UNIVERSITÉ DU LUXEMBOURG

PhD-FDEF-2024-024
Faculty of Law, Economics and Finance

DISSERTATION

Defence held on 10 December 2024 in Luxembourg
to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG EN SCIENCES ÉCONOMIQUES

by

FELIX STIPS
Born on April 14th 1992 in Rostock, Germany

ESSAYS IN THE ECONOMICS OF HUMAN MOBILITY AND LABOR MARKETS

Dissertation defence committee

Dr. Bertrand Verheyden, Dissertation supervisor
Luxembourg Institute of Socio-Economic Research (LISER)

Prof. Herbert Brücker
Professor, Humboldt University of Berlin

Prof. Luisito Bertinelli, Chairman
Professor, University of Luxembourg

Prof. Frédéric Docquier, Vice Chairman
Luxembourg Institute of Socio-Economic Research (LISER)

Prof. Giovanni Peri
Professor, University of California Davis

Acknowledgements

Completing a dissertation is a long and, at times, challenging process. I doubt I would have completed it without all the support I received from various people over the last four years.

First, I would like to express my profound gratitude towards my thesis supervision committee (CET), namely Luisito Bertinelli, Herbert Brücker, and, in particular, my principal supervisor, Bertrand Verheyden. The many discussions with you led me to gain new insights and perspectives, which enriched this thesis tremendously. You followed the process from the beginning until the end and believed in me even when I did not! And you always supported me generously with advice and opportunities. I am beyond thankful for that.

I am also very grateful to Andrea Albanese, Philipp Chapkovski, Frédéric Docquier, Christina Gathmann, Jose Vicente Guinot Saporta, Andreas Hauptmann, Giovanni Peri, and Ernesto Reuben, who, even though not formally part of my supervision committee, have supported me considerably at different stages of this journey. I appreciate every open ear, advice, and support you have given me over the last few years. Thank you for this privilege.

Most chapters are the result of collaborative efforts with my co-authors: Etienne Bacher, Frédéric Docquier, Nicolas Golenvaux, Mario Herrera-Rodriguez, Ernesto Reuben, Pierre Schaus, Bertrand Verheyden, and Siegfried Nijssen. I am deeply thankful for the productive collaboration and teamwork with you. Every project is a journey of academic and personal growth, and I learned lot from each one of you.

My sincere thanks also go to all my Ph.D. and post-doc colleagues for their support, motivation and friendship, in particular to Etienne Bacher, Beau Bressler, César Barreto, Alessandro Caiumi, Narcisse Cha'Ngom, Martín Fernández Sanchez, Julio Garbers, Terry

Gregory, Felix Grimm, Jules Linden, Lucas Vieira Magalhaes, Carlotta Montorsi, and Alexander Yarkin for making this journey more interesting and fun. I do hope that our paths will cross again in the future!

I appreciate the financial support from the National Research Fund of Luxembourg (FNR) through a PRIDE grant for the Analysis of cross-border human mobility (ACROSS) and the Luxembourg Institute of Socio-Economic Research (LISER) as well as the Faculty of Law, Economics and Finance, the Department of Economics and Management and the University of Luxembourg for the opportunity to undertake a Ph.D.. I would like to extend my thanks to Valérie Baran, Noémie Courtois, Anna Dober, and Roswitha Glorieux for generous and kind support with all the administrative aspects of my dissertaion.

Last but certainly not least, my heartfelt gratitude goes to my friends and family, especially Katharina, for the firm support, encouragement, and love. I cannot express how happy I am to have you by my side.

FELIX STIPS
Esch-Belval
October 2024

Contents

List of Figures

List of Tables

Abstracts

In Chapter 1, I examine how immigration following the EU Eastern Enlargement affected the German labor market. Drawing on a search-and-matching framework, I analyze how immigration influences the transitions of natives into and out of unemployment, job vacancies, and starting wages. The research strategy involves comparing commuting zones before and after the German labor market opened, addressing the potential endogeneity of migration across regions by employing two versions of a shift-share instrument. The findings reveal that immigration lowered native unemployment, increased the job-finding rate, decreased the job-separation rate, and raised starting wages, while leaving job vacancies unaffected. Delving into the mechanisms suggests that immigration spurred firm entry and productivity. The results support the common assumption of search-and-matching models that labor demand is highly elastic.

In Chapter 2, we examine how immigration affects wages when labor contracts are incomplete and native workers have reference-dependent wage expectations. We begin by providing empirical evidence of these wage expectations and then incorporate them into a monopsonistic labor market model. In our framework, native workers respond to wage reductions by decreasing their work effort, creating a trade-off for firms between minimizing wage costs and sustaining native worker motivation. At lower levels of migration, firms prioritize maintaining worker motivation, resulting in wage rigidity. However, as migration levels rise, firms increasingly focus on reducing wage costs, leading to downward wage adjustments. We test the model's predictions using a modified gift exchange experiment featuring migration episodes of different sizes. Our findings indicate significant wage reductions only in markets receiving large migration inflows. In markets with smaller inflows, migrant workers benefit without causing income losses to native workers or firms. In contrast, large migration inflows lead to smaller gains for migrants and substantial income losses for natives, while firms experience notable profit increases. These results align with efficiency wage theory and monopsonistic labor market dynamics, implying that the labor demand curve may not slope downward with small increases in labor supply.

In Chapter 3, we utilize a unique database on the mobility of Facebook users to examine the daily patterns of cross-border movements during the Covid-19 pandemic. To reduce censoring issues, our focus is on 45 pairs of European countries, documenting shifts in daily traffic throughout a full pandemic year. We employ both regression and machine learning models to assess the influence of infection risks and containment policies. Using permutation techniques, we compare the impact and predictive strength of these two variable categories. Unlike studies on within-border mobility, our models highlight the greater significance of containment policies in explaining fluctuations in cross-border traffic, as opposed to international travel restrictions and infection fears. These fears are represented by the number of Covid-19 cases and deaths at the destination. Though the ranking of coercive policies differs across modelling methods, containment measures at the destination (such as event cancellations, restrictions on internal movements, and limits on public gatherings), as well as school closures at the origin (which affect parental leave), have the most substantial effects on cross-border movements. While primarily descriptive, our findings carry important policy implications. Cross-border movement largely consists of labor commuting and business travel, which are only minimally affected by infection fears and travel bans, but are primarily driven by the stringency of internal containment policies and travel accessibility.

In Chapter 4, we replicate the study by Jones and Marinescu (2022), which examines the employment effects of a universal cash transfer in Alaska. Their use of a synthetic control method revealed no negative impact on employment. We reproduce their findings using their provided replication package and explore whether the results hold when applying a different software for analysis. Additionally, we employ alternative estimation techniques and conduct sensitivity checks to evaluate the robustness of the findings.

Our analysis shows some variations in the magnitude and significance of the average treatment effects on labor force participation and hours worked when using a different software (R) and various extensions of the synthetic control method. We also observe smaller coefficients for part-time employment when additional covariates are included. However, these differences do not alter the paper's primary conclusion.

Chapter 1

Immigration and Unemployment: Evidence from Germany after the EU Eastern Enlargement

Citation: Stips, F. (2024). Immigration and Unemployment: Evidence from Germany after the EU Eastern Enlargement. Unpublished mimeo.

1.1 Introduction

Rising numbers of international migrants worldwide have brought the question of how immigration affects receiving countries back to the forefront of policy discourse in many countries. An important aspect of this debate centers on the effects that immigration has on the labor market and on native workers. Immigrants can help to fill labor shortages in aging societies and bring new skills to fuel growth. However, fears of competition driving down wages and forcing incumbents into unemployment remain widespread. Immigration thus remains a controversial topic in receiving countries, as seen by its prominence in recent elections in the United States and Europe.

Most of the literature on the labor market effects of immigration has focused on the effects it has on employment and wages of natives (Borjas, 2003; Card, 2009; Ottaviano & Peri, 2012). In this paper, I instead assess the effect of immigration on unemployment and vacancies, transition rates between employment and unemployment, and starting wages after job transitions. These outcomes have been studied less, but are interesting for several reasons. First, unemployment is one of the key metrics that policymakers care about. Second, focusing on labor market transitions and tightness lends itself to a search-and-matching perspective on the labor market effects of immigration. Third, starting wages map closely to the theoretical counterpart wage offers, and are institutionally interesting since wages of stayers are highly rigid in the context of this study.

The empirical analysis is guided by recent theoretical work applying search-and-matching models to study the labor market effects of immigration. In most existing models, immigrants boost job creation because they are willing to accept lower wages, but firms can not only accept migrant workers (Albert, 2021; Battisti et al., 2018; Chassamboulli & Palivos, 2014; Chassamboulli & Peri, 2020). Native workers profit from more job opportunities. More job opportunities mean lower native unemployment, higher job-finding rates, and, sometimes, wage gains. On the other hand, Michaillat, 2023 argued that adding job rationing is important to reproduce observed business-cycle fluctuations (Michaillat, 2012; Michaillat & Saez, 2015; Shimer, 2005), and to allow for migration-induced increases in unemployment. In these models, immigration may increase welfare but will hurt native workers through lower job-finding rates and higher unemployment. The drastically different model results highlight the need for an empirical assessment.

I test the predictions using Germany's experience following the European Union's Eastern Enlargement as a natural experiment. This episode is well-suited to study the effects of immigration on native unemployment for at least three reasons. First, immigrant inflows were low in the years before Germany opened its labor market and increased rapidly afterward. Second, the relevant observation period between 2010 and 2014 is one without other major shocks, as Germany's labor market had already recovered from the global financial crisis, and the large inflow of asylum seekers began only in late 2015. It is also useful that Germany used the entire seven-year waiting period before granting full labor market access to new member states, as this creates a sufficient gap between product and labor market liberalization. Third, the German social security records are available at a daily frequency and are linked to detailed information about unemployment and job search, which allows me to follow workers in and out of unemployment.

The analysis utilizes data on all prime-age natives observed at least once in the universe of the Integrated Employment Biographies (IEB). I use the daily spell data to derive the monthly employment state and compute transition rates from changes in this state in between months (Jung & Kuhn, 2014). I aggregate outcomes to yearly averages at the commuting zone level (Kuhn et al., 2021). I also use data on job vacancies posted through the job portal of the Federal Employment Agency (BA). The main outcomes of interest are unemployment, labor market tightness, job-separation and job-finding rates, job-switching rates, and the entry wages of new workers entering from unemployment, inactivity, or switching between firms. The analysis aims to relate changes in these measures to changes in immigration from EU-13 countries.

The research design compares the evolution of local labor markets before and after Germany opened its labor market to EU-13 migrants in 2011, comparing regions that received considerable migrant inflows with regions that did not. Since migrants move to regions with favorable economic conditions, I use a shift-share instrumental variable strategy based on the presence of historical migrant networks in a region (Altonji & Card, 1991; Card, 2001). The identifying assumption is that, conditional on covariates, whether a region had more migrants from EU-13 countries twenty years earlier is as good as randomly assigned (Goldsmith-Pinkham et al., 2020). Importantly, the set of covariates controls for the total historical migrant share in a region. I also use a modified shift-share instrument that leverages the diversion of EU-13 migration from Southern Europe towards Germany after the global financial crisis to predict origin-specific inflows to Germany. This alternative instrument further relies on the exogeneity of these shocks for identification (Borusyak et al., 2022). First-stage regressions indicate that the instruments are strong predictors of actual EU-13 and total immigration, but not migration from other major sending regions or EU-13 migration before the policy change. Tests for differential pre-trends in key determinants of labor demand suggest that the instruments are not strongly correlated with these factors before the policy change.

My findings paint a rather positive image of the recent German immigration experience. A percentage point increase in the number of migrants from EU-13 countries relative to the baseline population leads to a decrease in the number of native unemployed between 6.6% and 8.3%. A simple back-of-the-envelope calculation would imply a decrease in native unemployment in a region by about 1500 individuals. Coefficient estimates on the unemployment rate are similar, though they do not pass significance tests at conventional levels. The estimated effects on the stock of vacancies, as well as on vacancy creation and vacancy filling, are positive but imprecise and not statistically significant. The effects on labor market tightness are positive and mainly driven by movements in unemployment rather than changes in vacancies. Delving into the transition rates, I find that immigration increased the job-finding rate from unemployment and decreased the job-separation rate, with similar elasticities on both. While not directly related to the predictions from search-and-matching models, I also find a reduction in job-to-job switches and no significant changes at the participation margin. An alternative hypothesis would be that wages adjusted to absorb the supply shift. I can not confirm this hypothesis, as the results for starting wages show consistently positive effects of immigration. Most of the results are robust to changes in the geographical unit, choice of covariates, semi-parametric estimation techniques, and other sample choices.

My findings suggest that unemployment decreased, tightness increased, and the job-finding rate from unemployment increased in response to immigration. These findings are at odds with most of the predictions of models with job rationing and instead suggest that labor demand was elastic enough to absorb new entries. Delving deeper into these mechanisms, I observed that output, productivity, and firm entry increased in response to the supply shock. Employment growth increased both within existing firms and through net new firm creation. This channel aligns with standard search-and-matching models: migrants accept lower wage offers, which increases productivity and leads to firm expansion. Other potential channels seem to be less relevant in this context. I do not find evidence that native internal migration attenuated the labor supply shock, nor that immigration increased the matching efficiency between workers and firms. Lastly, the effects of EU-13 immigration on the outcomes of immigrants are less positive, indicating, perhaps, moderate levels of imperfect substitutability between immigrants and natives. These mechanisms suggest that a model with job creation channels and imperfect substitutability between migrants and natives would likely be best to rationalize my results.

In addition to the main results, I generate a number of findings that may be interesting to theorists. First, I find that the job-separation rate from employment also seems to react to immigration. The job-separation rate is usually introduced as a parameter in search-and-matching models. My findings suggest that it may need to be endogenized. Second, I do not find significant movements along the participation margin. This finding suggests that two-state approximations without inactivity may be sufficient. At least, it does not provide evidence against it. Lastly, I observe that migrants in Germany seem to have higher job-separation rates, lower job-finding rates, and higher job-switching rates, a finding that is in line with previous research in European labor markets but markedly different to the United States (Chassamboulli et al., 2024). This confirms the modeling choice of Battisti et al., 2018 to use different job-separation rates as a key difference between migrants and natives. It also raises concerns about assuming perfect substitutability between immigrants and natives.

This paper contributes to recent efforts to understand the effects of immigration when labor markets are imperfect. A number of papers study immigration under monopsony and find some evidence that it can exacerbate firms' market power (Amior & Manning, 2020; Amior & Stuhler, 2024; Borjas & Edo, 2023; Gyetvay & Keita, 2023). Consensus remains to be reached on whether wage posting or wage bargaining is the more realistic assumption, in particular since both wage-setting mechanisms co-exist in practice (Brenzel et al., 2014; Caldwell et al., 2023; Hall & Krueger, 2012; Lachowska et al., 2022). Yet, search models have a long tradition in labor economics, and recent applications to migration show that they deliver interesting novel predictions and can be useful to simulate the effects of migration policies (Albert, 2021; Battisti et al., 2018; Chassamboulli & Palivos, 2014; Chassamboulli & Peri, 2020; Hart & Clemens, 2019). I contribute to this literature on the effects of immigration under matching frictions empirically by testing key assumptions and predictions of these models using a well-suited migration episode coupled with high-quality administrative data.

Perhaps the closest paper to mine is Anastasopoulos et al., 2021, who study the effect of the Mariel Boatlift on an index of newspaper job advertisements, unemployment, and job-finding rates. A strength of my paper relative to theirs is the data, which allows me to directly measure transition rates between employment and unemployment, observe the creation and filling of vacancies, and follow workers and establishments over time and space. This allows me to explore possible mechanisms such as internal migration or firm entry. Additionally, since I rely on large-scale administrative data, my results are not prone to the sampling and power issues surveys available at the time of the Mariel Boatlift have (Clemens & Hunt, 2019; Peri & Yasenov, 2019). Interestingly, my paper offers quite different results. Anastasopoulos et al., 2021 find that immigration reduces vacancies and the job-finding rate of natives, while I find a positive effect on both, albeit insignificant for vacancies. These differences may be due to differences in our data, our methodologies, or potentially also due to differences in our contexts. In particular, the U.S. was undergoing a recession during the period studied by Anastasopoulos et al., 2021, while Germany was in the midst of an economic boom during the period under study.

This paper also contributes to a literature evaluating the impact of increased immigration following the EU Eastern Enlargement on the German labor market. This event was economically and politically important for Germany, and it warrants careful assessment.

