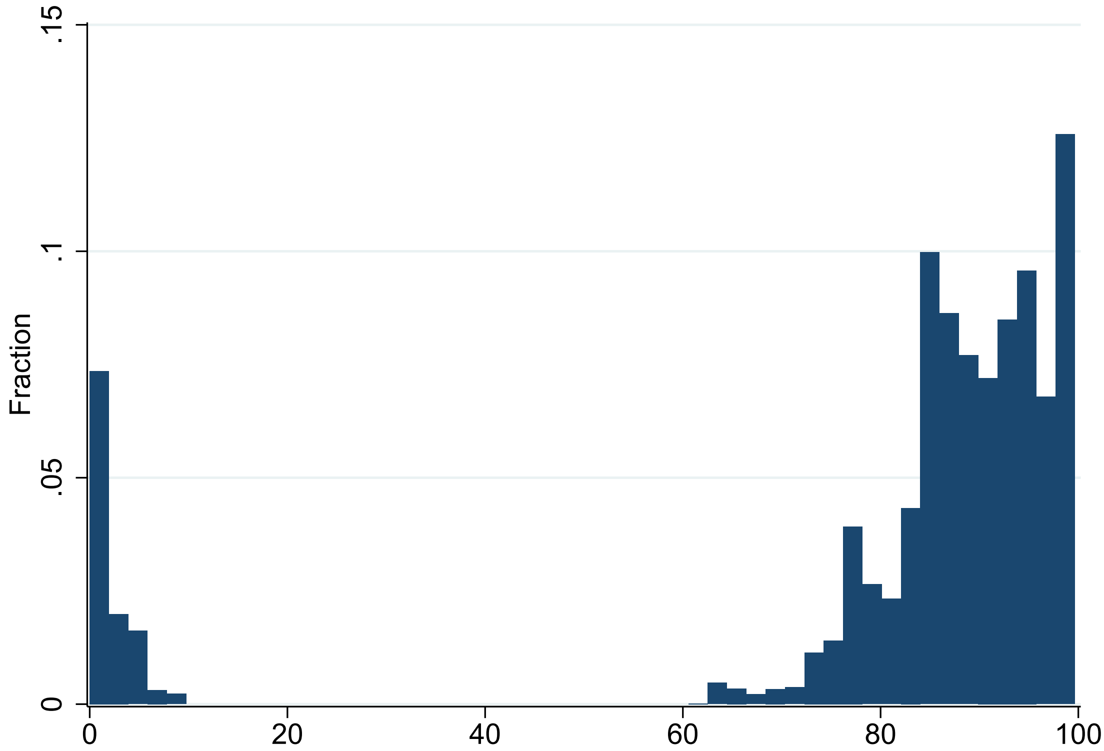
(F) Δ Employment vs. Wage Residuals - Migrants



Notes: This figure plots deciles of wage residuals from a Mincerian regression (x-axis) against workers' individual difference-in-differences outcomes (y-axis) (as in Borjas et al. [2019]). The difference-in-difference outcomes are further described in Equation (3). We compute wage residuals from a regression of log wages on 7 education group dummies, potential experience, experience squared, tenure, a dummy for full-time employment on June 30, age dummies, year dummies, and log firmsize. The regression includes both migrant and native displaced workers in the baseline year (t=-1). Source: IEB.

Figure 2.15: Distribution of the Share of Same-Nationality Working Age Population

(A) Distribution Share Same-Nationality Working Age Population in County in t=-1



Notes: This figure shows the distribution of the share of same-nationality working age population in a county in $t = -1$ for our sample of displaced workers. For migrants, the share ranges from 0-10%; for natives, it ranges from 60-100%. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. In $t = -1$, we observe 17,605 displaced migrants and 129,701 displaced natives. Source: Destatis.

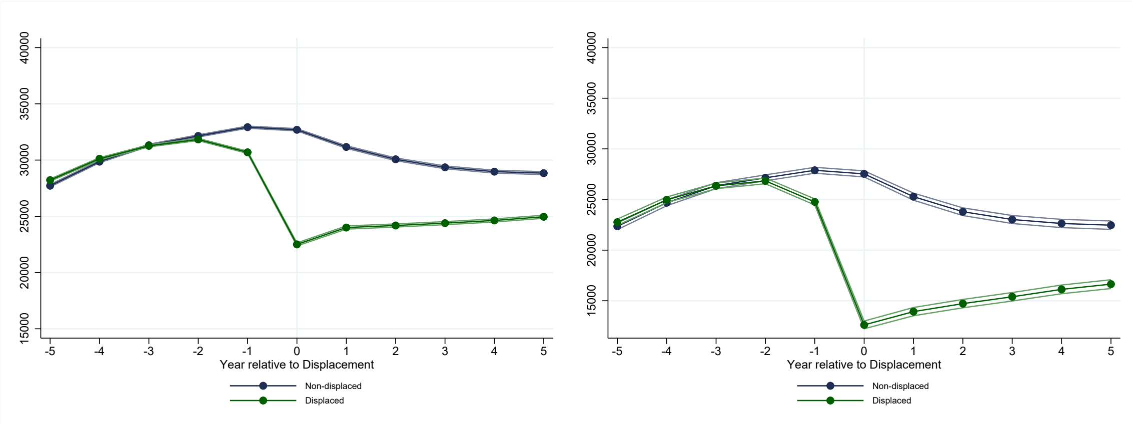
Appendix C: Main Results for Women

In Appendix C, we replicate the main results for a sample of women. Since men and women differ on a range of characteristics, we cannot directly compare the coefficients of the male vs. the female sample. What we can do, however, is to compare the migrant-native gap for men vs. women. We find that the raw earnings gap is somewhat larger for women: While migrant men have about 3000 EUR larger earnings losses than native men in the year after displacement, this gap is 5000 EUR for women (cf. Figure 2.17).

Decomposing the gap into wage and employment losses (Figure 2.18) shows that this is mainly because migrant women are substantially less likely to take up a new job after displacement. Without reweighting, migrant women are about 12 percentage points less likely to be re-employed in the year after displacement. If we reweight migrant women to native women, this gap shrinks to about 5 percentage points, but remains statistically significant. In particular, migrant women are substantially less likely than native women to work in full-time employment after job displacement.

Figure 2.16: Migrant and Native Workers' Earnings - No Controls - Women

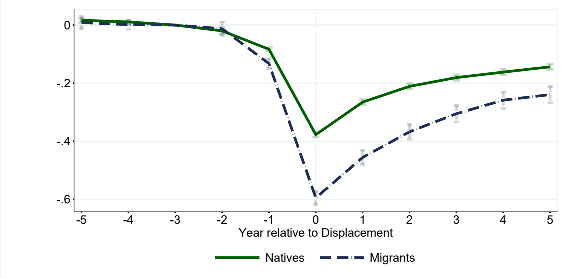
(A) Total Yearly Earnings in EUR - Natives (B) Total Yearly Earnings in EUR - Migrants



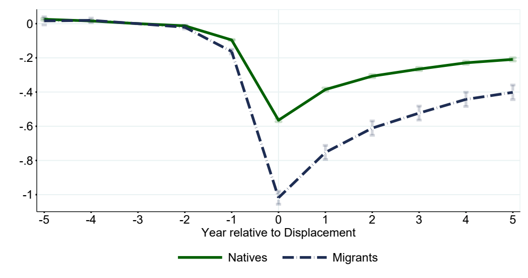
Notes: This figure plots raw earnings losses for displaced compared to nondisplaced workers and natives (Panel A) compared to migrants (Panel B). Sample of displaced women, only. The blue line shows earnings trajectories for nondisplaced workers, and the green line shows earnings trajectories for workers displaced between t=-1 and t=0. Displaced workers are matched to nondisplaced workers using propensity score matching. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. In t=-1, we observe 35,210 migrants and 259,402 natives. Source: IEB.

Figure 2.17: Earnings Rel. To t=-2, Log(Earnings), Log(Earnings+1), and Total Yearly Earnings by Migration Status - Women

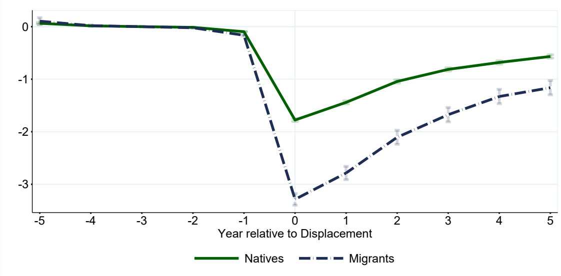
(A) Yearly Earnings Rel. To t=-2



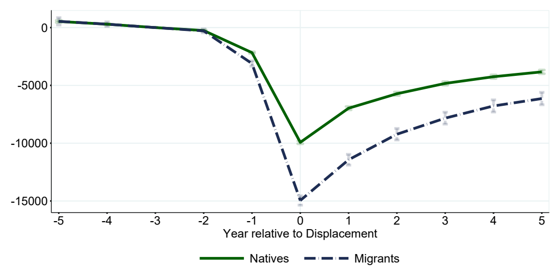
(B) Yearly Log(Earnings)



(C) Yearly Log(Earnings+1)



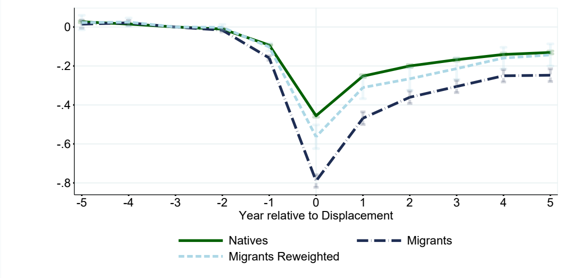
(D) Total Yearly Earnings in EUR



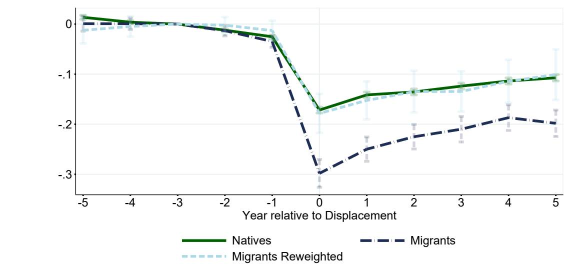
Notes: This figure shows losses in relative earnings (Panel A), yearly log(earnings) (Panel B), yearly log(earnings+1) (Panel C), and yearly earnings in EUR (Panel D) for displaced and nondisplaced workers. Sample of displaced women, only. The solid green line reports the results for our sample of native workers, and the dashed blue line reports the results for our sample of migrant workers. Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the individual level. Our regression controls for year fixed effects, year since displacement fixed effects, age polynomials, and worker fixed effects. We omit t=-3 as the reference category. Displaced workers are matched to nondisplaced workers using propensity score matching. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. Source: IEB.

Figure 2.18: Labor Market Outcomes by Migration Status- Women

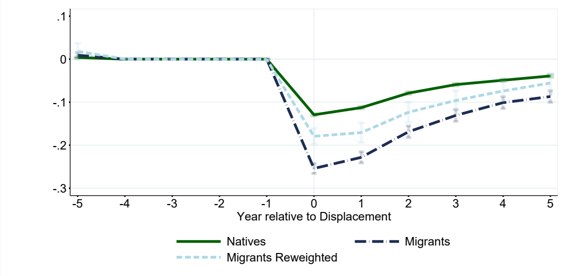
(A) Yearly Log(Earnings)



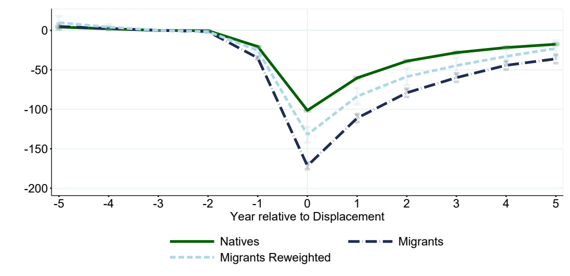
(B) Log Wage



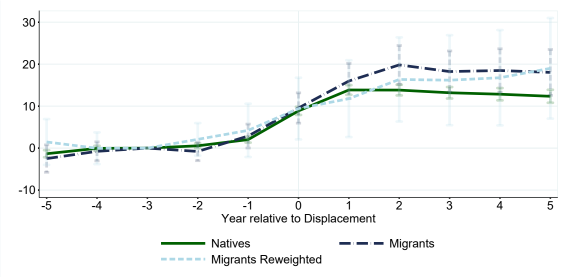
(C) Employment



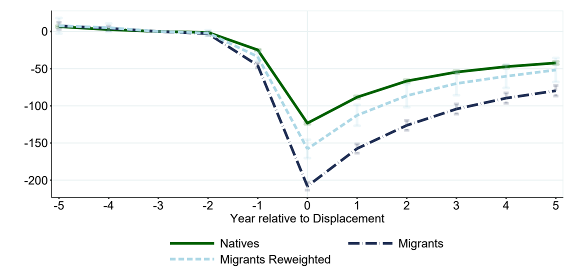
(D) Days Worked per Year



(E) Days Worked in Part-Time Employment



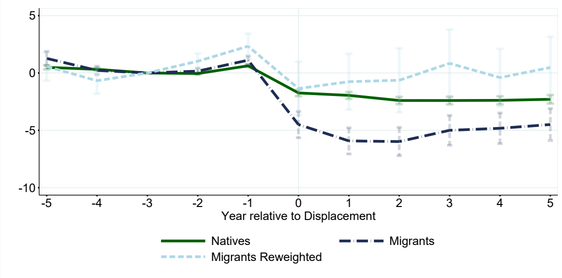
(F) Days Worked in Full-Time Employment



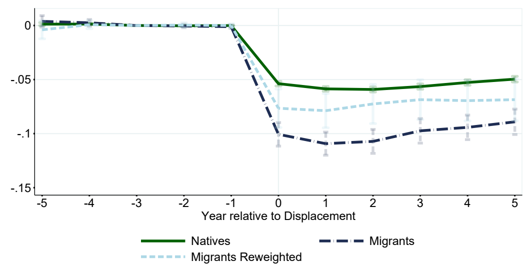
Notes: This figure shows losses in log (earnings+1) (Panel A), log wages (Panel B), employment probability (Panel C), yearly days worked (Panel D), days worked in part-time employment (Panel E), and days worked in full-time employment (Panel F) for displaced and nondisplaced workers. Sample of displaced women, only. The solid green line reports the results for our sample of native workers, the dashed blue line reports the results for our sample of migrant workers, and the light blue line reports the results for our sample of reweighted migrant workers. Reweighting characteristics are log wage (t=-3, t=-4), age (t=-1), years of education (t=-1), tenure (t=-1), city resident (t=-1), establishment size (t=-1), 1-digit industry (t=-1), and 1-digit occupation ($t = -1$). Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the individual level. Our regression controls for year fixed effects, year since displacement fixed effects, age polynomials, and worker fixed effects. We omit t=-3 as the reference category. Displaced workers are matched to nondisplaced workers using propensity score matching. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. Source: IEB.

Figure 2.19: Establishment Sorting after Displacement by Migration Status - Women

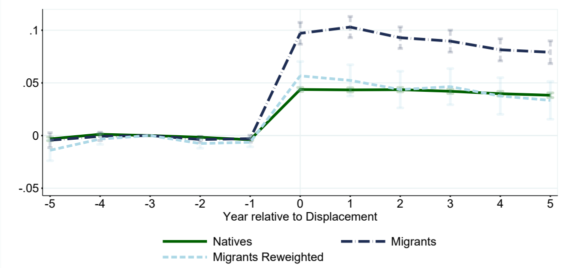
(A) Average Establishment Full-time Wage



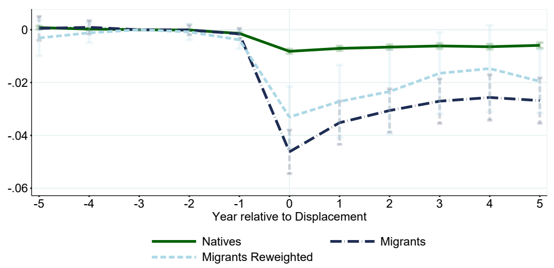
(B) AKM Establishment Effect



(C) Share of Marginally Employed Workers



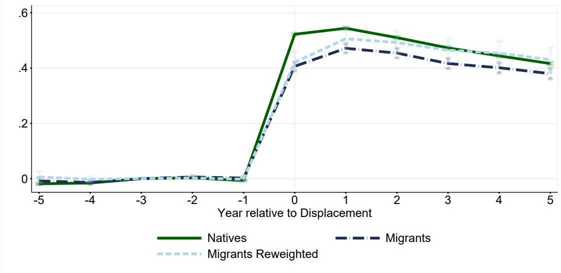
(D) Share of Migrant Workers



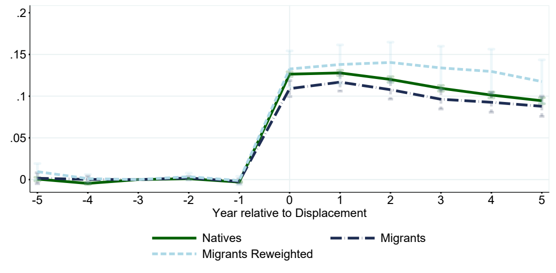
Notes: This figure shows average establishment full-time wages (Panel A), AKM-style establishment fixed effects (Panel B) (the AKM effect is a proxy for wage differentials across firms, based on Abowd et al. [1999]), the share of marginally employed workers in an establishment (Panel C), and the share of migrant workers in an establishment (Panel D, leave-one-out mean) for displaced and nondisplaced workers. Sample of displaced women, only. The solid green line reports the results for our sample of native workers, the dashed blue line reports the results for our sample of migrant workers, and the light blue line reports the results for our sample of reweighted migrant workers. Reweighting characteristics are log wage (t=-3, t=-4), age (t=-1), years of education (t=-1), tenure (t=-1), city resident (t=-1), establishment size (t=-1), 1-digit industry (t=-1), and 1-digit occupation (t=-1). Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the individual level. Our regression controls for year fixed effects, year since displacement fixed effects, age polynomials, and worker fixed effects. We omit t=-3 as the reference category. Displaced workers are matched to nondisplaced workers using propensity score matching. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. Source: IEB, BHP.

Figure 2.20: Geographic Mobility by Migration Status - Women

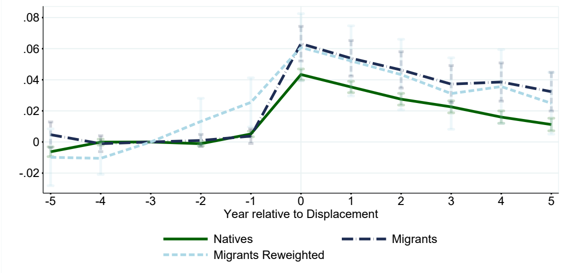
(A) Changed Workplace Municipality since t=-1



(B) Changed Workplace State since t=-1



(C) Commutes



Notes: This figure shows the propensity to change one's workplace to a different municipality from t=-1 (Panel A), the propensity to change one's workplace to a different federal state (Panel B), and the propensity to commute (Panel C). Sample of displaced women, only. The propensity to commute is defined as working and living in different municipalities. The solid green line reports the results for our sample of native workers, the dashed blue line reports the results for our sample of migrant workers, and the light blue line reports the results for our sample of reweighted migrant workers. Reweighting characteristics are log wage (t=-3, t=-4), age (t=-1), years of education (t=-1), tenure (t=-1), city resident (t=-1), establishment size (t=-1), 1-digit industry (t=-1), and 1-digit occupation (t=-1). Vertical bars indicate the estimated 95% confidence interval based on standard errors clustered at the individual level. Our regression controls for year fixed effects, year since displacement fixed effects, age polynomials, and worker fixed effects. We omit t=-3 as the reference category. Displaced workers are matched to nondisplaced workers using propensity score matching. Workers in our sample are displaced in the period 2001-2011, and they are observed from 1996 to 2017. Source: IEB.

Chapter 3

Coworkers, Neighbors, and Job-Search: Migrants' and Natives' Re-employment after a Layoff.

3.1 Introduction

Social networks play an important role in people's lives. Granovetter [1977] shows in his pioneering work that workers use personal networks for job searching. Finding a job after a layoff, therefore, depends not only on individual characteristics and the vacancies of establishments but also on social networks that influence job search behavior or transfer information about vacancies. This raises the question of how these networks influence individual labor market outcomes. While there has been substantial theoretical and empirical work conducted regarding social networks (see, e.g., Ioannides and Loury [2004], Jackson [2010], Topa and Zenou [2015], Glitz [2017]), we know little about how *different* kinds of networks affect labor market outcomes. Since individuals may belong to many different networks simultaneously, this important question is difficult to answer.

Why should different kinds of networks affect workers differently in their job search behavior? While, e.g., coworker networks include work-related contacts, neighbor networks are private contacts. Former coworkers, on the one hand, can assess the skills of the worker well and are more aware of suitable vacancies, either because they look for a job themselves or because they work for a company that searches for employees. Neighbors, on the other hand, compete less for jobs than do former coworkers who have worked in the same industry, have selected the same establishment before and have information about regional job offers. A priori, how a person's diverse networks affect their job search outcomes is unclear. Thus far, the main problem in dealing with the topic of diverse networks has been that the different network records of people were unavailable. However, looking at how diverse kinds of networks affect workers differently in their job search is important because people are always members of different and distinct social networks at the same time, and understanding these differences is crucial.

In this paper, I use administrative employment and newly available geo-coded residence records from Germany to estimate the effect of different networks on a worker's re-employment probability. The data set comprises the universe of workers in four large metropolitan areas (Cologne, Frankfurt, Hamburg, Munich)[1] as defined by Kropp and Schwengler [2016] from 1995-2015. For all male workers who were dis-

[1]Berlin is excluded as a special case due to reunification. Table 3.7 in the robustness section shows that the results are robust when including Berlin. Metropolitan areas include not only the cities themselves but also the surrounding areas within commuting distance.

placed as part of an establishment closure between 2005-2011, I compute two network types: coworker and neighbor networks. Focusing on workers who lost their job involuntarily due to an establishment closure guarantees that the workers in the sample are highly comparable and lost their job irrespective of their networks[2] but have different networks. The coworker network covers all persons with whom a person worked in the same establishment in the 5 years prior to layoff. I construct neighbor networks by geo-referencing the worker's places of residence within grids of 3 square kilometers in size; all workers who lived within this 3-times-3-kilometer square in the past five years, I then define as neighbors[3].

To investigate the effect of the employment share within a network on a worker's re-employment probability, in addition to using a rich set of control variables and fixed effects[4], I use an instrumental variable approach as first proposed by Glitz [2017]. The share of network members who also experienced a mass layoff within the past 5 years is the instrument for the network's employment share. The intuition is that, conditional on observable characteristics such as industry or county, the extent to which a mass layoff happens within a worker's network is independent of factors that determine a worker's re-employment probability.

The results section has three parts. Starting with descriptive analyses, I show that both network types positively affect re-employment probability. Second, I use IV estimation to test whether workers with a higher employment share in their networks have a higher probability of being employed one year after being laid off for

[2]Including all unemployed workers could lead to biased results if their job change or loss was correlated with their networks. Focusing on workers who lost their job due to a plant closure ensures that workers were displaced irrespective of their abilities and networks since all workers of that establishment were displaced. However, since these workers have different employment biographies and live in different locations, their networks are different; the empirical strategy uses this variation for identification.

[3]If defining a neighborhood as grids of 1 square kilometer in size, mass layoffs occur too rarely in the network to obtain a strong first stage and to conduct an IV estimation. However, the OLS estimation results show a positive correlation and provide suggestive evidence that close local networks also play an important role.

[4]I use different fixed effects for both types of networks, namely, displacement establishment fixed effects for the coworker networks and county fixed effects for the neighbor networks. In the case of coworker networks, this guarantees that I only compare workers who were displaced from the same establishment but had different network characteristics. Similarly, I only compare workers who were displaced within one county but had different neighbor networks. This is in line with Helm et al. [2021], who show the effect of mass layoffs for both displaced and nondisplaced workers within the same county where a displacement occurs. Thus, mass layoffs do not only affect displaced workers; there could also be varying differences between different counties that I control for by the fixed effects.

both network types. Coworker networks are more important for re-employment; for coworker networks, a 10 percentage point increase in the employment rate increases the probability of being employed after one year by 4.9 percentage points, while for neighbor networks, the employment probability rises by 0.7 percentage points.

Next, I proceed with heterogeneity analyses regarding the network members' and the displaced workers' characteristics. First, I test whether the displaced persons' own and their network members' migration status play a role. While a 10 percentage point increase in the employment rate within the coworker network has a positive effect on natives (5.0 percentage points), I find no significant effect on migrants. When restricting the coworker network to only migrants, however, I find increases in the probability of being employed after one year for migrants (5.2 percentage points) and natives (4.7 percentage points). This finding shows that for migrants, migrant coworker networks are especially important. For natives, migrants in their network are important but less important than they are for migrants. Regarding neighbor networks, the results reveal a positive impact of neighbors, irrespective of their migration status, for natives. For migrants, however, native neighbors positively affect their re-employment probability, while there is no significant effect of migrant neighbors. This is in line with the literature that highlights a competition effect between similar migrants (Beaman [2012], Albert et al. [2021]).

Second, regarding heterogeneity by migration and education status, the results reveal that the lower one's education status is, the more important the networks are. Neighbor networks are more important for workers without a university degree, irrespective of their migration status. The theoretical literature argues that different networks are of different quality in transmitting potential job offers and assessing workers' skills; thus, more generalized network information on jobs may be more helpful for low- and medium-educated people than for highly educated people. This holds true for both coworker and neighbor networks.

This paper contributes to the literature on network effects by taking into account different types of networks. Empirical work generally confirms that social networks have a positive impact on workers' labor market outcomes. Many of these previous studies use survey data to measure how informal hiring methods affect workers' labor market outcomes (e.g., Caliendo et al. [2011], Brown et al. [2016], Ioannides and Loury [2004], Topa [2011]). In the absence of surveys, prior literature also uses alternative network definitions to proxy or directly measure social interaction between agents. These include neighbors (Topa [2001], Weinberg et al. [2004], Bayer et al. [2008], Hellerstein et al. [2011], Damm [2014], Schmutte [2015]), individuals with the same (ethnic) origin (Munshi [2003], Edin et al. [2003], Beaman [2012], Dustmann et al. [2016]), friends (Cappellari and Tatsiramos [2015]), family members (Kramarz and Skans [2014]), hallmates (Marmaros and Sacerdote [2002]), coworkers (Glitz [2017], Saygin et al. [2014]), fellow war veterans (Laschever [2009], Costa et al. [2018]), and virtual contacts (Barwick et al. [2019], Bailey et al. [2018]). Overall, there is evidence for a positive role of social networks on different workers' outcomes. My findings contribute to this literature by providing some first insights into the effects of different types of networks on workers with heterogeneous characteristics.

This study addresses various difficult issues in regard to identifying social network effects. First, prior literature has examined either coworker or neighbor networks because of data limitations, which is why a direct comparison of the two network types has not been possible thus far. This article is the first to overcome this issue by combining establishment and individual geo-referenced data. Second, by using newly available geo-coded data from Germany, instead of using districts or municipalities to proxy a local neighborhood, I define neighbors as those persons living within a grid of 3 square kilometers centered on the displaced worker, irrespective of administrative borders.[5] This definition is an improvement over definitions that use administrative units, as it also makes it easier to approximate neighborhoods on the edge of an administrative unit. Last, in addition to general effects, this study focuses specifically on the effects of and on migrants. This is interesting because (i) displacements hit migrants more severely than natives (Illing and Koch [2021]) in terms of earnings losses and employment, and (ii) migrants and natives have different networks due to their selection into specific establishments and neighborhoods. This paper investigates how networks differently affect natives' and migrants' employment status after a displacement. It also examines the impact of migrants who are part of a network on all workers' labor market outcomes. These heterogeneity analyses identify different effects of various networks on specific groups and address the differences.

The remainder of the paper is as follows. Section 3.2 introduces a conceptual framework. Section 3.3 introduces the data and states the definition of mass layoffs in the sample. Section 3.4 describes the empirical strategy. Section 3.5 presents

[5]For an illustration of the grids, please take a look at figure 3.1 in the appendix.

the results, including descriptive evidence, insights into the effects of coworker and neighbor networks, and heterogeneity analyses regarding worker and network characteristics. Section 3.6 presents different robustness checks. Section 3.7 concludes the paper.

3.2 Conceptual Framework

This work builds on the theoretical literature by Calvo-Armengol and Jackson [2004] and Wahba and Zenou [2005]. In this theory, agents hear about potential job opportunities and the respective wage offers. The arrival rate of job offers is an exogenous probability. If an agent is unemployed or the job offer is better than her current job, she will accept the job. If she is employed and the new job is less attractive than her current job, she passes the information on to one (or several) of her unemployed contacts. Consequently, agents in a larger network with a higher employment rate should be more likely to find a job and should receive higher wages.

However, how do coworker and neighbor networks differ from a theoretical perspective? First, the arrival rate of job offers might be different in the two types of networks. Since former coworkers are probably better at referring suitable jobs than are neighbors, the arrival rate of suitable jobs is likely higher in coworker networks. Second, the number of other unemployed network members qualified to accept the job offer is different in the two networks. In theory, an employed agent passes information to one (or more) unemployed contact. On the one hand, coworkers are more likely to pass on a job offer to several people in their network because the network members are highly substitutable, since they have all previously worked in the same industry and the same establishment. Neighbor networks, on the other hand, are likely to be more heterogeneous. Thus, the competition for jobs could be lower in neighbor networks. Finally, the size of the network may be different in the two network types. Depending on whether an agent lives in an urban or rural area and whether she previously worked for a small or a large establishment, the size of the network varies greatly. A higher number of vacancies arrives in large networks, but within these networks, the competition for vacancies is likely higher. In summary, theory does not clearly answer the question of whether coworker or neighbor networks are more important for finding a job.

How do migrants' and natives' networks differ from a theoretical point of view? First, the employment rate within a network could differ between migrants and natives and consequently, so could the arrival rate of job offers. For instance, migrants could live in "worse neighborhoods" due to (in-)voluntary ghettoization. If migrants tend to live in neighborhoods with a higher unemployment rate, they face a lower arrival rate of job offers and at the same time could face higher competition within their networks, if there are more equally qualified, unemployed to whom the employed can forward the job offers (see, e.g., Beaman [2012], Albert et al. [2021]). Second, migrants and natives could select a priori into different kinds of firms. If sorting differs between migrants and natives, their networks and employment rate within these networks would differ. If migrants work for larger establishments, e.g., their network is larger, then they could obtain more job offers. Consequently, theory cannot clearly predict for whom networks are more important, i.e., migrants or natives.

3.3 Data

3.3.1 Administrative Data from Germany

For the empirical analyses, I use administrative data provided by the Institute for Employment Research (IAB). The data set *Integrated Employment Biographies (IEB)* covers individual-level data of all labor market participants in Germany (except for civil servants and the self-employed) from 1975 onwards on a day-to-day basis. The IEB contains worker characteristics such as education level, age, nationality, and wage information. Further, establishment characteristics are available, such as industry and number of employees. The newly available *IEB Geo* data extend the *IEB* data with the addition of address information in the form of geo-codes of the individual's residence location from 2000-2014. The possibility of combining these different data sets is a core advantage of the data. This combination enables me to simultaneously identify all displaced workers' prior (i) neighbors, (ii) coworkers, and (iii) shares of displacement due to mass layoffs within the different networks. I construct a panel dataset as of June 30 each year. To increase the validity of the data, I further implement two imputation procedures. First, I correct implausible education entries following Fitzenberger et al. [2006]. Second, I impute wages that are censored at the contribution assessment ceiling following Gartner [2005] and Dustmann et al. [2009].