Previous research has focused on the effect of EU-13 immigration on employment and found modest or no impacts (Gyetvay & Keita, 2023; Hammer & Hertweck, 2022; Illing, 2023).[1] I contribute to this literature by studying novel outcomes such as unemployment, vacancies, and job transitions, using a modified instrumental variable strategy, and exploring in-depth possible mechanisms for how regions adjusted to immigration from EU-13 countries. Indeed, this appears to be the first paper to document that output, productivity, and firm entry increased in response to EU-13 immigration.

Findings from post-2010 immigration to Germany also complement research on immigration to Germany during the 1990s (Amior & Stuhler, 2024; Dustmann et al., 2017, 2023). This influential line of research has found rather stark negative impacts on native workers. My work, along with other work on more recent immigration to Germany, in particular Gyetvay and Keita, 2023 and Hammer and Hertweck, 2022, come to a less bleak conclusion. One explanation that could reconcile these findings is, again, the business cycle. In the early 1990s, the German economy was in a deep recession, while in the early 2010s, it was in a boom. In a boom, when the labor market is inefficiently tight, an increase in labor supply increases welfare without displacing incumbents. This intuitive mechanism is developed theoretically by Michaillat, 2023 and may be an important source of heterogeneity to explore further. More broadly, this highlights the importance of moving away from thinking about *the* labor market effect of immigration to thinking about which factors shape the response across different contexts (Foged et al., 2022).

Lastly, two methodological aspects are also worth highlighting. First, I develop a novel variant of the commonly used shift-share instrument. The modification uses migrant inflows, which were diverted from Southern Europe because these countries were affected more by the global financial crisis (Bertoli et al., 2016). The identification relies on the exogeneity of the origin composition in Southern Europe to macroeconomic conditions in Germany. Second, since I use only immigration from EU-13 countries as endogenous variable and to construct the instruments, I can control for the lagged total migrant

[1] Illing, 2023 finds negative effects on employment and wage growth. However, she focuses on the impact of commuting in a sample of Czech border regions, and it is unclear if the results can generalized for all of Germany. Hammer and Hertweck, 2022 find short-term negative wage effects at the bottom of the distribution, but the overall effect across the whole wage distribution is null. Gyetvay and Keita, 2023 find that firms exposed to more EU-13 migration cut native wages and employment, but that native workers remain largely unaffected since they reallocate to different firms.

share in a region without issues of collinearity. Adding this control for the level of lagged immigration allows to leverage only variation in the origin composition. It, therefore, strengthens the case for instrument validity: conditional on the total migrant share in a region, having had more migrants from EU-13 countries rather than from other origins seems as good as randomly assigned from today's perspective. Extensive first-stage and falsification tests suggest that the empirical strategy works well in my context.

The remainder of this paper proceeds as follows. Section 1.2 describes the institutional details and provides an overview of immigration into Germany following the EU Eastern Enlargement. Section 1.3 describes the data sources and preparation. Section 1.4 presents key descriptive statistics. Section 1.5 outlines the empirical methodology. Section 1.6 discusses my results, and Section 1.7 concludes.

1.2 Setting

Thirteen new member states joined the European Union after the Eastern enlargements in 2004, 2007, and 2013. Politically, it marked the successful reunification of Europe following the end of the Cold War. Economically, it meant the largest expansion of the European labor market to date, and that between countries with stark wage differences. The income gap between new and old member states led some observers to raise concerns over a large immigrant influx and that unemployment in Germany would rise strongly if wages remained rigid (Burda, 2000; H.-W. Sinn, 2000; H.-.-W. Sinn, 2002).

Not least because of these fears, were old member states allowed to delay opening their labor markets for up to seven years. Germany, along with Austria, was the only country to exhaust the full waiting period. This means they were essentially forced to grant full labor market access to the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, Slovenia, Cyprus, and Malta in 2011, to Bulgaria and Romania in 2014, and to Croatia in 2015 (see Table 1..11 for the exact dates). Since then, workers from these countries essentially enjoy the same hiring conditions as persons with German citizenship (Illing, 2023).

The direction of EU-13 migration changed markedly after the end of the first transition phase in 2011. As the previous top destinations, Spain and Italy, were affected strongly by the burst of the housing bubble, Germany appeared as a feasible and attractive alternative destination (Bertoli et al., 2016). Figure 1..4 in the Appendix plots migration flows from EU-13 countries by origin and destination for the period 2011-2016 and thereby illustrates the central importance of Germany as a destination, especially for the large sending countries Bulgaria, Poland, and Romania. In fact, according to the data from Standaert and Rayp, 2022, 58% of total EU-13 migrant outflows during that period were directed to Germany. While these numbers should be interpreted with caution since the data relies on imputation, it does highlight the importance of this migration corridor during the early 2010s.

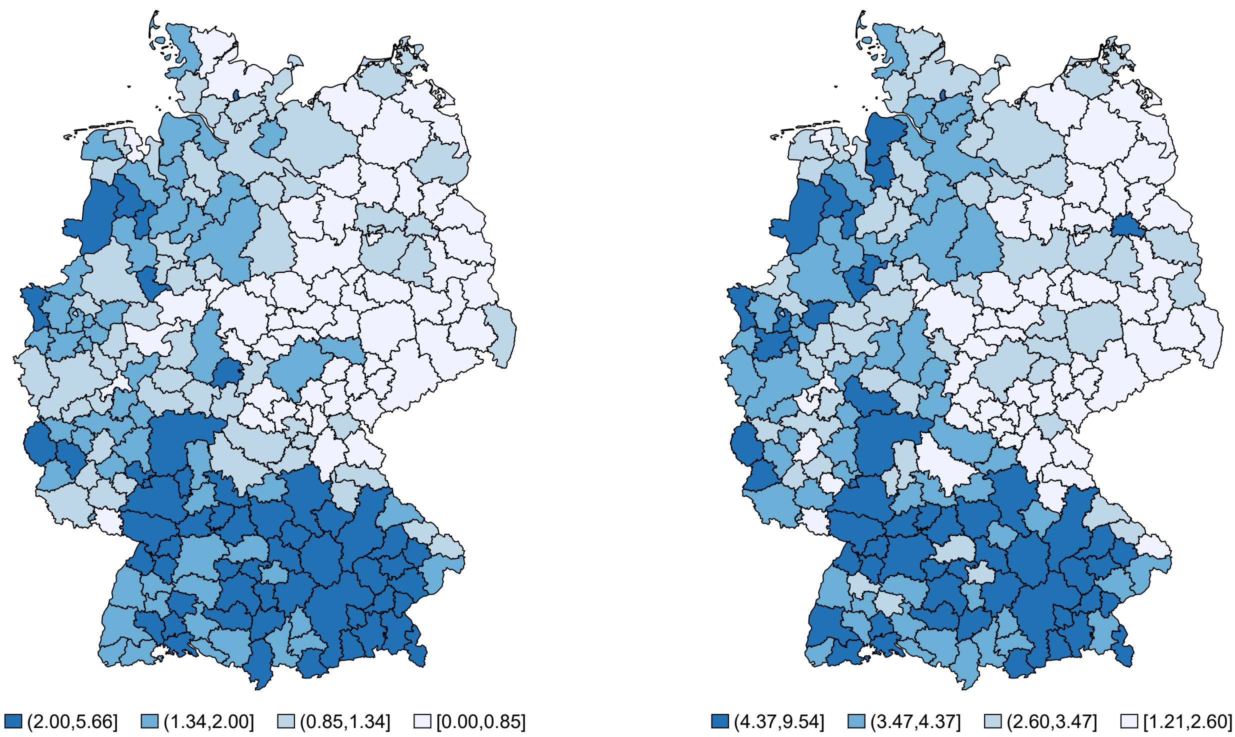
Germany witnessed a notable increase in migration over the following decade. Average yearly net inflows between 2010 and 2019 were at 450,000, compared to 100,000 in the previous decade (Adunts et al., 2022), and the share of foreigners in the total population increased from 8.7% to 12.2% over the same period (Eurostat, 2023). In employment shares, they even doubled from 6.5% to 12.9% between 2010 and 2020 (Gallegos Torres et al., 2022). While migration from asylum-sending countries gained much public attention, immigration from new EU member states was an equally if not more important contributor to this increase. With a net yearly inflow of 175,000 on average between 2011 and 2019 (Destatis, 2023), the share of EU-13 nationals increased markedly from 1.3% to 3.1% (Eurostat, 2023).

Migrants from EU-13 countries show high employment rates and are more likely to have completed vocational training than other migrant groups (Seibert & Wapler, 2020). Compared to German employees, they tend to be younger, predominantly male, and less educated on average (Gyetvay & Keita, 2023). The main industries were lower-paid sectors, particularly in services and agriculture, with the most common occupations being postal and warehouse positions, driving, cleaning services, construction, meat processing, and agricultural work (Gallegos Torres et al., 2022). The majority of EU-13 migrants thus entered occupations with less complex tasks and lower wages, an important segment of the Germany labor market in regard to inequality.

Figure 1.1 shows the geographical distribution of migrant inflows proxied as differences in stocks between 2010 and 2014 in percent of 2010 population size. Note that since these are in terms of population, increases are less pronounced than they would be in terms of working-age population. The pattern for EU-13 migration in panel (a) largely mirrors that of overall immigration in panel (b). The economically strong regions in the south of Germany experienced the largest increases in immigration, followed by regions in the West-North with stronger migrant networks. Eastern Germany has traditionally had a lower share of foreigners and also recently attracted fewer new migrants. Appendix Figures 1..5 to 1..7 show that the geographic distribution of immigrant inflows varies strongly depending on the country of origin. For example, in Figure 1..5, it can be seen that the distribution of Romanian is much more clustered in the South, while the inflow of Bulgarians involved much more in Western regions of Germany. The fact that migrants from different origins cluster in different regions will be an important component of the instrumental variable strategy in section 1.5.

Figure 1.1: Change in EU-13 and total migrant stock, % of 2010 population

(a) EU-13 (b) Total



Source: Destatis GENESIS-Online Table 12521-0040, own calculations. Figures plot the change in migrant stocks between 2010 and 2016 as a percentage of the 2010 population, on the left using EU-13 migration and on the right using total migration. Units are 223 commuting zones.

As mentioned by Gyetvay and Keita, 2023, this episode is well suited to study the labor market effects of immigration for at least three reasons. First, immigration to Germany was modest before 2011, such that concerns of serial correlation between inflows are a lesser concern in this context (Jaeger et al., 2018). Second, since Germany used the full waiting period, there is a seven-year gap between the removal of trade barriers and the labor market opening, safeguarding against confounding effects of the two shocks on the labor market. Third, the size of the shock generates enough variation in migration exposure across regions to identify potential effects.

This episode is also interesting since it is one during which the German labor market showed a remarkable recovery. Between 2005 and 2019, unemployment fell from 11.7% to 5% (OECD, 2024c), employment rose from 65.5% to 75.7% (OECD, 2024a), and inactivity decreased from 26.2% to 22% (OECD, 2024b). This may be linked to the Hartz reforms, a series of legislative changes in the early 2000s that liberalized the labor market, rising demand for German products from abroad, or perhaps even the immigration itself (Dustmann et al., 2014). Although the research design will compare the evolution of different regions in Germany, effectively canceling out national-level trends, it is important to remember that this paper assesses the impact of immigration on unemployment during an economic boom.

1.3 Data

The main data source for this study is the Integrated Employment Biographies (IEB V16.01.00) provided by the Institute for Employment Research (IAB). This dataset links administrative records from social security notifications, unemployment benefits, job search, and participation in activation measures for the entire German labor force except civil servants and self-employed. It includes daily information about employment, wages, unemployment, occupation, and demographic information such as age, gender, nationality, and education. I complement the main dataset with variables containing information on education and wages imputed by the data center (Drechsler et al., 2023; Thomsen et al., 2018). Using these variables to minimize concerns over censoring in the wage and missing values in the education variable is common practice when working with this dataset.

In the spirit of Jung and Kuhn, 2014, I then use the daily spell data to assign a monthly labor market status around a reference date. The classification is based on the hierarchical ordering of employment, unemployment, and inactivity. A person is employed in a month if she has at least one part-time or full-time job or works as an apprentice on the 15$^{\text{th}}$ of that month. As in Hartung et al., 2022, I try to count civil service using information from (de-)registrations, although the correction seems to matter less in my case. Note that, in line with the official definition in Germany, individuals who only have a marginal job[2] do not count as employed. A person is unemployed if they are not employed and registered as a job-seeker at the employment agency. This means, at least in theory, that they are actively searching for a job of at least 15 hours per week and are available for the placement effort of the employment agency (Bundesagentur für Arbeit, 2024b). The inactive is everybody else who is neither employed nor unemployed but in the data. These individuals are either registered at the employment agency but not available to work, receive means-tested welfare benefits, participate in activation measures, or only have a marginal job.

Labor market transitions are calculated as changes in the labor market status of an individual in between subsequent months. The job-finding rate is the share of unemployed who are employed the following month. The job-separation rate is the share of employed who are unemployed the month after. At the participation margin, I consider the share of unemployed who become inactive and the share of inactive who enter unemployment. The job-to-job transition rate is the share of employees who switch establishments between months. Wage offers are proxied by natural logarithm of main job's wage in the first month after transitioning from unemployment into full-time employment.[3] The restriction to full-time employment is necessary since the data do not contain information on working hours. I create a similar measure for the wages of job-to-job switchers. All monthly measures will be aggregated to yearly averages at the local labor market level. The granularity of the data allows for doing so separately by gender, broad education level, and nationality.

[2] Marginal employment or "Mini-job" is a job where individuals either work a maximum of 70 days a year or receive up to 538 Euro per month (450 during the period of observation). In this case, the work contract can be exempted from social security contributions (Bundesagentur für Arbeit, 2024c). Mini-jobs can be observed in the data.

[3] The definition of main job is based on the daily wage while giving preference to full-time employment.

As a sanity check, I compare national-level aggregates derived from my data over the period from 2007 to 2019 to official statistics from the Federal Employment Agency. Figure 1..8 and Figure 1..9 in the Appendix compare the monthly (un-)employment rates and stocks from both sources. I find that my data follows the movement of the official statistics closely, with mean absolute percentage errors in the range of 0.01 and 0.07. Employment seems to be slightly higher in the official statistics, while the unemployment rate is slightly lower. The correction approach to proxy civil servants in the labor force reduces the gap between the official unemployment rates.[4] However, the overall fit is very good. In Figure 1..10, I compare estimates of the labor force and the number of inactive. The left panel shows that my data does a good job at tracking the labor force in terms of dependent employment.[5] The fit between my data and a proxy for the number of inactive[6] in the right panel is less good and strongly improves once I include borderline cases, such as people in activation measures in the pool of inactive. Still, there is no doubt that I capture only part of the out-of-the-labor force population with this data source. Figure 1..11 compares the yearly job-finding and job-separation rates from my data to official rates from (Bundesagentur für Arbeit, 2024a), and finds that they are in the same ballpark. The fit for the job-finding rate is even better, but both seem close enough. Other transition rates, starting wages, and disaggregated transitions are not available from official statistics. Taken together, the exercise shows that my data can approximate official statistics reasonably well.

I aggregate the monthly data to yearly averages across 223 official commuting zones of the Federal Office for Building and Regional Planning (BBSR). These commuting zones are derived by minimizing the number of commuters across regions and restricting commuting to a maximum of 45 minutes (Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR), 2024). As such, the definition of a labor market based on commuting zones is better suited

[4] At least one difference between my data and the official statistics is that I use the 15$^{\text{th}}$ as reference data while official statistics use varying cutoff dates that tend to lie around the last days of the month. My approach follows Jung and Kuhn, 2014 and is comparable to the US concept of using the second week in the month. However, the close alignment between the figures indicates that the choice of reference may not matter too much in practice.

[5] The extended definition also includes self-employment, which is difficult to measure in my data. Note, however, that none of the two definitions is favored by the employment agency and that Hartung et al., 2022 also use the definition in terms of dependent employment.

[6] The employment agency does not provide official statistics on the number of inactive. Instead, I proxy them as the difference between the working-age (18-65) population and the extended labor force.

than using administrative boundaries such as districts. I will, however, also assess the robustness of my results using different geographical demarcations. The regional aggregates are calculated using all native individuals in the region who are aged 25 to 55 during the year in question. The focus on prime-age workers reduces measurement error due to transitions between education or inactivity, which are difficult to observe in the data. I assign workers to regions based on the workplace and impute regions for the nonemployed using past jobs and, if still missing, the place of residence. Natives are individuals whose first observed citizenship was German, a necessary definition since the data do not include information on country of birth. However, using the first observed citizenship is meant to approximate country of birth.