3.3.2 Displacement Events

For the identification strategy, I use different types of displacements, namely, total closures and mass layoffs (for the IV estimation). I follow Hethey-Maier and Schmieder [2013] to identify both types of mass layoffs since the reason for job losses is not available in the data. By using this identification technique, I ensure comparability with state-of-the-art literature from the US. To identify displaced workers, I focus on establishments with 5-50 full-time employees that closed down completely, as in Glitz [2017]. The benefit of the lower bound restriction is that very small establishments with little variation in network structure are excluded. The advantage of the upper bound is that the sample of displaced workers is still operational, while no excessively large networks are included. In contrast, for the IV estimation (share of displaced workers in the network), I use mass layoffs and only include establishments with more than 50 employees at the time of closure. The reason for this restriction is that larger mass layoffs are arguably more likely to be exogenous to displaced worker characteristics. An establishment is considered as having a mass layoff between June 30 in $t-1$ and June 30 in t if it is (i) either no longer included in the data and consequently closed down, (ii) or reduces its workforce by more than 30 percent within this period of time. All of these restrictions come at the cost of being not universally valid for all displaced employees. However, all of these restrictions are in line with prior literature; therefore, I ensure comparability with previous results.

Mergers, takeovers, spin-offs, and ID changes are problematic when identifying genuine mass layoffs in administrative data. Detecting these events is crucial, as their neglect leads to serious measurement error that biases the results to zero. In Illing and Koch [2021], we create a matrix of worker flows between establishments to identify these events, as in Hethey-Maier and Schmieder [2013]. According to the authors' definition, such an event occurs if more than 30 percent of the displaced workers move to the same successor.[6] Excluding these events guarantees that there is not simply a change in legal status, e.g., due to mergers and takeovers.

[6] Hethey-Maier and Schmieder [2013] provide additional information on employment outflows following a mass layoff in Germany.

3.3.3 Sample Restrictions

Next, I restrict the data to facilitate comparability with prior literature. In the first step, the sample only consists of displacements that took place between 2005 and 2011. Further, a worker in the sample is displaced (in $t-1$ or t) if the worker is both (i) no longer employed in the establishment in year $t-1$ or t and in any subsequent years ($t+1$, ..., $t+10$), and (ii) the establishment experiences a plant closure or mass layoff between June 30 in $t-1$ and t.

As in Glitz [2017] I define the network building phase as being 5 years prior to layoff. The network building phase therefore spans from 2000-2010, which is the earliest timeframe for which the geo-coded data are available. However, please note that I do not take into account the year of displacement for the network building phase. The reason for this is that coworkers who work together in the same establishment in the year of displacement are likely competitors on the labor market and thus should be unwilling to share job information with each other. For instance, if the worker is displaced in 2005, the network building phase thus ranges from 2000-2004. All workers with whom the displaced workers worked in this period are counted as network members.[7]

The sample further includes only workers who are (i) men, (ii) fulltime workers, (iii) between 25-50 years old, and (iv) working in one of the four largest metropolitan areas, as defined by Kropp and Schwengler [2016], in Germany at time of their displacement (Cologne, Frankfurt, Hamburg, and Munich[8]). However, neither the network members nor the displaced workers have to work in these areas after the displacement; regional mobility after layoff is therefore possible. Let me briefly justify these restrictions in three steps. First, prior literature focuses on men working full time; thus, to facilitate comparability, I restrict my sample accordingly. Second, the external options after job loss differ for women, and concentrating on men makes the interpretation of the results clearer. Third, the sample includes only metropolitan areas to comply with the sample size restrictions of the IEB.

Last, only those workers are included in the main estimation for whom high-quality geo-coded place of residence data are available. Since geo-codes are matched with register data, not all matches are of the same quality and sometimes include some

[7]Figure 3.2 in the appendix shows the timing of events graphically. Figure 3.3 gives a more intuitive explanation by providing an example.

[8]The sample excludes Berlin as it is a special case due to German reunification in 1990.

insecurity. I therefore restrict my sample to matches where the likelihood of true matching is very high.[9] However, in the robustness section, I show that the coworker results are indeed robust to using the larger sample of all displaced workers irrespective of the geo-code availability status.

3.4 Empirical Strategy

Estimating the effect of coworker and neighbor networks on individual labor market outcomes after a layoff is complex because networks are not random among workers. For instance, workers with better networks could have unobserved characteristics that simultaneously have a positive impact on their wages and employment status. This is critical because individual characteristics related to the network characteristics could bias the results. Moreover, the layoff itself could depend on personal characteristics and thus bias the estimates. To address these problems, I follow the identification strategy by Glitz [2017] in four steps but extend his approach to neighbor networks.

I start with a simple OLS regression using employment status one year after displacement as the dependent variable and employment share within a network in the 5 years prior to displacement as the explanatory variable. I then restrict the analysis to workers who are unemployed involuntarily and due to an establishment closure. This is necessary for two reasons: (i) differences in job search behavior and (ii) individual reasons for displacement. First, voluntarily unemployed workers may differ from involuntarily unemployed workers in their job search behavior, and their networks could differ in terms of size and quality because of anticipated unemployment. Using workers who are part of total closures ensures that they are involuntarily displaced and search unrelated to their networks. Second, displacement

[9] The IAB geo-data uses a four-digit letter combination for (i) postal code, (ii) city, (iii) street, and (iv) house number for address-hit quality. For each of the four address parts, the outline of the quality applies according to the following scheme: (A) specification can be found exactly in the reference database; (B) specification can only be found in the reference database by taking phonetic methods into account; (C) specification cannot be found in the reference database; (D) specification is missing; (E) only one street (section) center can be determined for this address, since no other house numbers can be found in the database; (G) Information can only be found with successive approximation; assignment only as good as possible, but not with certainty; (M) and information can only be found by comparison with different types of streets and by extending abbreviations. In this paper, I thus restrict the data to observations where the specification can be found in the reference database and is nonmissing.

due to a mass layoff is independent of (un)observed characteristics or performance. Therefore, all displaced workers (or workers who leave the establishment in the year prior to closure) within a closing establishment are highly comparable. This is a major advantage since a simple analysis of displaced workers could lead to severe selection bias if the reasons for their displacement correlate with their network characteristics. Regardless, the focus on displaced workers and the high comparability comes at the cost of the results not being universally valid if displaced workers differ from nondisplaced workers. The paper is in line with numerous articles [see von Wachter [2010] for a survey] that use mass layoffs to identify workers who become unemployed involuntarily not because of their personal abilities. I follow this literature by only including workers who are unemployed due to a total closure[10] in the analysis.

Second, the employment rate in one's coworker and neighbor network is not random to the network size. Therefore, I include the network size as an additional control to estimate the effect of the employment rate in the network on labor market outcomes after a mass layoff. By controlling for the network size, I ensure that the employment rate effect is measured rather than merely a size effect. I further include additional control variables in the regression.

Third, potential sources of network endogeneity are (i) workers' sorting in specific establishments or neighborhoods, (ii) shocks that particularly affect employees of one establishment or a specific region, or (iii) unobservable characteristics obtained during joint employment biographies or living in the same neighborhood. For instance, if high-ability workers are willing to change companies more often, e.g., to face new challenges and earn higher wages, they have a larger network, and the employment rate within the network differs from that of a person who always worked for one company. However, these high-ability workers also sort into specific establishments, thereby leading to a bias of both coefficients of the employment rate within a network and the network size in a simple OLS estimation. The direction of the biases is unpredictable and depends on the correlation between the network variables and the error term. Therefore, I include closing establishment and county fixed effects for each network analysis. By comparing only displaced workers from one establishment or county, I avoid any issues due to establishment- or neighborhood-specific endogeneity.

[10] Total closures are the most dramatic form of mass layoffs, namely, a 100 percent reduction of staff.

Last, I instrument the employment rate in the network by the occurrence of mass layoffs within the network. Even after adding different control variables and fixed effects, displaced workers from one establishment could share time-varying skills that likely affect workers' employment status and the employment rate of the network. This relationship would bias the results. To overcome this problem, I use the incidence of mass layoffs in a worker's network as an instrument for the employment rate within the network. The underlying assumption is that the displaced workers' (un)observable skills are unrelated to their network members' mass layoffs. The benefit of this assumption is that it is unlikely that a worker can impact whether network members working for different establishments are part of a mass layoff.

I examine the effect of the network employment rate on later labor market outcomes using the following estimation equation:

$$y_{it+1} = \alpha + \beta_1 UER_{it} + \beta_2 logNS_{it} + \gamma X_{fit} + \tau_f + u_{it+1}$$

The regressor y_{it+1} is a dummy variable that is equal to one if a displaced worker finds a job within one year after the mass layoff and zero otherwise. The coefficient of the unemployment rate in one's network, (UER_{it}), β_1, is the coefficient of interest. β_2 shows the impact of the network size, NS_{it}, at the time of the layoff. The network size, NS_{it}, only serves as a control variable herein; thus, β_2 cannot be interpreted causally. X_{fit} is a matrix of different individual and network control variables at time t. The individual characteristics include education, experience and its square, migration status, tenure in the closing establishment, the last log wage in the closing establishment, a dummy for the three-digit industry where the worker has worked in the past, the number of years employed, the average annual wage growth from $t-5$ to t, the average establishment size, and the number of different employers. The network characteristics that I calculate for each network type comprise the mean age of all former network members and its square, the share of former network members with medium and high education, the share of former female network members, and the share of foreign former network members. τ_f and u_{fit} represent the establishment or county fixed effects and the error term.

In the instrumental variable regression, UER_{it} is instrumented by Z_{it}. Z_{it} is the share of former network members who were themselves part of a mass layoff after separating from the displaced worker. More precisely, E_{it-s} is the number of former network members that are employed in year $t-s$, and M_{it-s} is the number of network members who, after separation from worker i, were in that year part of a mass layoff. t denotes the year of the establishment closure that pushes worker i into unemployment, and the instrumental variable is defined as follows:

$$Z_t = \sum_{s=1}^{5} \frac{M_{it-s}}{E_{it-s}}$$

As in Glitz [2017], I use the number of working network members (E_{it-s}) rather than their overall number to avoid any variation in the instrument stemming from differences in the network size. If a network member experiences more than one mass layoff after having worked in the same establishment as the displaced worker, only the last layoff counts. The identifying assumption is that the error term u_{fit} is uncorrelated to the instrument, conditional on the set of control variables.

3.5 Results

Section 3.5 presents the main results. I start with some descriptive analyses (section 3.5.1) that primarily show the sample composition and the first-stage correlations for both network types. The main results follow. First, the coworker and neighbor network estimates results are presented in section 3.5.2. Second, section 3.5.3 shows the heterogeneity in displaced workers' and network characteristics.

3.5.1 Descriptive Statistics

Table 3.1 presents workers' characteristics at the time of displacement. After the sample restrictions introduced in section 3.3, there are 17,760 displaced male workers in the sample; a share of 14 percent is foreign (n=2,432). Most layoffs take place in North Rhine-Westphalia (47 percent). Only 8 percent of the displaced workers hold a university degree, while 60 percent have vocational training, and 31 percent have no vocational training. The industry in which most displaced workers have worked is the wholesale sector.

Table 3.2 comprises the mean values of different worker and network characteristics at the time of layoff. Since the analyses focus on differences in network effects for migrants and natives, and both groups could differ already in their characteristics at time of layoff, the table shows the characteristics by migration status. The mean age of natives is 39 years, while migrants are, on average, younger. Moreover, migrants are less educated than natives and have worked for a shorter duration within the past 5 years. Migrants earn, on average, lower wages than natives and have worked for a higher number of different establishments in the past 5 years. Migrants have lived, on average, in a higher number of different neighborhoods.

The average coworker network size (counted during the coworker network phase ranging from 1 year to 5 years prior to layoff) is larger for migrants than for natives, which is plausible, as migrants change their workplaces more often than do natives. The employment rate in a worker's coworker network differs between migrants and natives, with natives having a higher employment share within their network at the time of layoff. Regarding the neighbor network, the table shows that migrants have larger neighbor networks than those of natives; this outcome is reasonable since migrants live in cities more often than do natives (Chiswick and Miller [2004], Bartel [1989]). However, the share of working neighbors is higher for natives than for migrants. This outcome is in line with prior findings that migrants tend to live in worse neighborhoods (see, e.g., Kling et al. [2007], Borgoni et al. [2019], Glas et al. [2019]). Overall, the table suggests that there are differences in the networks of migrants and natives that could explain diverse employment effects after a layoff.

3.5.2 Coworker Network Analysis

This section presents evidence on how the employment rate of former coworker and neighbor networks affects workers' re-employment probability one year after a displacement. Table 3.3 displays the main estimation results of equation (1) for the coworker networks. Column (1) shows the results of the OLS estimation, including all individual and network controls, but no establishment fixed effects. The coefficient of interest, β_1, indicates a positive relationship between the re-employment probability one year after a layoff and the employment share in a former coworker network. In column (2), the coefficient increases and is still significant when including establishment fixed effects in the equation. However, due to unobservable characteristics that affect not only the employment probability but also the employment share in a worker's coworker network, the estimated coefficient is likely biased. A priori, the direction of the bias is unpredictable, as different omitted variables could lead to an over- or underestimation of β_1.

To overcome this hurdle, column (3) contains the results when the coworker networks' employment rate is instrumented by the share of prior coworkers who were themselves part of a mass layoff. The first-stage results at the bottom of the table reveal that there is a strong negative correlation between the displacement share and the employment rate within a coworker network; i.e., a 10 percentage point increase in the share of displaced coworkers reduces the employment rate within the network by 1.2 percentage points, with an F-statistic of 15.3. Figure 3.4 depicts the negative first-stage relationship between the network employment rate and the share of coworkers who, after separation, are displaced because of a mass layoff. Using this variation in the employment rates due to mass layoffs for the IV estimation in column (3) shows a positive effect of the network employment rate on the re-employment probability; i.e., a 10 percentage point increase in the employment rate leads to a 4.9 percentage point higher probability of being employed one year after displacement.[11]

This coefficient is smaller than the one by Glitz [2017], who found that a 10 percentage point increase in the employment rate leads to a 7.5 percentage point higher probability of being employed. There are two possible reasons for this difference. First, Glitz used displacement data from 1995/1996, while this paper uses data from 2005-2011. however, the importance of a network varies over time, e.g., job seekers currently widely use the internet for job searches, which was uncommon 25 years ago, i.e., during Glitz's data period. Second, this analysis uses yearly panel data to construct each worker's network, in contrast to the process described in the paper by Glitz. By focusing on network members on 30 June of each respective year, the estimate likely underestimates the total network since coworkers who work for a establishment for a short period are not counted as network members. For instance, if a former coworker worked for an establishment from 1 January until 31 March, he or she is not counted as a member of the network since he or she did not work for an establishment on 30 June.

Similarly, table 3.4 shows the results for the neighbor networks. Column (1) presents a positive correlation between the employment share within a displaced worker's neighbor network and the re-employment probability for the OLS estimation without

[11] The IV approach estimates a larger coefficient in comparison to the OLS estimation. This could potentially be an indicator for a weak instrument problem. However, the magnitude of the IV coefficient is comparable to other findings in this field, and the F-statistic of 15.3 indicates a robust first stage. Furthermore, a bias of the OLS parameters is likely due to the endogenous employment rate and network variable.

fixed effects. The coefficient is remarkably smaller than the one in table 3.3 (0.040 vs. 0.005), which indicates neighbors' lower importance for re-employment compared to that of coworker networks.

Columns (2) and (3) provide the results when including fixed effects and using the IV estimation. There are no significant results for the OLS estimation when using fixed effects. In comparison to the coworker network, the first-stage results are less significant, and the coefficient is smaller than for the coworker network (F-Stat. 10.7); this indicates a weaker correlation between the displacement rate in a neighbor network and the employment rate within the network. The IV results therefore need to be interpreted with caution. Since mass layoffs take place more often in whole industries rather than regions, this partly explains the weaker instrument for neighbor networks. Therefore, the number of occurring layoffs within the neighbor network is lower than that in the coworker network. For the IV estimation, I find that a 10 percentage point higher employment share in the neighbor networks leads to a 0.7 percentage point higher re-employment probability one year after displacement.

The results in this section show that both coworker and neighbor networks positively contribute to a displaced worker's employment probability one year after layoff. The results show that coworkers play a more important role than that of neighbors in job searches after displacement[12]. This result is plausible since coworkers may have better information to pass on (see Calvo-Armengol and Jackson [2004]). On the one hand, coworkers can recommend prior displaced workers with whom they have worked in the past to employers who are searching for staff and thus reduce the uncertainty about a displaced worker's skills. On the other hand, they may forward highly suitable job offers and have information on positions that are not yet publicly available. Furthermore, prior coworkers recommending jobs to displaced workers also reduces displaced workers' uncertainty by sharing information about possible future employers. However, there are likely differences in the impact of networks on different workers. The next section, therefore, discusses the role of different networks on workers with various characteristics.

[12]This study focuses on displaced workers in metropolitan areas due to data restrictions. Therefore, the results are not universally valid. For instance, rural areas might differ in their network effects if neighbor networks play a more important role there.

3.5.3 Neighbor Network Analysis

Workers have different characteristics, as do their networks. This section shows how the results are heterogeneous according to the migration status and education status of the networks and workers. The literature offers different predictions on how similarity affects outcomes within a group. While preferences for homogeneity could explain the higher levels of cooperation within a group of similar people (e.g., Bell et al. [2011]), there are papers that emphasize the competition effect among very similar workers, especially among migrants (Beaman [2012], Albert et al. [2021]).

Table 3.5 contains the coworker and neighbor IV results for (i) migrants and natives, separately, and (ii) the native and migrant network, separately. Column (1) shows that the native and migrant coworker networks have a positive impact on native workers' re-employment probability. Even though the migrants' network effect is weaker, a 10 percentage point higher employment share in both networks indicates an approximately 5 percentage point higher re-employment probability one year after displacement. Interestingly, column (2) reveals that for migrants, the native network has a smaller effect than the migrant network, but both positively contribute to employment after a displacement. This finding is in line with the literature suggesting help within similar groups (Ioannides and Loury [2004], Beaman [2012]). Another explanation is that migrants tend to work in similar occupations, and therefore, their personal ties might be stronger.

Column (3) highlights that for natives, both migrants and native neighbor networks positively impact the re-employment probability. The effect of neighbors on natives is, however, 7 times smaller than the effect of coworkers. Turning to migrants in column (4), the table shows that the migrant neighbor network has no significant impact on the probability of re-employment, while the native network does have a significant impact. Strikingly, the effect of neighbors is larger on migrants than on natives. This outcome is in line with the literature that different ethnic groups use diverse channels for job search (Frijters et al. [2005], Battu et al. [2011]). The results indicate that migrants more intensively use their local neighbor networks. One explanation for the positive impact of the local native network but an insignificant effect of the migrant network is that the migrants within one neighborhood compete for similar jobs, especially if they are unskilled.

Overall, table 3.5 offers three key takeaways. First, for natives, coworker and neighbor networks contribute to employment after a layoff. Second, for migrants, we see a more complex picture; i.e., while migrants within their coworker networks are especially important for their re-employment, in the neighbor networks, natives are the influential individuals. Finally, native neighbor networks are more important for migrants than for natives.

Next, table 3.6 compares the effect of coworker and neighbor networks on workers with different educational levels by migration status. Columns (1)-(3) show the results for natives, and columns (4)-(6) show the results for migrants. When looking at the coworker network effect, the results reveal that the lower the educational status is, the more important coworker networks are, irrespective of the migration status. The same pattern is observed in neighbor networks. Neighbor networks have no significant impact on highly educated migrants and natives. However, due to the low number of observations, this conclusion must be drawn with caution. This result is in line with the theory by Calvo-Armengol and Jackson [2004] and Wahba and Zenou [2005]; i.e., different networks provide information of different quality. While the neighbor network likely provides information on routine tasks, highly specialized tasks are transmitted less frequently.[13] Therefore, highly educated individuals may benefit less from neighbors than from coworker networks.

Comparing migrants and natives, neighbor networks have a stronger effect on migrants than on natives, while the reverse is observable for coworker networks. Regarding differences in the network composition, table 3.5 reveals that for migrants, especially migrants in their coworker network, such differences are important. Since the overall migrant share within a coworker network is lower than the native share in most cases, it is reasonable that migrants benefit less from coworker networks than do natives. The higher impact of neighbor networks on migrants is likely due to composition effects, as the share of unskilled workers is higher among migrants, and unskilled workers benefit most from neighbor networks. Table 3.6 indicates two conclusions. First, the lower the educational status is, the higher the effect of networks is. Second, while natives benefit more from coworker networks than do migrants, the reverse is true for neighbor networks.

[13]For example, an engineer working in a highly specialized occupation such as battery advancement is less likely to get a job offer referred by neighbors than is a person working in the unskilled labor field who is open to working either at the local supermarket or in a restaurant.

3.6 Robustness Checks

In the following, I show that the main results from section 3.5 are robust if i) Berlin is included in the sample; ii) the regression is run without controls; iii) only network members who are both neighbors and coworkers are observed; iv) the instrument is defined differently; and v) not only workers for whom geo-coded information is available are included. Table 3.7 shows the results for both network types.

For the first robustness check, I include workers from Berlin in the sample. Thus far, I have excluded Berlin as a special case due to German unification in 1990. However, excluding the German capital and focusing on only western German cities could be selective[14]. As column (1) of table 3.7 shows, both network coefficients are slightly larger when including Berlin, but the main outcome is remarkably stable; i.e., coworker networks have a larger impact on the re-employment probability than that of neighbor networks, although both contribute positively. This may indicate that the network effects differ across Germany and that there are regional heterogeneities. To conclude, the result is robust to including Berlin, which suggests that the sample at hand provides rather lower bound effects.

The second robustness check includes an estimation without any worker- or network-specific controls (except network size and establishment (county) fixed effects, respectively, for coworker (neighbor) networks). If the instrument was systematically related to the controls and the estimation is run without controls, then a large change in the coefficient hints towards an endogeneity issue due to (un)observables. As column (2) shows, this is not the case; the coefficient barely changes in the absence of the control variables. This suggestive evidence supports the exogeneity assumption.

Third, there are workers within a network who are both coworkers and neighbors at the same time. To see if these network members are the drivers behind the results, I restrict the network to these workers. Consequently, in this robustness check, the coworker and neighbor networks are the same, and only the fixed effects within both estimations differ. In doing so, the number of network members decreases dramat-

[14]For this project, a large sample of German workers was drawn, including not only the metropolitan cities themselves but also the cities and villages within commuting distance. However, since not only the displaced workers but also their whole networks' employment biographies must be included, due to computational and data restriction reasons, I could not extend the sample to eastern German cities. A larger sample may deliver different results, and differences are unpredictable a priori.

ically; therefore, the standard errors increase. The magnitude of the insignificant coefficients in column (3), however, hints that these workers do not drive the results.

Fourth, I only use mass layoffs in large establishments with more than 50 employees for the computation of the instrument. The idea is to focus on mass layoffs that are most likely exogenous. In a next step, I extend this definition to mass layoffs in medium-sized establishments (with more than 10 employees). Column (4) shows that the results are robust to this different definition.

As a last robustness check, I estimate the coworker networks' impact with the full sample of available displaced workers. As the geo-coded data are not available for all workers and I restrict to high-quality matches (see section 3.3), many displaced workers are excluded from the sample. Running the estimation with the full sample and receiving similar estimates (see column (5)) suggests that the sample restriction is not systematically related to the controls. Therefore, the results are robust to a different sample selection.

3.7 Conclusion

This paper investigates how different types of networks, namely, neighbor and coworker networks, affect migrants' and natives' job search after a displacement. Previous literature shows that, in general, both network types positively affect the re-employment probability one year after a layoff (e.g., Ioannides and Loury [2004], Jackson [2010], Topa and Zenou [2015], Glitz [2017]). However, thus far, no study has focused on how different networks affect the same workers. As all workers are simultaneously members of different networks, this is an important, as yet unanswered, question.

As an additional contribution, this article examines how the two network types differ for migrants and natives. Using the occurrence of mass layoffs in one's network for an IV estimation, I investigate whether the re-employment probability one year after a mass layoff depends on how many people are employed in one's networks at the time of displacement. The results provide valuable insights into how diverse networks affect different worker groups.

The paper's main contribution is to quantify the effect of coworker or neighbor networks on the same workers' labor market outcomes one year after displacement. Due to data limitations, a direct comparison of the two network types was impossible. This article is the first to overcome this hurdle by combining high-quality administrative establishment and individual geo-referenced data from Germany. Second, by using newly available geo-coded data from Germany instead of using districts or municipalities to proxy a local neighborhood, I am able to define neighbors as those persons living within a grid of 3 square kilometers (centered around the displaced worker), irrespective of administrative borders. Last, in addition to general effects, this study highlights the effects on and of migrants. The heterogeneity analyses herein identify different effects of various networks on specific groups and address these differences.

In summary, the results of this paper indicate that both network types, namely, neighbor and coworker networks, are important for re-employment. This finding is in line with the literature highlighting that long-term unemployed individuals find jobs via networks more easily (Hirseland et al. [2019]). Using networks more intensely for job search, e.g., via mentoring programs, could thus be an interesting tool for policy-makers to use to help lead to higher employment rates after a displacement. Furthermore, these programs could focus on unskilled and/or migrant workers and their exchanges with employed workers.

This paper focuses on neighbor and coworker networks. However, currently, virtual formats and groups are often a relevant resource for information. For example, many individuals exchange information about job openings through social networks such as LinkedIn or Facebook. Analyzing what role friends and acquaintances play in re-entering the labor market, for example, by using cell phone data or platform usage data, is left for future research.

Table 3.1: Summary statistics - worker sample

	Workers
Total	17760
Natives	15328
Migrants	2432
Average age	37.3
Share in	
Hamburg	6.2
Cologne	46.5
Frankfurt	14.7
Munich	32.6
Education (in shares)	
Low education	31.0
Medium education	60.7
High education	8.3
Industry (in shares)	
Agriculture	1.4
Mining	0.1
Manufacturing Food	2.0
Manufacturing Textiles / Leather	0.2
Manufacturing Wood	0.5
Manufacturing Paper	0.5
Manufacturing Coke	1.8
Manufacturing Chemical	0.3
Manufacturing Rubber	0.3
Manufacturing Non-Metallic	0.2
Manufacturing Basic Metal	2.3
Manufacturing Machinery	0.6
Manufacturing Electric	0.6
Manufacturing Transport	0.1
Manufacturing Other	0.8
Electric	0.0
Construction	12.2
Service Wholesale	20.8
Service Hotel	16.5
Service Transport	10.1
Service Finance	0.7
Service Real Estate	3.2
Service Other Services	17.4
Service Public	6.1
Service Education	1.3

Notes: The table presents the descriptive statistics of the sample of male workers who become unemployed as the result of an establishment closure in the Cologne, Frankfurt, Hamburg, and Munich metropolitan area in the period of 2005-2011. Values are mean workers' characteristics at the time of layoff. Source: IEB. Own calculations.

Table 3.2: Summary statistics - worker and network characteristics

	(1)	(2)	(3)	(4)
	Natives		Migrants	
	Mean	Standard deviation	Mean	Standard deviation
Worker characteristics				
Age	38.6	12.2	35.7	10.6
Education	2.3	1.7	1.7	2.4
Number of years working (last 5 years)	4.3	2.1	4.1	2.7
Last wage in closing establishment	46.9	42.2	33.6	31.0
Last log wage in closing establishment	3.4	1.1	3.2	1.0
Network characteristics				
Number of establishments worked at (last 5 years)	2.5	1.0	3.2	1.3
Number of neighborhoods lived in (last 5 years)	1.1	1.7	1.5	1.6
Avg. coworker network size (last 5 years)	128.7	202.5	145.6	232.1
Share of coworkers working in t	60.3	20.1	55.1	30.0
Avg. neighbor network size (last 5 years)	2050.4	1910.3	3405.2	2104.4
Share of neighbors working in t	67.3	26.3	63.0	32.2

Notes: Table 3.2 presents mean workers' (network) characteristics at the time of layoff for natives and migrants. The sample comprises 17,760 displaced male workers. The instrumental variable is the share of coworkers (neighbors) who, after separation, worked in a large establishment (50 employees) and became unemployed due to a mass layoff. Source: IEB. Own calculations.

Table 3.3: Employment effects of coworker network

	(1) OLS Re-Emp Prob.	(2) OLS Re-Emp Prob. FE	(3) IV Re-Emp Prob.
Share employed in coworker network	0.0401**	0.0680*	0.4930**
	(0.0106)	(0.0367)	(0.2743)
Network size	0.0000	-0.0000	-0.019
	(0.0000)	(0.0001)	(0.0013)
Tenure job (years)	-0.0045**	0.0140***	0.0053
	(0.0018)	(0.0025)	(0.0046)
Tenure closing firm (years)	0.0114***	0.0070***	0.0064***
	(0.0011)	(0.0017)	(0.0022)
Mean firm size	0.0003***	0.0002***	0.0003***
	(0.0000)	(0.0000)	(0.0001)
Growth wage	-0.2979***	-0.1198**	-0.0721
	(0.0399)	(0.0507)	(0.0668)
Log imputed wage	0.0435***	0.0256***	0.0238***
	(0.0034)	(0.0050)	(0.0064)
Foreign	-0.0211***	-0.0088	-0.0102
	(0.0055)	(0.0063)	(0.0080)
Mean age network	-0.0026	0.0051	0.0225**
	(0.0025)	(0.0060)	(0.0102)
(Mean age network) squ.	0.0000	-0.0000	-0.0003**
	(0.0000)	(0.0001)	(0.0001)
Share medium skilled network	0.0347***	0.0246	-0.1891**
	(0.0103)	(0.0251)	(0.0888)
Share highly skilled network	-0.0502***	-0.1090*	-0.3688***
	(0.0188)	(0.0645)	(0.1297)
Share male network	0.0148	0.0307	-0.0429**
	(0.0093)	(0.0258)	(0.0181)
Share foreign network	-0.0090	-0.0173	0.0184
	(0.0107)	(0.0303)	(0.0407)
Establishment fixed effects		Yes	Yes
N	17760	17760	17760
First stage statistics - employment rate coworkers			
Share coworkers in mass layoffs			-0.1012**
			(0.0511)
F-stat first stage			15.3

Notes: Further controls included in columns (1)-(3) are the displaced worker's potential experience and its square, educational level, dummies for the number of different employers, and dummies for the main sector of activity during the network building phase. Column (1) reports the results for an OLS regression without fixed effects. Column (2) includes firm fixed effects for the estimation. Column (3) shows the results for an IV estimation in which the employment share is instrumented by the share of mass layoffs in a worker's coworker network. Source: IEB. Own calculations. Robust standard errors in parentheses; ***, ** and * refer to statistical significance at the 1, 5, and 10 percent level, respectively.