For vacancies, I rely on custom extracts from the Federal Employment Agency (BA), which aggregate information on all the vacancies posted through their job portal. The data is provided at the month-district level, and I aggregate it to yearly averages for each commuting zone. Using this data has many advantages,[7] for example, that vacancy creation and filling can be tracked. However, it should be noted that, as with all available vacancy data, the subset of vacancies in the data is not a representative sample from the universe of open jobs. For the BA data, it can be assumed that they over-proportionately feature low-skilled jobs. As such, they are different, and perhaps complementary, to the common approach of sourcing vacancies from private online job portals (Hershbein & Kahn, 2018, for example), as the latter would likely oversample high-skilled jobs. The focus on low-skilled jobs in my data aligns closely with the jobs EU-13 migrants most often worked in but complicates measuring complementarities on high-skilled jobs.

I complement the vacancy data with information on firms' job search activity using the IAB vacancy survey (Börschlein et al., 2023). The advantage of this unique survey dataset is that it provides detailed information on interesting outcomes such as job search duration, hiring difficulties, or applicant quality. The disadvantage is that it is designed to be representative at the national level and may lack statistical power at local labor market level. I pool data from two consecutive years together to mediate this concern and increase sample sizes. Then, I compute weighted averages across commuting zones using

[7] The primary advantage is that no other vacancy data is available during the period of observation. The commonly used Glassdoor data (Grasso & Tatsiramos, 2023) starts only in 2014.

the provided survey weights. The outcomes will be used in section 1.6.2 to understand the mechanisms through which immigration affects hiring.

Data on migrants by origin at the regional level is available from Destatis (AZR) or the IEB. The drawback of the IEB is that it does not include all individuals, as seen before. The drawback of the AZR is that it does not allow to select working-age population, classifies migrants based on current citizenship, provides no corresponding number for natives, and is available only from 1998 onwards. I choose the IEB since working-age migrants close to the labor market are the relevant population to identify and since it allows me to base the shift-share instrument in 1993. The latter is important since the longer lag of the base year relative to the shock safeguards against spurious results due to serial correlation between migration inflows (Christian & Barrett, 2022; Jaeger et al., 2018).[8] To compute the migration variable, I count the number of individuals observed in the IEB on the 30$^{\text{th}}$ of June in each year. Migrants are those whose first observed citizenship is not German. The analysis will consider only working-age migrants, that is, individuals aged above 18 and below 65 years.

In the Appendix, I compare the time series of immigration aggregates at the national level across different sources. Figure 1..12 shows that the total number of migrants in the IEB is somewhat lower than using census projections, and considerably lower than in the AZR. Note, however, that the numbers in the AZR is not the ground truth but likely too high. One important advantage of the IEB is that it provides a matching estimate for the number of natives.[9] This can be seen when comparing estimates in terms of the migrant share in Figure 1..13. The numbers in the IEB are rather similar to the migrant share from census projections, and even more so when considering working-age population. Since the analysis will consider changes in the number of migrants relative to baseline population, it is less important to get the level but rather to get the relative shares right. For this purpose, the IEB is an accurate and flexible source.

Besides the main data sources described above, I rely on a number of other datasets

[8] If measurement error is classical, and the AZR has less measurement error (which is not clear a priori), the measurement error from using the IEB will lead to attenuation bias, while the potential serial correlation may lead to spuriously significant findings. Choosing measurement error over serial correlation is thus more conservative.

[9] Using the AZR only for migrants and other sources such as the census projection for natives would lead to a significant overestimation of the migrant share since the latter is more conservative.

to generate controls and additional outcomes. I use the Establishment History Panel 7522 v1 (BHP) (Ganzer et al., 2022) to measure plant entry and exit as well job flows. Additional outcomes at the regional level such as real Gross Domestic Product per capita (GDP), rural classification, value added, tax revenue are from INKAR (Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR), 2022). Regional housing prices are retrieved from REDX RWI-GEO-REDX (Klick & Schaffner, 2019). I combine the three original indices for the cost of appartment rentals, appartment purchase, and house purchase into a single index using Principal Component Analysis (PCA) and normalize the support between zero and one. Most variables are provided at the district level and collapsed to the commuting zones using population weights for averages. Lastly, for modifications of the instrument based on the diversion of migration flows, I also use cross-country migration flow data from Standaert and Rayp, 2022.

1.4 Descriptives

Before moving to the causal analysis, I provide a number of descriptive analyses on tightness, labor market flows, and its relation to immigration: First, migrants have lower job-finding rates, higher job-separation rates, and higher job-switching rates. Second, these rates have been fairly constant over the period of observation. Third, labor market tightness has increased over the period of observation, and matching efficiency seems to be higher for natives.

Table 1.1 presents the average monthly rates of job-finding, job-separation, and job-switching for migrants and natives, broken down by gender and broad education groups. The data, expressed as percentages, covers prime-age individuals from 2007 to 2019.[10] The overall monthly job-finding rate from unemployment to employment was just below 6% on average.[11] It was higher for males and more educated unemployed. Natives had higher job-finding rates than migrants across all groups, with the disparity most pronounced

[10] The overhaul in unemployment reporting with the Hartz-IV reforms caused substantial under- and misreporting in 2005 and 2006. While correction procedures are available (Hartung et al., 2022), it seems more principled, for the purposes of this paper, to start the observation period in 2007.

[11] The job-finding rate and job-separation rates are also very close to the ones provided in Jung and Kuhn, 2014 Table 1. The job-switching rate is larger for my sample, which could be related to my focus on prime-age workers.

among females. The overall job-separation rate from employment to unemployment was much lower and averaged just above 0.5% overall. It was twice as high for migrants compared to natives and largest for migrants without any degree. Similarly, the job-switching rate, overall at around 3%, was almost double for migrants compared to natives and largest for migrants without a degree.

Table 1.1: Monthly transition rates in %

	Job-finding rate			**Job-separation rate**			**Job-switching rate**		
	All	Natives	Mig	All	Natives	Mig	All	Natives	Mig
All	5.850	6.237	4.500	0.529	0.488	0.840	3.089	2.871	4.712
Female	5.044	5.703	3.012	0.457	0.428	0.726	3.093	2.943	4.482
Male	6.494	6.654	5.815	0.588	0.540	0.909	3.085	2.809	4.853
No degree	2.745	2.537	2.984	0.865	0.729	1.062	4.436	3.896	5.225
Vocational	6.499	6.586	5.876	0.535	0.507	0.870	2.954	2.840	4.266
Tertiary	9.450	10.078	7.068	0.398	0.369	0.674	2.879	2.755	3.979

Notes: Data from Integrated Employment Biographies, years 2007 to 2019, individuals aged 25 to 55. The job-finding rate is the number of unemployed finding a job the following month as share of unemployed multiplied. The job-separation rate is the number of employed entering unemployment the following month as share of employed. The job-switching rate is the number of employed in this and the next month but whose main employer in the current month is not among the first three employers in the next month. All rates are yearly averages of monthly transitions rates and multiplied by one hundred to yield percentages.

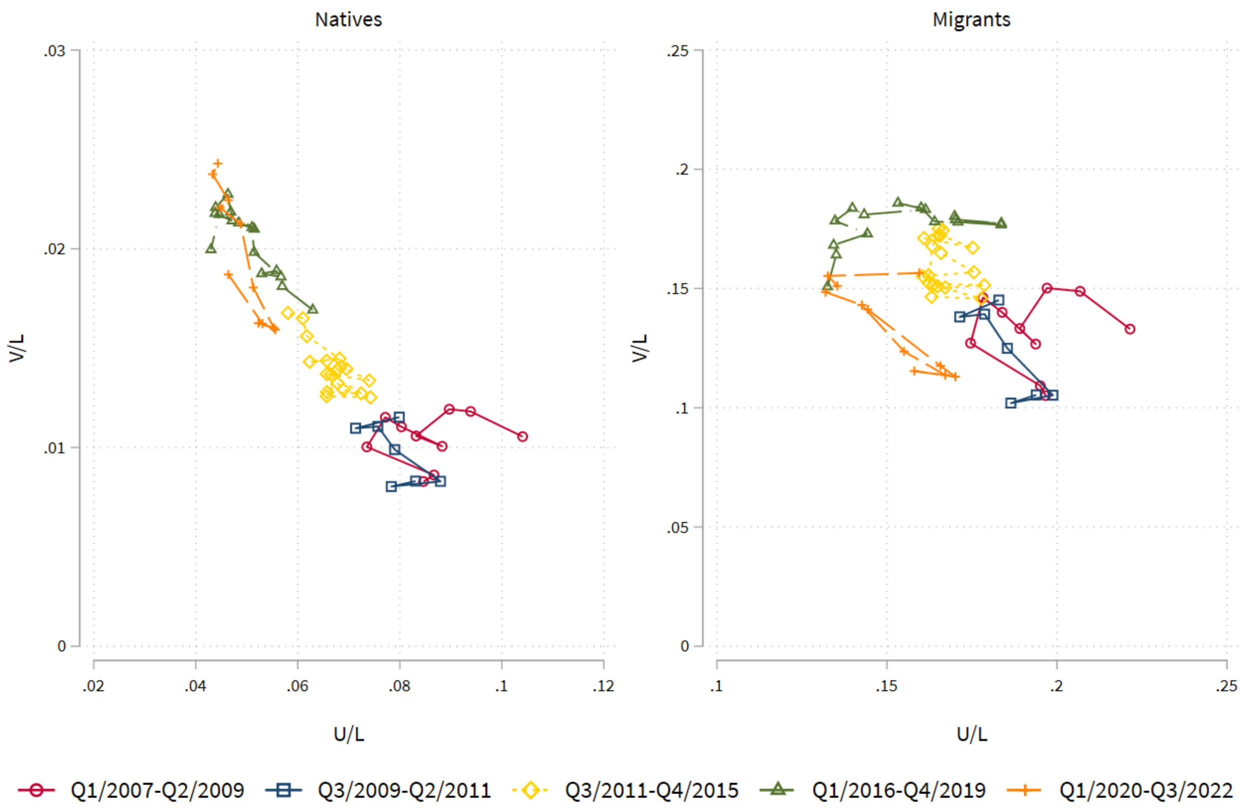
The findings align with previous evidence from Spain and France, showing lower job-finding rates, higher job-separation rates, and higher job-switching rates among migrants. This stands in contrast to the US, where transition rates are much more similar across groups (Chassamboulli et al., 2024). It is worth noting that our qualitative findings align despite using different data sources (administrative records versus labor force surveys) and computing transitions at different frequencies (monthly versus quarterly). They also confirm the modeling decision of Battisti et al., 2018 to assume larger job-separation rates as a key differentiator between migrant and native workers, at least for European labor markets.

The evolution of transition rates over time is shown in Figures 1..15 to 1..16 in the

Appendix. They show that the gap between natives and migrants remained largely stable over time. Since its peak in 2009, job-finding rates have been decreasing and job-separation rates have been increasing for all groups. Only the job-switching rate shows a slight diversion, as it has been increasing more for migrants compared to natives.

The favorable development in the labor market is also evident in the Beveridge curves presented in Figure 1.2. The figure plots the relationship between the vacancies and unemployment, as a ratio of the labor force, separately for migrants and natives. Both curves display an upward movement along the curve between 2007 and 2019. This indicates that the labor market for migrants and natives has become tighter over the observation period. However, it seems that the upward movement was less strong for migrants. Also, the position of the Beveridge curve is more outward, indicating a lower matching efficiency between jobs and migrant job-seekers. These differences could be explained by labor market segmentation, which would indicate a low degree of substitutability between migrants and natives in Germany.

Figure 1.2: Beveridge Curves



Source: Federal Employment Office (BA), own calculations. Monthly numbers averaged to quarterly frequency. Figures plot the ratio of native or migrant unemployed over the workforce, respectively, on the X-axis and the ratio of vacancies over the workforce on the Y-axis.

1.5 Empirics

The aim of this paper is to estimate the elasticity of unemployment, vacancies, transition rates, and wage offers with respect to changes in immigration in a labor market. The research design leverages variation in immigrant inflows across local labor markets to compare regions that experienced significant EU-13 migration after 2011 to regions that did not. In the spirit of Edo, 2020, I implement it using the following difference-in-difference regression with continuous treatment

$$\Delta Y_{r,2010-2014} = \alpha + \beta \cdot \Delta M_{r,2010-2014} + \theta X_{r,2010} + \varepsilon_r \tag{1.1}$$

where $\Delta Y_{r,2010-2014} = Y_{r,2014} - Y_{r,2010}$ is the long difference of each outcome from 2010 to 2014 across 223 official commuting zones indexed by r. I choose 2010 as baseline since this is the last year before the end of the first transition phase for the EU-10 countries. The year 2014 is well-suited as post-treatment period since the transition period for Bulgaria and Romania ended in that year, accompanied by a notable increase in immigration to Germany. Extending the post-treatment period further is possible but could lead to contamination from the 2015 minimum wage introduction and the asylum seeker influx in 2015/16. The explanatory variable of interest is $\Delta M_{r,2010-2014} = \frac{M_{r,2014} - M_{r,2010}}{L_{r,2010}} \cdot 100$ which captures the working-age migrant inflow from EU-13 countries, proxied by changes in stocks, as a share of the 2010 working-age population. The specification implicitly controls for region fixed effects due to the differencing. Additional controls $X_{r,2010}$ in the preferred specification include indicators for east and rural, as well as indicators for quintiles of the 1993 total migrant share. For more extensive controls, I also add dummies for regions with above median GDP per capita, rental price level, or share of highly educated workers in the baseline period. All observations are weighted by native population size in the base year and the error term ε_r is clustered at the level of 50 broad commuting zones from Kropp and Schwengler, 2016.[12]

The parameter of interest is β, which captures the effect of changes in EU-13 immigration density on the growth of outcomes such as unemployment or vacancies. Since $\Delta M_{r,2010-2014}$ measures changes in the total number of migrants across all skill and occupation groups, estimates will deliver total rather than relative effects of immigration (Dustmann et al., 2016). Unlike other recent papers studying labor market effects on workers (Castellanos, 2023; Delgado-Prieto, 2024; Foged & Peri, 2016), my analysis is at the commuting zone level, and thus β identifies aggregate effects of immigration on a region's labor market rather than effects on workers (Dustmann et al., 2023). The labor market-wide effect seems to be the appropriate object to identify, as it maps closely to the theoretical counterparts.

A drawback of the difference-in-difference strategy in equation 1.1 is that immigrants move to regions that experience positive labor demand shocks as wages will grow faster there. If true, this would violate the parallel trends assumption. To mitigate this concern,

[12] The specification in differences does not allow for clustering at the treatment level, as this is the unit of observation. Clustering at the federal-state level would have led to a small number of clusters without being conceptually superior, since demarcations based on worker flows likely capture intra-group correlations better than administrative boundaries.

I employ a shift-share instrument common in this literature (Altonji & Card, 1991; Card, 2001). The idea is to isolate variation in migrant location choice driven by the presence of migrant networks in a region. Since networks function as amenities or reductions to migration costs, they increase migration to a region irrespective of local labor demand. I construct the predicted migration inflow as follows

$$\Delta\widehat{M} = \frac{1}{L_{r,2010}} \cdot \sum_{o=1}^{8} \frac{M_{or,1993}}{M_{o,1993}} \cdot \Delta M_{o,2010-2014} \tag{1.2}$$

where $\frac{1}{L_{r,2010}}$ is the labor market size in the baseline year 2010, $\frac{M_{or,1993}}{M_{o,1993}}$ is the historical share of migrants from origin o who resided region r in 1993 relative to all migrants from origin o in 1993, and $\Delta M_{o,2010-2014}$ is the current inflow of migrants from origin group o at the national level.[13] The instrument redistributes nationwide inflows to areas according to the origin composition in 1993. Conditional on covariates, the exclusion assumption requires that the lagged shares of each origin are as-good-as-randomly assigned and independent of current unobserved labor demand shocks (Goldsmith-Pinkham et al., 2020). Importantly, and unlike most papers in this literature, the set of covariates includes the lagged total share of migrants in a region. This covariate is important since it is plausible that most violations of the exclusion restriction will stem from level effects. For example, an economically strong region like Stuttgart was attractive for migrants in 1993, and is attractive for migrants today since it still offers a vibrant economy with many jobs and relatively high wages. However, conditional on the total migrant share, whether Stuttgart had more migrants from EU-13 countries or, instead, from other origins in 1993 is much more idiosyncratic. In other words, by controlling for the historical *level* of immigration, we can leverage the *composition* of immigration; and the composition is much more likely to be as good as randomly assigned from today's perspective.[14]

I also use a slightly modified version of the shift-share instrument described above. The motivation for this modification is that a part of the migration flows to Germany during the 2010s were diverted from Southern Europe as these countries were affected more by the global financial crisis (Bertoli et al., 2016). I use this to construct a different,

[13] I group to ten origins by putting small adjacent origins together, for example Lithuania, Estonia, and Latvia. See Table 1..12 in the Appendix for a list of the groups.