Table 3.4: Employment effects of neighbor network

	(1) OLS Re-Emp Prob.	(2) OLS Re-Emp Prob. FE	(3) IV Re-Emp Prob.
Share employed in neighbor network	0.0051**	0.0068	0.0071*
	(0.0026)	(0.0052)	(0.0043)
Network size	-0.0001	-0.0000	0.0113
	(0.0001)	(0.0002)	(0.0305)
Tenure job (years)	0.0007	0.0219***	0.0582
	(0.0039)	(0.0070)	(0.1096)
Tenure closing firm (years)	0.0076***	0.0052	0.0023
	(0.0026)	(0.0052)	(0.0374)
Mean firm size	0.0002***	0.0002	0.0018
	(0.0001)	(0.0002)	(0.0045)
Growth wage	-0.0213	-0.0059	0.1784
	(0.0897)	(0.1299)	(1.0464)
Log imputed wage	0.0107	0.0202	-0.0326
	(0.0077)	(0.0136)	(0.1713)
Mean age network	0.0056	-0.0079	-0.3506
	(0.0053)	(0.0159)	(0.9267)
(Mean age network) squ.	-0.0001	0.0002	0.0051
	(0.0001)	(0.0002)	(0.0133)
Share medium skilled network	0.0273	0.0090	1.4938
	(0.0209)	(0.0614)	(4.0095)
Share highly skilled network	0.0143	0.3670*	0.5455*
	(0.0476)	(0.2230)	(0.3300)
Share male network	0.0009	0.0721	4.1318
	(0.0202)	(0.0657)	(10.9078)
Share foreign network	0.0443***	-0.0552	-0.1728
	(0.0165)	(0.0626)	(0.5451)
County fixed effects		Yes	Yes
N	17760	17760	17760
First stage statistics - employment rate coworkers			
Share coworkers in mass layoffs			-0.0527*
			(0.0315)
F-stat first stage			10.7

Notes: Further controls included in columns (1)-(3) are the displaced worker's potential experience and its square, educational level, dummies for the number of different employers, and dummies for the main sector of activity during the network building phase. Column (1) reports the results for an OLS regression without fixed effects. Column (2) includes county fixed effects for the estimation. Column (3) shows the results for an IV estimation in which the employment share is instrumented by the share of mass layoffs in a worker's coworker network. Source: IEB. Own calculations. Robust standard errors in parentheses; ***, ** and * refer to statistical significance at the 1, 5, and 10 percent level, respectively.

Table 3.5: Heterogeneity I: Employment effects of diverse coworker and neighbor networks

	(1) Coworker Natives	(2) Coworker Migrants	(3) Neighbor Natives	(4) Neighbor Migrants
Share employed in network	0.5030*** (0.1870)	0.3929 (0.3244)	0.0071* (0.0043)	0.0084* (0.0046)
Share employed in migrant network	0.4730** (0.2377)	0.5168** (0.2311)	0.0077* (0.0047)	0.0076 (0.0102)
Individual controls	Yes	Yes	Yes	Yes
Network controls	Yes	Yes	Yes	Yes
Establishment FE	Yes		Yes	
County fixed effects		Yes		Yes
N	15328	2432	15328	2432

Notes: Column (1) reports the IV results for the effect of coworker networks on natives. Column (2) presents the IV results for the effect of coworker networks on migrants. Column (3) and (4), show the results of the IV estimation of neighbor networks on employment probability for migrants and natives, respectively. Source: IEB. Own calculations. Robust standard errors in parentheses; ***, ** and * refer to statistical significance at the 1, 5, and 10 percent level, respectively.

Table 3.6: Heterogeneity II: Employment effects of coworker and neighbor networks by education and migration status

	(1)	(2)	(3)	(4)	(5)	(6)
		Natives			Migrants	
	No voc. training	Voc. training	University	No voc. training	Voc. training	University
Share employed in coworker network	0.5230** (0.1246)	0.4921** (0.2409)	0.4701* (0.2538)	0.4819* (0.2646)	0.4608* (0.2581)	0.3007 (0.4967)
Share employed in neighbor network	0.0071** (0.0035)	0.0069* (0.0038)	0.0052 (0.0043)	0.0093 (0.0061)	0.0079* (0.0044)	-0.0325 (0.0351)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Network controls	Yes	Yes	Yes	Yes	Yes	Yes
Establishment / County FE	Yes	Yes	Yes	Yes	Yes	Yes
N	4928	9189	1211	723	1566	143

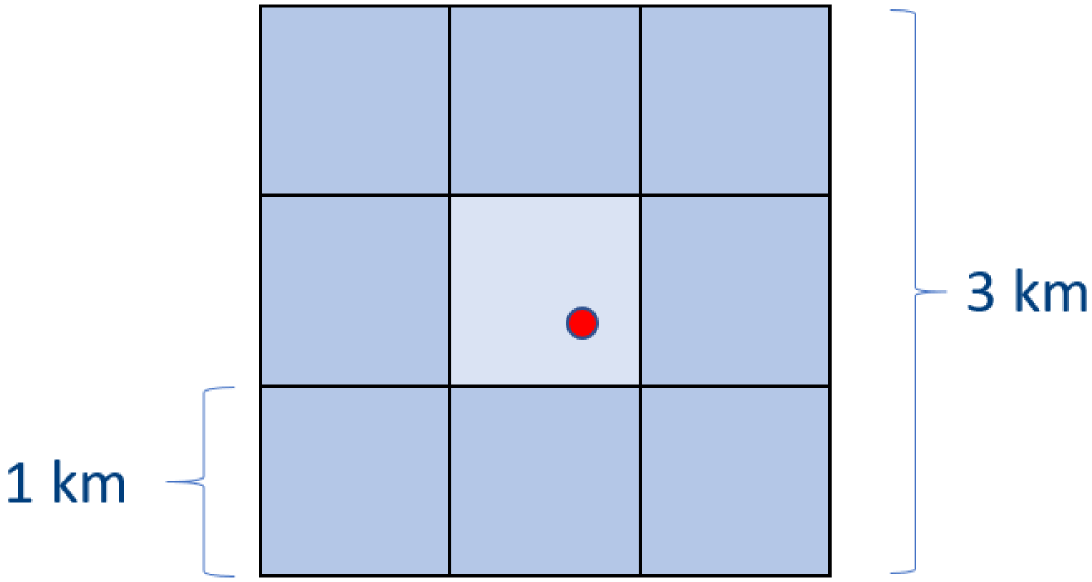
Notes: The coefficients in this table show the effect of the employment share within a network during the network building phase (1 to 5 years prior to displacement), subdivided by the educational status of the displaced workers. Line 1 (2) shows the effect of the employment share within a coworker (neighbor) network. The coefficients presented in the first row of tables 3.3 and 3.4 shows the effects for the whole sample (not divided by educational status). Column (1) reports the IV results for the effect of networks on native workers with no vocational training. Column (2) presents the IV results for the effect of networks on native workers with vocational training. Column (3) show the results of the IV estimation of the effect of networks on native workers with a university degree. Columns (4)-(6) show the results for migrant workers with the same educational degree/training as those mentioned above (i.e., no vocational training, vocational training, and a university degree, respectively). Source: IEB. Own calculations. Robust standard errors in parentheses; ***, ** and * refer to statistical significance at the 1, 5, and 10 percent level, respectively.

Table 3.7: Robustness checks

	(1) Berlin	(2) Without controls	(3) In both networks	(4) Include Mass Layoffs in medium est.	(5) Whole sample
Share employed in coworker network	0.5678* (0.3119)	0.4125* (0.2440)	0.3771 (0.7275)	0.4798* (0.2713)	0.437* (0.1673)
Share employed in neighbor network	0.0077* (0.0044)	0.0081* (0.0045)	0.0069 (0.0311)	0.0072* (0.0038)	
Individual controls	Yes		Yes	Yes	Yes
Network controls	Yes		Yes	Yes	Yes
Establishment / County FE	Yes	Yes	Yes	Yes	Yes
N	21962	17760	17760	17760	37351

Notes: Column (1) reports the IV results for the effect of networks on native workers with no vocational training. Column (2) presents the IV results for the effect of networks on native workers with vocational training. Column (3) show the results of the IV estimation for the effect of networks on native workers with university degree. Columns (4)-(6) show the results for migrant workers with the same educational degree/training as those mentioned above (i.e., no vocational training, vocational training, and a university degree, respectively). Source: IEB. Own calculations. Robust standard errors in parentheses; ***, ** and * refer to statistical significance at the 1, 5, and 10 percent level, respectively.

Figure 3.1: Illustration of geo-coded residence data



The figure shows how the geo-coded neighborhood data are computed. The red dot is the displaced worker, and I identify his location within a 1-km x 1-km grid (light blue grid). Afterwards, I identify all neighbors living in grids of 3 km x 3 km around the displaced worker (dark blue grids). For this purpose, I include neighbors who are covered by the German registry data in each of the 5 years of the network phase, i.e., only employees subject to social security contributions. Persons not included in this definition (pensioners, children, self-employed, civil servants) are not included in the networks.

Figure 3.2: Timing of events

Time	t-5	t-4	t-3	t-2	t-1	t	t+1
Event	Network phase					Layoff	Employ ment status

The figure shows the timing of events within this analysis. The network building phase covers the 5 years prior to displacement. All coworkers (neighbors) who are covered by the social security data are then included in the network analysis. The year of the layoff is not included, as all workers who worked in the same establishment and experienced a total closure are likely competitors on the labor market. The variable of interest for the analysis is the displaced worker's employment status (employed / unemployed) one year after displacement.

Figure 3.3: Timing of events: Example

Time	2000	2001	2002	2003	2004	2005	2006
Event	Network phase					Layoff	Employment status
Establishment (Est.)	Est. A		Est. B	-	Est. C		Est. D
Residence	Cologne				Munich		

The table includes an example for a worker who was displaced in 2005 from establishment C. In the analysis, I want to determine whether his employment status in 2006 (employed or unemployed) depends on the employment share within the displaced worker's network. Therefore, I compute the displaced worker's (coworker and neighbor) network in the 5 years prior to displacement.

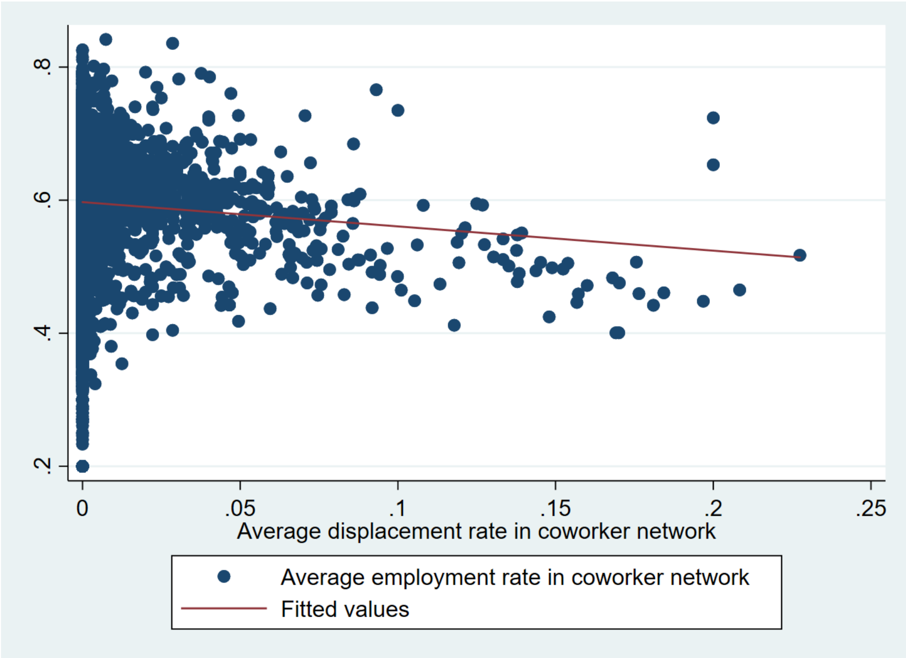
In this example, the displaced worker worked for the establishments (Est.) A, B, and C during the network phase (blue boxes). After the displacement, i.e., in 2006, the worker works in establishment D. In 2003, this worker was unemployed and therefore had no coworker network.

The white boxes show the neighbor network. For the neighbor network, I include all workers who lived in a 3-km x 3-km grid around the displaced worker in the cities of Cologne (2000, 2001, 2002, 2003) and Munich (2004). For a graphical illustration of this procedure, please examine Figure 3.1.

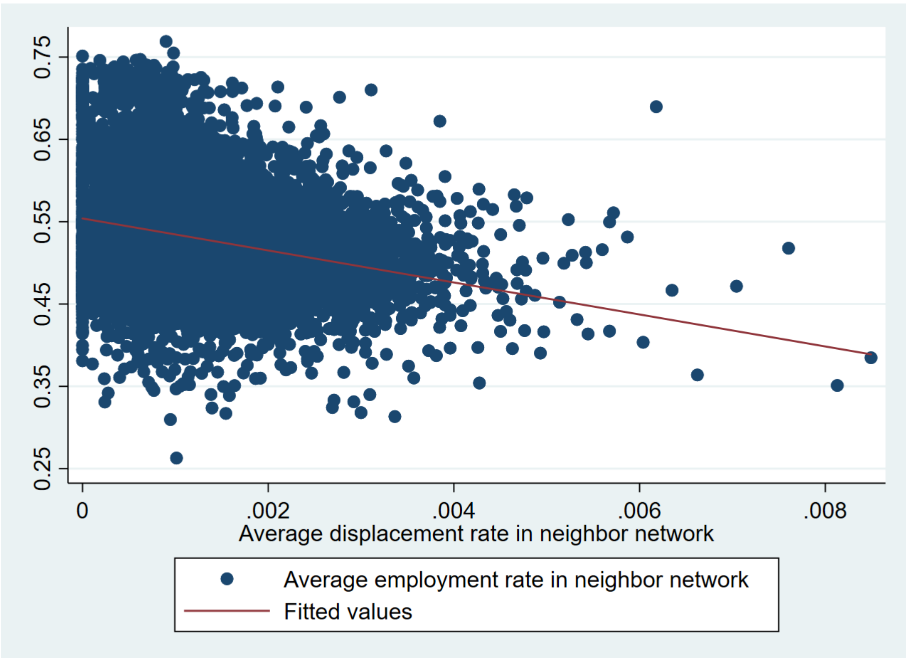
What does the coworker network look like? I include all workers who worked at the same time as the displaced worker in establishment A, i.e., from 2000-2001, in the network. Further, I include all workers who worked in 2002 for establishment B. Last, I include all workers who worked in 2004 for establishment C. After having identified all prior coworkers, I then identify all prior coworkers who were themselves part of a mass layoff within the network phase. Thus if, e.g., a coworker from establishment A got displaced in 2004, I would count him or her as displaced and use this categorization to compute the instrument. For workers who were displaced in 2005, I do not take them into account, as 2005 is outside of the network building phase (2000-2004).

For the neighbor networks, I use the same strategy. I identify (i) all workers within the 3-km x 3-km grids within the network building phase and (ii) all workers within the network who also experienced a mass layoff.

Figure 3.4: Relevance of the instrument - correlation between displacement rate within networks and the networks' employment share



(A) First stage, coworkers (Correlation in main specification table 3.3: -0.1012 (0.0511), F-stat.=15.3)



(B) First stage, neighbors (Correlation in main specification table 3.4: -0.0527 (0.0315), F-stat.=10.7)

The figure shows the relationship between the employment share within a network and the displacement rate within a network for coworker and neighbor networks. Source: IEB. Own calculations.

Chapter 4

Female inventors.

Joint work with Panu Poutvaara.

4.1 Introduction

Innovation is a key driver of technological change and economic growth (Romer [1990], Aghion and Howitt [1992]). However, Romer [1990] shows that too little human capital is devoted to research in equilibrium. A newly emerging body of literature analyzes the factors that determine who becomes an inventor. This literature seeks to find new ways to foster innovation by increasing the number of inventors, especially within underrepresented groups such as women. Bell et al. [2019] find that women have a higher probability of becoming inventors in a particular field if they grew up in an area with more female inventors in that particular field. Furthermore, they reveal that women are underrepresented among star inventors (i.e., those who are most productive), as they are among all inventors. The authors explain these central findings in terms of missing female role models and networks (exposure effects). The founders of BioNTech, Özlem Türeci and Uğur Şahin, are prominent examples of the potential of inventors who belong to underrepresented groups: both are Germans with immigrant backgrounds, one of whom is a female. Although this is anecdotal evidence, it illustrates the major welfare gains available from the efficient use of the talent of potential inventors by encouraging "lost Einsteins" to innovate. In the spirit of Bell et al. [2019], the unused potential of female inventors may have cost the world many "lost Marie Curies".

To date, there is little evidence on how the share of female inventors differs around the world, how it has developed, and how female inventors differ from male inventors in terms of the number of patents and citations that they obtain. We help to close this gap by investigating newly available patent and employment register data on inventors in Germany who applied for a patent at the European patent office between 2000 and 2010.[1] This dataset enables us to paint a comprehensive portrait of female inventors in Germany, and compare them with male inventors. Germany is an interesting country to investigate since it is among the leading patenting countries in terms of overall numbers of patents and its industry includes diverse research fields. Our contribution to the prior literature is the identification of the characteristics of women who have already decided to become inventors. To date, there has been

[1]The *Linked Inventor Biography Data 1980-2014* (INV-BIO) are provided by the Institute for Employment Research (IAB) and the Max Planck Institute. The dataset combines inventor and patent information obtained from patent register data with administrative labor market data on individuals and their employing establishments. Section 4.3 provides further information on this dataset.

little economics research on female inventors, most likely due to data restrictions. Identifying the characteristics of female inventors might be a first step to avoiding "lost Marie Curies" in the future.

To put Germany into a wider context, figure 4.1 shows how the percentage of female inventors has developed from 2005 to 2015 in all European Union (EU) countries, as well as in Iceland, Norway, and Switzerland. Countries are ordered according to the share of female inventors in 2005, ranging from 2 percent in Cyprus and Slovakia to 37 percent in Croatia. The share of female inventors increased from 2005 to 2015 in most countries, with the largest increases in Latvia and Romania. The share of female inventors is lower in Germany (10 percent in 2015) than in France, Italy, Spain, and the UK. Table 4.17 compares the share of female inventors in the major patenting nations around the world (the USA, Japan, Germany, China, South Korea, France, and the UK) from 1990 to 2018. The share of female inventors has increased considerably in all countries, more than doubling in Germany, South Korea, and the UK. Nevertheless, the share of female inventors in Germany lags behind not only that of France and the UK but also that of the United States and China.

Bell et al. [2019] show that fewer women with good math scores (a proxy for high potential) become inventors than men. Furthermore, in many European countries (including Germany), the majority of university graduates are female in all subjects except for the STEM subjects (see, e.g., https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Tertiary_education_statistics#Graduates [last access: 12-20-2021]). Figures 4.10 and 4.11 in the appendix present the number of female and male graduates, respectively, in STEM subjects in Germany for all students and foreign students only. The figures reveal that fewer women than men graduate from STEM programs.[2] A low share of female in STEM subjects can thus largely explain a low share of female inventors.

Our paper contributes to the literature on the "leaky pipeline", which refers to the phenomenon in which a progressively lower share of women reach each stage of higher education. Figure 4.2 shows the proportion of women and men in each stage of higher education, further research, and inventorship worldwide from 1998 to 2017. While women account for more than half of all students in higher education,

[2]However, the share of women among foreign students is higher than that among German students.

only 43% of Ph.D. graduates are women. Less than one-third of all researchers are women. The proportion of women who go on to apply for patents as inventors is even lower (11 percent). This confirms prior findings (see, e.g., Carrell et al. [2010], Mansour et al. [2021], Buckles [2019], Canaan and Mouganie [2021]) that demonstrate the existence of a leaky pipeline for women. Figure 4.3 shows that the same pattern holds for Germany: even though more than half of all graduates are women, only approximately 40% of those who obtain a Ph.D. are women. The share of women is even smaller among researchers (2008: 34%; 2018: 40%) and professors (2008: 17%; 2018: 25%). However, we observe a clear trend toward greater gender equality across all academic levels from 2008 to 2018.

In the first part of this article, we present descriptive evidence on the personal characteristics of inventors. Comparing inventors to a random sample of employees in Germany ("average" workers in the German labor market) reveals that inventors are on average older, more educated, and earn higher wages. Women earn lower wages than men among both inventors and noninventors. In line with prior literature, we find that female inventors tend to work in fields such as pharmaceuticals and biotechnology rather than mechatronics. When pooling all years and separately analyzing 34 technological fields, we find that the share of women among inventors is lowest at 0.02 in basic communication processes, machinetools, and mechanical elements and highest in biotechnology (0.26), pharmaceuticals (0.24) and organic chemistry (0.20). Second, we contrast female and male inventors in terms of patent characteristics. The results show that female inventors have, on average, fewer patent citations, applications, and rejections than their male colleagues. Furthermore, there are fewer star inventors with exceptionally high numbers of patents among female inventors than among male inventors. When pooling all years, the average number of patent applications by female inventors is found to be lower than the average number by male inventors in all but 2 of the 34 technological fields, and the average number of patents granted is also lower for female inventors in 32 fields. Last, the share of inventors with a top-coded income is lower among females than males.[3]

[3]The dataset contains records of employee gross daily wages. These gross daily wages are calculated from the fixed-period wages reported by the employer and are thus highly reliable. Daily wages are documented in euros. Earnings exceeding the upper earnings limit for statutory pension insurance are reported only up to this limit and therefore are top-coded. This information stems from the following source that further provides an overview of these limits and thresholds: https://fdz.iab.de/de/FDZ_Individual_Data/IAB_Employment_Samples/IAB_Employment_Samples_Working_Tools.aspx [last access: 01-25-2022].

Next, we use multivariate OLS regressions to check whether the descriptive results also hold when taking additional explanatory variables into account. These explanatory variables include (along with sex) education, migration status, and a dummy for being a mother conditional on being a female inventor. The OLS regression further includes establishment size, industry, year, and age group fixed effects. The OLS regression estimates show that female inventors have a higher patent rejection rate and fewer granted patents and patent applications (all measured at the individual level). Regarding citations between the different patent offices, the picture is mixed: in Germany, female inventors have fewer citations, while in the European and US patent offices, they have more citations, when taking different control variables into account. Four possible explanations for this last result are (i) that women tend to work on teams rather than alone (Jung and Ejermo [2014], Mauleón and Bordons [2010], Naldi et al. [2004]) and therefore profit from the larger international networks of larger teams and are therefore more often cited, (ii) that those women who are granted international patents are a highly productive group of high-ability women since international patents are harder to obtain, (iii) countries other than Germany could be more progressive such that women experience less discrimination in countries with higher shares of female inventors. Although one would expect that patent citations are reasonably objective, Jensen et al. [2018] provide intriguing evidence that is suggestive of gender discrimination in patent citations in the US, and (iv) different technological fields in patenting could play a role here. Women tend to work more often in biotechnology and related fields that are relatively stronger in many other countries than Germany, while men work in machinery and mechatronics, where Germany is among the leading countries. For this reason, women's patents may be relatively more relevant for research done abroad and men's patents for research done in Germany.

In the last results section, we investigate the role of motherhood. Overall, female inventors are less likely to become mothers than the members of a random sample of female employees in Germany ("average" workers in the German labor market). Among female inventors, the likelihood of becoming a mother is higher when the female inventor is highly productive (as measured by number of granted patents and citations in Germany). When looking at parental leave durations, we find that female inventors take longer parental leaves than the "average" woman in Germany. Completely exiting the labor market is also more common among female inventors. When we look at labor market re-entry after a birth, female inventors are less likely to return to either part-time or full-time work than the "average" woman in Germany. This might be an indicator that working as an inventor demands being present full-time or that R&D jobs are less family-friendly than other jobs.

What tools are available to increase the number of female inventors? Bell et al. [2019] find that girls who grow up in areas with a high share of female inventors are more likely to become inventors themselves (the exposure effect). Therefore, a steady increase in the number of female inventors and fighting the "leaky pipeline" should contribute to higher female participation in the R&D sector. Furthermore, the lower rate of returning to part-time and full-time positions after giving birth among female inventors (relative to "average" women) could indicate a possible incompatibility between family life and a career as an inventor. Addressing these two factors in the future and thus making the inventor profession more attractive to young women is the task of politicians and the entrepreneurs.

The remainder of this paper proceeds as follows. Section 4.2 provides an overview of the literature related to female inventors. Section 4.3 describes the data and shows summary statistics and general trends in the data. Section 4.4.1 presents the differences in the personal and occupational characteristics of male and female inventors, while section 4.4.2 focuses on differences in patenting characteristics. Section 4.4.3 reveals differences in the number of patent citations between male and female inventors. In section 4.4.4, we explore the extent to which motherhood affects female inventorship. Section 4.5 concludes.

4.2 Related literature

Our results build on and contribute to different strands of literature. First, we contribute to the literature on the question of who becomes an inventor and in which fields and geographical areas the share of female inventors is lowest and why. Bell et al. [2019] examine who becomes an inventor in the United States and find that exposure to innovation is a critical factor in the choice of this career path. However, children's chances of becoming inventors vary sharply with their background characteristics such as their race, gender, and parents' socioeconomic class. Women are more likely to invent within a specific technology class if they grew up in an area with more women (but not men) who invent in that class. Mechanisms such as role model or network effects most likely drive these gender and technology class–specific exposure effects. Bell et al. [2019] indicate that there are many "lost Einsteins"- individuals who could have become inventors if they had been exposed to innovation in childhood but did not - especially among women, minorities, and children from low-income families. In line with this finding, there is also a large body of literature on the effectiveness of role models in encouraging women to study STEM subjects (see, e.g., Carrell et al. [2010], Mansour et al. [2021], Buckles [2019], Canaan and Mouganie [2021], Lim and Meer [2017], Lim and Meer [2020]). Unfortunately, our data are not linked to information on inventors' family backgrounds or the locations where they grew up. Instead, our data enable us to identify mothers and to zoom in on the establishments for which female inventors work.

Several papers have found that the share of female inventors is very low in Germany relative to that in other European countries (Hunt et al. [2013], Busolt and Kugele [2009], Naldi et al. [2004]). Prior literature further reveals that female inventors are underrepresented within the industries with the highest total patenting rates (electrical and mechanical engineering) and instead tend to work in chemistry and pharmaceuticals [see, e.g., Hunt et al. [2013], Busolt and Kugele [2009], Naldi et al. [2004], Frietsch et al. [2009]]. By looking at the historical perspective, Sugimoto et al. [2015] highlight that female inventors' patenting activities have increased considerably across all sectors in the past 40 years in the US. Nählinder [2010] uses a group experiment and qualitative interviews to seek the reasons why women innovate less. The author argues that women are less self-confident and shows through a project experiment with nurses that with special instructions, nurses can become more innovative.

Second, we analyze differences in patenting behavior by sex. Different papers, mostly from the US, have found that women are less likely to write single-authored patents but rather tend to work in groups [e.g., Naldi et al. [2004], Sugimoto et al. [2015], Mauleón and Bordons [2010]].[4] In this context, Wang et al. [2019] show that if a woman runs an inventing team, more women will be on the team, indicating that differences in composition depend on the characteristics of the head of the team. Furthermore, women and men work in different kinds of organizations: if women patent, they are more likely to patent in academic institutions than in corporations or in

[4]Table 4.27 in the appendix shows the number of inventors per patent in the INV-BIO dataset.

governmental institutions (Sugimoto et al. [2015]).[5] Additionally, firm type plays an important role according to Whittington and Smith-Doerr [2008]: For instance, in biotechnology firms that are known for their flatter, more flexible network-based organizational structures, female scientists are more likely to hold patents than in more hierarchically arranged organizational settings. Whittington and Smith-Doerr [2008] also find that in larger firms, women are more likely to patent since application processes are standardized. Additionally, women tend to work in fields with fewer patent applications (in absolute terms) (Mauleón and Bordons [2010]).[6] Relatedly, it is important to note that some female-dominated fields (e.g., humanities) tend to publish differently, e.g., in books rather than in international journals, and therefore, their work is less frequently cited (Naldi et al. [2004]). That women's patents are cited less often than men's is also confirmed by other papers (e.g., Sugimoto et al. [2015]).

Jensen et al. [2018] further show in a descriptive analysis that female inventors are apparently disadvantaged in the process of obtaining and maintaining patents relative to men. Comparing grants and citations for men and women with common and uncommon names, Jensen et al. [2018] find that women with common female names (such as "Mary") are less likely to be cited and granted patents than women with names that do not suggest gender. Aneja et al. [2021] further show that there are gender differences in patent application behavior: women who submit patent applications (that are comparable to those of men) are less likely to continue the patent process after receiving an early-stage (but not final) rejection of a patent claim. In terms of patent commercialization, Cook and Kongcharoen [2010] note that female and male inventors now commercialize their patents in similar ways and that the large difference that existed 40 years ago has disappeared over time.

Third, there is an important strand of literature on the returns to invention for workers and firms. Aghion et al. [2018] find that inventors receive only 8 percent of total private returns from invention, while entrepreneurs obtain over 44 percent of the total profits and also (nonpatenting) coworkers benefit from invention. Moreover, entrepreneurs begin with negative returns before filing patents, but their returns later become very positive. Kline et al. [2019] find that an initial approval of an ex

[5] Table 4.18 in the appendix presents the worldwide number of patent applications by employer, gender, and year, which confirms this result.

[6] Table 4.30 in the appendix provides an overview of different establishment characteristics within the INV-BIO dataset pooled across 2000-2010 by sex.

ante valuable (low-value) patent leads to substantial (no) increases in firm productivity and worker compensation. They further find that patent grants increase firm employment while entry wages and workforce composition remain stable. Regarding monetary incentives, workers capture approximately 30 percent of the patent-related surplus in higher earnings. This proportion is notably higher among workers who have been present since the year of application, probably because experienced employees are harder to replace. These income effects are most evident for men and workers in the top half of the earnings distribution. Toivanen and Väänänen [2012] also show that returns to patenting are heterogeneous over time: inventors receive a temporary reward of 3% of their annual earnings for a patent grant plus a longer-lasting premium of 30% of earnings three years later but only for highly cited patents.

Fourth, our paper is also related to the wider literature on the importance of a more balanced gender distribution in different societal domains. Miller [2008] reveals that introducing voting rights for women in the United States led to an increase in health campaigns focusing on child welfare. Chattopadhyay and Duflo [2004] show that female representation changes public goods provision in Indian villages, and Bhalotra and Clots-Figueras [2014] find that women's political activity at the state level leads to better provision of antenatal and childhood health services in India. Matsa and Miller [2013] find that increased participation by women on corporate boards leads to fewer workforce reductions, increases in relative work costs and employment levels, and a reduction in short-term profits in Norway.

Fifth, there is an emerging literature that investigates the impact of gender on invention and health outcomes. Koning et al. [2021] use a text analysis of all biomedical patents in the US and find that patents from female-only inventor teams are 35% more likely to focus on women's health than patents from male-only teams. This is in line with other papers that find that women scientists, inventors, and entrepreneurs are more likely to develop ideas, inventions, and products that are beneficial to women (Nielsen et al. [2017], Einiö et al. [2019], Kozlowski et al. [2022], Murray [2021]). Furthermore, most studies have been found to focus on males, while women are underrepresented in studies of disease mechanisms and treatment (https://www.nature.com/articles/550S18a). Women are therefore more likely to be prescribed ineffective or dangerous medications. Moreover, Greenwood et al. [2018] find that women have lower chances of surviving a heart attack when treated by a male physician and that the likelihood of recovery increases when male doctors work more frequently in teams with female doctors or have treated more female patients in the past.

4.3 Data

We use a newly available dataset from the Institute for Employment Research (IAB) and the Max Planck Institute on inventors who applied for a patent at the European Patent Office (EPO) between 1999 and 2011, the *Linked Inventor Biography Data 1980-2014* (henceforth: INV-BIO), for our empirical analysis. This dataset combines inventor and patent information obtained from patent register data with administrative labor market data on individuals and their employing establishments. Such high-quality linked data have not been previously available for investigations of inventor careers. In addition, Germany's national innovation system provides an interesting context for the study of invention processes, as Germany is among the leaders in numbers of patent applications filed and patents granted.

For an inventor to be in the dataset, that inventor must be (i) listed on patent filings at the EPO and residing in Germany from 1999-2011 and (ii) an employee of an establishment during that period (thus, neither self-employed nor a civil servant). The INV-BIO dataset comprises 152,350 complete employment biographies of unique inventors who patented between 1999 and 2011 and combines inventor patent documents with labor market biographies from high-quality social security data for this period. For those inventors identified in the dataset, employment biographies and patent activities (at the European and German Trademark Office) from 1980 to 1999 are also available in the dataset.

Importantly for our study, the social security data include the inventors' nationality, gender, education level, age and information on the employing establishments in a panel structure. The employing establishment data include information on the industrial sector, the number of employees, the average wage of all full-time employees, and the location of the workplace at the federal state and district levels. The data enable analyses at the daily level.

The INV-BIO data include 643,856 patent families[7] (as defined by the DOCDB[8] of the EPO), representing approximately 71.4 percent of all inventions (evaluated at the patent family level) in Germany from 1999 until 2011[9]. Dorner et al. [2018] provide a more detailed overview of this dataset.

Table 4.1 provides an overview of the summary statistics of the inventor dataset in comparison with a 2 percent random sample of workers in Germany. The INV-BIO dataset comprises 133,396 male and 11,005 female inventors (columns (1) and (3)). Regarding age, we see that male inventors (column (1)) are on average older than a random sample of male workers (column (2)), while there is no difference in age between female inventors and noninventors (columns (3) and (4)). When looking at daily wages (or, in cases of unemployment, benefit receipts), we see a sharp contrast between inventors and the random sample of workers, with both female and male inventors receiving significantly higher daily wages. Men have higher daily wages than women in the random sample and among inventors. Unsurprisingly, the share of university graduates is significantly higher among inventors. In summary, when comparing inventors with noninventors, there are some differences: Inventors have higher mean daily wages (men: 155.62 vs. 58.01; women: 134.78 vs. 37.27) and have, on average, higher educational attainment.[10]

Figures 4.1 and 4.2 reveal the trends in the total number of female and male inventors and in the share of female inventors in Germany from 2000 to 2010. Figure 4.1 shows that there has been a steady increase in the total number of female and male inventors from 2000 to 2008. However, the total number of male inventors is more than ten times as high as the number of female inventors. From 2009

[7]A patent family is a collection of patent applications covering the same or similar technical content. Further information is available here: https://www.epo.org/searching-for-patents/helpful-resources/first-time-here/patent-families.html [last access: 11-25-2021].

[8]The DOCDB is the EPO's master documentation database and has worldwide coverage. It contains bibliographic data, abstracts, citations and the DOCDB simple patent family structure but no full text or images. For further information, visit: https://www.epo.org/searching-for-patents/data/bulk-data-sets/docdb.html [last access: 01-14-2022].

[9]The inventions of self-employed individuals, civil servants, or inventors not residing in Germany are not linkable to the German register data.

[10]Table 4.25 provides an overview of employed workers only, while table 4.26 presents summary statistics for employed workers with university degrees.

onward, there has been a sharp decrease in the number of inventors in Germany.[11] There are likely two reasons for this. First, the Patent Cooperation Treaty (PCT) was reformed several times. The PCT is an international treaty on patent law. It establishes a standardized process for the submission of patent applications for the protection of inventions in each of its contracting states. A patent application submitted under the PCT is referred to as an international application or PCT application. Due to the reforms, the number of patents (and thus inventors) is not comparable across the entire time period, as there was a large decrease in the number of patent applications.[12] Second, the financial crisis led to decreases in R&D sector investments. This is another explanation for the decreasing number of inventors from 2008 onward.

Due to major differences between former East Germany and former West Germany in terms of female labor force participation (Wyrwich [2017], Trappe and Rosenfeld [1998], Trappe [1996], Braun et al. [1994], Bachmann et al. [2018], Cooke [2006], Künzler et al. [2001], Rosenfeld et al. [2004]) and the industrial structure (Mertens and Müller [2020], Wolf [2010], Schnabel [2016], Brenke [2014], Röhl [2014], Paqué [2009], Burda and Severgnini [2018]), in figure 4.4 (4.5), we analyze the number of female and male inventors (the share of female inventors) separately for these two regions. While we see a clear increase in the share of female inventors in former West Germany, there is no clear trend in former East Germany. This could reflect both a higher beginning female share and a small number of observations in former East Germany (the lowest number of female inventors in former East Germany was 111 in 2000 and the highest number was 195 in 2007). When analyzing the share of female inventors at the state level, excluding Berlin, as it was divided between East and West Germany, the share of female inventors in former West German states ranges from 0.06 to 0.14, and in former East German states from 0.10 to 0.13 (see

[11]Note that this decreasing trend in inventorship is in line with evidence from other datasets from Germany. Additionally, in data from the Statistical Office of Germany, the total number of inventors decreases from 2008 onward (https://www-genesis.destatis.de/genesis/online?operation=previous&levelindex=&step=&titel=&levelid=&acceptscookies=false). The common trend within both datasets is remarkably similar; small differences are likely because the INV-BIO dataset does not include self-employed inventors or inventors in the public sector, such as professors.

[12]Further information on the PCT and its reforms is available online: https://en.wikipedia.org/wiki/Patent_Cooperation_Treaty#Statistics [last access: 01-26-2022] and https://www.wipo.int/pct-reform/en/ [last access: 03-15-2022].

table 4.32).[13] In summary, both the number and the share of female inventors in Germany increased between 2000 and 2010, with a peak in 2008. However, female inventors are still a small fraction (less than 10 percent) of all inventors in Germany.

4.4 Results

4.4.1 How do female and male inventors differ?

This section describes the key personal characteristics of inventors and the technical fields in which they work with a special focus on gender differences. Personal characteristics include professional trajectories and age structures. We take a first step toward evaluating the role of gender in regard to the prerequisites for invention: education and professional experience. Furthermore, we focus on whether differences in fields with varying levels of inventing activity explain the gender gap in innovation.

Table 4.2 provides an overview of all inventors who applied for a patent (irrespective of whether it was granted or not) by (i) age group, (ii) sex, and (iii) year (2000, 2005, and 2010).[14] It turns out that the percentage of inventors younger than 30 was lower for women (6 percent) than for men (12 percent) in 2000 and in 2005 and 2010. Furthermore, we observe that the percentage of inventors in the age group 41 to 50 increased from 2000 to 2010 for both genders. This is likely because of the general aging trend in the labor force in Germany.

Next, we look at the different technological fields in which female inventors work (figure 4.6) and how the number of patents varies by technological field (figure 4.7). The three fields with the highest share of female inventors are biotechnology, pharmaceuticals, and organic chemistry. The three fields with the lowest share are basic communication processes, machine tools, and mechanical elements. Figure 4.7 shows that the share of women is highest in the field with a low number of patents per inventor (Biotechnology). However, the share of females is also comparatively high in the field with the most patenting activity (organic chemistry). Therefore,

[13] Please note that the shares are quite volatile, especially in East Germany, due to the low number of female inventors. Therefore, table 4.33 provides an overview of the share of female inventors in different states in 2000, 2005, and 2010 to provide a more comprehensive picture of female inventorship.

[14] A more detailed overview of the shares of the different age groups among female and male inventors can be found in figures 4.12 and 4.13 in the appendix.

female selection into specific (low-patenting) fields can only partly explain their lower number of patents.

We also compare the share of female inventors and the average number of patents applied for and granted per inventor across 34 technological fields.[15] When pooling all years, the share of female inventors ranged from 0.021 in basic communication processes to 0.263 in biotechnology. Male inventors applied for more patents per inventor in each technological field, with the gender gap ranging from 0.25 in materials & chemistry to 22.15 in engines/pumps/turbines (figure 4.7). The only two technological fields in which female inventors received more patents grants per inventor were materials & chemistry and materials & metallurgy, in which the average number of granted patents for women exceeded that for men by 0.26 and 0.45 (see figure 4.14 in the appendix). In other fields, the gender gap in the average number of patents in favor of male inventors ranged from 0.79 in other machines to 6.17 in semiconductors. In each technological field, the inventor with most patents was a male. When analyzing the different years separately, we observe cases in which the average number of patent applications by or patents granted to female inventors exceeded that for male inventors in some technological fields (examples are available upon request).[16]

4.4.2 How does patenting activity differ between female and male inventors?

In this section, we study how men and women differ in their patenting activity. First, we start with a table of descriptive statistics for patent characteristics by sex. Then, we proceed with the number of star inventors[17] and share of top-coded inventors. Finally, we dig deeper and see how different variables such as gender, age, and migration status contribute to the diversity in inventor and patent characteristics (we include, among others, the probability of an inventor's patent application being

[15] Table 4.31 in the appendix shows the share of patenting activity in different technological areas according to the inventors' education by sex.

[16] In the main specification, we do not take into account inventor group size, as several inventors often apply for patents together. In figures 4.15 and 4.16, we show a graph that accounts for team size. For instance, we assign an inventor only 0.5 "patent points" if two inventors file a patent together. The overall numbers are smaller by construction, but the pattern is the same for both definitions.

[17] We define star inventors as inventors with more than 15 granted patents who are therefore among the top 5% of all inventors in terms of number of patents granted.

rejected, the number of patents granted, and the number of forward citations) with OLS regressions.

In table 4.3, we provide an overview of patents and citations by sex. Male inventors have, on average, more granted patents than female inventors (1.74 vs. 1.07) and apply for more patents (5.33 vs. 3.78). The number of rejected patents is also higher for men (3.38 vs. 2.59). Forward citations are an indicator of the novelty and impact of a patent. This indicator counts the number of forward citations that the invention (DOCDB family) received from patent applications at the German/EU/US Patent and Trademark Offices after 2 or 10 years from the earliest publication date. The table reveals that men are cited more frequently than women in Germany, but less at the European and the US patent offices. This pattern holds irrespective of the time period (2 or 10 years after patent grant). Tables 4.19 and 4.20 in the appendix show the mean number of patents for inventors who were never or ever top-coded by sex. Tables 4.28 and 4.29 contain the mean number of patents per inventor in former East Germany and former West Germany.

Analyzing male and female inventors in former West Germany and former East Germany separately reveals interesting similarities and differences. Tables 4.28 and 4.29 show that the gender gap in the number of patents applied for and granted is almost the same in the two regions. Additionally, the number of patents granted to male inventors is almost the same in former East Germany and former West Germany, as is the number of patents granted to female inventors. However, the number of patent applications and rejections and the total number of patent citations are higher in former West Germany than in former East Germany for both men and women.

Figure 4.8 presents the number of star inventors by sex and the share of women separately for each year. We define star inventors as those inventors with at least 15 granted patents (and, thus, those who are among the top 5% of inventors in terms of granted patents). As the number of granted patents is likely a good proxy for productivity, this is a very interesting measure in regard to differences between women and men. When looking at the absolute numbers, we see that female star inventors are rare. Over the whole period, there are 4,324 unique male and 77 unique female star inventors. Star inventors are somewhat older than other inventors among both men and women and somewhat younger among women than among men (table 4.21). Importantly, female inventors are more likely to work part-time than male inventors, and women are less likely to work in technological areas with high levels of patenting activity (see figure 4.7). Therefore, the smaller number of female star inventors is not surprising. The share of female star inventors increased from 0.01 in 2000 to 0.02 in 2010.

Figure 4.9 shows how the share of top-coded employees among inventors and among a random sample of employees in Germany developed over time by sex. Note that daily wages above the German social security limit are censored in the INV-BIO dataset and the random two-percent sample of workers in Germany (SIAB). Therefore, the share of top-coded inventors is an indicator for highly paid inventors and could be an indicator for work quality if we think of wages as reflecting the marginal product of work. Figure 4.9 offers three takeaways. First, the share of top-coded workers is higher among inventors than among the random sample of employees in Germany. Second, the graphs show that the share of top-coded inventors is lower for women than for men. This gender gap is also observable among noninventors. Third, while the top-coded share among inventors has decreased over time, no such trend is identifiable among noninventors. What are the possible explanations for these findings? First, the payment of bonuses is more common in the R&D sector than in other sectors in Germany, and the decrease in the share of top-coded inventors could be due to the trend of paying bonuses for successes rather than increasing regular wages.[18] Second, we find a rather strong decrease in the top-coding of inventors after the financial crisis in 2007/2008. One potential explanation is that the financial crisis hit the innovation sector (and expenditures in the risky R&D sector) more severely than other sectors or the economy as a whole. Last, inventors are a special, highly educated, and selected group. Therefore, having higher shares of top-coding among inventors than among a random sample of employees should be no surprise.

Table 4.4 shows the estimates from an OLS regression with the number of patents applications as the dependent variable and different control variables. We cannot find any statistically significant relationship between the number of patent applications per inventor per year and being a woman. Column (1) reveals that women apply for fewer patents per year than men; however, the coefficient of interest is insignificant and thus must be interpreted with caution. In column (2), the coefficient changes its sign when controlling for education and migration status. Remarkably, foreign inventors also apply for fewer patents. When including different fixed ef-

[18]Bonuses are not reported in the social security data, only daily wages.

fects (establishment size, industry, year, and age group), the coefficient of interest increases considerably. Column (7) identifies whether women are mothers. Interestingly, mothers apply for more patents than other women, although the difference is not statistically significant. Therefore, we do not find evidence that motherhood is related to having fewer patent applications granted. Table 4.34 in the appendix presents the results for the cumulative number of patent applications. We return to the analysis of the relationship between motherhood and patenting in section 4.4.3.

Table 4.5 shows that the pattern we find in table 4.4 is stronger and more statistically significant for granted patents: women are granted fewer patents per year even after including numerous controls in the OLS regression (column (6)). There is a small and statistically insignificant positive correlation between the number of granted patents and motherhood, as shown in table 4.4. Table 4.35 in the appendix presents the results for the cumulative number of patents granted across all years.

Table 4.6 presents the results of an OLS regression with the probability of rejection when applying for a patent as the dependent variable. Columns (1)-(7) include different controls in a stepwise manner. Column (1) includes only a dummy for being female; we see that female inventors' patent applications are more likely to be rejected. This is a somewhat surprising finding given that previous research on gender differences in competition has found that women are more likely to avoid competition (Niederle and Vesterlund [2007], Buser et al. [2014]). If women tend to shy away from applying for patents, then we would expect women's applications to be less likely to be rejected.

Column (2) further includes a dummy for being of foreign origin as well as educational dummies. Interestingly, patent applications by inventors of foreign origin are less likely to be rejected. One possible explanation for this is positive self-selection, i.e., if highly productive foreign inventors decide to come to Germany to work and patent. Furthermore, we see that the patents of inventors with university degrees are more likely to be rejected. This suggests that those inventors with no university degree are strongly (self-)selected. There are two possible explanations for this: First, this result is also in line with research based on the idea of being "on the shoulders of giants". Most likely, inventors with no university degree are likely to patent in areas that are more innovative and less specialized (and thus need less prior knowledge). Second, it could be that inventors without a university degree are more careful about not submitting an application until they are confident that it will be accepted. Including establishment, industry, and year fixed effects decreases the coefficient of interest only slightly. However, including age fixed effects (column (6)) leads to a decrease in the coefficient on the female dummy by approximately one-third, indicating that age groups explain part of the difference in female and male rejection rates. In column (7), we further control for motherhood to determine whether being a mother negatively contributes to the rejection of patents; however, the coefficient is insignificant. Overall, we find that women do not shy away from applying for patents. Rather, they submit applications that are rejected more often than those of men even after controlling for industry, establishment size, and year fixed effects.[19]

4.4.3 How often are the patents of female and male inventors cited?

Table 4.7 presents the estimates from an OLS regression of the number of forward citations that the invention (DOCDB family) received in patent applications at the German Patent and Trademark Office 2 years after its earliest publication date, separately for each patent, on different control variables. There is a truncation of count variables at the 99th percentile of the citations distribution for each year and patent. Forward citations are regarded as reflecting the technological relevance of a patent for later developments (e.g., Trajtenberg [1990]) and are therefore a good proxy for productivity and importance. The dependent variable is the average number of citations for each inventor across all patents. Column (1) indicates a negative coefficient, meaning that women have on average fewer citations in Germany within 2 years. Differences in establishment size, industry, and age group in addition to education and migration status can explain nearly half of the gap between male and female inventors. However, the remaining difference is still statistically significant. Again, column (7) shows that mothers receive more citations than other female inventors. Summing the coefficients on the dummy for being woman and the dummy for being a mother shows that mothers have more citations than men after including other controls, while women who are not mothers have fewer citations.

Tables 4.8 and 4.9 contain the estimation results for patent citations within 2 years

[19] Tables 4.36 and 4.37 in the appendix present the results for the cumulative number and the total number of rejected patents per year.

in Europe and the US. Unlike the results in table 4.7, we see a strong positive coefficient indicating that female inventors are cited more frequently in Europe and the US than male inventors. This is especially true for inventors who are mothers (column (7)). Table 4.22, 4.23 and 4.24 in the appendix show that the same pattern is also observable when looking at patent citations within 10 years for Germany, Europe, and the US, respectively. Why are female inventors cited more frequently in Europe and the US but less frequently in Germany? There are four possible explanations for this last result: (i) Women tend to work in teams rather than alone (Jung and Ejermo [2014], Mauleón and Bordons [2010], Naldi et al. [2004]) and therefore profit from the larger international networks of larger teams and are therefore cited more often, (ii) women who are granted international patents are a highly productive (and selected) group of high-ability women since international patents are harder to obtain; probably this group of women is a stronger self-selected group than internationally patenting male inventors, (iii) countries other than Germany may be more progressive such that women experience less discrimination in countries with higher shares of female inventors, and (iv) even though we control for different industries in the regression by using fixed effects, different technological fields in patenting could still play a role here.[20] Women tend to work more often in biotechnology and related fields that are relatively stronger in many other countries than Germany, while men work in machinery and mechatronics, where Germany is among the leading countries. For this reason, women's patents may be relatively more relevant for research done abroad and men's patents for research done in Germany.

4.4.4 Inventing and motherhood

Even though gender inequalities in the labor market resulting from parenthood have decreased in recent decades, there is still substantial gender inequality in all countries. For instance, Kleven et al. [2019] find substantial labor market differences after childbirth for Sweden, Denmark, Austria, Germany, the UK, and the US. They conclude that childcare policies (see, e.g., Olivetti and Petrongolo [2017] for a review

[20] Tables 4.38 and 4.39 in the appendix support this conjecture. The tables show the last two columns of tables 4.7/4.8/4.9 but with technology area instead of industry fixed effects. The estimated negative coefficient for women for citations in Germany reduces remarkably once technology area fixed effects are included as controls in columns (1) and (2) of table 4.38, while the estimated positive coefficient for women goes close to zero in table 4.38 (columns (3) and (4)) and switches sign to negative (but statistically insignificant) in table 4.38 (columns (5) and (6)) and in table 4.39 (column(3)-column(6)).

and Collischon et al. [2022] for a recent application to the German context) only partly explain the gap in labor market outcomes between men and women in different countries. Another possible explanation is different gender norms and cultures, but the literature offers little causal evidence on this mechanism (see Bertrand [2009] for a review and Collischon et al. [2020] for a study in Germany). In light of this gender gap, we next focus on the consequences of motherhood for female inventors. We start by investigating whether female inventors differ from female noninventors in their likelihood of having a child. We then go one step further by exploring whether the likelihood of having a child differs among female inventors according to, e.g., the number of patents that they are granted. This last subsection examines differences in parental leave durations, labor market leaves, and returns to full-time or part-time work after giving birth.

First, table 4.10 describes the likelihood of female inventors having a child relative to a randomly selected group of women in the German labor market. To answer this question, we combine the INV-BIO dataset with the "Sample of Integrated Labour Market Biographies" (SIAB) dataset[21], a random 2-percent sample of all workers in Germany. We identify mothers in the manner first proposed by Müller and Strauch [2017][22], in which information on indirect identifiers specifying the reason for cancellation/notification/termination of an employment or unemployment spell are used to identify family-related breaks in the administrative data. We focus on

[21]The Sample of Integrated Labour Market Biographies (SIAB) is a 2 percent random sample drawn from the Integrated Employment Biographies (IEB) of the Institute for Employment Research (IAB). The IEB makes it possible to track the employment status of a person to the exact day. The IEB consists of all individuals in Germany who are characterized by at least one of the following employment statuses: employment subject to social security (in the data since 1975), marginal part-time employment (in the data since 1999), benefit receipt according to the German Social Code III or II (SGB III since 1975, SGB II since 2005), officially registered as job-seeking at the German Federal Employment Agency or (planned) participation in programs of active labor market policies (in the data since 2000). These data, which come from different sources, are merged in the IEB. This information stems from the following source, which contains even more in-depth details: https://fdz.iab.de/en/FDZ_Individual_Data/integrated_labor_market_biographies.aspx [last access: 12-14-2021]

[22]Administrative data from the Federal Employment Agency are an important database for labor market research. However, the Federal Employment Agency's tasks restrict what kind of information is recorded. E.g. due to data security reasons, some information is missing that is relevant to various research objectives. One example is information on childbirth, which is necessary for analyzing female employment biographies. It is still mainly mothers who have an employment break to care for children. Müller and Strauch [2017] present one way to identify family-related breaks using indirect identifiers in the administrative data. This information stems from the following source, which contains even more in-depth details: https://fdz.iab.de/187/section.aspx/Publikation/k171220304 [last access: 12-14-2021]

women here, as information on male inventors who are fathers is not identifiable in the dataset. Table 4.10 reveals that female inventors are less likely to become mothers even after controlling for education, migration status, and different fixed effects (establishment size, industry, year, age cohort). One possible explanation is that working as an inventor is not compatible with family life, so women may decide a priori not to become inventors or, if they do, not to have children.

We next analyze how the decision to have children and patenting activity are related among female inventors and, thus, among women who self-select into the R&D sector. Table 4.11 and 4.12 show how the likelihood of becoming a mother varies with different characteristics, especially the cumulative number of granted patents and citations in Germany per year prior to giving birth. For women who are never identified as mothers during the observation period, we count the cumulative number of patents/citations over the years before the age of 40. This is in line with our identification of mothers, since we restrict our sample to women who give birth to their first child before the age of 40. The estimates indicate that female inventors with more cumulative granted patents and citations in Germany are more likely to become mothers. This shows that either highly productive inventors are more likely to become mothers or that female inventors first decide to achieve their professional goals before starting a family.

Tables 4.13 and 4.14 focus on the period after giving birth. While table 4.13 displays the gap in the number of days before working again after giving birth and how it differs between female inventors and noninventors, table 4.14 estimates the likelihood of not returning to the labor market after giving birth. Table 4.13 reveals that compared with a random sample of women in the German labor market (as in table 4.10), female inventors take longer parental leaves after childbirth. Column (7) focuses on female inventors only, but no significant relationship between the gap after giving birth and inventor productivity (as measured by the number of granted patents prior to giving birth) is found. What lessons can we take away from table 4.14? The results suggest that female inventors are less likely to completely exit the labor market after giving birth when they have a higher number of granted patents (see table 4.14 (column(7)). However, table 4.14 reveals that among females in general, female inventors are more likely to completely exit the German labor market after childbirth than women in the random sample. One possible explanation for this result is that working as an inventor is a full-time job and that women face problems when re-entering the labor market (e.g., because they are assigned tasks other than innovation due to their part-time employment). Another explanation is that taking a longer leave after childbirth and even exiting the labor market completely might be more affordable for female inventors than for the "average" woman in the German labor market.

Table 4.15 (table 4.16) displays the results of a test for whether female inventors are more or less likely to return to full-time (part-time) employment than randomly selected women in the German labor market. We find that female inventors are less likely to return to their job after childbirth in either a part-time or a full-time capacity than female noninventors. In line with the reasoning that inventors mostly have full-time positions, we see that the likelihood of returning in a part-time capacity is lower for female inventors than the likelihood of returning to full-time employment.

4.5 Conclusion

In this paper, we address a topic that has received little attention thus far: female inventors. The literature on who becomes an inventor is still in its infancy, and we contribute to this literature by providing the first evidence on female inventors in Germany. For our empirical analysis, we use a newly available dataset (the INV-BIO dataset) that combines patent office data with highly reliable administrative data from Germany. We show that inventors, irrespective of their sex, on average earn higher wages and are better educated than a random sample of employees in Germany. However, female inventors earn lower wages than their male colleagues. We further find evidence for the widely discussed "leaky pipeline": the share of women decreases dramatically from the bachelor's to the master's to the Ph.D. level. The share of women among inventors is 11% lower than the share of women among all researchers. These women tend to work in fields such as pharmaceuticals and biotechnology rather than mechatronics. Additionally, women tend to work in fields with fewer total patents. When looking at patent characteristics, the results reveal that female inventors have, on average, fewer patent applications, citations, and rejections. Female inventors are also less frequently cited, and star inventors make up a smaller share of female inventors than of male inventors. In line with this result, we also find that female inventors' wages are also top-coded less often.

We further find that the probability that a female inventor's patent is rejected is higher than for male inventors after accounting for different personal and establishment characteristics. We find mixed results for the number of citations: in Germany, female inventors are cited less frequently, while in the European and US patent offices, they have a higher number of citations per inventor. Another important issue when investigating female inventors is the role of motherhood. Generally, female inventors are less likely to become mothers than the members of a random sample of women in Germany ("average" female workers in the German labor market). Among inventors, it is the highly productive women (in terms of granted patents/citations in Germany prior to giving birth) who are most likely to become mothers. Our results further show that inventors take longer parental leaves and are more likely to exit the workforce after giving birth than "average" women. Inventor mothers are also less likely to return to work part-time (full-time) than "average" mothers in Germany.

Since female inventors exit the labor market more often and are less likely to work part-time after giving birth, there is likely potential to create a more family-friendly environment for female inventors. Making childcare easily available and helping female inventors establish work-life balance (e.g., by allowing home office work whenever possible) could be a tool to encourage female participation in the innovation sector. However, as prior findings suggest, policymakers need to adopt a complete set of measures that target earlier stages in life, such as mentoring programs in school or advertising positive role models such as BioNTech founder Özlem Türeci, if they want to avoid the "leaky pipeline" in the future (see, among others, Breda et al. [2020], Buckles [2019], Canaan and Mouganie [2021], Carrell et al. [2010], Mansour et al. [2021]). Evaluating benefits and costs of these measures is an important question for future research.

Table 4.1: Summary Statistics (2000-2010)

	Men		Women	
	(1)	(2)	(3)	(4)
	Inventors	2 percent random sample of workers in Germany	Inventors	2 percent random sample of workers in Germany
Number of workers	133,396	939,347	11,005	818,483
Age				
Mean	42.23	37.70	37.68	37.96
Median	41	37	36	38
Age (2000)				
Mean	41.27	37.45	36.94	37.77
Median	40	36	35	37
Age (2005)				
Mean	42.20	37.47	37.62	37.58
Median	41	37	36	38
Age (2010)				
Mean	43.83	38.14	39.07	38.4
Median	43	38	39	39
Daily wage				
Mean	155.62	58.01	134.78	37.27
Median	167.67	47.41	145.65	24.95
Education (imputed)				
No vocational training	0.03	0.48	0.04	0.50
Vocational training	0.27	0.45	0.31	0.45
University	0.69	0.07	0.64	0.05
Share of ever unemployed during 2000-2010	0.02	0.34	0.04	0.37

Notes: The daily wage is a record of the employee's gross daily wage. It is calculated by using the fixed-period wages reported by the employer and the duration of the (unsplit) original notification period in calendar days. Daily wages are reported in euros. Earnings exceeding the upper earnings limit for statutory pension insurance are reported only up to this limit. If an inventor (or a worker within the random 2-percent sample) is unemployed, the daily wage variable records the daily benefit rate converted into euros. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. Data source: INV-BIO.

Table 4.2: Age structure of inventors in 2000, 2005, and 2010

	2000		2005		2010	
	Women	Men	Women	Men	Women	Men
	Percentage					
Below 30	6	12	5	14	5	14
31-40	44	57	39	52	28	41
41-50	30	24	36	24	41	33
51-60	16	6	16	8	20	11
Above 60	4	0	4	1	6	1
Total Number	3404	61545	4267	69883	3180	56361

Notes: Table 4.2 shows the percentage of male and female inventors by age group in 2000, 2005, and 2010. Data source: INV-BIO.

Table 4.3: Average number of patents and citations per inventor (2000-2010), by sex

	Men (N=133,213)				Women (N=10,974)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Patent applications per year (2000-2010)	1.71	2.16	0	92	1.53	2.00	0	59
Patents granted per year (2000-2010)	0.61	1.03	0	31	0.46	0.82	0	11
Patents rejected per year (2000-2010)	1.11	1.84	0	82	1.06	1.80	0	54
Patent applications per year (2000-2010, considering team size)	0.43	0.45	0	14	0.33	0.39	0	21
Patents granted per year (2000-2010, considering team size)	0.16	0.29	0	6	0.11	0.20	0	4
Patents rejected per year (2000-2010, considering team size)	0.26	0.37	0	14	0.22	0.36	0	21
Patent applications (cumulative)	5.33	9.29	1	572	3.78	7.32	1	227
Patents granted (cumulative)	1.74	3.33	0	120	1.07	2.18	0	49
Patents rejected (cumulative)	3.38	7.24	0	529	2.59	6.06	0	178
Patent applications (cumulative, considering team size)	1.07	1.37	0	48	0.66	0.92	0	25
Patents granted (cumulative, considering team size)	0.40	0.67	0	24	0.21	0.36	0	10
Patents rejected (cumulative, considering team size)	0.67	1.04	0	44	0.46	0.79	0	25
Total forward citations in Germany within 2 years	1.10	3.40	0	190	0.62	2.33	0	90
Total forward citations in Europe within 2 years	1.62	5.16	0	294	1.79	6.88	0	321
Total forward citations in the US within 2 years	5.97	23.19	0	1157	6.05	24.55	0	1161
Total forward citations in Germany within 10 years	7.03	17.66	0	878	3.72	11.41	0	354
Total forward citations in Europe within 10 years	10.32	26.13	0	974	10.51	30.83	0	1398
Total forward citations in the US within 10 years	24.80	75.00	0	4008	21.35	69.10	0	2947

Notes: The table shows the average inventors' patent characteristics in the sample in 2000-2010, by gender. An inventor who is observed in more than one year is included only in the latest year. Granted (applied, rejected) patents gives the average number among all inventors. Granted / applied / rejected (considering team size) patents gives the average number among all inventors, when accounting for the number of inventors that cooperated for a patent. Total forward citations in Germany (Europe, USA) gives the number of citations of (applied and/or granted) patents. Count variables for each year are truncated at the 99th percentile of the distribution of citations. Respectively within 2 or 10 years after the first application. Data source: INV-BIO.