[14] The approach is similar to the re-centering idea of Borusyak et al., 2022.

perhaps more exogenous, shift component for the shift-share instrument using

$$\Delta\widehat{M}_{div} = \frac{1}{L_{r,2010}} \cdot \sum_{o=1}^{8} \frac{m_{or,1993}}{m_{o,1993}} \cdot \sum_{d=1}^{2} \sum_{t=2011}^{2014} I_{o,d,t} - I_{o,d,2006} \tag{1.3}$$

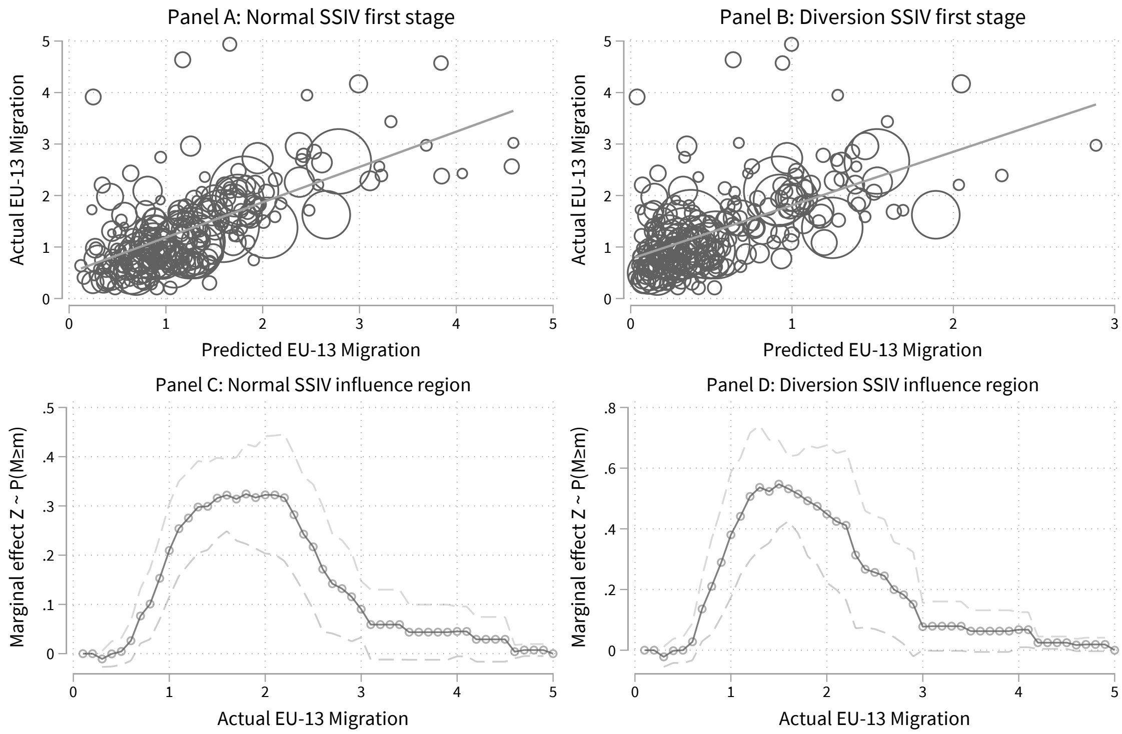
where d stands for the destinations Spain and Italy and $I_{o,d,t}$ are yearly bilateral migration inflows from Standaert and Rayp, 2022.[15] The term $\sum_{d=1}^{2}\sum_{t=2011}^{2014} I_{o,d,2006} - I_{o,d,t}$ is the cumulative gap of migrant inflows from EU-13 countries to Spain and Italy in the years 2011 to 2014. These are measured as the difference between the actual flows in each year and the inflows in 2006, which is prior to the housing crisis. The gap represents the *missing flows* which were diverted from Spain and Italy to Germany due to the fact that the former countries were affected more strongly by the housing crisis. The identification argument is that the 2006 origin composition of migration flows to Spain and Italy is plausibly random from Germany's perspective. For example, Spain has had larger inflows of Romanians, while Polish migrants preferred the United Kingdom. The missing flows' shifter uses the fact that Spain, rather than the United Kingdom, had a strong recession and thus predicts a larger diversion of Romanians to Germany and a lesser inflow of Polish migrants. If successful, this modification allows me to base instrument validity not only on the exogeneity of the shares but also on the exogeneity of the shocks as in Borusyak et al., 2022.

For the instruments to be relevant, predicted EU-13 migration needs to be strongly associated with actual EU-13 migration. Panel (a) and (b) of Figure 1.3 provide a first visual assessment of whether this assumption holds in practice. They plot the predicted migration change on the X-axis versus the actual change in migration on the Y-axis, on the left using the standard shift-share instrument from equation 1.2 and on the right using the diversion instrument from equation 1.3. Two findings arise from the figure. First, both instruments predict the actual change in migration relatively well. The scatter plots hinge on a relationship, which is confirmed by the strongly positive slope of the fitted line. The correlation is around 0.6 with a mean absolute percentage error of 0.4 and 0.58. It

[15] Spain and Italy are the most relevant Southern European destinations. The other relevant destination, the United Kingdom, was not affected as strongly by the recession and did not experience a comparable drop in immigration in the late 2000s. The IV procedure is robust to calculating missing flows using other southern European destinations. Results are available upon request.

also seems that the regions with the largest deviations tended to have smaller 2010 native population sizes, as indicated by the size of the marker, and will thus receive less weight in the estimation. Second, the standard shift-share instrument is a better predictor, in particular, as the diversion instrument systematically underestimates immigration: The mean of actual migration change was 1.37pp. The mean of the standard shift-share comes close at 1.26pp while the mean of the diversion instrument is only 0.58pp. Note, however, that this is not unexpected: the idea of the modification is to isolate the migration flows that were diverted to Germany due to the housing crisis in Southern Europe. This will be a subpart of total migration, and the levels ought to be lower. More importantly, the fitted line is strongly positive, which indicates that the instrument can capture differences between regions.

Figure 1.3: First stage relationship



Source: Integrated Employment Biographies (IEB), own calculations. Figures plot the predicted EU-13 migration change on the X-axis and actual change in EU-13 migration between 2010 and 2014 on the Y-axis, on the left using the standard shift-share instrument from eq. 1.2 and on the right using the diversion instrument from eq. 1.3. Units are 223 commuting zones. The size of the markers is the 2010 native population size. The fitted lines ar from unconditional regressions of the predicted on the actual change in EU-13 migration.

Instrumental variable regressions with a continuous treatment estimate an average causal response (ACR) across different doses with weights proportional to the strength of each marginal first-stage relationship (Angrist & Imbens, 1995). It is thus informative to assess at which values of the endogenous variable the instrument is strongest. I do so in Panels (c) and (d) of Figure 1.3 by estimating the first stage regression for each marginal increase in EU-13 immigration

$$\mathbb{I}[\Delta M_{r,2010-2014} \geq m] = \alpha + \beta \cdot \Delta \widehat{M} + \theta X_{r,2010} + \varepsilon_r \tag{1.4}$$

where $\mathbb{I}[\Delta M_{r,2010-2014} \geq m]$ is an indicator if the inflows of EU-13 migrants in a commuting zone were above the threshold m and the remaining variables are as before. I increase threshold m in steps of 0.1pp to trace out at which points in the distribution the instrument marginally shifts commuting zones to receive more migration. Two findings arise from the exercise. First, the two instruments operate broadly in the same region of the actual immigration shock. This is reassuring since it implies that their results are roughly comparable.[16] Second, the region of largest influence is at moderate levels of immigration between 1pp and 2.3pp increases relative to baseline population. In Appendix Figure 1..17 I repeat the exercise using baseline native population weights. I find that the region of highest influence is broadly similar and extends perhaps further until 2.6pp using a somewhat arbitrary but intuitive cutoff of $\beta = 0.3$. In the next step, I define the 111 regions with an EU-13 immigration inflow in this range as "influential" and compare their characteristics to other regions and the entire sample. The summary statistics in Appendix Table 1..14 show that the influential commuting zones are slightly more populated and economically prosperous but broadly similar to the full sample or the remaining regions. They are, however, much less likely to be in East Germany and, consequently, had more immigrants in 1993. One should, therefore, be cautious when drawing policy conclusions for Eastern Germany based on the results of this study. As part of the robustness checks, I will, however, re-run the regressions excluding the influential commuting zones.

Relatedly, Goldsmith-Pinkham et al., 2020 show that under certain assumptions, shift-share IV regressions are equivalent to weighted IV regressions with each origin

[16] Comparing results from different instruments can be challenging under treatment effect heterogeneity since differences may be due to the fact that complier populations differ between the instruments.

share as a separate instrument.[17] They suggest running a diagnostic to assess the risk of negative weights placed on the origin-specific instruments. I implement their so-called Bartik decomposition in Appendix Tables 1..15 and 1..16 for the main outcomes and both instruments. The results show that virtually all origin-specific shares receive positive weights α, and the only negatively-weighted origin shares have weights very close to zero. This means they will have virtually no influence on the aggregated estimation results. Interestingly, the diversion instrument puts 86% of the weight on Romania. This makes sense since Romania was the most important origin of EU-13 immigrants to Spain and Italy during the 2000s. The standard shift-share, on the other hand, is more balanced across the main origins. This shows that the identification of both instruments differs somewhat, which means that using both is useful to assess the robustness of the findings. Less crucial but also reassuring is that for both instruments and across virtually all outcomes, the important origins yield second-stage effects β of the same sign. Overall, the results from this exercise imply that negative weighting issues are unlikely to drive my results.

Table 1.2 investigates the relevance of the instruments more systematically by showing first-stage regression results of the shift-share instruments on changes in immigration across commuting zones. Panels A and C provide unconditional correlations, while panels B and D give results conditional on the preferred set of pre-treatment controls. Column (1) shows the association between the instruments and the endogenous variable of interest, that is, the 2010 to 2014 change in migration from EU-13 countries. The coefficients imply around 0.7 percentage point increases in actual immigration for each percentage point increase of predicted immigration for the normal shift-share and 1 for the diversion instrument. The small standard errors mean that we can rule out null coefficients at very high statistical significance. As expected, the Montiel Olea and Pflueger, 2013 effective F-statistic is higher for the normal shift-share instrument, but both exceed the commonly used rule-of-thumb value of ten against weak-instrument bias (Staiger & Stock, 1997).

[17] Recent methodological contributions on shift-share-like instruments have argued that identification can come either from homogeneity of the shifts (Borusyak et al., 2022) or homogeneity of the shares (Goldsmith-Pinkham et al., 2020). In the case of the immigrant networks instrument, the share exogeneity seems more plausible since the shifts are few and likely correlated across origins and with destination macroeconomic conditions (Anelli et al., 2023).

The results for total immigration in column (3) are broadly similar. The coefficients have meaningful magnitudes and high statistical significance. The F-statistics are, as expected, lower than for EU-13 migration, and in the range of 10. While the regression will use immigration from EU-13 countries as explanatory variable of interest, it is reassuring that this is a meaningful component of total immigration during the period of observation.

Table 1.2: First stage regression results

	(1) $\Delta M_{EU13,post}$	(2) $\Delta M_{EU13,pre}$	(3) $\Delta M_{Total,post}$	(4) $\Delta M_{Asylum,post}$	(5) $\Delta M_{Balkan,post}$
Panel A. SSIV					
$\Delta \widehat{M}$	0.742	0.043	1.292	0.014	0.063
	(0.107)	(0.024)	(0.185)	(0.011)	(0.017)
MP F-Stat	47.650	3.372	48.627	1.783	12.903
Panel B. SSIV + controls					
$\Delta \widehat{M}$	0.745	0.017	0.927	-0.029	0.028
	(0.142)	(0.032)	(0.281)	(0.014)	(0.023)
MP F-Stat	27.650	0.305	10.876	4.145	1.441
Panel C. Diversion SSIV					
$\Delta \widehat{M}_{div}$	1.007	0.060	1.722	0.007	0.097
	(0.200)	(0.034)	(0.326)	(0.017)	(0.027)
MP F-Stat	25.359	3.063	27.851	0.151	13.071
Panel D. Diversion SSIV + controls					
$\Delta \widehat{M}_{div}$	1.011	0.029	1.352	-0.055	0.056
	(0.249)	(0.038)	(0.418)	(0.027)	(0.032)
MP F-Stat	16.429	0.584	10.455	4.236	2.960
$\bar{Y}$	1.368	0.129	2.149	0.105	0.096
N	223	223	223	223	223

Cluster-robust standard errors in parentheses. *Post*: 2010 - 2014, *Pre*: 2007 - 2010. Outcomes stand for EU-13 migration in the *post* period (1), EU-13 migration in the *pre* period (2), total migration in the *post* period (3), migration from asylum-sending countries in the *post* period (4), and migration from west-balkan countries in the *post* period (5). All outcomes expressed as percentage point changes relative to baseline population. N is the number of observations. $\bar{Y}$ is the mean of the dependent variable. $\Delta\widehat{M}$ is the shift-share instrument as described in equation 1.2. $\Delta\widehat{M}_{div}$ is the modified diversion shift-share instrument as described in equation 1.3. Regressions in Panel A and C include the instrument and a constant. Regressions in Panel B and D also include dummies for quintiles of the total migrant share in 1993 as well as indicators for rural region and eastern Germany. Observations weighted by 2010 native population. *MP F-Stat* is the Montiel Olea and Pflueger (2013) effective first-stage F-Statistic implemented using *weakivtest* from Pflueger and Wang (2015).

The association between the predicted and actual migration flows is much less strong in the pre-treatment period and for migration from other important origins. Column (2) shows that the association between the instrument and the change in migration between 2007 and 2010 is smaller and statistically insignificant. The overall explanatory power of the regression is small, as indicated by the low values of the F-statistic. This is expected since migration from EU-13 countries was low before 2011, but it is reassuring for the validity of the differences-in-differences design. The instrument also does not seem to predict immigration from the two other main sending regions. This is evident from the association between the instrument and immigration from the main asylum-sending countries (Eritrea, Iraq, Iran, Nigeria, Pakistan, Somalia, Syria, and Ukraine) in column (4) as well as with immigration from western Balkan countries (Albania, Bosnia and Herzegovina, Serbia, Montenegro, Northern Macedonia, and Kosovo) in column (5). The coefficients are rather small, do not produce strong F-statistics, partly insignificant and partly of the wrong sign, at least conditional on covariates. Taken together, the evidence from Table 1.2 confirms that the shift-share instruments indeed predict the increase in immigration in the post-treatment period through its increase in immigration from new EU member states.

The exclusion restriction for instrument validity requires that the historical presence of migrants from EU-13 countries affects the outcomes only by stimulating more current EU-13 migration to regions with larger networks. An important case in which this assumption is violated is when general equilibrium adjustments to previous migration episodes create a serial correlation between historical and current migrant inflows (Jaeger et al., 2018). As mentioned, this is unlikely to be the case in this context since the last important immigration episode occurred in the 1990s, and a number of important shocks, such as the Hartz IV reforms or the global financial crisis, occurred in between. It is also unlikely since the composition of migrant flows changed markedly between the 1990s and the 2010s (see Appendix Table 1..13 for a comparison of the main origins).

More generally, conditional on covariates, the instrument should not be related to unobserved time-varying determinants of labor demand. While this assumption is not directly testable, it is possible to test it indirectly by assessing the correlation between the instrument and important determinants of labor demand before 2011. Since immigration from EU-13 was low before 2011, we would expect the instrument to not be strongly associated with changes in these variables during this time period. Table 1.3 conducts placebo tests for pre-trends tests on a series of relevant variables. All outcomes are differences between the last baseline and the first sample period $\Delta Y_{r,pre} = Y_{r,2010} - Y_{r,2007}$.[18]

Table 1.3: Pre-trend regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\Delta\, Edu_{pre}$	$\Delta\, VA_{Mfg,pre}$	$\Delta\, ln\ Inc_{pre}$	$\Delta\, Price_{pre}$	$\Delta\, ln\ Pop_{nat,pre}$	$\Delta\, ln\ V_{pre}$	$\Delta\, ln\ U_{nat,pre}$
Panel A. OLS							
$\Delta\, \widehat{M}$	0.227	0.349	-0.003	0.007	0.015	-0.078	0.040
	(0.065)	(0.210)	(0.008)	(0.002)	(0.003)	(0.028)	(0.016)
Panel B. SSIV + controls							
$\Delta\, \widehat{M}$	-0.037	0.625	-0.002	0.010	0.003	-0.050	-0.020
	(0.036)	(0.226)	(0.008)	(0.003)	(0.004)	(0.036)	(0.010)
Panel C. Diversion SSIV							
$\Delta\, \widehat{M}_{div}$	0.287	0.460	-0.004	0.006	0.020	-0.099	0.074
	(0.070)	(0.363)	(0.010)	(0.004)	(0.005)	(0.043)	(0.030)
Panel D. Diversion SSIV + controls							
$\Delta\, \widehat{M}_{div}$	-0.007	0.941	-0.003	0.012	0.004	-0.067	-0.004
	(0.050)	(0.408)	(0.009)	(0.003)	(0.007)	(0.051)	(0.021)
$\bar{Y}$	0.787	-0.278	0.001	0.000	-0.008	0.005	-0.179
N	223	223	223	223	223	223	223

Cluster-robust standard errors in parentheses. *Pre*: 2007 - 2010. Outcomes are the share of high-educated workers (1), the share of value in manufacturing from INKAR (2), the natural logarithm of real GDP per capita (3), the housing price index (4), the natural logarithm of native population (5), the natural logarithm of vacancies (6), and the natural logarithm of the stock of native unemployed (7). All outcomes expressed as differences between 2007 and 2010, except (4) where the earliest available data is 2008. $\Delta\widehat{M}$ is the shift-share instrument as described in equation 1.2. $\Delta\widehat{M}_{div}$ is the modified diversion shift-share instrument as described in equation 1.3. Regressions in Panel A and C include the instrument and a constant. Regressions in Panel B and D also include dummies for quintiles of the total migrant share in 1993 as well as indicators for rural region and eastern Germany. Observations weighted by 2007 native population. N is the number of observations. $\bar{Y}$ is the mean of the dependent variable.