Table 4.4: OLS Regression – Number of patent applications per inventor per year

	(1) # applications	(2) # applications	(3) # applications	(4) # applications	(5) # applications	(6) # applications	(7) # applications
Woman	-0.0543	0.0573	0.0754	0.0781	0.275	0.0889	0.0855
	(0.325)	(0.323)	(0.330)	(0.331)	(0.314)	(0.312)	(0.322)
No voc. training		1.495***	1.596***	1.605***	1.658***	1.352**	1.352**
		(0.439)	(0.453)	(0.452)	(0.455)	(0.442)	(0.442)
University		1.569***	1.524***	1.525***	1.491***	1.191***	1.191***
		(0.116)	(0.116)	(0.117)	(0.109)	(0.101)	(0.101)
Foreign origin		-0.240	-0.209	-0.202	-0.155	-0.330	-0.330
		(0.226)	(0.225)	(0.225)	(0.219)	(0.218)	(0.218)
Mother							0.0784
							(0.878)
Constant	4.303***	3.103***	2.308	2.236	16.18***	16.61***	2.168
	(0.0796)	(0.0617)	(1.533)	(1.533)	(4.283)	(4.236)	(1.314)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	770715	770668	770668	770668	765799	765799	765796
Adjusted R^2	0.000	0.011	0.012	0.013	0.037	0.060	0.060

Notes: Table 4.4 displays the results of an OLS regression indicating correlations between the number of patent applications per year and different control variables. These control variables include dummy variables for being female, of foreign origin, or a mother and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Patent applications as a whole are analyzed, independent of the patent office (Germany, Europe). Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.5: OLS Regression - Number patents granted per inventor per year

	(1) # granted	(2) # granted	(3) # granted	(4) # granted	(5) # granted	(6) # granted	(7) # granted
Woman	-0.309***	-0.294***	-0.296***	-0.278***	-0.215***	-0.217***	-0.221***
	(0.0524)	(0.0526)	(0.0542)	(0.0532)	(0.0436)	(0.0437)	(0.0447)
No voc. training		0.424**	0.439**	0.486**	0.462**	0.453**	0.453**
		(0.155)	(0.158)	(0.158)	(0.159)	(0.157)	(0.157)
University		0.248***	0.235***	0.236***	0.208***	0.200***	0.200***
		(0.0264)	(0.0263)	(0.0260)	(0.0250)	(0.0251)	(0.0251)
Foreign origin		-0.0113	-0.00771	0.0211	0.0188	0.00956	0.00930
		(0.0610)	(0.0614)	(0.0605)	(0.0586)	(0.0582)	(0.0582)
Mother							0.0873
							(0.177)
Constant	1.212***	1.015***	2.195*	1.589	6.113***	6.090***	1.620
	(0.0160)	(0.0192)	(0.989)	(1.010)	(1.506)	(1.506)	(1.033)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	770715	770668	770668	770668	765799	765799	765796
Adjusted R^2	0.001	0.004	0.005	0.022	0.041	0.042	0.042

Notes: Table 4.5 shows the estimates from an OLS regression of the number of patents granted per year as the dependent variable on different control variables. These control variables include dummy variables for being female, of foreign origin, or a mother and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Patent grants as a whole are analyzed, independent of the patent office (Germany, Europe). Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.6: Logit Regression - Probability of rejection

	(1) Rejection patent	(2) Rejection patent	(3) Rejection patent	(4) Rejection patent	(5) Rejection patent	(6) Rejection patent	(7) Rejection patent
Woman	0.148***	0.166***	0.144***	0.134***	0.163***	0.114***	0.114***
	(0.0222)	(0.0220)	(0.0221)	(0.0224)	(0.0226)	(0.0217)	(0.0221)
Foreign origin		-0.0220	-0.0238	-0.0422	-0.0182	-0.0602*	-0.0603*
		(0.0242)	(0.0243)	(0.0243)	(0.0240)	(0.0237)	(0.0237)
No voc. training		0.175***	0.162***	0.134***	0.165***	0.0831**	0.0832**
		(0.0320)	(0.0322)	(0.0314)	(0.0319)	(0.0315)	(0.0315)
University		0.248***	0.228***	0.229***	0.237***	0.152***	0.152***
		(0.0118)	(0.0118)	(0.0119)	(0.0121)	(0.0117)	(0.0117)
Mother							0.0183
							(0.0857)
Constant	0.488***	0.301***	-0.947***	-0.560*	-0.910***	-0.113	-0.113
	(0.00565)	(0.00992)	(0.254)	(0.260)	(0.255)	(0.270)	(0.270)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	770715	770668	770668	770668	765799	765799	765796

Notes: Table 4.6 displays the results of a logit regression indicating the correlations between the probability of a patent being rejected and different control variables. These control variables include dummy variables for being female, of foreign origin, or a mother and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Patent rejections as a whole are analyzed, independent of the patent office (Germany, Europe).Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.7: OLS Regression - Number of total forward citations per inventor in Germany within 2 years

	(1) # citations GER	(2) # citations GER	(3) # citations GER	(4) # citations GER	(5) # citations GER	(6) # citations GER	(7) # citations GER
Woman	-2.151***	-1.959***	-1.787***	-1.813***	-0.971***	-1.192***	-1.304***
	(0.303)	(0.296)	(0.300)	(0.302)	(0.286)	(0.290)	(0.278)
No voc. training		2.461	2.841*	2.781*	2.507	2.106	2.110
		(1.340)	(1.381)	(1.375)	(1.375)	(1.351)	(1.351)
University		2.062***	2.017***	2.016***	1.774***	1.361***	1.360***
		(0.240)	(0.234)	(0.234)	(0.224)	(0.211)	(0.211)
Foreign origin		-1.449***	-1.311***	-1.354***	-1.180***	-1.396***	-1.403***
		(0.318)	(0.318)	(0.320)	(0.313)	(0.325)	(0.325)
Mother							2.446
							(1.502)
Constant	4.922***	3.379***	-1.196	-0.712	-3.065*	-3.021*	-1.808
	(0.159)	(0.151)	(1.269)	(1.236)	(1.321)	(1.371)	(1.433)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	770715	770668	770668	770668	765799	765799	765796
Adjusted R^2	0.002	0.009	0.012	0.013	0.036	0.048	0.048

Notes: Table 4.7 shows an OLS regression indicating relationships between the number of citations per inventor across all of his / her patents in the German Patent and Trademark Office after 2 years and different control variables. Number of forward citations that the invention(s) received from patent applications at the German Patent and Trademark Office after 2 years from the earliest publication date, respectively for each patent, is taken for calculating the overall number of citations per inventor. Count variables for each year are truncated at the 99th percentile of the citation distribution. The dependent variable is the mean number of citations across all patents per inventor. The control variables include dummy variables for being female, of foreign origin, or a mother and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 (the reference category) comprises those who have completed in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.8: OLS Regression - Number of total forward citations per inventor in Europe within 2 years

	(1) # citations EUR	(2) # citations EUR	(3) # citations EUR	(4) # citations EUR	(5) # citations EUR	(6) # citations EUR	(7) # citations EUR
Woman	4.012*	4.313*	4.824*	4.836*	4.180*	3.584	3.212
	(1.985)	(1.993)	(2.045)	(2.043)	(1.983)	(1.934)	(1.956)
No voc. training		1.416	2.450*	2.519*	3.202**	2.222*	2.236*
		(1.095)	(1.122)	(1.122)	(1.139)	(1.101)	(1.100)
University		4.413***	4.381***	4.384***	4.606***	3.613***	3.609***
		(0.460)	(0.475)	(0.475)	(0.457)	(0.391)	(0.391)
Foreign origin		0.108	0.464	0.483	0.302	-0.222	-0.247
		(0.989)	(0.975)	(0.972)	(0.949)	(0.966)	(0.968)
Mother							8.086
							(8.615)
Constant	7.617***	4.292***	1.782	-0.0223	79.32***	80.38***	-2.228
	(0.287)	(0.259)	(2.386)	(2.354)	(23.89)	(23.85)	(1.713)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	770715	770668	770668	770668	765799	765799	765796
Adjusted R^2	0.002	0.012	0.016	0.018	0.071	0.095	0.095

Notes: Table 4.8 shows an OLS regression indicating relationships between the number of citations per inventor across all of his / her patents in the European Patent and Trademark Office after 2 years and different control variables. Number of forward citations that the invention(s) received from patent applications at the European Patent and Trademark Office after 2 years from the earliest publication date, respectively for each patent, is taken for calculating the overall number of citations per inventor. Count variables for each year are truncated at the 99th percentile of the citation distribution. The dependent variable is the mean number of citations across all patents per inventor. The control variables include dummy variables for being female, of foreign origin, or a mother and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 (the reference category) comprises those who have completed in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.9: OLS Regression - Number of total forward citations per inventor in the US within 2 years

	(1) # citations US	(2) # citations US	(3) # citations US	(4) # citations US	(5) # citations US	(6) # citations US	(7) # citations US
Woman	5.744	6.969	8.098	8.146	7.597	6.082	5.241
	(6.579)	(6.604)	(6.828)	(6.827)	(6.778)	(6.639)	(6.818)
No voc. training		3.062	5.967***	6.276***	7.572***	5.006**	5.037**
		(1.668)	(1.773)	(1.785)	(1.860)	(1.882)	(1.882)
University		18.63***	18.19***	18.20***	18.45***	15.74***	15.74***
		(1.583)	(1.601)	(1.598)	(1.587)	(1.417)	(1.418)
Foreign origin		2.004	2.949	3.031	2.877	1.657	1.601
		(3.037)	(3.032)	(3.032)	(3.061)	(3.139)	(3.138)
Mother							18.29
							(22.00)
Constant	27.74***	13.73***	10.02	1.118	96.59***	97.49***	-6.169
	(0.994)	(0.876)	(8.149)	(8.073)	(28.33)	(28.14)	(6.392)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	770715	770668	770668	770668	765799	765799	765796
Adjusted R^2	0.000	0.011	0.015	0.017	0.025	0.036	0.036

Notes: Table 4.9 shows an OLS regression indicating relationships between the number of citations per inventor across all of his / her patents in the US Patent and Trademark Office after 2 years and different control variables. Number of forward citations that the invention(s) received from patent applications at the US Patent and Trademark Office after 2 years from the earliest publication date, respectively for each patent, is taken for calculating the overall number of citations per inventor. Count variables for each year are truncated at the 99th percentile of the citation distribution. The dependent variable is the mean number of citations across all patents per inventor. The control variables include dummy variables for being female, of foreign origin, or a mother and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 (the reference category) comprises those who have completed in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.10: Logit Regression - Likelihood of having a child

	(1) Likelihood having a child	(2) Likelihood having a child	(3) Likelihood having a child	(4) Likelihood having a child	(5) Likelihood having a child	(6) Likelihood having a child
Inventor	-4.980***	-4.730***	-5.016***	-5.006***	-5.006***	-5.123***
	(0.0841)	(0.0842)	(0.0842)	(0.0842)	(0.0842)	(0.0847)
No voc. training		-0.401***	-0.191***	-0.177***	-0.177***	-0.182***
		(0.00664)	(0.00715)	(0.00721)	(0.00721)	(0.00853)
University		-0.505***	-0.734***	-0.738***	-0.738***	-0.666***
		(0.0129)	(0.0135)	(0.0136)	(0.0136)	(0.0136)
Foreign origin		-0.438***	-0.575***	-0.612***	-0.612***	-0.538***
		(0.0152)	(0.0164)	(0.0167)	(0.0167)	(0.0162)
Constant	-0.734***	-0.608***	-6.131***	-5.946***	-5.946***	-5.943***
	(0.00419)	(0.00454)	(0.168)	(0.168)	(0.168)	(0.141)
Age FE			Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes
Industry FE					Yes	Yes
Establ. size FE						Yes
Observations	13276121	13224476	13224476	13224476	13224476	7289475

Notes: Table 4.10 shows a logit regression indicating relationships between the likelihood of having a child and different control variables. The dependent variable is the likelihood of having a child. The control variables include: dummy variables of being female / of foreign origin / mother, and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the individual level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.11: Logit Regression - Likelihood of becoming a mother

	(1) Likelihood becoming a mother	(2) Likelihood becoming a mother	(3) Likelihood becoming a mother	(4) Likelihood becoming a mother	(5) Likelihood becoming a mother	(6) Likelihood becoming a mother
# gran. patents pr.	0.0422*	0.0415*	0.0631**	0.0616**	0.0580**	0.0601**
birth	(0.0217)	(0.0219)	(0.0260)	(0.0251)	(0.0245)	(0.0237)
No voc. training		-1.5024***	-1.2532***	-1.2904***	-1.2464***	-1.2821***
		(0.4485)	(0.4589)	(0.4629)	(0.4561)	(0.4575)
University		0.2548	0.2216	0.2083	0.1967	0.1650
		(0.1694)	(0.1736)	(0.1728)	(0.1774)	(0.1771)
Foreign origin		0.2208	0.2247	0.1720	0.1793	0.1691
		(0.2519)	(0.2496)	(0.2459)	(0.2477)	(0.2512)
Constant	-3.6355***	-3.8075***	-13.2166***	-13.7493***	-13.0824***	-12.5424***
	(0.0723)	(0.1311)	(0.6489)	(0.4383)	(0.8926)	(0.8043)
Age FE			Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes
Industry FE					Yes	Yes
Establ. size FE						Yes
Observations	28696	28696	28421	28421	28416	28416

Notes: Table 4.11 displays the results of a logit regression indicating the relationships between the likelihood of becoming a mother and different control variables. The dependent variable is the likelihood of becoming a mother. For female inventors who are not mothers, the cumulative number of patents granted per year before the age of 40 are counted. The control variables include dummy variables for being female or of foreign origin and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger and 30-39. Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.12: Logit Regression - Likelihood of becoming a mother

	(1) Likelihood becoming a mother	(2) Likelihood becoming a mother	(3) Likelihood becoming a mother	(4) Likelihood becoming a mother	(5) Likelihood becoming a mother	(6) Likelihood becoming a mother
# citations in GER	0.0299*	0.0285	0.0239	0.0224	0.0196	0.0199
pr. birth	(0.0182)	(0.0181)	(0.0184)	(0.0186)	(0.0190)	(0.0189)
No voc. training		-1.5747***	-1.2591***	-1.2969***	-1.2737***	-1.3004***
		(0.4478)	(0.4485)	(0.4519)	(0.4508)	(0.4507)
University		0.2293	0.1738	0.1634	0.1555	0.1275
		(0.1686)	(0.1672)	(0.1679)	(0.1731)	(0.1736)
Foreign origin		0.2055	0.2304	0.1691	0.1758	0.1591
		(0.2529)	(0.2521)	(0.2502)	(0.2520)	(0.2564)
Constant	-3.5443***	-3.6927***	-13.3547***	-13.6631***	-13.7405***	-13.1885***
	(0.0880)	(0.1263)	(0.5897)	(0.8047)	(0.8446)	(0.7647)
Age FE			Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes
Industry FE					Yes	Yes
Establ. size FE						Yes
Observations	28696	28696	28421	28421	28416	28416

Notes: Table 4.12 displays the results of a logit regression indicating the relationships between the likelihood of becoming a mother and different control variables. The dependent variable is the likelihood of becoming a mother. For female inventors who are not mothers, the cumulative number of patents granted per year before the age of 40 are counted. The control variables include dummy variables for being female or of foreign origin and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger and 30-39. Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.13: OLS Regression - Parental leave duration

	(1) Gap after giving birth	(2) Gap after giving birth	(3) Gap after giving birth	(4) Gap after giving birth	(5) Gap after giving birth	(6) Gap after giving birth	(7) Gap after giving birth
Inventor	218.3***	288.5***	275.7***	253.0***	269.0***	281.4***	0
	(39.80)	(39.91)	(39.91)	(38.32)	(39.08)	(39.23)	(.)
No voc. training		-70.72***	-65.93***	-72.11***	-26.28***	-20.54***	41.68
		(6.334)	(6.340)	(6.164)	(6.132)	(6.138)	(165.6)
University		-117.3***	-107.1***	-79.51***	-43.68***	-36.48***	-93.32
		(5.706)	(5.786)	(5.701)	(5.670)	(5.716)	(104.3)
Foreign origin			29.02***	31.48***	18.22*	21.48**	-84.42
			(7.754)	(7.481)	(7.265)	(7.256)	(114.2)
# gran. patents pr. birth							-23.35
							(13.41)
Constant	391.8***	406.9***	325.6***	113.5**	361.4***	282.2***	-365.9
	(1.924)	(2.154)	(36.31)	(36.41)	(45.47)	(46.19)	(233.1)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	38415	38415	38408	38408	32764	32764	209
Adjusted R^2	0.002	0.011	0.028	0.091	0.077	0.081	0.117

Notes: Table 4.13 displays the results of an OLS regression indicating the relationships between the parental leave duration of mothers and different control variables, conditional on being a mother. The dependent variable is the duration of parental leave after giving birth among mothers in days. The coefficient of interest is that on "Inventor". The control variables include dummy variables for being an inventor or of foreign origin and education. Education is split into 3 categories. Category 1 includes all mothers who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other mothers with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the individual level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.14: Logit Regression - Likelihood of exiting the labor market

	(1) Exit after giving birth	(2) Exit after giving birth	(3) Exit after giving birth	(4) Exit after giving birth	(5) Exit after giving birth	(6) Exit after giving birth
Inventor	0.111	0.0891	0.0921	0.353**	0.406***	0
	(0.110)	(0.114)	(0.115)	(0.113)	(0.114)	(.)
No voc. training	0.158***	0.0954***	0.109***	0.127***	0.135***	-0.440
	(0.0100)	(0.0106)	(0.0107)	(0.0124)	(0.0129)	(0.486)
University	-0.237***	-0.233***	-0.230***	-0.154***	-0.133***	0.0585
	(0.0358)	(0.0371)	(0.0376)	(0.0300)	(0.0298)	(0.233)
Foreign origin		0.219***	0.213***	0.227***	0.233***	-0.0731
		(0.0262)	(0.0262)	(0.0266)	(0.0267)	(0.261)
# gran. patents pr. birth						-0.0756
						(0.0421)
Constant	-1.904***	-1.754***	-1.621***	-1.557***	-1.648***	-1.141**
	(0.00851)	(0.0288)	(0.0323)	(0.0377)	(0.0444)	(0.431)
Age FE			Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes
Industry FE					Yes	Yes
Establ. size FE						Yes
Observations	5393732	5326589	5326589	2979671	2979671	2973

Notes: Table 4.14 displays the results of a logit regression indicating the relationships between the likelihood of exiting the labor market after birth and different control variables, conditional on being a mother. The dependent variable is the likelihood of exiting the labor market after birth measured by the reason for cancellation of the working contract and not returning from 2000 to 2010. The coefficient of interest is that on "Inventor". The control variables include dummy variables for being an inventor or of foreign origin and education. Education is split into 3 categories. Category 1 includes all mothers who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other mothers with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the individual level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.15: Logit Regression - Likelihood that a mother returns in a full-time capacity

	(1) Return full time	(2) Return full time	(3) Return full time	(4) Return full time	(5) Return full time	(6) Return full time
Inventor	0.824***	0.541***	0.489***	0.480***	-0.489***	-0.456***
	(0.119)	(0.124)	(0.134)	(0.138)	(0.122)	(0.120)
No voc. training		-0.131***	-0.0357***	-0.0569***	-0.0528***	-0.0455***
		(0.0103)	(0.0105)	(0.0106)	(0.0112)	(0.0113)
University		0.367***	0.339***	0.337***	0.0633***	0.0783***
		(0.0154)	(0.0155)	(0.0155)	(0.0159)	(0.0160)
Foreign origin		-0.275***	-0.275***	-0.266***	-0.00965	-0.00300
		(0.0160)	(0.0164)	(0.0164)	(0.0167)	(0.0168)
Constant	-1.690***	-1.665***	-5.159***	-5.208***	-5.152***	-5.153***
	(0.00393)	(0.00438)	(0.144)	(0.144)	(0.146)	(0.147)
Age FE			Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes
Industry FE					Yes	Yes
Establ. size FE						Yes
Observations	5393732	5376057	5376057	5376057	3008888	3008888

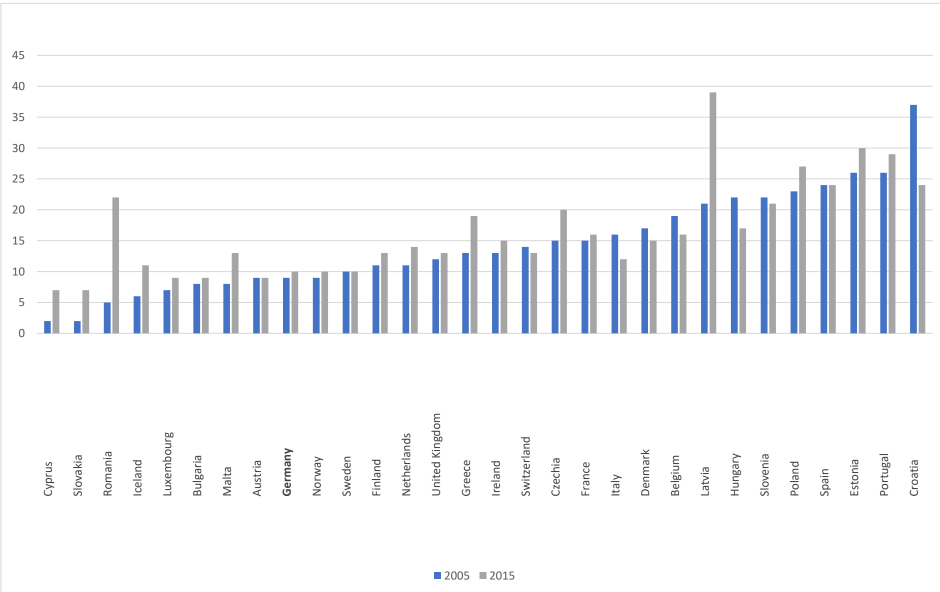
Notes: Table 4.15 displays the results of a logit regression indicating the relationships between the likelihood of returning to full-time work after giving birth and different control variables, conditional on being a mother. The dependent variable is the likelihood of returning to full-time work after giving birth from 2000 to 2010. The coefficient of interest is that on "Inventor". The other control variables are a dummy variable for being of foreign origin and education. Education is split into 3 categories. Category 1 includes all mothers who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other mothers with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the individual level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.16: Logit Regression - Likelihood that a mother returns in a part-time capacity

	(1) Return part time	(2) Return part time	(3) Return part time	(4) Return part time	(5) Return part time	(6) Return part time
Inventor	-0.781***	-0.683***	-0.737***	-0.738***	-0.992***	-0.985***
	(0.155)	(0.155)	(0.167)	(0.167)	(0.180)	(0.181)
No voc. training		-0.382***	-0.272***	-0.279***	-0.179***	-0.177***
		(0.00994)	(0.0105)	(0.0106)	(0.0113)	(0.0114)
University		-0.189***	-0.236***	-0.240***	-0.714***	-0.712***
		(0.0159)	(0.0161)	(0.0161)	(0.0167)	(0.0169)
Foreign origin		-0.406***	-0.408***	-0.404***	-0.111***	-0.110***
		(0.0160)	(0.0166)	(0.0166)	(0.0183)	(0.0183)
Constant	-1.060***	-0.969***	-6.502***	-6.389***	-5.902***	-5.976***
	(0.00389)	(0.00421)	(0.304)	(0.304)	(0.305)	(0.306)
Age FE			Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes
Industry FE					Yes	Yes
Establ. size FE						Yes
Observations	5393732	5376057	5376057	5376057	3008888	3008888

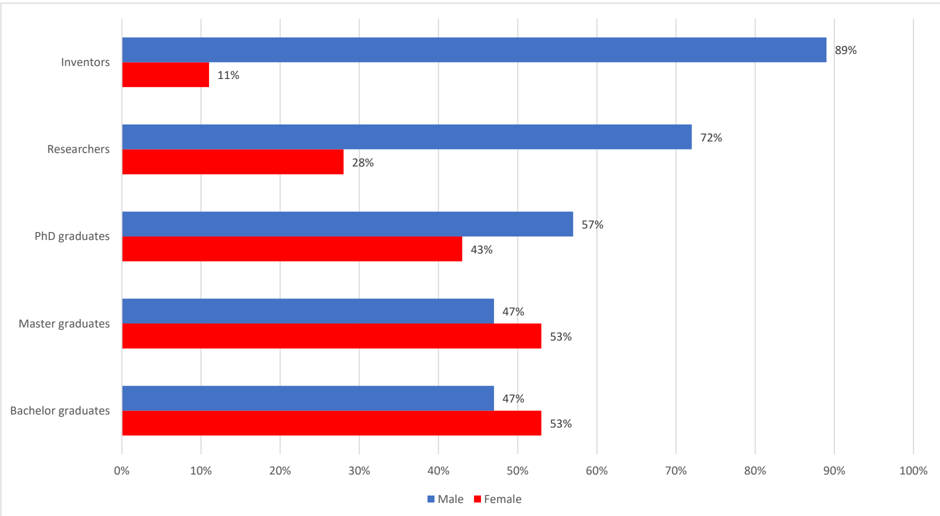
Notes: Table 4.15 displays the results of a logit regression indicating the relationships between the likelihood of returning to part-time work after giving birth and different control variables, conditional on being a mother. The dependent variable is the likelihood of returning to part-time work after giving birth from 2000 to 2010. The coefficient of interest is that on "Inventor". The other control variables are a dummy variable for being of foreign origin and education. Education is split into 3 categories. Category 1 includes all mothers who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other mothers with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the individual level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Figure 4.1: Percentage of female inventors in different European countries, 2005 and 2015



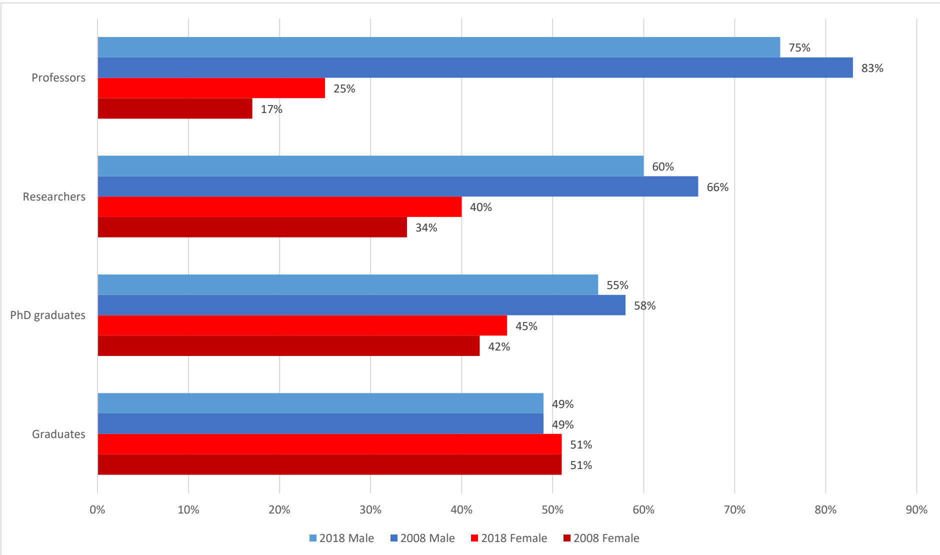
Data source: WIPO Statistics Database. Indicator: Share of women inventors (yearly statistics).

Figure 4.2: Share of female and male university graduates, researchers, and inventors worldwide, 1998-2017



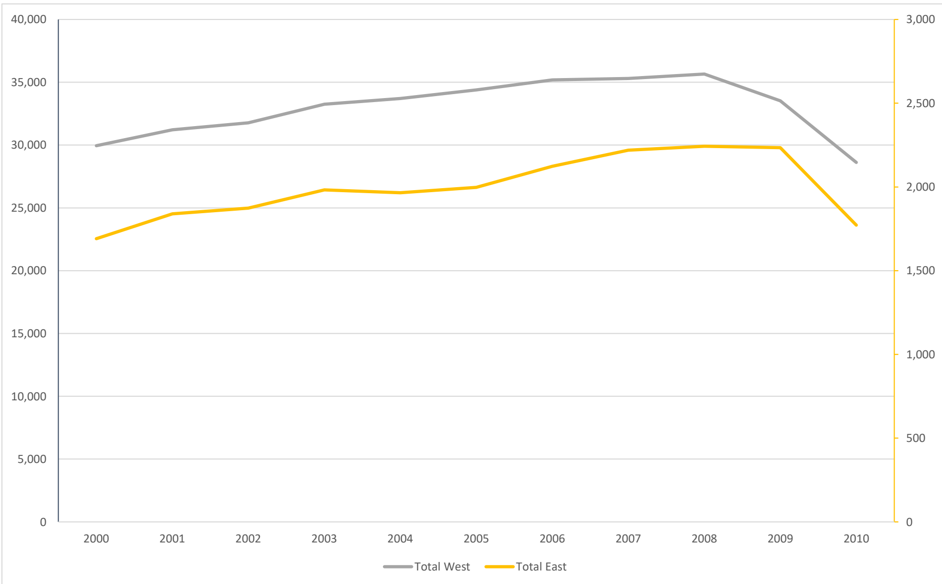
Data source: UK IPO (https://www.gov.uk/government/publications/gender-profiles-in-worldwide-patenting-an-analysis-of-female-inventorship-2019-edition).

Figure 4.3: Share of female and male university graduates, researchers, and professors in Germany, 2008 and 2018



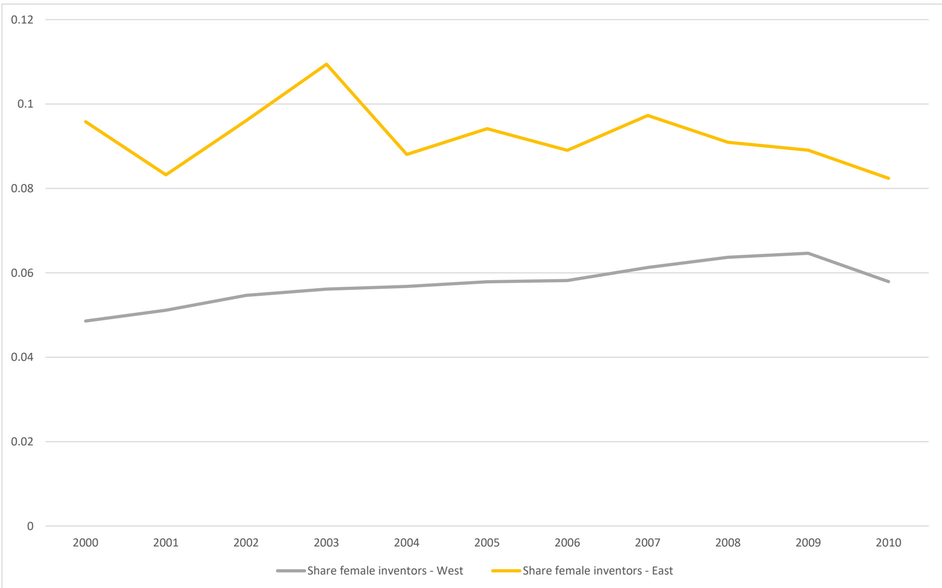
Data source: The German Statistical Office DESTATIS (https://www.destatis.de/DE/Service/Statistik-Campus/Datenreport/Downloads/datenreport-2021-kap-3.pdf?__blob=publicationFile (p. 18)).

Figure 4.4: Number of female and male inventors in former East and West Germany, 2000-2010



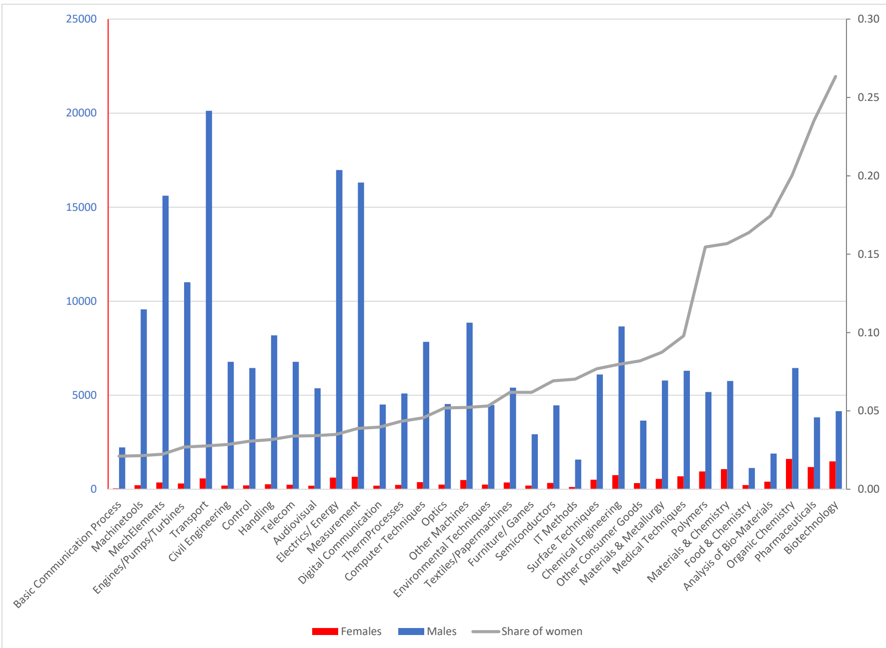
Notes: Figure 4.4 shows the total number of inventors by year and sex. Data source: INV-BIO.

Figure 4.5: Share of female inventors in former East and West Germany, 2000-2010



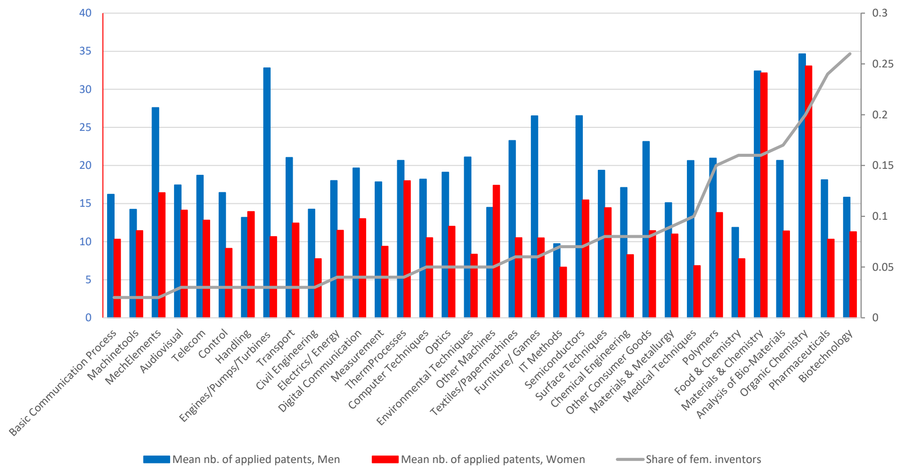
Notes: Figure 4.5 shows the share of female inventors among all inventors by year separately for former East and West Germany. Given the details of German history and Berlin's reunification in 1989, we exclude Berlin from this figure due to its status as a special case. Data source: INV-BIO.

Figure 4.6: Female and male inventors within different technological areas



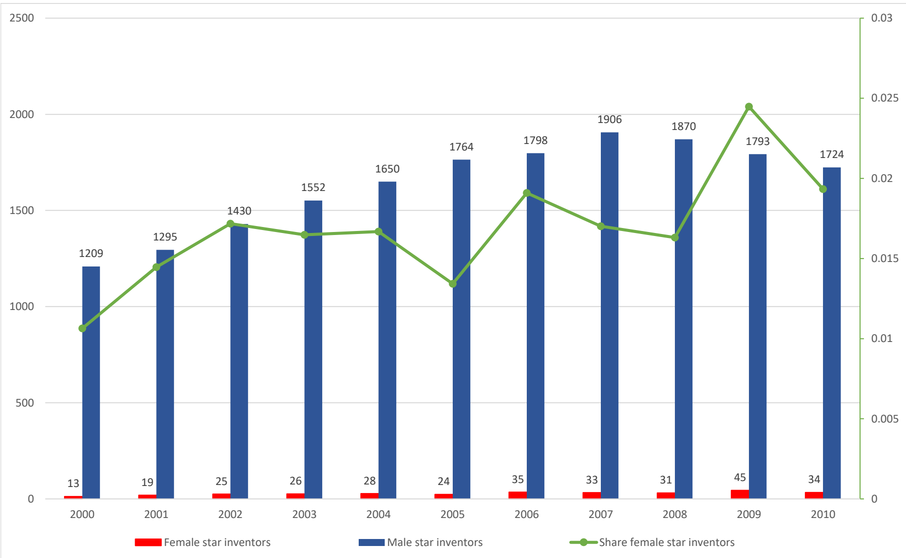
Notes: Number of female and male inventors within different industries sorted by share of female inventors. The classification of technological areas is based on the WIPO classification; Schmoch [2008] provides further information. Data source: INV-BIO.

Figure 4.7: Mean number of patent applications per inventor and the share of female inventors within different technological areas



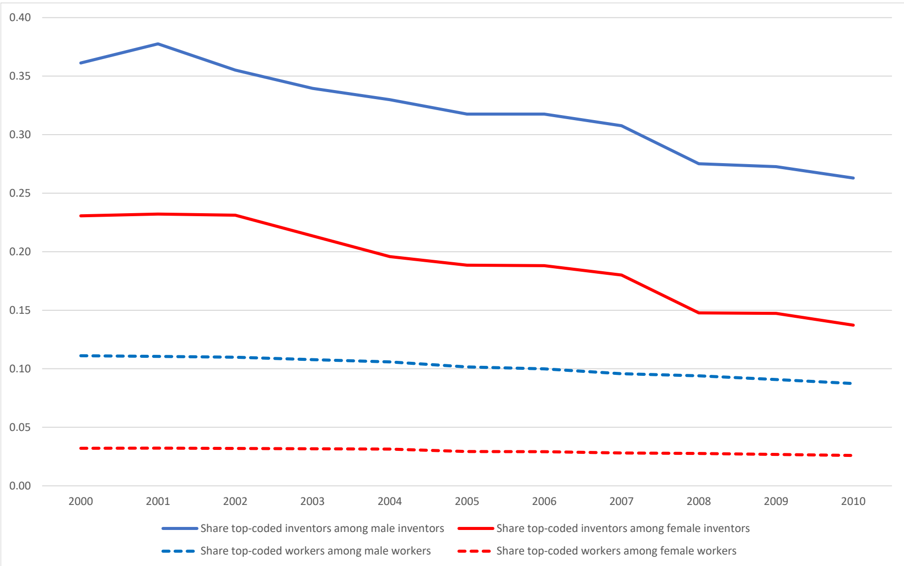
Notes: Number of patent applications per inventor in different technological areas and the share of female inventors. The classification of technological areas is based on the WIPO classification; Schmoch [2008] provides further information. Data source: INV-BIO.

Figure 4.8: Number of male and female star inventors and the share of female star inventors by year



Notes: Figure 4.8 shows the number of star inventors by sex and year (left axis) and the share of women among the star inventors (right axis). Star inventors are inventors with at least 15 granted patents (top 5% of all patenting inventors). Data source: INV-BIO.

Figure 4.9: Share of top-coded inventors (workers) among female and male inventors (workers), by year



Notes: Figure 4.9 shows the share of top-coded inventors among female (red) and male (blue) inventors by year. Figure 7 also displays the share of top-coded female (red dashed line) and male (blue dashed line) workers in a random 2-percent sample of workers in Germany. Data source: SIAB and INV-BIO.

Online Appendix

Table 4.17: Share of female inventors in selected countries, 1990 to 2015

Filing year	Country of residence						
	US	Japan	Germany	China	South Korea	France	UK
1990	6.31%	3.72%	3.14%	10.80%	2.84%	11.37%	4.42%
1991	6.61%	3.83%	3.29%	10.71%	2.62%	11.52%	4.89%
1992	6.88%	4.22%	3.40%	10.74%	3.58%	11.83%	5.96%
1993	7.16%	4.43%	3.61%	10.32%	5.01%	12.48%	6.41%
1994	7.46%	4.87%	3.79%	10.18%	6.35%	13.75%	6.69%
1995	7.91%	5.14%	4.26%	10.44%	7.31%	14.81%	7.20%
1996	7.83%	5.38%	4.66%	9.92%	5.87%	15.26%	6.77%
1997	8.28%	5.29%	4.72%	10.03%	7.61%	13.95%	7.82%
1998	8.85%	5.19%	5.25%	10.40%	5.26%	15.38%	7.71%
1999	9.09%	4.95%	5.63%	9.99%	5.19%	15.89%	8.46%
2000	9.43%	5.16%	5.92%	9.74%	4.91%	16.03%	8.83%
2001	10.26%	5.40%	6.38%	10.18%	4.59%	16.55%	8.77%
2002	10.77%	5.35%	6.41%	10.70%	5.35%	16.74%	8.85%
2003	10.04%	5.49%	6.70%	10.79%	6.17%	16.76%	9.59%
2004	9.78%	5.36%	6.68%	11.26%	5.43%	16.32%	8.95%
2005	9.60%	5.38%	6.89%	11.36%	5.64%	15.58%	9.43%
2006	9.81%	5.94%	7.01%	11.84%	6.00%	15.89%	10.01%
2007	9.78%	6.03%	7.09%	13.56%	5.77%	15.82%	9.68%
2008	9.86%	6.24%	7.43%	13.61%	5.72%	15.91%	9.72%
2009	10.20%	6.28%	7.74%	14.39%	5.59%	16.74%	9.51%
2010	10.20%	6.67%	8.17%	13.17%	5.61%	16.61%	9.23%
2011	10.05%	6.84%	7.97%	14.55%	6.18%	16.46%	9.19%
2012	10.09%	6.62%	7.96%	15.16%	5.88%	16.23%	9.55%
2013	10.26%	6.30%	7.95%	14.10%	6.59%	15.84%	9.49%
2014	10.32%	6.13%	8.40%	14.07%	6.14%	15.92%	9.87%
2015	10.44%	6.02%	8.56%	14.62%	5.89%	15.82%	10.51%

Notes: Table 4.17 shows the share of female inventors based on self-declared residence information among all patent applications. Because declarations of residency are not required by all patent offices, the PATSTAT data included in the UK IPO dataset does not have full coverage (19.7 million inventors out of 49.6 million (40%) have no such information). Data source: UK IPO. (https://www.gov.uk/government/publications/gender-profiles-in-worldwide-patenting-an-analysis-of-female-inventorship-2019-edition).

Table 4.18: Worldwide patent applications by employer, gender, and year

	Company applicants			University applicants		
	Male inventors	Female inventors	Proportion of female inventors	Male inventors	Female inventors	Proportion of female inventors
1990	1454442	69896	4.59%	22234	2614	10.52%
1991	1438849	71418	4.73%	20358	2867	12.34%
1992	1445076	75237	4.95%	21914	2817	11.39%
1993	1530672	82587	5.12%	25148	3586	12.48%
1994	1567633	88760	5.36%	29502	4745	13.86%
1995	1718533	103778	5.69%	36704	6004	14.06%
1996	1794512	109662	5.76%	36605	5798	13.67%
1997	2001032	122431	5.77%	42814	7230	14.45%
1998	2040892	133002	6.12%	48791	8393	14.68%
1999	2135949	143590	6.30%	51171	9048	15.03%
2000	2384469	170796	6.68%	57907	9934	14.64%
2001	2570902	195866	7.08%	65466	11837	15.31%
2002	2468216	192727	7.24%	64769	12109	15.75%
2003	2599007	206038	7.35%	78904	14845	15.83%
2004	2657942	207711	7.25%	81832	14478	15.03%
2005	2828734	226900	7.43%	111410	20229	15.37%
2006	2846110	240821	7.80%	130463	23135	15.06%
2007	2914085	252875	7.98%	156933	28353	15.30%
2008	2928948	260556	8.17%	171610	32705	16.01%
2009	2750396	257508	8.56%	190046	37508	16.48%
2010	2840692	280049	8.97%	210596	42746	16.87%
2011	2949708	302516	9.30%	231257	50894	18.04%
2012	3188298	340884	9.66%	265216	58204	18.00%
2013	3183030	359776	10.16%	285502	67197	19.05%
2014	2943809	349180	10.60%	258176	63415	19.72%
2015	2768039	342437	11.01%	247480	65124	20.83%

Notes: Table 4.18 presents the number of female inventors associated with patents filed between 1990-2018 for which the applicant listed in PATSTAT is a company and for which the applicant listed is a university. The PATSTAT database included in the UK IPO dataset includes information on whether a patent identifier relates to (among others) an individual, a company or a university. Not all applicant or inventor identifiers are attributed to a category, but those that are should give some indication regarding how inventorship is split by sex in academia or by industry. Data source: UK IPO (https://www.gov.uk/government/publications/gender-profiles-in-worldwide-patenting-an-analysis-of-female-inventorship-2019-edition).

Table 4.19: Average number of patents and citations per inventor (2000-2010) if inventor is never top-coded in terms of earnings, by gender

	Men (n=100,997)				Women (n=9,559)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Patent applications per year (2000-2010)	1.60	1.83	0	82	1.45	1.67	0	22
Patents granted per year (2000-2010)	0.58	0.92	0	31	0.45	0.77	0	38
Patents rejected per year (2000-2010)	1.02	1.60	0	82	1.00	1.54	0	11
Patent applications per year (2000-2010, considering team size)	0.43	0.43	0	14	0.33	0.40	0	21
Patents granted per year (2000-2010, considering team size)	0.16	0.29	0	6	0.11	0.21	0	4
Patents rejected per year (2000-2010, considering team size)	0.26	0.36	0	14	0.22	0.37	0	21
Patent applications (cumulative)	4.39	7.49	1	572	3.25	5.40	1	134
Patents granted (cumulative)	1.47	2.70	0	107	0.95	1.82	0	49
Patents rejected (cumulative)	2.78	5.96	0	529	2.20	4.51	0	130
Patent applications (cumulative, considering team size)	0.95	1.16	0	48	0.62	0.81	0	22
Patents granted (cumulative, considering team size)	0.36	0.59	0	18	0.20	0.34	0	10
Patents rejected (cumulative, considering team size)	0.59	0.89	0	44	0.42	0.70	0	22
Total forward citations in Germany within 2 years	0.91	2.81	0	165	0.56	2.14	0	90
Total forward citations in Europe within 2 years	1.30	4.11	0	258	1.46	4.49	0	143
Total forward citations in US within 2 years	4.62	18.09	0	1157	4.83	15.60	0	341
Total forward citations in Germany within 10 years	5.84	14.41	0	744	3.34	10.15	0	354
Total forward citations in Europe within 10 years	8.36	20.46	0	974	8.73	20.87	0	548
Total forward citations in US within 10 years	19.17	56.82	0	4008	17.29	47.37	0	1274

Notes: Earnings exceeding the upper earnings limit for statutory pension insurance are reported only up to this limit and are therefore top coded. If an inventor is unemployed, the daily wage variable records the daily benefit rate converted into euros. The number of forward citations is the number of citations that the invention (DOCDB family) received in patent applications at the German/European/US Patent and Trademark Offices within 2 and 10 years of the earliest publication date. The citation counts are corrected for equivalencies across patent authorities and include citations by the applicant. Count variables for each year are truncated at the 99th percentile of the citation distribution. Forward citations are assumed to measure the technological importance of the invention for subsequent developments (e.g., Trajtenberg [1990]). Data source: INV-BIO.

Table 4.20: Average number of patents and citations per inventor (2000-2010) if inventor is ever top-coded in terms of earnings, by gender

	Men (n=32,216)				Women (n=1,415)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Patent applicationsper year (2000-2010)	2.04	2.90	0	92	1.95	3.24	0	59
Patents granted per year (2000-2010)	0.70	1.28	0	26	0.54	1.02	0	10
Patents rejected per year (2000-2010)	1.35	2.40	0	81	1.41	2.83	0	54
Patent applications per year (2000-2010, considering team size)	0.43	0.52	0	13	0.34	0.39	0	5
Patents granted per year (2000-2010, considering team size)	0.16	0.31	0	6	0.10	0.19	0	2
Patents rejected per year (2000-2010, considering team size)	0.27	0.41	0	13	0.24	0.35	0	5
Patent applications(cumulative)	8.24	12.97	1	271	6.73	13.42	1	227
Patents granted (cumulative)	2.56	4.64	0	120	1.71	3.48	0	49
Rejected patents (cumulative)	5.16	9.91	0	253	4.73	11.02	0	178
Patent applications (cumulative, considering team size)	1.42	1.81	0	39	0.94	1.36	0	25
Patents granted (cumulative, considering team size)	0.51	0.85	0	24	0.27	0.43	0	5
Patents rejected (cumulative, considering team size)	0.91	1.36	0	36	0.67	1.18	0	25
Total forward citations in Germany within 2 years	1.67	4.70	0	190	0.93	3.16	0	61
Total forward citations in Europe within 2 years	2.56	7.38	0	294	3.64	13.90	0	321
Total forward citations in US within 2 years	10.01	33.83	0	988	12.79	50.32	0	1161
Total forward citations in Germany within 10 years	10.58	24.61	0	878	5.84	16.59	0	304
Total forward citations in Europe within 10 years	16.16	37.76	0	962	20.28	60.66	0	1398
Total forward citations in US within 10 years	41.63	111.35	0	2952	43.77	134.71	0	2947

Notes: Earnings exceeding the upper earnings limit for statutory pension insurance are reported only up to this limit and are therefore top coded. If an inventor is unemployed, the daily wage variable records the daily benefit rate converted into euros. The number of forward citations is the number of citations that the invention (DOCDB family) received in patent applications at the German/European/US Patent and Trademark Offices within 2 and 10 years of the earliest publication date. The citation counts are corrected for equivalencies across patent authorities and include citations by the applicant. Count variables for each year are truncated at the 99th percentile of the citation distribution. Forward citations are assumed to measure the technological importance of the invention for subsequent developments (e.g., Trajtenberg [1990]). Data source: INV-BIO.

Table 4.21: Inventor characteristics (2000-2010) by productivity category and sex

		Star inventors (>14 patents granted)	Regular (1-14 patents granted)	Potential (0 patents granted)
Share of women		0.03	0.06	0.08
Mean age	Women	42.19	37.47	36.56
	Men	46.05	42.42	41.42
Share of foreign origin	Women	0.093	0.094	0.15
	Men	0.047	0.047	0.078
Mean no. of applicants per patent	Women	1.02	1.1	1.1
	Men	1.05	1.07	1.09
Tech. area where most inventors of category patent are in	Women	Organic Chemistry	Organic Chemistry	Organic Chemistry
	Men	Transport	Transport	Transport
Mean no. employees in establishment	Women	3830.55	3698.06	3824.02
	Men	5329.58	3767.78	4296.71

Notes: Star inventors are defined as inventors with at least 15 patents granted, regular inventors are defined as inventors with 1-14 patents granted, and potential inventors are inventors with 0 granted patents (thus far). Data source: INV-BIO.

Table 4.22: OLS Regression - Number of total forward citations per inventor in Germany within 10 years

	(1) # citations GER	(2) # citations GER	(3) # citations GER	(4) # citations GER	(5) # citations GER	(6) # citations GER	(7) # citations GER
Woman	-13.62***	-12.55***	-11.33***	-11.42***	-6.642***	-7.879***	-8.634***
	(1.465)	(1.417)	(1.443)	(1.453)	(1.363)	(1.376)	(1.212)
No voc. training		11.24*	13.73*	13.56*	11.94*	9.665	9.692
		(5.571)	(5.720)	(5.701)	(5.687)	(5.579)	(5.579)
University		11.79***	11.66***	11.66***	10.27***	7.927***	7.919***
		(1.240)	(1.208)	(1.207)	(1.139)	(1.057)	(1.057)
Foreign origin		-7.333***	-6.457***	-6.613***	-5.620**	-6.854***	-6.904***
		(1.916)	(1.918)	(1.927)	(1.897)	(1.965)	(1.967)
Mother							16.39
							(11.16)
Constant	29.43***	20.65***	-0.885	-0.851	-10.45	-10.44	-8.334
	(0.838)	(0.736)	(5.825)	(5.711)	(6.429)	(6.915)	(6.732)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	770715	770668	770668	770668	765799	765799	765796
Adjusted R^2	0.003	0.011	0.015	0.015	0.043	0.058	0.058

Notes: Table 4.22 shows an OLS regression indicating relationships between the number of citations per inventor across all of his / her patents in the German Patent and Trademark Office after 10 years and different control variables. Number of forward citations that the invention(s) received from patent applications at the German Patent and Trademark Office after 10 years from the earliest publication date, respectively for each patent, is taken for calculating the overall number of citations per inventor. Count variables for each year are truncated at the 99th percentile of the citation distribution. The dependent variable is the mean number of citations across all patents per inventor. The control variables include dummy variables for being female, of foreign origin, or a mother and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.23: OLS Regression - Number of total forward citations per inventor in Europe within 10 years

	(1) # citations EUR	(2) # citations EUR	(3) # citations EUR	(4) # citations EUR	(5) # citations EUR	(6) # citations EUR	(7) # citations EUR
Woman	11.86	13.59	16.58*	16.71*	14.67	11.70	9.695
	(8.177)	(8.196)	(8.395)	(8.386)	(8.292)	(8.072)	(8.191)
No voc. training		6.502	12.37*	12.90**	16.20**	11.17*	11.25*
		(4.828)	(4.924)	(4.933)	(5.011)	(4.826)	(4.825)
University		23.83***	23.83***	23.85***	24.98***	19.93***	19.91***
		(1.986)	(2.058)	(2.056)	(2.025)	(1.733)	(1.734)
Foreign origin		-1.757	0.234	0.433	0.375	-2.376	-2.507
		(4.411)	(4.358)	(4.344)	(4.319)	(4.409)	(4.412)
Mother							43.48
							(33.72)
Constant	44.11***	26.30***	12.36	0.711	230.6***	235.1***	-11.25
	(1.246)	(1.143)	(10.32)	(10.15)	(69.06)	(68.92)	(8.132)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	770715	770668	770668	770668	765799	765799	765796
Adjusted R^2	0.001	0.014	0.021	0.023	0.051	0.081	0.081

Notes: Table 4.23 shows an OLS regression indicating relationships between the number of citations per inventor across all of his / her patents in the European Patent and Trademark Office after 10 years and different control variables. Number of forward citations that the invention(s) received from patent applications at the European Patent and Trademark Office after 10 years from the earliest publication date, respectively for each patent, is taken for calculating the overall number of citations per inventor. Count variables for each year are truncated at the 99th percentile of the citation distribution. The dependent variable is the mean number of citations across all patents per inventor. The control variables include dummy variables for being female, of foreign origin, or a mother and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.24: OLS Regression - Number of total forward citations per inventor in the US within 10 years

	(1) # citations US	(2) # citations US	(3) # citations US	(4) # citations US	(5) # citations US	(6) # citations US	(7) # citations US
Woman	-0.815	4.022	8.964	9.747	13.54	7.974	4.330
	(17.09)	(17.08)	(17.64)	(17.62)	(17.43)	(17.06)	(17.27)
No voc. training		17.78**	29.33***	31.76***	32.76***	23.12**	23.26***
		(6.593)	(6.857)	(6.857)	(7.028)	(7.040)	(7.039)
University		68.31***	67.17***	67.22***	66.03***	56.17***	56.13***
		(5.069)	(5.017)	(5.006)	(4.850)	(4.292)	(4.294)
Foreign origin		-1.871	1.866	3.116	3.366	-1.653	-1.891
		(9.268)	(9.285)	(9.280)	(9.365)	(9.660)	(9.666)
Mother							79.18
							(72.96)
Constant	108.2***	56.97***	30.94	-10.37	273.3**	278.1**	-27.36
	(3.380)	(2.637)	(27.15)	(27.20)	(103.6)	(102.9)	(22.46)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	770715	770668	770668	770668	765799	765799	765796
Adjusted R^2	-0.000	0.014	0.019	0.023	0.031	0.046	0.046

Notes: Table 4.24 shows an OLS regression indicating relationships between the number of citations per inventor across all of his / her patents in the US Patent and Trademark Office after 10 years and different control variables. Number of forward citations that the invention(s) received from patent applications at the US Patent and Trademark Office after 10 years from the earliest publication date, respectively for each patent, is taken for calculating the overall number of citations per inventor. Count variables for each year are truncated at the 99th percentile of the citation distribution. The dependent variable is the mean number of citations across all patents per inventor. The control variables include dummy variables for being female, of foreign origin, or a mother and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.25: Summary statistics (2000-2010), employed only

	Men		Women	
	(1) Inventors	(2) 2 percent random sample of workers in Germany	(3) Inventors	(4) 2 percent random sample of workers in Germany
Number of workers	132,601	535,747	10,862	507,743
Age				
Mean	42.22	38.48	37.68	38.6
Median	41	38	36	39
Daily wage				
Mean	155.84	65.62	135.12	40.45
Median	167.67	61.65	145.79	31.14
Education (imputed)				
No vocational training	0.03	0.44	0.04	0.47
Vocational training	0.27	0.48	0.32	0.49
University	0.70	0.08	0.64	0.05

Notes: This table is restricted to employed workers. The daily wage is a record of the employee's gross daily wage. It is calculated by using the fixed-period wages reported by the employer and the duration of the (unsplit) original notification period in calendar days. The daily wage is reported in euros. Earnings exceeding the upper earnings limit for statutory pension insurance are reported only up to this limit. If an inventor (or a worker within the random 2-percent sample) is unemployed, the daily wage variable records the daily benefit rate converted into euros. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. Data source: INV-BIO.

Table 4.26: Summary statistics (2000-2010), employed with a university degree only

	Men		Women	
	(1)	(2)	(3)	(4)
	Inventors	2 percent random sample of workers in Germany	Inventors	2 percent random sample of workers in Germany
Number of workers	89,217	35,543	6,687	20,969
Age				
Mean	41.91	43.96	37.83	41.31
Median	41	43	37	41
Daily wage				
Mean	158.43	130.00	142.64	89.04
Median	167.67	144.43	147.95	92.32

Notes: This table is restricted to employed workers with a university degree. The daily wage is a record of the employee's gross daily wage. It is calculated by using the fixed-period wages reported by the employer and the duration of the (unsplit) original notification period in calendar days. The daily wage is reported in euros. Earnings exceeding the upper earnings limit for statutory pension insurance are reported only up to this limit. If an inventor (or a worker within the random 2-percent sample) is unemployed, the daily wage variable records the daily benefit rate converted into euros. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. Data source: INV-BIO.

Table 4.27: Overview: Number of inventors per patent

Number of applicants		
Observations	1,185,889	
Mean	1.06	
Std. Deviation	0.61	
Min	1	
Max	62	

Number of applicants	Freq.	
1	1,133,920	95.62
2	46,640	3.93
3	3,915	0.33
4	719	0.06
5	288	0.02
6	73	0.01
7	149	0.01
8	46	0
9	17	0
10	13	0
11	6	0
13	7	0
61	46	0
62	50	0

Notes: Table 4.27 provides an overview on the number of inventors per patent and the frequency of patents with that number of inventors within the dataset. Data source: INV-BIO.

Table 4.28: Average number of patents and citations per inventor (2000-2010) if former West Germany, by gender

	Men (N=121,565)				Women (N=9,408)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Patent applications per year (2000-2010)	1.74	2.21	0	92	1.59	2.12	0	59
Patents granted per year (2000-2010)	0.61	1.03	0	31	0.47	0.83	0	11
Patents rejected per year (2000-2010)	1.13	1.88	0	82	1.11	1.92	0	54
Patent applications per year (2000-2010, considering team size)	0.43	0.45	0	14	0.34	0.42	0	21
Patents granted per year (2000-2010, considering team size)	0.17	0.30	0	6	0.11	0.21	0	4
Patents rejected per year (2000-2010, considering team size)	0.27	0.38	0	14	0.23	0.38	0	21
Patent applications (cumulative)	5.47	9.54	1	572	3.99	7.82	1	227
Patents granted (cumulative)	1.77	3.38	0	120	1.09	2.27	0	49
Patents rejected (cumulative)	3.49	7.46	0	529	2.77	6.50	0	178
Patent applications (cumulative, considering team size)	1.09	1.40	0	48	0.69	0.98	0	25
Patents granted (cumulative, considering team size)	0.40	0.68	0	24	0.21	0.37	0	10
Patents rejected (cumulative, considering team size)	0.69	1.06	0	44	0.48	0.85	0	25
Total forward citations in Germany within 2 years	1.13	3.48	0	190	0.67	2.48	0	90
Total forward citations in Europe within 2 years	1.66	5.29	0	294	1.95	7.44	0	321
Total forward citations in US within 2 years	6.18	24.04	0	1157	6.61	26.52	0	1161
Total forward citations in Germany within 10 years	7.28	18.17	0	878	3.97	12.05	0	354
Total forward citations in Europe within 10 years	10.61	26.81	0	974	11.28	33.14	0	1398
Total forward citations in US within 10 years	25.64	77.62	0	4008	23.07	73.99	0	2947

Notes: The table shows the average inventors' patent characteristics in 2000-2010 for West Germany, by gender. Berlin is excluded due to German reunification in 1989. An inventor who is observed in more than one year is included only in the latest year. Granted (applied, rejected) patents gives the average number among all inventors. Total forward citations in Germany (Europe, the US) gives the number of citations of (applied and/or granted) patents. Respectively within 2 or 10 years after the first application. Data source: INV-BIO.