The first three columns of Table 1.3 assess differences in the economic potential across regions. In column (1), both instruments show a positive unconditional association of the instruments with changes in the share of highly educated workers, but this relation vanishes once controls are introduced. In column (2), I find that the instrument predicts more migration for regions with a negative trend in the share of manufacturing in value-added, but that, again, this relationship is not statistically significant at conventional levels once controls are introduced. In column (3), a negative and insignificant coefficient for the change in real income per capita. In columns (4) and (5), I check whether the instrument correlated with pre-trends of agglomeration forces. Reassuringly, the positive correlation between changes in rental prices in column (4) and the change in native population (5) may also be zero, given the large standard errors. Lastly, looking at two

[18] The change in the price level in column (4) is from 2008 to 2010 since the data is only available from 2008 onwards.

key outcomes in columns (6) and (7), I find that the instrument does not strongly correlate with either the change in vacancies or the change in native unemployed before the policy change. Taken together, while not perfect, the results indicate that the instrument does not correlate strongly with many important indicators of regional labor demand growth or important outcomes in the pre-treatment period. While the tests before are reassuring that the instrument is not purely driven by reverse causality, failing to reject pre-trends that the instrument is valid. As part of the robustness checks, I will use the Union of Confidence Intervals (UCI) method of Conley et al., 2012 to allow for a direct effect of the instrument on the outcomes and report the extent to which the exclusion needs to be violated for the 95% confidence bounds to include zero.

Lastly, Blandhol et al., 2022 point out that under treatment effect heterogeneity the linear specification must well approximate the conditional mean of the instrument given covariates. Following their suggestions, I implement a RESET misspecification test for the first stage regression in Appendix Table 1..17. The null hypothesis of no misspecification is rejected for one of the two instruments at the 10% significance level. I nonetheless keep the linear functional form as main specification for both instruments since it is easier to interpret, does not require arbitrary hyperparameter choices, and is probably more robust in modestly sized samples. However, as part of the robustness checks, I implement the Chernozhukov et al., 2018 double-debiased machine learning approach suggested by Mogstad and Torgovitsky, 2024 to deal with potential misspecification of the confounders. Assessing the robustness of my findings to using a semi-parametric model allows me to judge the likely importance of misspecification bias in my setting.

1.6 Results

1.6.1 Main Results

This section documents the main findings of the paper, that is, the effect of immigration on regional unemployment, vacancies, tightness, labor market transitions, and wage offers as proxied by starting wages. To start with, Table 1.4 presents the two-stage least squares regression results for unemployment, tightness, and vacancies controlling for east, rural, and the 1993 migrant share. The top panel shows OLS estimates which measure the conditional correlation, while the panels below give the second-stage coefficient estimates using the standard as well as the modified shift-share instrument.

Table 1.4: Results for native unemployment and vacancies

	(1) $\Delta\, ln\, U_{nat}$	(2) $\Delta\, Urate_{nat}$	(3) $\Delta\, ln\, V$	(4) $\Delta\, ln\, V_{fill}$	(5) $\Delta\, ln\, V_{new}$	(6) $\Delta\, V/U_{nat}$
Panel A. OLS						
ΔM_{EU13}	-0.017	0.089	-0.012	-0.012	-0.024	0.025
	(0.009)	(0.052)	(0.018)	(0.013)	(0.015)	(0.007)
ϵ	-0.030	0.152	-0.020	-0.021	-0.040	0.044
Panel B. 2SLS, normal IV						
ΔM_{EU13}	-0.066	-0.160	0.014	0.003	0.003	0.043
	(0.025)	(0.118)	(0.026)	(0.018)	(0.018)	(0.014)
ϵ	-0.113	-0.274	0.024	0.005	0.004	0.074
Panel C. 2SLS, diversion IV						
ΔM_{EU13}	-0.083	-0.128	0.068	0.043	0.040	0.079
	(0.034)	(0.114)	(0.047)	(0.039)	(0.037)	(0.028)
ϵ	-0.142	-0.220	0.117	0.074	0.068	0.136
$\bar{Y}_{2010}$	9.991	8.657	7.935	7.137	7.191	0.147
$\Delta\bar{Y}$	-0.178	-1.463	0.349	0.047	0.023	0.100
N	223	223	223	223	223	223

Cluster-robust standard errors in parentheses. Outcomes abbreviations stand for the natural logarithm of native unemployment (1), the native unemployment rate (2), the natural logarithm of open vacancies (3), the natural logarithm of the number filled vacancies (4), the natural logarithm of newly posted vacancies (5) and the ratio of vacancies over native unemployment (6). All outcomes are expressed as differences between yearly averages in 2014 and 2010. ΔM_{EU13} is the change in working age EU-13 migrant stock between 2014 and 2010 relative to the 2010 working age population expressed as percentage points. Controls include indicators for quintiles of the total migrant share in 1993, rural region, and eastern Germany. Observations weighted by 2010 native population. N is the number of observations. $\bar{Y}_{2010}$ is the weighted mean level outcome in 2010. $\Delta\bar{Y}$ is the weighted mean of the dependent variable (in differences). ϵ is the elasticity at baseline mean values, calculated as $\epsilon \approx \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \cdot 100$ for logarithmic and $\epsilon = \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \frac{1}{\bar{Y}_{2010}} \cdot 100$ for other outcomes. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016.

The results in columns (1) and (2) point towards a negative relationship between immigration from EU-13 countries and native unemployment. In column (1), both the OLS and IV coefficients are negative in the range between -0.017 and -0.083, the IV estimates being larger and highly statistically significant. The semi-elasticities of the IV regressions imply that a percentage point increase in immigration relative to the 2010 population decreases the growth rate in native unemployment by between 6.8% and 8.3%. The results for the native unemployment rate in column (2) have similar or even larger magnitude, but lack statistical precision to rule out null effects.[19] To map my coefficients more closely to its theoretical counterparts, I back out the elasticity at 2010 mean values using the approximation[20]

$$\epsilon \approx \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \cdot 100 \tag{1.5}$$

and find an elasticity of native unemployment with respect to immigration in the range of -0.116 and -0.142. A back-of-the-envelope calculation may yield more intuition for the magnitude of these numbers: average native unemployment in a region was 21,829 in 2010, and it fell on average by 3,559 individuals over the period from 2010 to 2014. In regions with one percentage point more EU-13 immigration over that time period, the reduction in native unemployment would be, ceteris paribus, between 1,435 and 1,739 larger.[21] This is a very meaningful contribution to the *German labor market miracle*.

In columns (3) to (5) I assess the effect of immigration on vacancies and find positive yet insignificant coefficients throughout. The IV coefficients in column (3) show changes in the growth rate of vacancies of 2.2% and 6.9% for 1pp increase in immigration. The coefficients are rather small and would not pass hypothesis tests at conventional significance levels, possibly due to the fact that the vacancy data is more noisy.[22] The null effect differs from Anastasopoulos et al., 2021, who find a negative relationship between immigration and vacancies. The authors hypothesize this may be due to increases in the vacancy filling rate. To investigate this channel, I estimate the effect of immigration on changes in vacancy creation and filling in columns (4) and (5). The results yield, again, positive coefficients of small magnitude and with considerable imprecision, such that I can not rule out null effects. The last column (6) brings together the previous results by investigating

[19] The difference between clustering and not clustering the standard errors makes an important difference for this outcome. With heteroscedasticity-robust standard errors, all results are statistically significant at 5% level.

[20] For outcomes in levels, I use $\epsilon = \frac{\Delta Y}{\Delta M}\frac{M}{Y} = (\frac{Y_{2014}-Y_{2010}}{Y_{2010}})/(\frac{M_{2014}-M_{2010}}{M_{2010}}) = (\frac{Y_{2014}-Y_{2010}}{Y_{2010}})/(\frac{M_{2014}-M_{2010}}{L_{2010}}\frac{L_{2010}}{M_{2010}}\frac{100}{100}) \approx \beta \cdot \frac{M_{2010}}{L_{2010}}\frac{1}{Y_{2010}} \cdot 100$ since $\beta \approx (\frac{Y_{2014}-Y_{2010}}{Y_{2010}})/(100 \cdot \frac{M_{2014}-M_{2010}}{L_{2010}})$. The bars then indicate that the elasticity is evaluated at the mean of these variables. For outcomes in logarithms I also use that $\Delta \ln Y \approx \frac{Y_{2014-2010}}{Y_{2010}}$ for small changes in Y so that $\epsilon = \frac{dY}{dM}\frac{M}{Y} \approx \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \cdot 100$ with a similar derivation as before. The approach is akin to Borjas, 2003, p.1349 but accounts for the different definition of the migration variable.

[21] I use $e^{\bar{Y}_{2010}}$ to get the level average, $(e^{\Delta \bar{Y}} - 1) \cdot \bar{Y}_{2010}$ to get average level change, and $(e^{\beta} - 1) \cdot \bar{Y}_{2010}$ to get the contribution of immigration to the level change.

[22] The vacancy data captures only part of the labor market, and firms may double-post new or fail to remove filled vacancies.

the effect of immigration on native tightness, that is, the fraction of vacancies over native unemployment. The results suggest immigration from EU-13 countries has led to an increase in tightness, with elasticities between 0.084 and 0.136. The previous discussion would imply that the increase in tightness is driven primarily by reductions in unemployment.

In Appendix Tables 1..18 and 1..19, I assess the robustness of these findings to a wide variety of specification choices. In Panels A and B, I find that the results remain stable when using either 400 administrative districts or 105 broader commuting zones as unit of observation.[23] In Panels C and D, I find that the unconditional IV regressions deliver somewhat different results, but, more importantly, adding further controls for GDP per capita, price levels, and the share of highly skilled workers does not qualitatively affect the results. The unweighted results in Panel E are robust for vacancies and unemployment in stocks, but not in terms of the unemployment rate. The same holds in Panel F when excluding the 111 most influential CZ with a migrant inflow between 1 and 2.3pp (see Figure 1.3). The magnitude on unemployment is smaller but still statistically significant. This suggests that the previous results are not only due to the regression putting large weights on regions with low immigration inflows. In Panel G, I use a partially linear IV model with double-debiased machine learning as suggested by Mogstad and Torgovitsky, 2024 and find that the results on unemployment, vacancies and labor market tightness remain intact. This implies that misspecification bias is unlikely to drive my results. In Appendix Table 1..20, I find that bounds on unemployment and tightness remain above zero for violations of the exclusion restriction between 32% and 51% of the reduced form effect. Overall, with the exception of the unemployment rate, the findings in Table 1.4 are robust to a very demanding set of robustness checks.

It is interesting to compare the results from Table 1.4 to theoretical predictions from the search-and-matching literature. Proposition (1) in Michaillat, 2023 predicts a positive relationship between immigration and unemployment and a negative relationship between immigration and labor market tightness. Taken at face value, my results reject

[23] The point that estimated effects of immigration are more negative the larger the geographical unit of observation was first made by Borjas et al., 1996 for the US. It could be an indication of bias from spillover effects in regressions using smaller geographic units. This point also reinforces the argument for using commuting zones.

this hypothesis, as unemployment decreased and tightness increased. Any negative competition effects seem to have been more than offset by positive job creation effects (Albert, 2021). On the other hand, models in Chassamboulli and Palivos, 2014 or Battisti et al., 2018 would predict a positive effect on the stock of vacancies and, in particular, more vacancy creation. The results here are less conclusive. I do find a positive relationship between immigration and vacancy creation as well as the stock of vacancies, but the coefficient sizes are small and not statistically significant.

Table 1.5 below presents the regression results using a series of native transition rates as outcomes. The first two columns help rationalize the negative coefficients on unemployment previously found: immigration is linked to an increase in the native job-finding rate, and to a decrease in the native job-separation rate. In column (1), the average job-finding rate from unemployment was 6.5% in 2010, and it decreased slightly over the period of observation. Regions that had more immigration would witness a less strong decrease in the job-finding rate of around 0.1pp for each percentage point increase in EU-13 migration. This effect is not overly large for a four-year period, but it is meaningful compared with the average change and highly statistically significant, at least when using the standard shift-share instrument. Note, however, that it may partly be driven by decreases in the denominator. In column (2), I find that immigration was associated with a reduction in the native job-separation rate into unemployment. The coefficients appear to be smaller, yet have similar elasticities to those of the job-finding rate and are significant throughout. This is because the base level is much lower, as only around 0.7% of employees transition to unemployment each month. Column (3) provides the results for regressions of immigration on the native job-switching rate, that is, how many native employees switch establishments in between months. I find negative, partly significant coefficients with elasticities in the range of -0.024 and -0.038. While this measure does not have a clear normative interpretation on its own, jointly with column (2), it seems to suggest that workers not only find a job more often but also leave their current job less often in response to immigration. Surprisingly, immigration seems to have increased native job stability.

The last two columns of Table 1.5 show results for transition rates between unemploy-

Table 1.5: Results for native transition rates

	(1) $\Delta\, UE_{nat}/U_{nat}$	(2) $\Delta\, EU_{nat}/E_{nat}$	(3) $\Delta\, JJ_{nat}/E_{nat}$	(4) $\Delta\, UI_{nat}/U_{nat}$	(5) $\Delta\, IU_{nat}/I_{nat}$
Panel A. OLS					
ΔM_{EU13}	-0.005	-0.011	-0.056	-0.023	0.040
	(0.040)	(0.004)	(0.012)	(0.013)	(0.041)
ϵ	-0.001	-0.027	-0.038	-0.033	0.017
Panel B. 2SLS, normal IV					
ΔM_{EU13}	0.160	-0.023	-0.054	-0.018	0.056
	(0.059)	(0.010)	(0.018)	(0.018)	(0.077)
ϵ	0.042	-0.058	-0.037	-0.027	0.024
Panel C. 2SLS, diversion IV					
ΔM_{EU13}	0.092	-0.016	-0.036	-0.026	0.104
	(0.070)	(0.007)	(0.020)	(0.022)	(0.074)
ϵ	0.024	-0.038	-0.025	-0.038	0.044
$\bar{Y}_{2010}$	6.492	0.695	2.528	1.179	4.017
$\Delta\bar{Y}$	-0.084	-0.131	-0.108	0.151	0.559
N	223	223	223	223	223

Cluster-robust standard errors in parentheses. Outcomes are the share of employed natives who loose their job the following month in percentage points (1), the share of unemployed natives who find a job the following month (2), the share of employed natives who switch employers (3), the share of unemployed natives who become inactive the following month (4), and the share of native inactive who become unemployed the following month (5). All outcomes are expressed as differences between yearly averages in 2014 and 2010. ΔM_{EU13} is the change in EU-13 migrant stock between 2014 and 2010 relative to the 2010 population in percentage points. Controls include the total migrant share in 1993 as well as indicators for rural region and eastern Germany. Regressions are weighted by 2010 native population. $\bar{Y}$ is the mean of the dependent variable. ϵ is the elasticity at mean values, calculated as $\epsilon \approx \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \cdot 100$ for logarithmic outcomes and $\epsilon = \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \frac{1}{\bar{Y}_{2010}} \cdot 100$ for outcomes in levels. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016.

ment and inactivity. The participation margin is an important source of adjustment, as stressed by Carrillo-Tudela et al., 2021 in relation to search-and-matching models and by Borjas and Edo, 2021 in relation to the wage effects of immigration. Theoretically, if more immigration would reduce the job-finding rate and search costs remain stable, we would expect fewer natives to search for a job. Yet, in column (4), I find a negative but insignificant relationship between immigration and the unemployment to inactivity transition rate of natives. Similarly, in column (5), the coefficients on the transition rate from inactivity to unemployment are negative but insignificant across all specifications. Note, however, that only part of inactive individuals can be observed in the data[24], meaning that this

[24] An inactive person must be somehow in contact with the employment office to be in the data, that is, either receiving benefits or taking part in measures. Inactive individuals whose partner is working and therefore do not require benefits will, for example, likely not be observed.

analysis is more prone to measurement error and, potentially, attenuation bias.