Table 4.29: Average number of patents and citations per inventor (2000-2010) if former East Germany, by gender

	Men (N=8,308)				Women (N=1,047)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Patent applications per year (2000-2010)	1.50	1.59	0	34	1.29	1.18	0	12
Patents granted per year (2000-2010)	0.63	0.99	0	19	0.51	0.78	0	10
Patents rejected per year (2000-2010)	0.88	1.28	0	34	0.79	1.02	0	12
Patent applications per year (2000-2010, considering team size)	0.38	0.39	0	10	0.28	0.26	0	3
Patents granted per year (2000-2010, considering team size)	0.16	0.26	0	4	0.12	0.19	0	2
Patents rejected per year (2000-2010, considering team size)	0.22	0.33	0	10	0.17	0.22	0	2
Patent applications (cumulative)	4.33	6.48	1	155	2.82	3.37	1	43
Patents granted (cumulative)	1.72	2.89	0	49	1.11	1.71	0	22
Patents rejected (cumulative)	2.46	4.62	0	120	1.60	2.36	0	21
Patent applications (cumulative, considering team size)	0.90	1.13	0	26	0.54	0.52	0	5
Patents granted (cumulative, considering team size)	0.37	0.54	0	8	0.23	0.32	0	3
Patents rejected (cumulative, considering team size)	0.53	0.86	0	21	0.31	0.38	0	3
Total forward citations in Germany within 2 years	0.92	2.64	0	44	0.41	1.14	0	11
Total forward citations in Europe within 2 years	1.09	2.76	0	63	0.96	2.28	0	21
Total forward citations in US within 2 years	4.28	14.45	0	289	3.01	8.65	0	108
Total forward citations in Germany within 10 years	5.55	12.32	0	218	2.96	6.73	0	89
Total forward citations in Europe within 10 years	7.27	16.26	0	487	6.56	13.26	0	151
Total forward citations in US within 10 years	17.94	49.98	0	1226	12.30	30.64	0	374

Notes: The table shows the average inventors' patent characteristics in 2000-2010 for East Germany, by gender. Berlin is excluded due to German reunification in 1989. An inventor who is observed in more than one year is included only in the latest year. Granted (applied, rejected) patents gives the average number among all inventors. Total forward citations in Germany (Europe, the US) gives the number of citations of (applied and/or granted) patents. Respectively within 2 or 10 years after the first application. Data source: INV-BIO.

Table 4.30: Establishment characteristics in 2000-2010 by sex

	Men	Women
Number of establishments	25208	4135
Number of observations	722790	43281
Year the establishment number first appears	1983.2	1984.16
Total no. of employees	5237.3	4848.19
No. full-time employees	4396.46	3887.31
No. marginal part-time workers	28.69	52.57
No. female employees	1017.62	1181.27
No. regular workers	4682.78	4242.46
No. trainees/apprentices	166.92	165.44
No. employees in partial retirement	302.4	333.92
No. part-time (regular workers + others)	328.42	394.72
No. full-time female employees	673.54	743.78
No. part-time female employees	238.52	294.01
No. full-time regular workers	4386.9	3880.5
No. unskilled employees	539.66	501.04
No. qualified employees	3248.19	3079.2
No. highly qualified employees	1439.62	1257.01
No. full-time unskilled employees	295.9	240.49
No. full-time qualified employees	2772.21	2530.16
No. full-time highly qualified employees	1322.22	1109.75
No. employees aged 15-19	114.61	102.5
No. employees aged 20-24	299.78	281.21
No. employees aged 25-29	468.85	422.24
No. employees aged 30-34	673.96	584.21
No. employees aged 35-39	852.68	761.05
No. employees aged 40-44	882.52	822.21
No. employees aged 45-49	748.24	713.08
No. employees aged 50-54	598.71	574.48
No. employees aged 55-59	468.83	449.18
No. employees aged 60-64	126.47	135.3
No. employees aged 65 or older	2.61	2.71
Mean age of all employees	40.19	40.01
No. engineers and natural scientists	704.26	576.53
No. Germans	4727.98	4462.01
No. full-time Germans	3965.33	3574.16
No. EU-Europeans (excluding Germany)	235.43	191.9
No. agricultural occupations	3.49	7.85
No. unskilled manual occupations	1093.83	980.5
No. unskilled services	200.56	166.17
No. unskilled commercial and admin. occupations	140.07	165.17
No. skilled manual occupations	799.33	566.26
No. skilled services	37.61	49.72
No. skilled commercial and admin. occupations	779.74	826.43
No. technicians	779.53	902.79
No. semiprofessionals	31.81	57.3
No. engineers	842.26	667.62
No. professionals	97.22	144.81
No. managers	289.99	220.65
No. employees with top-coded wages	1724.77	1462.45
Median wage, all full-time employees	128.94	128.21
P25 wage, all full-time employees	109.77	107.3
P25 wage, all full-time employees	147.35	147.24
Mean imp. wage, all full-time employees	140.62	135.7

Notes: Table 4.30 provides an overview of establishment characteristics, pooled across 2000-2010, by sex. Data source: INV-BIO.

Table 4.31: Share of patenting activity in different technological areas by educational attainment and sex

	Without vocational training				With vocational training				With university degree			
	Women		Men		Women		Men		Women		Men	
	N=466		N=4,041		N=3,300		N=35,338		N=6,752		N=89,524	
Technological area	Freq.	Percent	Freq.	Percent	Freq.	Percent	Freq.	Percent	Freq.	Percent	Freq.	Percent
Electrics/ Energy	46	3	1961	10	412	4	11060	8	904	3	36879	7
Audiovisual	11	1	337	2	74	1	2095	2	252	1	8641	2
Telecom	18	1	488	2	48	0	1669	1	452	2	14550	3
Digital Communication	17	1	391	2	50	0	909	1	287	1	9597	2
Basic Communication Processes	6	0	175	1	12	0	410	0	52	0	3802	1
Computer Techniques	35	2	902	5	76	1	1891	1	617	2	14947	3
IT Methods	6	0	98	0	13	0	290	0	133	0	1762	0
Semiconductors	110	7	905	5	129	1	1457	1	657	2	14733	3
Optics	20	1	343	2	70	1	1427	1	472	2	11956	2
Measurement	45	3	1082	5	201	2	6182	4	898	3	32367	6
Analysis of Bio-Materials	16	1	65	0	107	1	386	0	515	2	2963	1
Control	17	1	407	2	65	1	1873	1	241	1	9490	2
Medical Techniques	51	3	489	2	326	3	4103	3	894	3	14897	3
Organic Chemistry	319	22	369	2	2227	21	4772	3	5708	21	29371	6
Biotechnology	62	4	331	2	524	5	966	1	2149	8	8261	2
Pharmaceuticals	29	2	112	1	428	4	1115	1	1750	6	8612	2
Polymers	64	4	245	1	910	9	3329	2	1493	6	13843	3
Food & Chemistry	6	0	52	0	118	1	650	0	210	1	1522	0
Materials & Chemistry	77	5	178	1	1169	11	3470	3	2084	8	15222	3
Materials & Metallurgy	33	2	228	1	292	3	2218	2	841	3	10382	2
Surface Techniques	30	2	412	2	230	2	2598	2	649	2	10603	2
Chemical Engineering	38	3	402	2	355	3	4057	3	754	3	13884	3
Environmental Techniques	9	1	281	1	112	1	2004	1	288	1	8716	2
Handling	19	1	466	2	163	2	6466	5	256	1	11566	2
Machinetools	20	1	637	3	146	1	7341	5	263	1	15553	3
Engines/Pumps/Turbines	65	4	1630	8	174	2	7567	5	481	2	32921	6
Textiles/Papermachines	25	2	513	3	220	2	5503	4	468	2	13488	3
Other Machines	41	3	577	3	235	2	6135	4	632	2	14057	3
ThermProcesses	23	2	335	2	113	1	3459	3	492	2	11156	2
MechElements	26	2	1281	6	394	4	12238	9	436	2	36891	7
Transport	126	9	3017	15	443	4	16713	12	889	3	54235	11
Furniture/ Games	17	1	302	2	129	1	3432	2	239	1	5897	1
Other Consumer Goods	29	2	385	2	282	3	2757	2	407	2	7630	2
Civil Engineering	16	1	522	3	200	2	7573	5	169	1	9903	2
Total nb. of observations (percent)	1472	100	19918	100	10447	100	138115	100	27032	100	510297	100
Total nb. of applied patents	1,451		18,861		9,580		110,455		23,390		294,330	

Notes: Table 4.31 shows the frequency and percentage of inventions within each technological area by inventor educational status. Data source: INV-BIO.

Table 4.32: Mean share of female inventors in different German states (2000-2010)

State	Number of male inventors	Number of female inventors	Total number of inventors	Share of female inventors
Schleswig-Holstein	2,222	151	2,373	0.06
Hamburg	3,027	493	3,520	0.14
Lower Saxony	8,700	604	9,304	0.06
Bremen	916	70	986	0.07
North Rhine-Westphalia	27,027	2,396	29,423	0.08
Hesse	11,796	1,164	12,960	0.09
Rhineland-Palatinate	6,366	656	7,022	0.09
Baden-Wuerttemberg	33,789	1,977	35,766	0.06
Bavaria	32,233	2,129	34,362	0.06
Saarland	874	91	965	0.09
Berlin	4,701	637	5,338	0.12
Brandenburg	1,200	154	1,354	0.11
Mecklenburg-Western Pomerania	477	74	551	0.13
Saxony	3,649	390	4,039	0.1
Saxony-Anhalt	886	160	1,046	0.15
Thuringia	2,078	250	2,328	0.11
Total	139,941	11,396	151,337	

Notes: Table 4.32 presents the mean share of female inventors across the different German states, pooled across 2000-2010. Data source: INV-BIO.

Table 4.33: Mean share of female inventors in different German states (2000, 2005, and 2010)

State	2000 Total number of inventors	2000 Share of female inventors	2005 Total number of inventors	2005 Share of female inventors	2010 Total number of inventors	2010 Share of female inventors
Schleswig-Holstein	367	0.04	480	0.05	506	0.04
Hamburg	656	0.12	851	0.15	668	0.16
Lower Saxony	1,897	0.03	2,165	0.06	1,781	0.06
Bremen	146	0.03	204	0.06	177	0.04
North Rhine-Westphalia	6,655	0.06	6,889	0.07	5,737	0.07
Hesse	2,930	0.07	3,192	0.08	2,281	0.08
Rhineland-Palatinate	1,797	0.06	1,986	0.09	1,323	0.1
Baden-Wuerttemberg	7,858	0.03	9,592	0.04	8,542	0.04
Bavaria	7,546	0.04	8,863	0.05	7,481	0.05
Saarland	176	0.04	215	0.08	173	0.06
Berlin	1,024	0.09	1,212	0.1	925	0.1
Brandenburg	223	0.11	234	0.1	227	0.08
Mecklenburg-Western Pomerania	75	0.15	119	0.09	108	0.14
Saxony	765	0.08	907	0.07	823	0.07
Saxony-Anhalt	173	0.13	173	0.15	168	0.12
Thuringia	455	0.09	564	0.11	449	0.07
Total	32,743		37,646		31,369	

Notes: Table 4.33 contains the mean share of female inventors across the different German states separately for 2000, 2005, and 2010. Data source: INV-BIO.

Table 4.34: OLS Regression - Cumulative number of patent applications per inventor

	(1) # applica-tions	(2) # applica-tions	(3) # applica-tions	(4) # applica-tions	(5) # applica-tions	(6) # applica-tions	(7) # applica-tions
Woman	-3.671*	-2.926	-1.810	-1.810	-0.478	-1.636	-1.683
	(1.758)	(1.728)	(1.741)	(1.742)	(1.677)	(1.632)	(1.692)
No voc. training		5.471**	7.636***	7.681***	7.761***	5.785***	5.787***
		(1.728)	(1.745)	(1.747)	(1.748)	(1.697)	(1.697)
University		9.230***	9.222***	9.225***	8.972***	7.038***	7.038***
		(0.866)	(0.853)	(0.853)	(0.781)	(0.675)	(0.676)
Foreign origin		-2.801*	-2.049	-2.055	-1.713	-2.858*	-2.861*
		(1.230)	(1.200)	(1.202)	(1.180)	(1.231)	(1.230)
Mother							1.028
							(3.370)
Constant	20.62***	13.71***	0.735	-0.882	-5.870*	-4.111	76.87***
	(0.609)	(0.419)	(2.174)	(2.189)	(2.589)	(3.113)	(19.76)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	764809	764763	764763	764763	760049	760049	760046
Adjusted R^2	0.001	0.015	0.021	0.022	0.044	0.077	0.077

Notes: Table 4.34 shows an OLS regression indicating correlations between the cumulative number of applied patents per year and different control variables. These control variables include: dummy variables of being female / of foreign origin / mother, and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Patent applications as a whole are analyzed, independent of the patent office (Germany, Europe). Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.35: OLS Regression - Cumulative number of patents granted

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	#	#	#	#	#	#	#
	granted	granted	granted	granted	granted	granted	granted
Woman	-2.026***	-1.898***	-1.539***	-1.512***	-1.117***	-1.225***	-1.262***
	(0.345)	(0.345)	(0.347)	(0.345)	(0.321)	(0.316)	(0.326)
No voc. training		1.635**	2.295***	2.380***	2.245***	2.018***	2.019***
		(0.605)	(0.605)	(0.605)	(0.607)	(0.601)	(0.601)
University		1.667***	1.692***	1.694***	1.532***	1.305***	1.305***
		(0.180)	(0.178)	(0.178)	(0.171)	(0.167)	(0.167)
Foreign origin		-0.491	-0.263	-0.224	-0.207	-0.349	-0.351
		(0.348)	(0.347)	(0.346)	(0.342)	(0.342)	(0.342)
Mother							0.797
							(0.740)
Constant	5.641***	4.375***	2.293*	0.809	21.87***	21.75***	0.199
	(0.103)	(0.133)	(0.922)	(0.914)	(6.309)	(6.302)	(1.078)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	770715	770668	770668	770668	765799	765799	765796
Adjusted R^2	0.003	0.009	0.017	0.022	0.050	0.056	0.056

Notes: Table 4.35 shows the estimates of an OLS regression with the cumulative number of granted patents per year as dependent variable and different control variables. These control variables include: dummy variables of being female / foreign / mother, and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Patent grants as a whole are analyzed, independent of the patent office (Germany, Europe). Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.36: OLS Regression - Cumulative number of rejected patents per inventor

	(1) Rejection patent	(2) Rejection patent	(3) Rejection patent	(4) Rejection patent	(5) Rejection patent	(6) Rejection patent	(7) Rejection patent
Woman	-1.646	-1.033	-0.327	-0.357	0.594	-0.448	-0.463
	(1.505)	(1.478)	(1.488)	(1.491)	(1.444)	(1.407)	(1.459)
Foreign origin		-2.304*	-1.795	-1.844	-1.515	-2.507*	-2.508*
		(1.043)	(1.017)	(1.019)	(1.002)	(1.053)	(1.052)
No voc. training		3.892**	5.333***	5.286***	5.510***	3.775**	3.776**
		(1.292)	(1.312)	(1.311)	(1.319)	(1.279)	(1.279)
University		7.586***	7.516***	7.517***	7.429***	5.725***	5.725***
		(0.759)	(0.748)	(0.748)	(0.684)	(0.585)	(0.585)
Mother							0.342
							(2.864)
Constant	14.85***	9.197***	-1.513	-1.708	35.70**	37.91**	-3.683
	(0.542)	(0.341)	(1.609)	(1.689)	(13.30)	(13.46)	(2.721)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	770715	770668	770668	770668	765799	765799	765796
Adjusted R^2	0.000	0.013	0.018	0.018	0.034	0.071	0.071

Notes: Table 4.36 shows a regression indicating correlations between the (cumulative) number of rejected patents per inventor and different control variables. These control variables include: dummy variables of being female / foreign / mother, and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Patent rejections as a whole are analyzed, independent of the patent office (Germany, Europe). Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.37: OLS Regression - Number of rejected patents per inventor per year

	(1) Rejected patents	(2) Rejected patents	(3) Rejected patents	(4) Rejected patents	(5) Rejected patents	(6) Rejected patents	(7) Rejected patents
Woman	0.253	0.349	0.369	0.354	0.488	0.304	0.304
	(0.299)	(0.298)	(0.304)	(0.306)	(0.296)	(0.294)	(0.305)
Foreign origin		-0.228	-0.201	-0.223	-0.173	-0.339	-0.339
		(0.198)	(0.197)	(0.197)	(0.192)	(0.192)	(0.192)
No voc. training		1.073***	1.159***	1.121***	1.198***	0.901**	0.901**
		(0.318)	(0.329)	(0.327)	(0.330)	(0.320)	(0.320)
University		1.320***	1.288***	1.288***	1.283***	0.992***	0.992***
		(0.105)	(0.105)	(0.105)	(0.0994)	(0.0906)	(0.0906)
Mother							0.00134
							(0.745)
Constant	3.097***	2.094***	0.112	0.645	10.06***	10.52***	0.548
	(0.0718)	(0.0535)	(0.598)	(0.595)	(2.992)	(2.982)	(0.448)
Age FE			Yes	Yes	Yes	Yes	Yes
Year FE				Yes	Yes	Yes	Yes
Industry FE					Yes	Yes	Yes
Establ. size FE						Yes	Yes
Observations	770715	770668	770668	770668	765799	765799	765796
Adjusted R^2	0.000	0.013	0.018	0.018	0.034	0.071	0.071

Notes: Table 4.37 shows a regression indicating correlations between the number of rejected patents per inventor per year and different control variables. These control variables include: dummy variables of being female / foreign / mother, and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, industry fixed effects, and establishment size fixed effects are included in a stepwise manner. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Patent rejections as a whole are analyzed, independent of the patent office (Germany, Europe).Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.38: Different OLS Regressions - Number of citations at the German / European / US patent office, after 2 years with technology area FE

	(1) # citations GER	(2) # citations GER	(3) # citations EUR	(4) # citations EUR	(5) # citations US	(6) # citations US
Woman	-1.077***	-1.190***	0.0733	-0.263	-7.384	-8.041
	(0.321)	(0.310)	(1.988)	(2.021)	(6.706)	(6.914)
No voc. training	2.221	2.226	2.422*	2.436*	4.756*	4.782*
	(1.430)	(1.430)	(1.178)	(1.178)	(1.916)	(1.915)
University	1.623***	1.622***	3.233***	3.230***	13.27***	13.26***
	(0.232)	(0.233)	(0.421)	(0.421)	(1.583)	(1.584)
Foreign origin	-1.324***	-1.332***	-0.474	-0.498	-0.930	-0.975
	(0.330)	(0.330)	(0.959)	(0.960)	(2.992)	(2.992)
Mother		2.505		7.390		14.50
		(1.563)		(8.739)		(22.63)
Constant	-6.279***	-6.261***	-6.893*	-6.841*	-7.760	-7.659
	(1.726)	(1.726)	(2.705)	(2.685)	(7.292)	(7.260)
Age FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Tech. area FE	Yes	Yes	Yes	Yes	Yes	Yes
Establ. size FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	770715	770668	721830	721825	721830	721825
Adjusted R^2	0.000	0.013	0.118	0.118	0.085	0.085

Notes: Table 4.38 shows different OLS regression indicating relationships between the number of citations per inventor across all of his / her patents in the German / European / US Patent and Trademark Office after 2 years and different control variables. Number of forward citations that the invention(s) received from patent applications at the German / European / US Patent and Trademark Office after 2 years from the earliest publication date, respectively for each patent, is taken for calculating the overall number of citations per inventor. Count variables for each year are truncated at the 99th percentile of the citation distribution. The dependent variable is the mean number of citations across all patents per inventor. The control variables include dummy variables for being female, of foreign origin, or a mother and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, technology area fixed effects, and establishment size fixed effects are included. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Table 4.39: Different OLS Regressions - Number of citations at the German / European / US patent office, after 10 years with technoloy area FE

	(1) # citations GER	(2) # citations GER	(3) # citations EUR	(4) # citations EUR	(5) # citations US	(6) # citations US
Woman	-6.108***	-6.906***	-5.445	-7.270	-29.46	-32.66
	(1.483)	(1.329)	(8.087)	(8.245)	(17.19)	(17.48)
No voc. training	10.59	10.63	11.79*	11.86*	21.51**	21.64**
	(5.898)	(5.898)	(5.140)	(5.138)	(7.256)	(7.253)
University	9.816***	9.809***	17.86***	17.84***	49.00***	48.97***
	(1.184)	(1.184)	(1.837)	(1.838)	(4.915)	(4.916)
Foreign origin	-5.955**	-6.009**	-4.252	-4.377	-9.840	-10.06
	(1.980)	(1.982)	(4.294)	(4.298)	(9.297)	(9.305)
Mother		17.58		40.15		70.28
		(11.69)		(33.93)		(75.27)
Constant	-29.87***	-29.75***	-33.07**	-32.79**	-27.30	-26.80
	(7.925)	(7.922)	(10.24)	(10.15)	(21.75)	(21.60)
Age FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Tech. area FE	Yes	Yes	Yes	Yes	Yes	Yes
Establ. size FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	721830	721825	721830	721825	721830	721825
Adjusted R^2	0.074	0.074	0.128	0.128	0.090	0.090

Notes: Table 4.39 shows different OLS regression indicating relationships between the number of citations per inventor across all of his / her patents in the German / European / US Patent and Trademark Office after 10 years and different control variables. Number of forward citations that the invention(s) received from patent applications at the German / European / US Patent and Trademark Office after 10 years from the earliest publication date, respectively for each patent, is taken for calculating the overall number of citations per inventor. Count variables for each year are truncated at the 99th percentile of the citation distribution. The dependent variable is the mean number of citations across all patents per inventor. The control variables include dummy variables for being female, of foreign origin, or a mother and education. Education is split into 3 categories. Category 1 includes all inventors who have not completed any vocational training. Category 2 comprises those who received in-firm vocational training/external vocational training or attended a full-time vocational school (Berufsfachschule). All other inventors with a degree from a university (of applied science) are in category 3. In addition, age fixed effects, year fixed effects, technology area fixed effects, and establishment size fixed effects are included. The age groups are 29 or younger, 30-39, 40-49, 50-59, and 60 or older. Standard errors are clustered at the inventor level. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: INV-BIO.

Figure 4.10: German students in STEM fields in Germany per winter term by sex

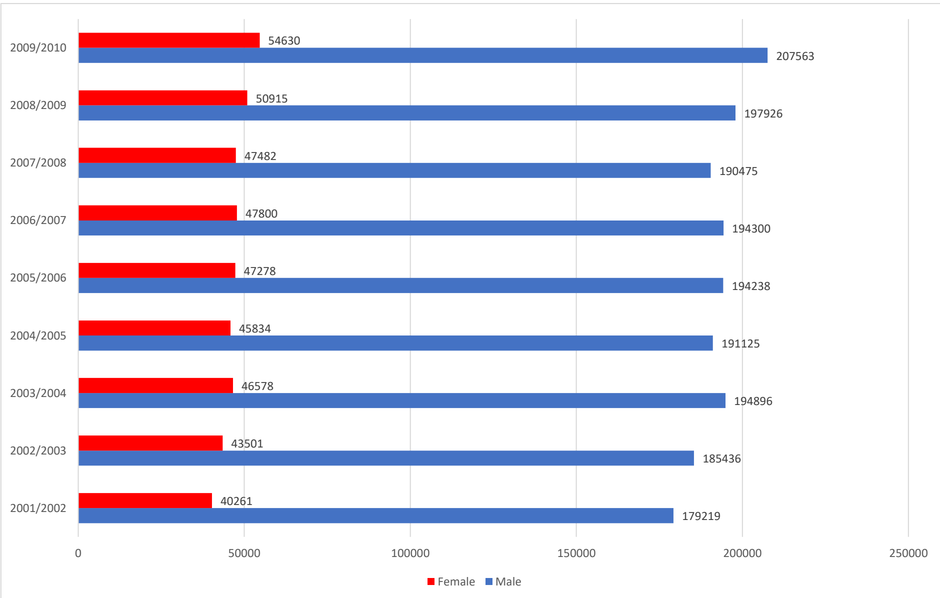


Figure 4.11: Foreign students in STEM fields in Germany per winter term by sex

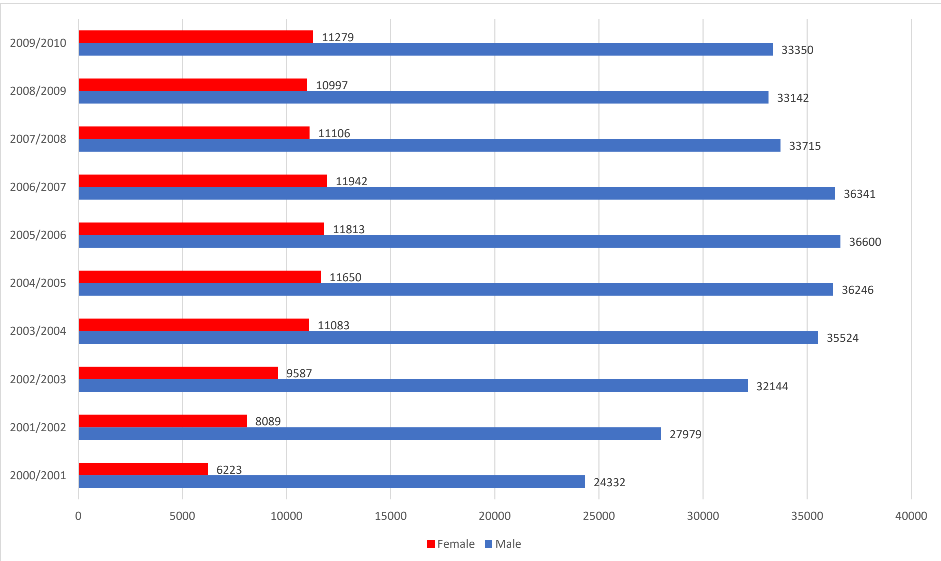
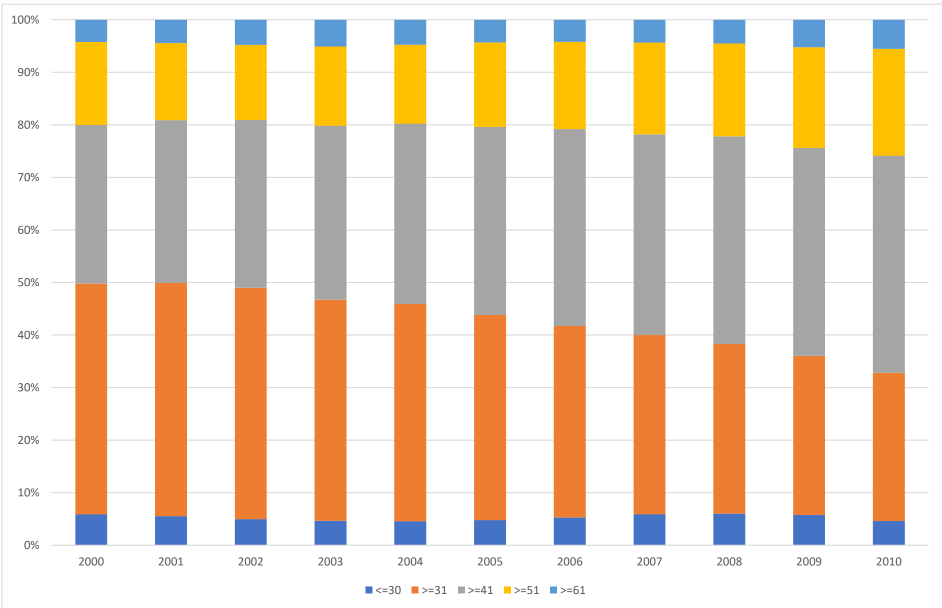
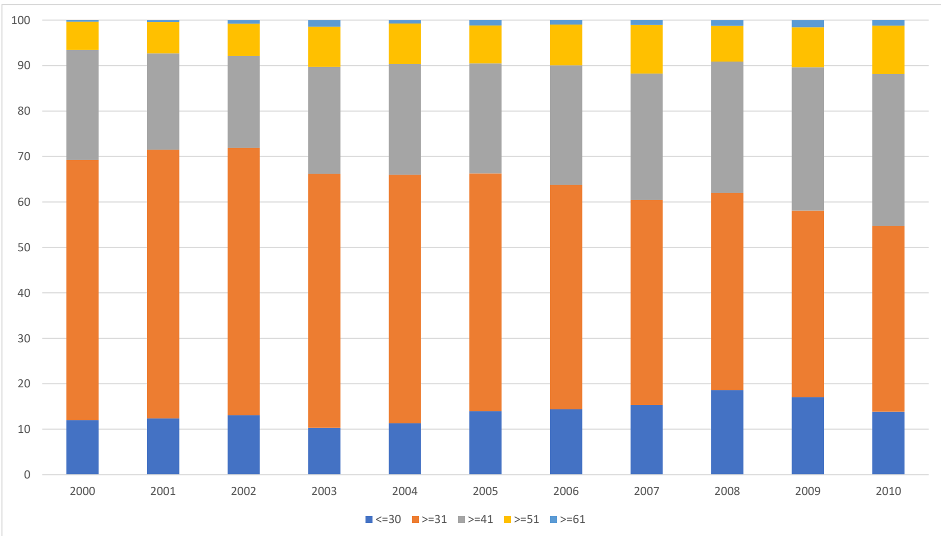


Figure 4.12: Share of female inventors in different age groups by year



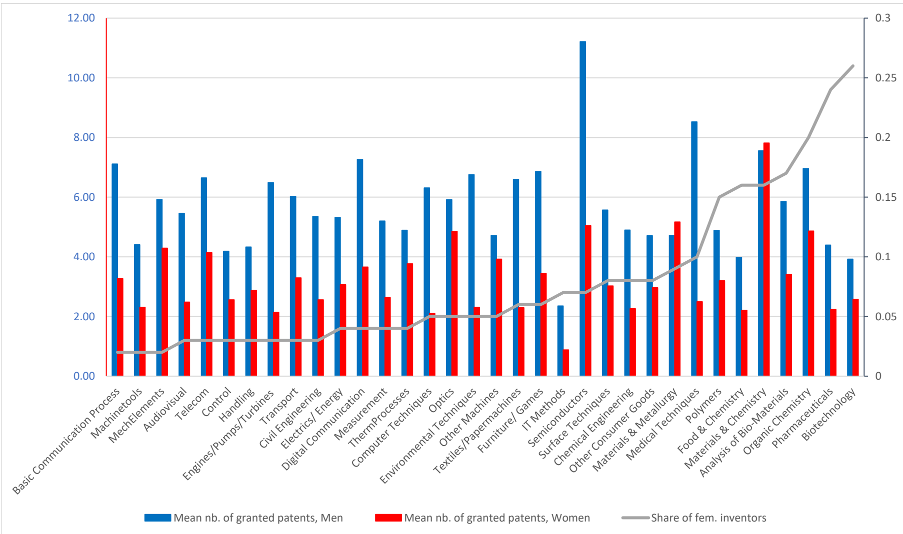
Notes: Figure 4.12 displays the share of female inventors in different age groups by year. The age groups are 29 or younger, 30-39. 40-49, 50-59, 60 or older. Data source: INV-BIO.

Figure 4.13: Share of male inventors in different age groups by year



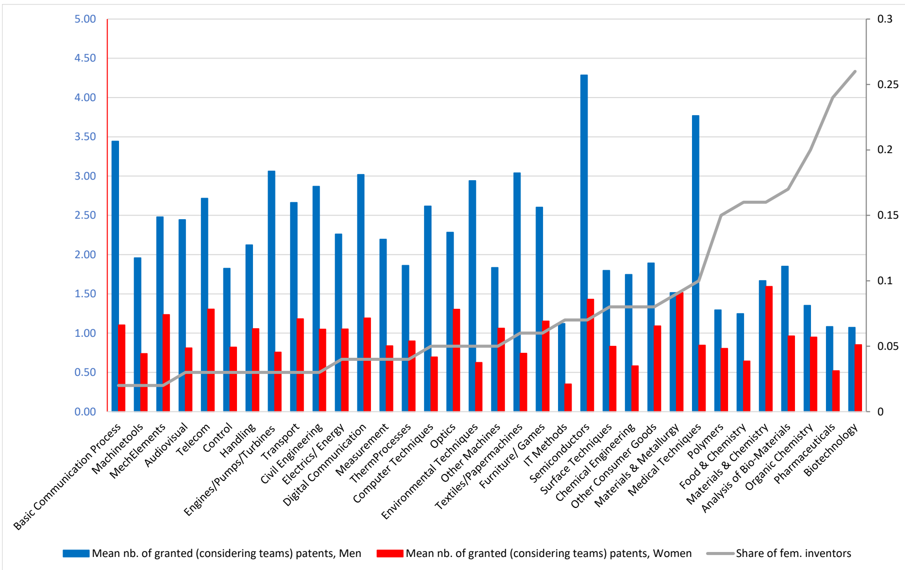
Notes: Figure 4.13 displays the share of male inventors in different age groups by year. The age groups are 29 or younger, 30-39. 40-49, 50-59, 60 or older. Data source: INV-BIO.