Turning to the robustness of these findings in Appendix Tables 1..21 and 1..22, I find that the results are somewhat less stable than the results for unemployment and vacancies. Reassuringly, the choice of geographic unit again does not seem to drive the results. Removing all covariates flips the sign for the job-separation and job-switching rate, and running unweighted regressions turns the effect on the job-finding rate statistically insignificant. However, more importantly, adding more controls or removing commuting zones with moderate immigration inflows does not change the results in virtually all specifications. The double-debiased machine learning estimates preserve the same sign in all specifications and, with the exception of the job-finding rate, retain statistical significance. Thus, the results seem moderately stable overall as most of the specifications align with the main results, and those that do not align, such as specifications without covariates, seem less trustworthy in the first place. In Table 1..23, I find that the bounds on the statistically significant coefficients do not intersect with zero until violations of the exclusion restriction of around a third of the reduced form effect.

Comparing the previous results from Table 1.5 to the theoretical predictions, again, they seem to favor the positive predictions of models in the Chassamboulli and Palivos, 2014 tradition. Job-rationing models as in Michaillat, 2023 would instead predict a negative elasticity of native job-finding rates with respect to immigration. The findings are, however, consistent with the previously found effects on unemployment. The negative relationship between migration and the job-separation rate is also interesting from a modeling perspective. In most search-and-matching models, the job-separation rate is exogenous, and all adjustments occur through the job-finding rate (Battisti et al., 2018; Chassamboulli & Palivos, 2014; Michaillat, 2023). This would have implied null effects with respect to changes in immigration in the empirical application. My findings suggest that the job-separation rate needs to be endogenized as well. While the other columns are not directly related to theoretical predictions, the insignificant effects at the participation margin may, cautiously, indicate that two-state search-and-matching models may be sufficient. At least, I do not find strong evidence in favor of adding this dimension.

Table 1.6 below shows the two-stage least square regression results for changes in the natural logarithm of daily starting wages of full-time native workers deflated to 2015 Euros.[25] As mentioned, using starting wages is meant to proxy wage offers in search-and-matching models. It is also institutionally relevant since wages of incumbent employees are highly rigid in Germany. In this sense, testing for effects using starting wages is a kind of acid test for the effect of immigration on native wages. Since the previous results suggest that tightness did not respond to migration, the alternative hypothesis would be that adjustments through wages absorbed the increase in labor supply. My prior was, therefore, a negative relationship between EU-13 immigration and the starting wages of native workers.

Table 1.6: Results for native starting wages

	(1) $\Delta\, ln\, Wage$	(2) $\Delta\, ln\, Wage(UE)$	(3) $\Delta\, ln\, Wage(IE)$	(4) $\Delta\, ln\, Wage(JJ)$
Panel A. OLS				
ΔM_{EU13}	0.009	0.015	0.016	0.012
	(0.004)	(0.003)	(0.006)	(0.005)
ϵ	0.016	0.026	0.028	0.020
Panel B. 2SLS, normal IV				
ΔM_{EU13}	0.019	0.023	0.039	0.019
	(0.005)	(0.005)	(0.016)	(0.005)
ϵ	0.033	0.039	0.067	0.032
Panel C. 2SLS, diversion IV				
ΔM_{EU13}	0.025	0.026	0.050	0.023
	(0.007)	(0.007)	(0.023)	(0.006)
ϵ	0.043	0.044	0.086	0.040
$\bar{Y}_{2010}$	4.357	4.153	3.968	4.491
$\Delta\bar{Y}$	0.074	0.100	0.045	0.035
N	223	223	223	223

Cluster-robust standard errors in parentheses. Outcomes are the natural logarithm of the starting wage of all native job-switchers who start a full-time job (1), the natural logarithm of the starting wage of all native unemployed who start a full-time job (2), the natural logarithm of native inactive who start a full-time job (3), the natural logarithm of native employed who start a full-time job (4). ΔM_{EU13} is the change in EU-13 migrant stock between 2014 and 2010 relative to the 2010 population in percentage points. Controls include the total migrant share in 1993 as well as indicators for rural region and eastern Germany. Regressions are weighted by 2010 native population. $\bar{Y}$ is the mean of the dependent variable. ϵ is the elasticity at mean values, calculated as $\epsilon \approx \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \cdot 100$ for logarithmic outcomes and $\epsilon = \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \frac{1}{\bar{Y}_{2010}} \cdot 100$ for outcomes in levels. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016.

In Table 1..24 and Table 1..25 in the Appendix, I assess the robustness of the results on wages to specification choices and find that they remain virtually unchanged across all

[25] Since the data does not include working hours, I restrict the analysis to full-time jobs. This is the standard approach for papers working with German administrative wage data.

regressions. The only exception is using no controls for all starting wages (column 1), in which case the estimate preserves the positive sign but looses its statistical significance. However, again, the unconditional regression should not be considered the benchmark specification. The more credible robustness checks do not alter the conclusions from the main text. In Table 1..26 I find that bounds on all starting wages, except from inactivity, remain above zero for violations of the exclusion restriction of around 40% of the reduced form effect.

This section documented the effects of EU-13 immigration on tightness, unemployment, vacancies, transition probabilities, and starting wages of natives across commuting zones. Unlike the theoretical predictions guiding the analysis, I did not find evidence for negative spillovers on native unemployment or job-finding rates. I did also not find evidence that native wages adjusted. The remainder of this paper aims to explain these perhaps surprising results. In the next section, I will assess other possible mechanisms that could rationalize the positive effects of EU-13 immigration on native unemployment.

1.6.2 Mechanisms

The previous findings are hard to rationalize when immigration is viewed purely as an increase in the number of job-seekers. The improvement of job prospects for unemployed natives would suggest that other adjustments occurred, which more than compensated for the potentially negative competition effects. The aim of this section is to understand what these adjustments might have been.

The internal migration response of natives is a well-studied mechanism through which receiving economies adjust to immigration (Borjas, 2006; Monras, 2020). In column (1) of Table 1.7, I test whether it was also important in my case. The regression is as before but uses the logarithmic growth of the native working age population as outcome variable.[26] I find that immigration is linked to positive population growth with large coefficients relative to the average growth in the observation period. This finding is at odds with an internal migration response to immigration. In columns (2) and (3), I decompose the

[26] Note that, due to practical constraints, the outcomes for this regression are based on a 10% of the IEB. The results will be updated in future drafts.

native population change into inflows and outflows[27] to find that a reduction in outflows was the main driver. Interestingly, this is opposite to the decomposition in Dustmann et al., 2023, who find that a reduction in inflows was the main driver. Overall, the results suggest that native internal migration was unlikely to be the main adjustment mechanism.

Table 1.7: Immigration and native population response

	(1) Δ ln Native population	(2) ln Native inflows	(3) ln Native outflows
Panel A. OLS			
ΔM_{EU13}	0.013	-0.154	-0.196
	(0.003)	(0.081)	(0.076)
ϵ	0.022	-0.264	-0.336
Panel B. 2SLS, normal IV			
ΔM_{EU13}	0.013	-0.239	-0.289
	(0.005)	(0.143)	(0.131)
ϵ	0.023	-0.409	-0.494
Panel C. 2SLS, diversion IV			
ΔM_{EU13}	0.018	-0.376	-0.442
	(0.004)	(0.201)	(0.193)
ϵ	0.031	-0.643	-0.756
$\bar{Y}_{2010}$	3.554		
$\Delta\bar{Y}$	0.015	9.296	9.257
N	223	223	223

Cluster-robust standard errors in parentheses. Outcomes are the difference in the natural logarithm of working-age native population (1), the natural logarithm of native working-age individuals observed in a region in 2014 but not in 2010 (2), and the natural logarithm of native working-age individuals observed in a region in 2010 but not in 2014 (3). All outcomes are from IEB. ΔM_{EU13} is the change in EU-13 migrant stock between 2014 and 2010 relative to the 2010 population in percentage points. Controls include the total migrant share in 1993 as well as indicators for rural region and eastern Germany. Regressions are weighted by 2010 native population. $\bar{Y}$ is the mean of the dependent variable. ϵ is the elasticity at mean values, calculated as $\epsilon \approx \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \cdot 100$ for logarithmic outcomes and $\epsilon = \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \frac{1}{\bar{Y}_{2010}} \cdot 100$ for outcomes in levels. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016.

Another explanation relates to demand effects. If the presence of EU-13 workers spurred consumption (Berbée et al., 2022) or led to productivity gains (Peri, 2016), output growth could shift labor demand to more than outweigh the supply shift. I investigate this hypothesis in Table 1.8. The regressions are as in the previous chapter, but use outcomes that relate to output and productivity. In column (1), I assess the effect of immigration on

[27] Inflows are calculated using individuals who were in a region in 2014 but not in 2010. Outflows are calculated by counting individuals who were in a region in 2010 but not anymore in 2014. I include this outcome in levels in the regression since the quantities already represent a difference.

the change of real GDP per capita. I find a positive, statistically significant relationship, with coefficients implying 3.4% and 4.5% real output growth for a percentage point increase in immigration. In columns (2) and (3), I measure local consumption and profits using deflated tax revenues, a necessary proxy since these metrics are not available at this level of geographic disaggregation. Both turn out positive across all specifications. Interestingly, the elasticity is larger in the case of profit/tax revenues than for sales-tax revenues, an indication that either sales growth is driven by export demand or that productivity increased. While both explanations are plausible, only the latter can be tested with my research design. I do so in column (4), where I assess the effect of EU-13 immigration on real value added per employee and find a strong positive relationship between productivity and immigration. This result is in line with a large body of research from other countries documenting the positive effects of immigration on productivity (Mitaritonna et al., 2017; Ottaviano et al., 2018; Paserman, 2013; Peri, 2012), but appears novel for the case of Germany.

Finally, I assess the role of firm entry as proxied by the creation of new establishments.[28] More specifically, the outcome in column (5) is the change in the net establishment creation, that is the difference between plant closures and plant openings in a year.[29] I find a strong positive response across all specifications, with magnitudes of around 100 new establishments per percentage point increase in immigration. This stands in marked contrast to the average slowdown in regional plant growth over the period of observation. In Appendix Table 1..27 columns (1) to (3), I decompose net establishment growth and find it is driven primarily by more plant entries rather than fewer closures. In Columns (4) to (6), I further show that employment growth per plant also increased with immigration, which, again is driven primarily by more job creation and not by less job destruction. Similar to Dustmann and Glitz, 2015, the results show that the adjustment to immigration occurred both through firm entry and through within-firm expansion. The magnitudes suggest, however, that in my case new firm entry played a more important

[28] The German administrative data does not allow distinguishing different plants pertaining to the same company. It is common to conceptually treat plants as firms, see for example Card et al., 2013.

[29] As is common practice, I ignore plant closures and entries that likely stem from mergers or identifier changes (Hethey & Schmieder, 2010). Note that I cannot use logs here, since the variable includes values below zero.

Table 1.8: Immigration and output

	(1) Δ ln GDP	(2) Δ ln Sales − tax revenue	(3) Δ ln Profit − tax revenue	(4) Δ ln Value added/worker	(5) Δ Net firm creation
Panel A. OLS					
ΔM_{EU13}	0.026	0.015	0.086	0.010	60.646
	(0.006)	(0.004)	(0.021)	(0.004)	(17.992)
ϵ	0.044	0.026	0.148	0.016	103.849
Panel B. 2SLS, normal IV					
ΔM_{EU13}	0.031	0.015	0.121	0.015	93.937
	(0.009)	(0.007)	(0.026)	(0.006)	(27.152)
ϵ	0.053	0.026	0.208	0.026	160.855
Panel C. 2SLS, diversion IV					
ΔM_{EU13}	0.045	0.021	0.146	0.025	107.833
	(0.009)	(0.006)	(0.030)	(0.006)	(31.939)
ϵ	0.077	0.036	0.249	0.043	184.650
$\bar{Y}$	23.815	17.242	19.490	10.982	250.201
$\Delta\bar{Y}$	0.066	0.056	0.128	0.029	-65.873
N	223	223	223	223	223

Cluster-robust standard errors in parentheses. Outcomes are the natural logarithm of the gross domestic product deflated to 2015 prices (1), the natural logarithm of sales tax revenue deflated to 2015 prices (2), the natural logarithm of profit tax revenue deflated to 2015 prices (3) the natural logarithm value added per employee in 2015 prices (4), the natural logarithm of the net number of new plants, that is the number of plant entries minus the number of plant closures in a year (5). Outcomes (1) to (5) are from INKAR, outcome (5) is from BHP. ΔM_{EU13} is the change in EU-13 migrant stock between 2014 and 2010 relative to the 2010 population in percentage points. Controls include the total migrant share in 1993 as well as indicators for rural region and eastern Germany. Regressions are weighted by 2010 native population. $\bar{Y}$ is the mean of the dependent variable. ϵ is the elasticity at mean values, calculated as $\epsilon \approx \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \cdot 100$ for logarithmic outcomes and $\epsilon = \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \frac{1}{\bar{Y}_{2010}} \cdot 100$ for outcomes in levels. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016.

role.

Increases in output, additional firm entry, and productivity growth help rationalize the positive effects of immigration on the job prospects of unemployed natives. But then, why did regions with more EU-13 immigration experience this expansion? Viewed through the lens of search-and-matching models, an intuitive explanation is that immigration increases in matching efficiency. Urban economists have stressed that agglomeration economies can reduce hiring costs and improve match quality (Dauth et al., 2022; Moretti & Yi, 2024). Immigration mechanically increases market size. It may also strengthen positive assortative matching (PAM) if migrants exhibit larger variance in unobserved ability (Orefice & Peri, 2020).

Table 1.9: Results for matching efficiency

	(1) Δ Search duration	(2) Δ Difficult searches	(3) Δ Failed searches	(4) Δ Share qualified applicants	(5) Δ Positive assortative matching
Panel A. OLS					
ΔM_{EU13}	-0.597	-0.006	-0.058	3.070	-0.002
	(0.481)	(0.182)	(0.040)	(2.178)	(0.002)
ϵ	-1.023	-0.010	-0.100	5.256	-0.003
Panel B. 2SLS, normal IV					
ΔM_{EU13}	-0.089	-0.275	0.070	5.458	0.001
	(1.063)	(0.239)	(0.087)	(3.503)	(0.004)
ϵ	-0.153	-0.470	0.120	9.346	0.001
Panel C. 2SLS, diversion IV					
ΔM_{EU13}	-0.084	-0.250	0.156	3.963	-0.002
	(1.322)	(0.184)	(0.125)	(4.591)	(0.004)
ϵ	-0.144	-0.429	0.267	6.786	-0.003
$\bar{Y}_{2010}$	8.940	1.329	0.768	50.536	0.267
$\Delta\bar{Y}$	0.232	-0.165	0.257	2.030	-0.021
N	223	219	223	223	223

Cluster-robust standard errors in parentheses. Outcomes are the weeks the last successful job search lasted (1), the number of job searches with hiring difficulties in the previous twelve months (2), the number of failed job searches in the previous twelve months (3) the share of qualified among total applicants for the last job search (4), the correlation between the worker and firm fixed effect (5). Outcomes (1) to (4) are from the IAB vacancy survey, outcome (5) is from Bellmann et al., 2020. All outcomes are expressed as differences. ΔM_{EU13} is the change in EU-13 migrant stock between 2014 and 2010 relative to the 2010 population in percentage points. Controls include the total migrant share in 1993 as well as indicators for rural region and eastern Germany. Regressions are weighted by 2010 native population. $\bar{Y}$ is the mean of the dependent variable. ϵ is the elasticity at mean values, calculated as $\epsilon \approx \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \cdot 100$ for logarithmic outcomes and $\epsilon = \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \frac{1}{\bar{Y}_{2010}} \cdot 100$ for outcomes in levels. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016.

I investigate this channel in Table 1.9, first using information about firms' job search activity from the IAB vacancy survey. The first three outcomes are meant to proxy for hiring costs: they include the number of weeks the last successful job search took (1), the number of job searches for which the firm experienced difficulties finding an appropriate candidate in the previous twelve months (2) and, similarly, the number of failed job searches in the previous twelve months (3). The results yield mostly negative coefficients, which would indicate that hiring costs decreased, yet none of them are statistically significant at conventional levels; I must refrain from drawing further conclusions. In the last two columns, I use proxies for the quality of matches between firms and workers. In column (4), still using the IAB vacancy survey, I find a positive relationship between immigration in a region and the share of qualified applicants for the firm's last job search. This would imply that the quality of the average applicant increased. The findings are, however, again not statistically significant at usual thresholds. In column (5), I assess whether immigration led to more positive assortative matching using worker and firm fixed effects derived from Abowd et al., 1999 AKM regressions.[30] Following Orefice and Peri, 2020, I use the Spearman correlation between worker and firm fixed effects in a commuting zone setting to zero all correlations with p-values above 0.1. I find insignificant effects with magnitudes very close to zero throughout. Taken together, the results from Table 1.9 provide little evidence that matching efficiency increased due to immigration.

Another possible explanation to rationalize the lack of adverse effects on native workers is imperfect substitutability between migrants and natives (Manacorda et al., 2012; Ottaviano & Peri, 2012). If migrant workers are imperfect substitutes for natives, complementarity effects will raise the marginal product of natives. Testing for imperfect substitutability is challenging, and it is beyond the scope of this paper to deliver conclusive evidence in this regard. Yet, as a first indication, it is interesting to compare the effects of EU-13 immigration on native workers to the effects it had on migrants. To that end, Table 1.10 estimates the effects of immigration on unemployment, transitions, and starting wages, but this time using prime-age migrants only. The results are somewhat less positive than for natives. In column (1), I find a negative but insignificant relationship between EU-13 immigration and the stock of migrant unemployed. This is despite finding a strong positive effect on the job-finding rate from unemployment in column (2) and a null effect on job-separation in column (3). While the change in the stock need not match the changes in transition rates since they are measured at different frequencies[31], the discrepancy might indicate that other movements, such as return migration, may be important here. Finally, I find small and insignificant coefficients on the change of migrant starting wages, both when considering any transition in column (4) and when focusing on the transitions from unemployment to employment (5). Taken together, the results show

[30] The current version of the paper relies on pre-calculated fixed effects from Bellmann et al., 2020 for the period 2007 to 2013 and aggregates them regionally using the 2% sample of the administrative data. This will be updated in due course.

[31] Remember that transitions are measured as yearly averages of monthly transitions while the stock is the difference between 2010 and 2014. This discrepancy is necessary; transition rates cannot be measured at a four-year lag since virtually all job-seekers will have left unemployment by then.

modest differences between the effects of immigration on natives and migrants. This implies that some degree of imperfect substitutability is likely present in the German labor market. Yet, it is unlikely to be overtly large since, in that case, one would have expected to find more negative effects of immigration on migrant workers.

Table 1.10: Effects on migrants

	(1) $\Delta\, ln\, U_{mig}$	(2) $\Delta\, UE_{mig}/U_{mig}$	(3) $\Delta\, EU_{mig}/E_{mig}$	(4) $\Delta\, ln\, Wage_{mig}$	(5) $\Delta\, ln\, Wage_{mig}(UE)$
Panel A. OLS					
ΔM_{EU13}	0.033	0.336	-0.005	-0.022	-0.010
	(0.010)	(0.075)	(0.009)	(0.008)	(0.005)
ϵ	0.056	0.575	-0.009	-0.038	-0.017
Panel B. 2SLS, normal IV					
ΔM_{EU13}	-0.029	0.597	-0.022	-0.008	0.001
	(0.029)	(0.109)	(0.019)	(0.008)	(0.007)
ϵ	-0.049	1.023	-0.038	-0.014	0.001
Panel C. 2SLS, diversion IV					
ΔM_{EU13}	-0.053	0.634	-0.003	-0.008	-0.006
	(0.041)	(0.077)	(0.011)	(0.007)	(0.009)
ϵ	-0.091	1.086	-0.006	-0.014	-0.010
$\bar{Y}$	8.205	4.665	1.075	4.181	4.030
$\Delta\bar{Y}$	0.048	0.373	-0.094	0.082	0.111
N	223	223	223	223	223

Cluster-robust standard errors in parentheses. Outcomes are the natural logarithm of migrant unemployment (1), the share of unemployed migrants who find a job the following month (2), the share of employed migrants who are unemployed the following month (3) starting wage of all native job-switchers who start a full-time job (4), and the natural logarithm of the starting wage of all native unemployed who start a full-time job (5). All outcomes based on IEB, workers aged 25 to 55. ΔM_{EU13} is the change in EU-13 migrant stock between 2014 and 2010 relative to the 2010 population in percentage points. Controls include the total migrant share in 1993 as well as indicators for rural region and eastern Germany. Regressions are weighted by 2010 native population. $\bar{Y}$ is the mean of the dependent variable. ϵ is the elasticity at mean values, calculated as $\epsilon \approx \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \cdot 100$ for logarithmic outcomes and $\epsilon = \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \frac{1}{\bar{Y}_{2010}} \cdot 100$ for outcomes in levels. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016.

This section explored different mechanisms that could explain the lack of negative effects immigration from EU-13 countries had on native unemployed. I did not find a strong internal migration response or large increases in matching efficiency. Instead, I found some indication that migrants and natives are imperfect substitutes and strong evidence that output, productivity, and firm creation were boosted by the arrival of new migrants. The results emphasize the importance of including productivity responses in search-and-matching models of immigration.

1.7 Conclusion

This paper assesses the impact of immigration to Germany following the EU Eastern Enlargement era. Using two variants of a shift-share instrument, I compare the evolution of unemployment, vacancies, transition rates, and starting wages across commuting zones before and after the enlargement. These outcomes are less studied in the literature on the labor market effects of immigration, but are important policy metrics and map closely to theoretical counterparts in the search-and-matching literature.

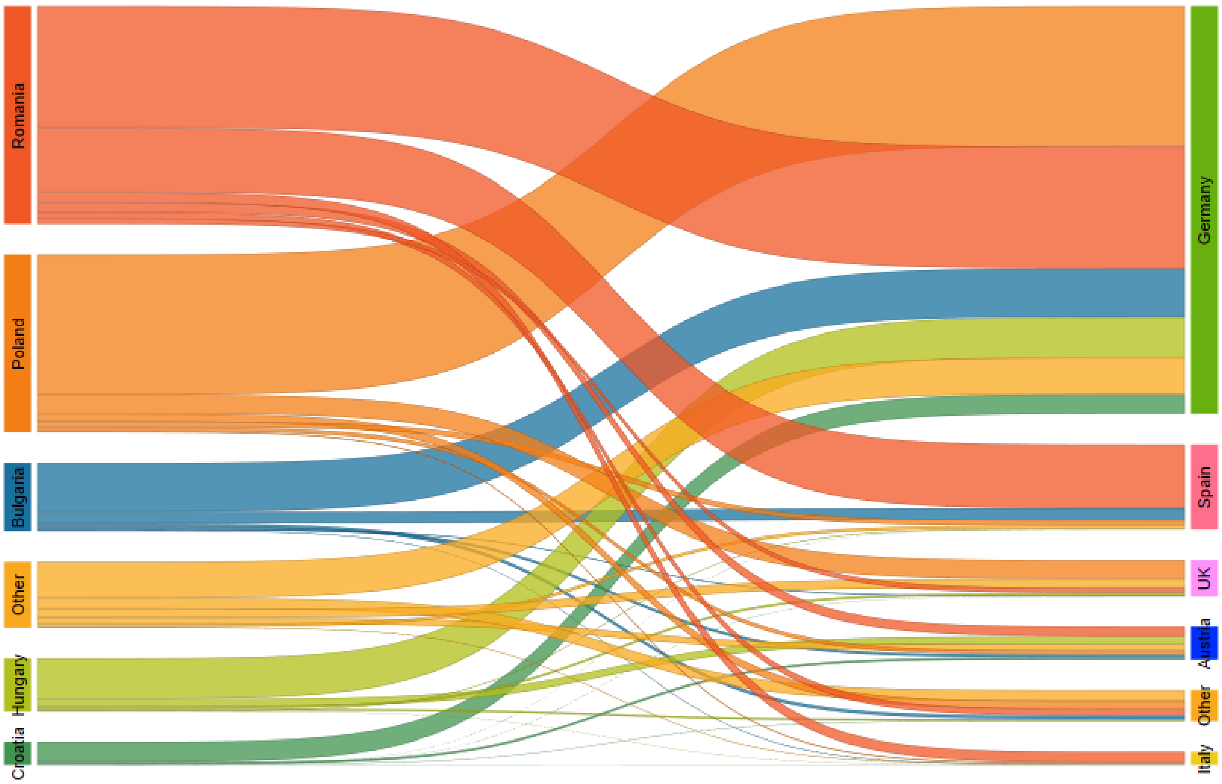
The results indicated that increased immigration led to a substantial reduction in native unemployment through a drop in the job-separation rate and an increase in the job-finding rate, as well as positive effects on starting wages while having no discernible impact on vacancies. Most results are robust to important specification choices such as covariate selection or geographical unit. These findings have several implications for the design of search-and-matching models as applied to studying the labor market effects of immigration. The negative predictions of the simple search-and-matching model with job rationing and perfect substitution between natives and migrants of could not be confirmed. The evidence on mechanisms suggests that productivity responses to immigration and imperfect substitution between immigrants and natives are interesting features that could be added to the model.

A potential avenue for future research is to use the reduced-form elasticities from this study to calibrate spatial equilibrium models with unemployment as in Kline and Moretti, 2013, Bilal, 2023, or Kuhn et al., 2021. This would be in the spirit of Monras, 2020, who uses Rosen, 1979 and Roback, 1982 type of spatial equilibrium models to extrapolate the local effects of immigration to the national level and to simulate general equilibrium effects of migration policies. Extending his framework to include a frictional labor market with unemployment could be an interesting avenue for future work. The results from the present study could be of use for simulations of the German labor market.

Another interesting area of research is the effect of migrants' access to social welfare on the labor market outcomes of natives. The outside option is a key metric under imperfect competition, and many migrants face worse outside options in receiving countries. Using monopsony, Borjas and Edo, 2023 show that improving the outside option of migrants will raise their reservation wages and reduce negative impacts on native wages. Similar arguments can be made in search-and-matching models, where the outside option enters the decision to accept a job. In both cases, improving the outside option of migrants is likely to have positive spillover effects on incumbents.

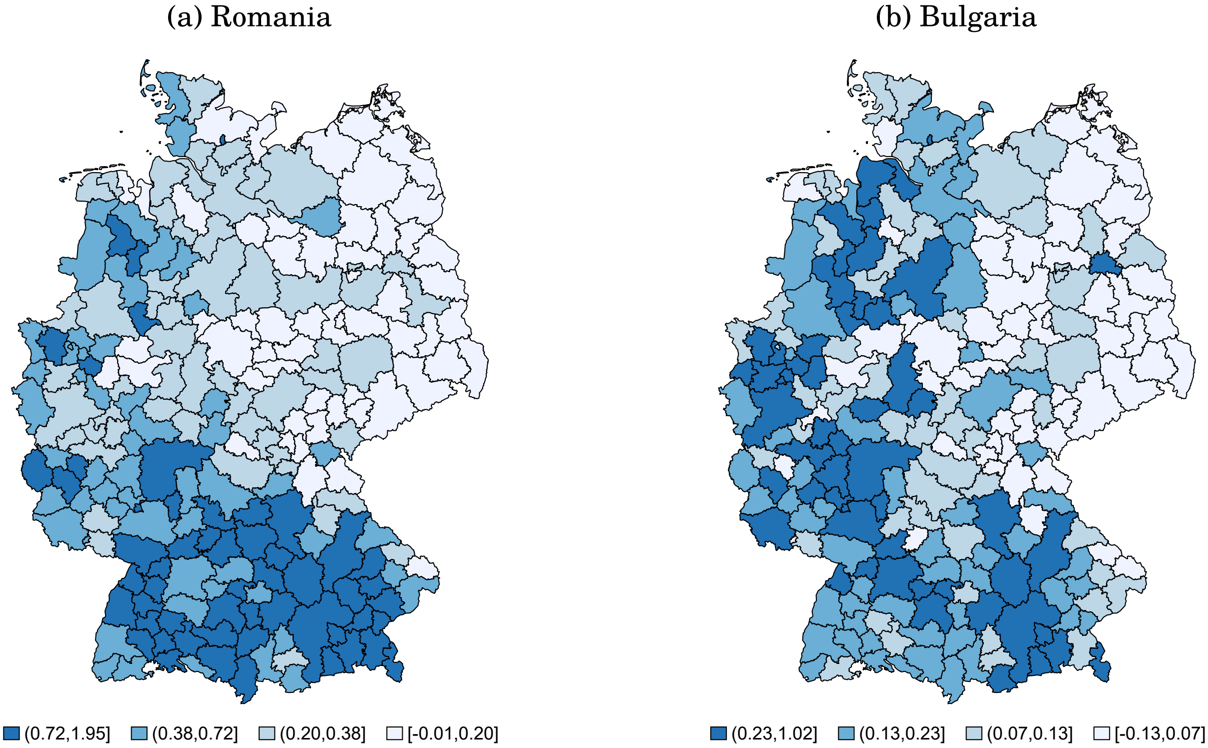
Appendix

Figure 1..4: EU-13 Migration Flows, 2011-2016



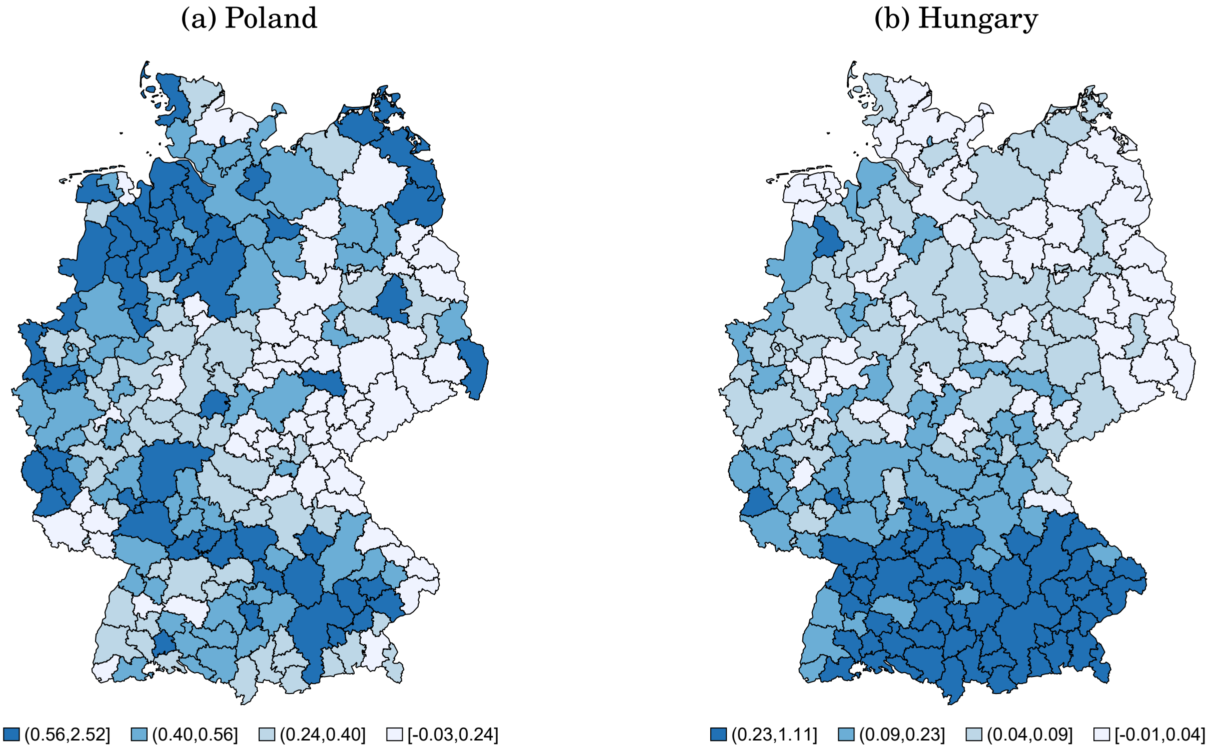
Source: Standaert and Rayp, 2022, own calculations. Total 2011-2016 outflows by origin and destination.

Figure 1..5: Change in Romanian and Bulgarian migrant stock, % of 2010 population



Source: Destatis GENESIS-Online Table 12521-0040, own calculations. Figures plot the change in migrant stocks between 2010 and 2016 as a percentage of 2010 population, on the left using EU-13 migration, on the right using total migration. Units are 223 commuting zones.