Figure 4.14: Mean number of total patents granted per inventor by tech. area and share of female inventors



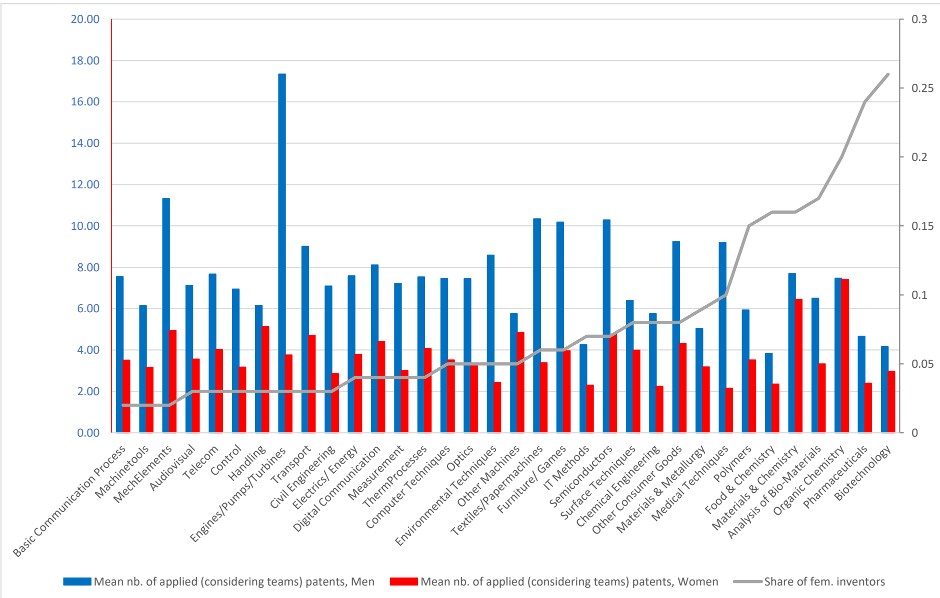
Notes: Number of granted patents per inventor in each tech. area and share of female inventors in each tech. area. The classification of technological areas is based on the WIPO classification; Schmoch [2008] provides further information. Data source: INV-BIO.

Figure 4.15: Mean number of total granted patents per inventor (accounting for team size) by tech. area and share of female inventors



Notes: Number of granted patents per inventor (accounting for team size) in each tech. area and share of female inventors in each tech. area. Team size is accounted for as follows: If a patent is filed by 2 inventors, each inventor is assigned a value of 0.5 for that patent. The classification of technological areas is based on the WIPO classification; Schmoch [2008] provides further information. Data source: INV-BIO.

Figure 4.16: Mean number of total patent applications per inventor (accounting for team size) by tech. area and share of female inventors



Notes: Number of applied patents per inventor (accounting for team size) in each tech. area and share of female inventors in each tech. area. Team size is accounted for as follows: If a patent is filed by 2 inventors, each inventor is assigned a value of 0.5 for that patent. The classification of technological areas is based on the WIPO classification; Schmoch [2008] provides further information. Data source: INV-BIO.

Bibliography

Bibliography

J. Abowd, F. Kramarz, and D. Margolis. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, 1999.

P. Aghion and P. Howitt. A model of growth through creative destruction. *Econometrica*, 60(2):323–351, 1992.

P. Aghion, U. Akcigit, A. Hyytinen, and O. Toivanen. On the returns to invention within firms: Evidence from Finland. *AEA Papers and Proceedings*, 108:208–212, 2018.

C. Albert, A. Glitz, and J. Llull. Labor market competition and the assimilation of immigrants. Working Paper 14641, Institute of Labor Economics (IZA), 2021.

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

J. Altonji and R. Blank. *Race and gender in the labor market.*, volume 3.C. Handbook of Labor Economics, 1999.

A. Aneja, O. Reshef, G. Subramani, and W. S. Louis. Differential responses to rejection & the gender innovation gap. Unpublished paper, 2021.

O. Ashenfelter. Estimating the effect of training programs on earnings. *The Review of Economics and Statistics*, 60(1):47–57, 1978.

R. Bachmann, R. Felder, and M. Tamm. Labour market participation and atypical employment over the life cycle: A cohort analysis for Germany. Working Paper 12010, 2018.

M. Bailey, R. Cao, T. Kuchler, and J. Stroebel. The economic effects of social networks: Evidence from the housing market. *Journal of Political Economy*, 126 (6):2224–2276, 2018.

A. Bartel. Where do the new US immigrants live? *Journal of Labor Economics*, 4 (7):371–391, 1989.

P. J. Barwick, Y. Liu, E. Patacchini, and Q. Wu. Information, mobile communication, and referral effects. Working Paper 25873, National Bureau of Economic Research, 2019.

M. Battisti, A. Romiti, and G. Peri. Dynamic effects of co-ethnic networks on immigrants' economic success. *The Economic Journal*, 641(132):58–88, 2021.

H. Battu, P. Seaman, and Y. Zenou. Job contact networks and the ethnic minorities. *Labour Economics*, 1(18):48–56, 2011.

P. Bayer, S. Ross, and G. Topa. Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy*, 116 (6):1150–1196, 2008.

L. Beaman. Social networks and the dynamics of labor market outcomes: Evidence from refugees resettled in the U.S. *Review of Economic Studies*, 79(1):128–161, 2012.

F. Bedaso. The labor market integration of refugees and other migrants in Germany. Working paper 884, Global Labor Organization (GLO), 2021.

A. Bell, R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen. Who becomes an inventor in America? The importance of exposure to innovation. *The Quarterly Journal of Economics*, 134(2):647–713, 2019.

S. T. Bell, A. J. Villado, M. A. Lukasik, L. Belau, and A. L. Briggs. Getting specific about demographic diversity variable and team performance relationships: A meta-analysis. *Journal of management*, 3(37):709–743, 2011.

M. Bertrand. *New Perspectives on Gender.*, volume 4B. Handbook of Labor Economics, 2009.

M. Bertrand and S. Mullainathan. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, 94(4):991–1013, 2004.

S. Bhalotra and I. Clots-Figueras. Health and the political agency of women. *American Economic Journal: Economic Policy*, 2(6):164–197, 2014.

O. J. Blanchard and L. F. Katz. Regional evolutions. *Brookings Papers on Economic Activity*, 23(1):1–61, 1992.

U. Blien, W. Dauth, and D. Roth. Occupational routine intensity and the adjustment to job loss: Evidence from mass layoffs. *Labour Economics*, 68, 2020.

H.-P. Blossfeld. Labor-market entry and the sexual segregation of careers in the Federal Republic of Germany. *American Journal of Sociology*, 93(1):89–118, 1987.

S. Bonhomme, T. Lamadon, and E. Manresa. A distributional framework for matched employer employee data. *Econometrica*, 87(3):699–739, 2019.

R. Borgoni, G. Degli Antoni, M. Faillo, and A. Michelangeli. Natives, immigrants and social cohesion: Intra-city analysis combining the hedonic approach and a framed field experiment. *International Review of Applied Economics*, 33(5):697–711, 2019.

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

G. J. Borjas. Does immigration grease the wheels of the labor market? *Brookings Papers on Economic Activity*, 1:69–133, 2001.

G. J. Borjas and H. Cassidy. The fall and rise of immigrant employment during the covid-19 pandemic. Unpublished paper, Harvard University Working Paper, 2021.

G. J. Borjas and K. B. Doran. The collapse of the Soviet Union and the productivity of American mathematicians. *The Quarterly Journal of Economics*, 3(127):1143–1203, 2012.

G. J. Borjas, I. Kauppinen, and P. Poutvaara. Self-selection of emigrants: Theory and evidence on stochastic dominance in observable and unobservable characteristics. *The Economic Journal*, 617(129):143–171, 2019.

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

M. Braun, J. Scott, and D. F. Alwin. Economic necessity or self-actualization? Attitudes toward women's labour-force participation in East and West Germany. *European Sociological Review*, 1(10):29–47, 1994.

T. Breda, J. Grenet, M. Monnet, and C. Van Effenterre. Can female role models reduce the gender gap in science? Evidence from French high schools. Working Paper 13163, Institute of Labor Economics (IZA), 2020.

K. Brenke. Eastern Germany still playing economic catch-up. *DIW Economic Bulletin*, 4(11):6–23, 2014.

M. Brown, E. Setren, and G. Topa. Do informal referrals lead to better matches? Evidence from a firm's employee referral system. *Journal of Labor Economics*, 34 (1):161–209, 2016.

H. Brücker, A. Glitz, A. Lerche, and A. Romiti. Occupational recognition and immigrant labor market outcomes. *Journal of Labor Economics*, 2(39), 2021.

K. Buckles. Fixing the leaky pipeline: Strategies for making economics work for women at every stage. *Journal of economic perspectives*, 1(33):43–60, 2019.

M. C. Burda and B. Severgnini. Total factor productivity convergence in German states since reunification: Evidence and explanations. *Journal of Comparative Economics*, 1(46):192–211, 2018.

T. Buser, M. Niederle, and H. Oosterbeek. Gender, competitiveness and career choices. *Quarterly Journal of Economics*, 3(129):1409–1447, 2014.

U. Busolt and K. Kugele. The gender innovation and research productivity gap in Europe. *International Journal of Innovation and Sustainable Development*, 3 (4.2):109–122, 2009.

B. Cadena and B. Kovak. Immigrants equilibrate local labor markets: Evidence from the Great Recession. *American Economic Journal: Applied Economics*, 8 (1):257–290, 2016.

S. Caldwell and O. Danieli. Outside options in the labor market. Technical report, Unpublished manuscript, 2021.

M. Caliendo, R. Schmidl, and A. Uhlendorff. Social networks, job search methods and reservation wages: Evidence for Germany. *International Journal of Manpower*, 32(7):796–824, 2011.

A. Calvo-Armengol and M. Jackson. The effects of social networks on employment and inequality. *American Economic Review*, 94(3):426–454, 2004.

S. Canaan and P. Mouganie. The impact of advisor gender on female students' stem enrollment and persistence. *Journal of Human Resources*, pages 0320–10796R2, 2021.

L. Cappellari and K. Tatsiramos. With a little help from my friends? Quality of social networks, job finding and job match quality. *European Economic Review*, 78:55–75, 2015.

D. Card, J. Heining, and P. Kline. Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly Journal of Economics*, 128(3):967–1015, 2013.

S. E. Carrell, M. E. Page, and J. E. West. Sex and science: How professor gender perpetuates the gender gap. *The Quarterly Journal of Economics*, 3(125):1101–1144, 2010.

R. Chattopadhyay and E. Duflo. Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica*, 5(72):1409–1443, 2004.

B. R. Chiswick and P. W. Miller. Where immigrants settle in the United States. *Journal of Comparative Policy Analysis: Research and Practice*, 2(6):185–197, 2004.

B. R. Chiswick and P. W. Miller. Do enclaves matter in immigrant adjustment? *City & Community*, 4(1):5–35, 2005.

W. A. Clark and A. I. Drever. Residential mobility in a constrained housing market: Implications for ethnic populations in Germany. *Environment and Planning A*, 32(5):833–846, 2000.

M. Collischon, A. Eberl, and M. Reichelt. Structural legacies and the motherhood penalty: How past societal contexts shape mothers' employment outcomes in reunified Germany. Unpublished paper, 2020.

M. Collischon, D. Kühnle, and M. Oberfichtner. Who benefits from cash-for-care? The effects of a home care subsidy on maternal employment, childcare choices, and children's development. *The Journal of Human Resources*, 2022.

L. D. Cook and C. Kongcharoen. The idea gap in pink and black. Working Paper 16331, National Bureau of Economic Research, 2010.

L. P. Cooke. Policy, preferences, and patriarchy: The division of domestic labor in East Germany, West Germany, and the United States. *Social Politics: International Studies in Gender, State & Society*, 1(13):117–143, 2006.

D. L. Costa, M. E. Kahn, C. Roudiez, and S. Wilson. Persistent social networks: Civil war veterans who fought together co-locate in later life. *Regional science and urban economics*, 70:289–299, 2018.

K. Couch and D. Placzek. Earnings losses of displaced workers revisited. *American Economic Review*, 100(1):572–589, 2010.

A. Damm. Neighborhood quality and labor market outcomes: Evidence from quasi-random neighborhood assignment of immigrants. *Journal of Urban Economics*, 79:139–166, 2014.

W. Dauth and J. Eppelsheimer. Preparing the sample of integrated labour market biographies (SIAB) for scientific analysis: A guide. *Journal for Labour Market Research*, 54(1):1–14, 2020.

S. Davis and T. von Wachter. Recessions and the costs of job loss. *Brookings papers on economic activity*, 2011(2):1–72, 2011.

J. DiNardo, N. Fortin, and T. Lemieux. Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, 64(5): 1001–1044, 1996.

M. Dorner, D. Harhoff, F. Gaessler, K. Hoisl, and F. Poege. Linked inventor biography data 1980-2014. *FDZ-Datenreport*, 2018.

C. Dustmann, J. Ludsteck, and U. Schönberg. Revisiting the German wage structure. *The Quarterly Journal of Economics*, 124(2):843–881, 2009.

C. Dustmann, T. Frattini, and G. Lanzara. Educational achievement of second-generation immigrants: An international comparison. *Economic Policy*, 27(69): 143–185, 2012.

C. Dustmann, A. Glitz, U. Schönberg, and H. Brücker. Referral-based job search networks. *Review of Economic Studies*, 83(2):514–546, 2016.

P.-A. Edin, P. Fredriksson, and O. Åslund. Ethnic enclaves and the economic success of immigrants: Evidence from a natural experiment. *The Quarterly Journal of Economics*, 118(1):329–357, 2003.

E. Einiö, J. Feng, and X. Jaravel. Social push and the direction of innovation. Unpublished paper, 2019.

Expert Council on Integration and Migration. Unter Einwanderungsländern: Deutschland im internationalen Vergleich, Jahresgutachten 2015, 2015. https://www.svr-migration.de/publikationen/jahresgutachten-2015/, Last accessed on 2022-02-25.

Expert Council on Integration and Migration. Normalfall Diversität? Wie das Einwanderungsland Deutschland mit Vielfalt umgeht, 2021. https://www.svr-migration.de/wp-content/uploads/2021/04/SVR_Kernbotschaften_2021.pdf, Last accessed on 2022-02-25.

R. W. Fairlie, K. Couch, and H. Xu. The impacts of covid-19 on minority unemployment: First evidence from april 2020 CPS microdata. Working Paper 27246, National Bureau of Economic Research, 2020.

Federal Ministry for Economic Affairs and Climate Action. Skilled professionals for Germany, 2022. https://www.bmwi.de/Redaktion/EN/Dossier/skilled-professionals.html, Last accessed on 2022-02-25.

B. Fitzenberger, A. Osikominu, and R. Völter. Imputation rules to improve the education variable in the IAB employment subsample. *Schmollers Jahrbuch: Journal of Applied Social Science Studies/Zeitschrift für Wirtschafts-und Sozialwissenschaften*, 126(3):405–436, 2006.

E. Forsythe and J.-C. Wu. Explaining demographic heterogeneity in cyclical unemployment. *Labour Economics*, (69):101955, 2021.

R. B. Freeman, R. A. Gordon, D. Bell, and R. E. Hall. Changes in the labor market for black Americans, 1948-72. *Brookings Papers on Economic Activity*, (1):67–131, 1973.

R. Frietsch, I. Haller, M. Funken-Vrohlings, and H. Grupp. Gender-specific patterns in patenting and publishing. *Research policy*, 4(38):590–599, 2009.

P. Frijters, M. A. Shields, and S. W. Price. Job search methods and their success: A comparison of immigrants and natives in the UK. *The Economic Journal*, 507 (115):F359–F376, 2005.

J. Fuchs, D. Söhnlein, and B. Weber. Projektion des Erwerbspersonenpotenzials bis 2060: Arbeitskräfteangebot sinkt auch bei hoher Zuwanderung. *IAB Brief Report*, (6), 2017.

H. Gartner. The imputation of wages above the contribution limit with the German IAB employment sample. *FDZ Methodenreport*, 2(2005), 2005.

C. Gathmann, I. Helm, and U. Schönberg. Spillover effects of mass layoffs. *Journal of the European Economic Association*, 18(1):427–468, 2020.

German Council of Economic Experts (GCEE). Annual report 2017/18 – towards a forward-looking economic policy, 2017. https://www.sachverstaendigenrat-wirtschaft.de/en/publications/annual-reports/previous-annual-reports/annual-report-201718.html, Last accessed on 2022-02-25.

German Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR). Laufende Stadtbeobachtung - Raumabgrenzungen, 2021. https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deutschland/gemeinden/StadtGemeindetyp/StadtGemeindetyp.html, Last accessed on 2021-05-30.

K. Gërxhani and Y. Kosyakova. The effect of social networks on migrants' labor market integration: A natural experiment. Working Paper No. 20203, Institute for Employment Research (IAB), Nürnberg, 2020.

R. Gibbons and L. F. Katz. Layoffs and lemons. *Journal of Labor Economics*, 9(4): 351–380, 1991.

I. Glas, G. Engbersen, and E. Snel. The street level and beyond: The impact of ethnic diversity on neighborhood cohesion and fear of crime among Dutch natives and nonnatives. *Journal of Urban Affairs*, 6(41):737–755, 2019.

A. Glitz. Coworker networks in the labour market. *Labour Economics*, 44:218–230, 2017.

M. Granovetter. The strength of weak ties. In *Social networks*, pages 347–367. Elsevier, 1977.

B. N. Greenwood, S. Carnahan, and L. Huang. Patient–physician gender concordance and increased mortality among female heart attack patients. *Proceedings of the National Academy of Sciences*, 34(115):8569–8574, 2018.

P. Haller and D. F. Heuermann. Opportunities and competition in thick labor markets: Evidence from plant closures. *Regional Science*, 60(2):273–295, 2019.

I. Hardoy and P. Schøne. Does the impact of plant closure on labour market attachment differ between immigrants and native workers across the business cycle? *Empirical Economics*, 46(1):229–252, 2014.

J. Hellerstein, M. McInerney, and D. Neumark. Neighbors and co-workers: The importance of residential labor market networks. *Journal of Labor Economics*, 29 (4):659–695, 2011.

I. Helm, A. Kügler, and U. Schönberg. Displacement effects in manufacturing. Technical report, University College London: Mimeo, 2021.

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

A. Hirseland, L. Kerschbaumer, I. Küsters, and M. Trappmann. Langzeitleistungsbeziehende im SGB II: Unerwartete übergänge in bedarfsdeckende Arbeit (no. 20/2019). *IAB-Kurzbericht*, 2019.

H. Hoynes, D. Miller, and J. Schaller. Who suffers during recessions? *Journal of Economic Perspectives*, 3(26):27–48, 2012.

J. Hunt, J. P. Garant, H. Herman, and D. J. Munroe. Why are women underrepresented amongst patentees? *Research Policy*, 4(42):831–843, 2013.

K. Huttunen, J. Møen, and K. G. Salvanes. Job loss and regional mobility. *Journal of Labor Economics*, 36(2):479–509, 2018.

H. Illing and T. Koch. Who suffers the greatest loss? Costs of job displacement for migrants and natives. *IAB Discussion Paper*, 08/2021, 2021.

H. Illing, J. F. Schmieder, and S. Trenkle. The gender gap in earnings losses after job displacement. Technical Report w29251, National Bureau of Economic Research, 2021.

Y. Ioannides and L. Loury. Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature 42*, 2004.

M. O. Jackson. *An overview of social networks and economic applications*, volume 1B. Elsevier, 2010.

L. S. Jacobson, R. J. LaLonde, and D. G. Sullivan. Earnings losses of displaced workers. *American Economic Review*, 83(4):685–709, 1993.

K. Jensen, B. Kovács, and O. Sorenson. Gender differences in obtaining and maintaining patent rights. *Nature Biotechnology*, 4(36):307–309, 2018.

T. Jung and O. Ejermo. Demographic patterns and trends in patenting: Gender, age, and education of inventors. *Technological Forecasting and Social Change*, (86):110–124, 2014.

L. Kahn. The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics*, 17(2):303–316, 2010.

H. Kleven, C. Landais, J. Posch, A. Steinhauer, and J. Zweimuller. Child penalties across countries: Evidence and explanations. *AEA Papers and Proceedings*, 109: 122–126, 2019.

P. Kline, N. Petkova, H. Williams, and O. Zidar. Who profits from patents? Rent-sharing at innovative firms. *The quarterly journal of economics*, 3(134):1343–1404, 2019.

J. Kling, J. Liebman, and L. Katz. Experimental analysis of neighborhood effects. *Econometrica, 75(1), 83-119*, 2007.

S. Klinger and J. Fuchs. *Effects of Population Changes in the Labour Market: An Analysis of six European Countries*, volume Country Report: Germany. Ministry of Economic Affairs and Employment of Finland, 2019.

A. Kondo. Differential effects of graduating during a recession across gender and race. *Journal of Labor Economics*, 4(1):1–24, 2015.

R. Koning, S. Sampsa, and J.-P. Ferguson. Who do we invent for? Patents by women focus more on women's health, but few women get to invent. *Science*, 6548(372):1345–1348, 2021.

D. Kozlowski, V. Larivière, C. R. Sugimoto, and T. Monroe-White. Intersectional inequalities in science. *Proceedings of the National Academy of Sciences*, 2(119), 2022.

F. Kramarz and O. Skans. When strong ties are strong: Family networks and youth labor market entry. *Review of Economic Studies*, 81(3):1164–1200, 2014.

T. Krebs and M. Scheffel. Structural reform in Germany. Working Paper 9787, Institute of Labor Economics (IZA), 2016.

P. Kropp and B. Schwengler. Three-step method for delineating functional labour market regions. *Regional Studies*, 50(3):429–445, 2016.

J. Künzler, W. Walter, E. Reichart, and G. Pfister. Gender division of labour in unified Germany. Technical report, Tilburg University, 2001.

M. Lachowska, A. Mas, and S. A. Woodbury. Sources of displaced workers' long-term earnings losses. *American Economic Review*, 110(10):3231–3266, 2020.

R. Laschever. The doughboys network: Social interactions and the employment of World War I veterans. Unpublished paper, Available at SSRN 1205543, 2009.

J. Lim and J. Meer. The impact of teacher-student gender matches: Random assignment evidence from South Korea. *Journal of Human Resources*, 4(52):979–997, 2017.

J. Lim and J. Meer. Persistent effects of teacher-student gender matches. *Journal of Human Resources*, 3(55):809–835, 2020.

P. Lundborg. Refugees' employment integration in Sweden: Cultural distance and labor market performance. *Review of international economics*, 2(21):219–232, 2013.

H. Mansour, D. I. Rees, B. M. Rintala, and N. N. Wozny. The effects of professor gender on the postgraduation outcomes of female students. *ILR Review*, 2021.

D. Marmaros and B. Sacerdote. Peer and social networks in job search. *European Economic Review*, 46(4-5):870–879, 2002.

D. A. Matsa and A. R. Miller. A female style in corporate leadership? Evidence from quotas. *American Economic Journal: Applied Economics*, 3(5):136–169, 2013.

E. Mauleón and M. Bordons. Male and female involvement in patenting activity in Spain. *Scientometrics*, 3(83):605–621, 2010.

J. Meekes and W. H. Hassink. Fired and pregnant: Gender Differences in Job Flexibility Outcomes after Job Loss. Working Paper 13779, Institute of Labor Economics (IZA), 2020.

M. Mertens and S. Müller. The East-West German gap in revenue productivity: Just a tale of output prices? Working Paper 14/2020, 2020.

E. Miguélez and R. Moreno. What attracts knowledge workers? The role of space and social networks. *Journal of Regional Science*, 1(54):33–60, 2014.

G. Miller. Women's suffrage, political responsiveness, and child survival in American history. *The Quarterly Journal of Economics*, 3(123):1287–1327, 2008.

L. Montenovo, X. Jiang, F. L. Rojas, I. M. Schmutte, K. I. Simon, B. A. Weinberg, and C. Wing. Determinants of disparities in covid-19 job losses. Working Paper 27132, National Bureau of Economic Research, 2020.

D. Müller and K. Strauch. Identifying mothers in administrative data. *FDZ-Methodenreport*, (13), 2017.

K. Munshi. Networks in the modern economy: Mexican migrants in the U.S. labor market. *The Quarterly Journal of Economics*, 118(2):549–599, 2003.

F. Murray. Mothers of invention. *Science*, 6548(372):1260–1262, 2021.

J. Nählinder. Where are all the female innovators?: Nurses as innovators in a public sector innovation project. *Journal of technology management & innovation*, 1(5): 13–29, 2010.

F. Naldi, D. Luzi, A. Valente, and I. Vannini Parenti. *Scientific and technological performance by gender.* Handbook of quantitative science and technology research, 2004.

M. Niederle and L. Vesterlund. Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, 122(3):1067–1101, 2007.

M. W. Nielsen, J. P. Andersen, L. Schiebinger, and J. W. Schneider. One and a half million medical papers reveal a link between author gender and attention to gender and sex analysis. *Nature human behaviour*, 11(1):791–796, 2017.

C. Olivetti and B. Petrongolo. The economic consequences of family policies: Lessons from a century of legislation in high-income countries. *Journal of Economic Perspectives*, 1(31):205–230, 2017.

Organization for Economic Cooperation and Development (OECD). Recruiting immigrant workers: Germany, 2013. https://www.oecd-ilibrary.org/social-issues-migration-health/recruiting-immigrant-workers-germany_9789264189034-en, Last accessed on 2022-02-25.

Organization for Economic Cooperation and Development (OECD). Health at a glance, 2019. https://doi.org/10.1787/19991312, Last accessed on 2022-03-09.

K.-H. Paqué. *Die Bilanz: eine wirtschaftliche Analyse der deutschen Einheit.* Hanser Verlag, 2009.

K.-H. Röhl. 25 jahre nach dem Mauerfall: Bilanz und Herausforderungen. *IW-Trends-Vierteljahresschrift zur empirischen Wirtschaftsforschung*, 3(41):73–90, 2014.

P. M. Romer. Endogenous technological change. *Journal of political Economy*, 5 (98):71–S102, 1990.

R. A. Rosenfeld, H. Trappe, and J. C. Gornick. Gender and work in Germany: Before and after reunification. *Annual Review of Sociology*, 1(30):103–124, 2004.

D. Rouzet, A. C. Sánchez, T. Renault, and O. Roehn. Fiscal challenges and inclusive growth in ageing societies. *OECD Economic Policy Paper*, (27), 2019.

P. Saygin, A. Weber, and M. Weynandt. Coworkers, networks, and job search outcomes. Discussion Paper 10003, Centre for economic policy research, 2014.

J. Schmieder, T. von Wachter, and J. Heining. The costs of job displacement over the business cycle and its sources: Evidence from Germany. Technical report, Boston University: Mimeo, 2020.

U. Schmoch. Concept of a technology classification for country comparisons. *Final report to the world intellectual property organisation (wipo)*, 2008.

I. Schmutte. Job referral networks and the determination of earnings in local labor markets. *Journal of Labor Economics*, 33(1):1–32, 2015.

C. Schnabel. United, yet apart? A note on persistent labour market differences between Western and Eastern Germany. *Jahrbücher für Nationalökonomie und Statistik*, 2(236):157–179, 2016.

J. Song, D. Price, F. Guvenen, N. Bloom, and T. von Wachter. Firming up inequality. *The Quarterly Journal of Economics*, 134(1):1–50, 2019.

J. Speer. Wages, hours, and the school-to-work transition: The consequences of leaving school in a recession for less-educated men. *The BE Journal of Economic Analysis and Policy*, 16(1):97–124, 2016.

C. R. Sugimoto, C. Ni, J. D. West, and V. Lariviére. The academic advantage: Gender disparities in patenting. *PloS one*, 5(10):e0128000, 2015.

O. Toivanen and L. Väänänen. Returns to inventors. *Review of Economics and Statistics*, 4(94):1173–1190, 2012.

G. Topa. Social interactions, local spillovers and unemployment. *Review Economic Studies*, 68(2):261–295, 2001.

G. Topa. *Labor markets and referrals*, volume 1. Handbook of Social Economics (1 ed.), 2011.

G. Topa and Y. Zenou. *Neighborhood versus network effects*, volume 5. Elsevier, 2015.

M. Trajtenberg. A penny for your quotes: Patent citations and the value of innovations. *The Rand Journal of Economics*, 1(21):172–187, 1990.

H. Trappe. Work and family in women's lives in the German Democratic Republic. *Work and Occupations*, 4(23):354–377, 1996.

H. Trappe and R. A. Rosenfeld. A comparison of job-shifting patterns in the former East Germany and the former West Germany. *European Sociological Review*, 4 (14):343–368, 1998.

T. von Wachter. Summary of the literature on job displacement in the US and EU: What we know and what we would like to know. *Wage structures, employment adjustments and globalization: Evidence from linked and firm-level panel data*, 2010.

T. von Wachter, J. Song, and J. Manchester. Long-term earnings losses due to mass layoffs during the 1982 recession: An analysis using US administrative data from 1974 to 2004. Unpublished paper, Columbia University, 2011.

J. Wahba and Y. Zenou. Density, social networks and job search methods: theory and application to Egypt. *Journal of Development Economics*, 2(78):443–473, 2005.

Y. Wang, Z. Yang, and X. Wang. Gender bias in invention. *Proceedings of the Association for Information Science and Technology*, 1(56):796–797, 2019.

E. Weber. *Zukunftsfeste Rente. Neue Impulse für die Alterssicherung.* Herder, 2016.

B. Weinberg, P. Reagan, and J. Yankow. Do neighborhoods affect work behavior? Evidence from the NLSY79. *Journal of Labor Economics*, 22(4):891–924, 2004.

K. B. Whittington and L. Smith-Doerr. Women inventors in context: Disparities in patenting across academia and industry. *Gender & Society*, 2(22):194–218, 2008.

H. Wolf. German economic unification twenty years later. *German Politics and Society*, 2(28):71–81, 2010.

M. Wyrwich. Woman and the labour market in East and West Germany: Socialist legacy and pre-socialist tradition. Working Paper 2017-015, Friedrich Schiller University Jena, 2017.

Eidesstattliche Versicherung

Ich versichere hiermit eidesstattlich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie mir gegebene Anregungen sind als solche kenntlich gemacht. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Sofern ein Teil der Arbeit aus bereits veröffentlichten Papers besteht, habe ich dies ausdrücklich angegeben.

Nürnberg, 18.03.2022

Maria Theresa Koch