Figure 1..6: Change in Polish and Hungarian migrant stock, % of 2010 population

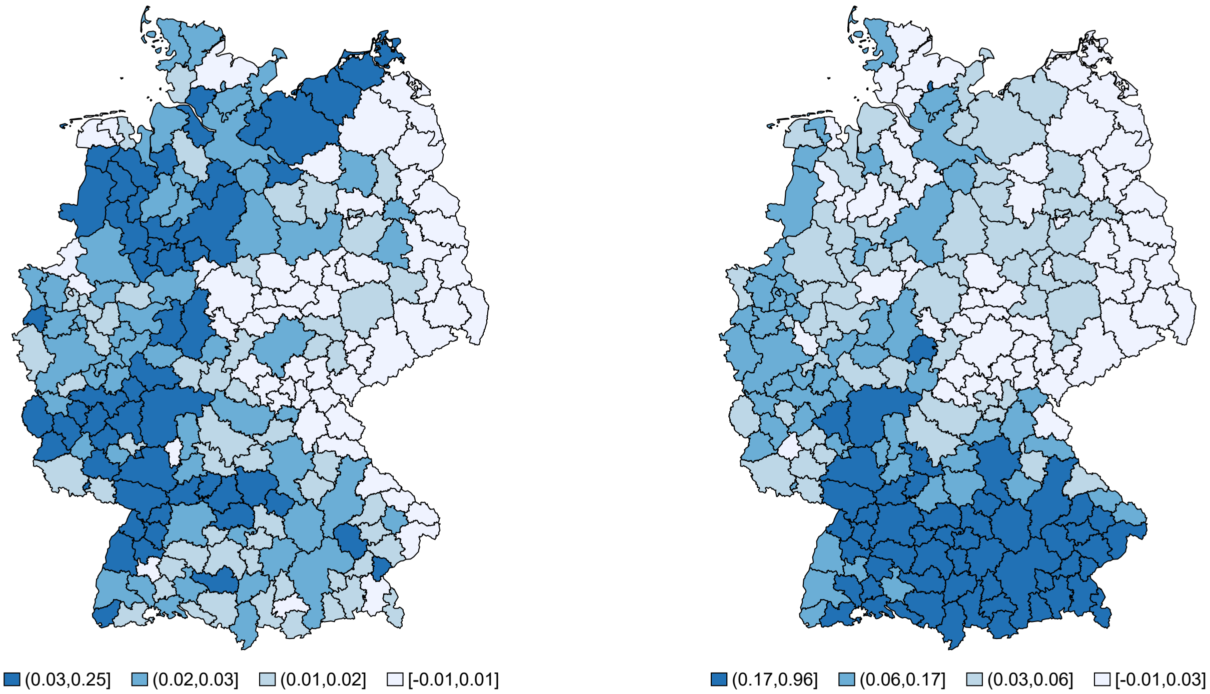


Source: Destatis GENESIS-Online Table 12521-0040, own calculations. Figures plot the change in migrant stocks between 2010 and 2016 as a percentage of 2010 population, on the left using EU-13 migration, on the right using total migration. Units are 223 commuting zones.

Figure 1..7: Change in Lithuanian and Croatian migrant stock, % of 2010 population

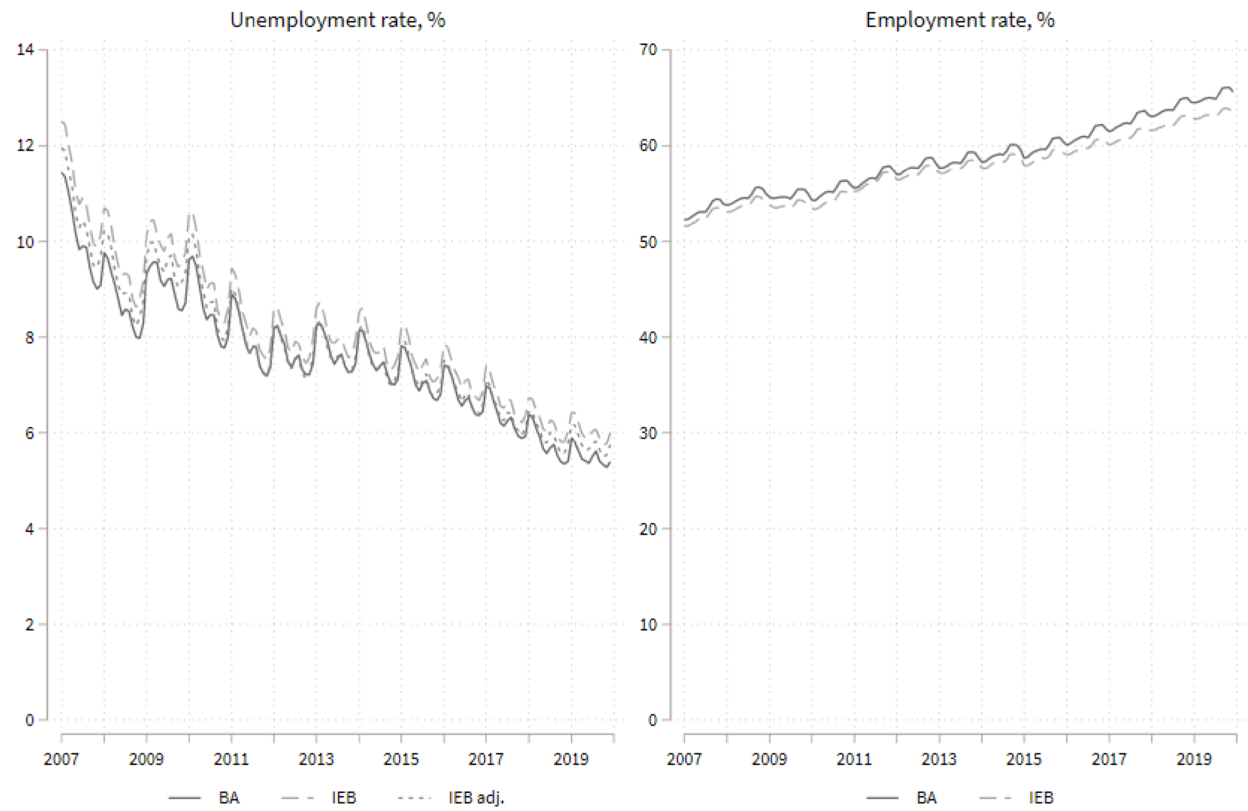
(a) Lithuania

(b) Croatia



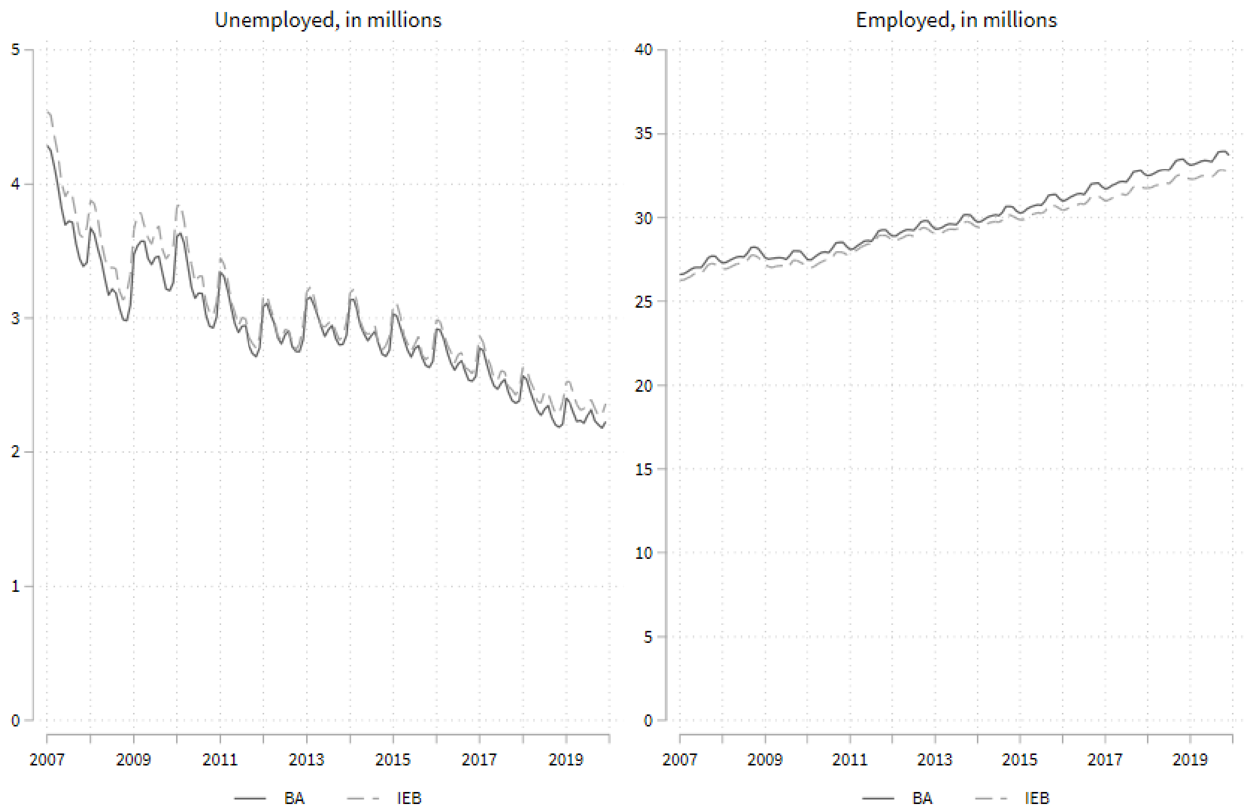
Source: Destatis GENESIS-Online Table 12521-0040, own calculations. Figures plot the change in migrant stocks between 2010 and 2016 as a percentage of 2010 population, on the left using EU-13 migration, on the right using total migration. Units are 223 commuting zones.

Figure 1..8: (Un-)employment rate over time



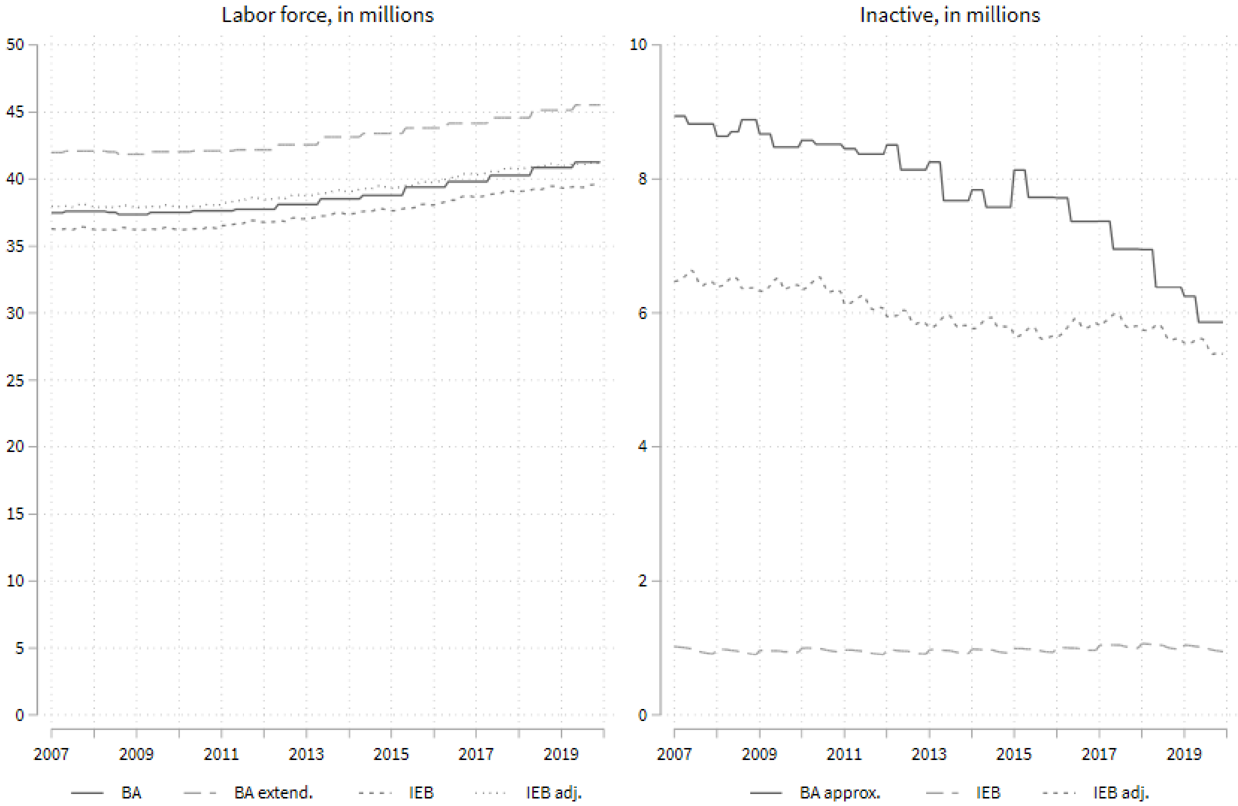
Source: IEB, BA, INKAR, own calculations. Left panel plots the monthly unemployment rate, which is the number of unemployed over the labor force in percentage points. In the legend, "BA" indicates the official unemployment rate from the Federal Employment Agency using the concept "zivile abhängige Erwerbspersonen" as a measure of the labor force. "IEB" provides the measure of the unemployment rate derived from the Integrated Employment Biographies, while "IEB adj." tries to correct the labor force for civil servants. The right panel plots the employment rate, which is the number of employed over working-age (18-65) population in percentage points. Working age opulation is from INKAR.

Figure 1..9: Employment and unemployment over time



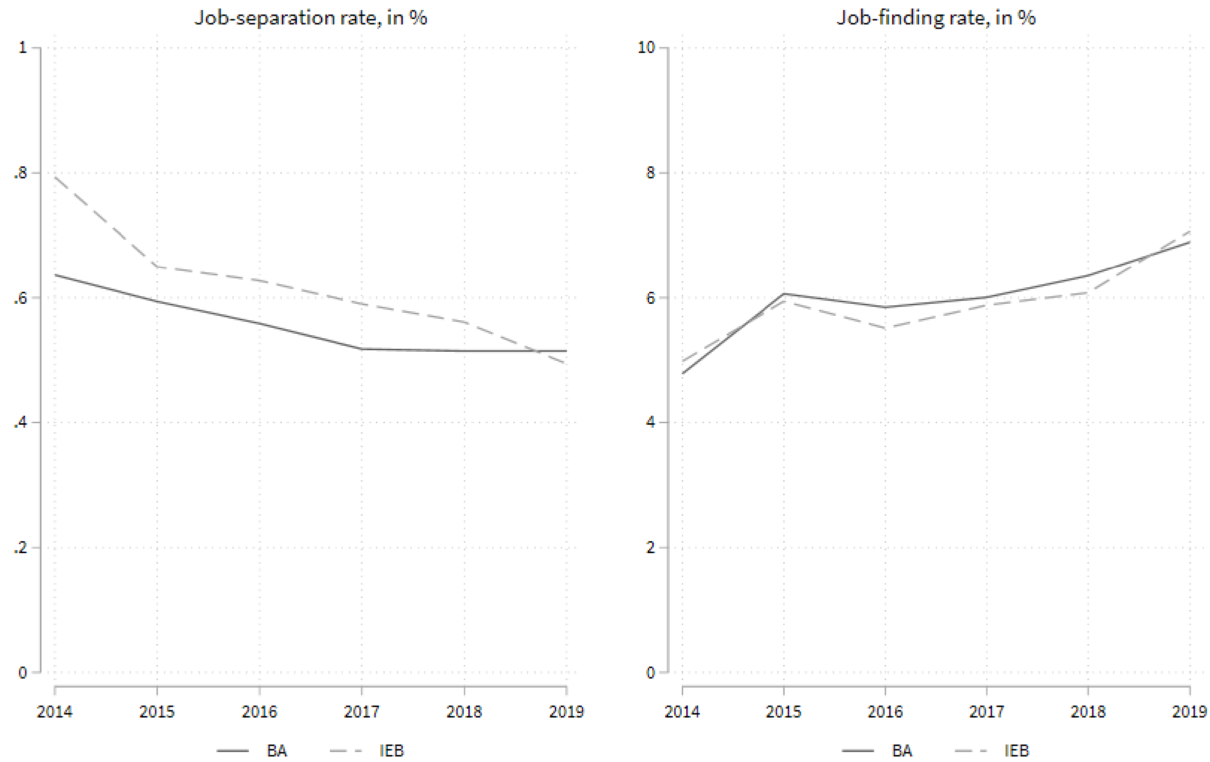
Source: IEB, BA, INKAR, own calculations. Left panel plots the monthly unemployment stock, the right panel plots the monthly stock of employed subject to social security. In the legend, "BA" indicates the official number of unemployed. "IEB" provides the measure of the unemployment rate derived from the Integrated Employment Biographies, while "IEB adj." tries to correct the labor force for civil servants.

Figure 1..10: Labor force and inactivity over time



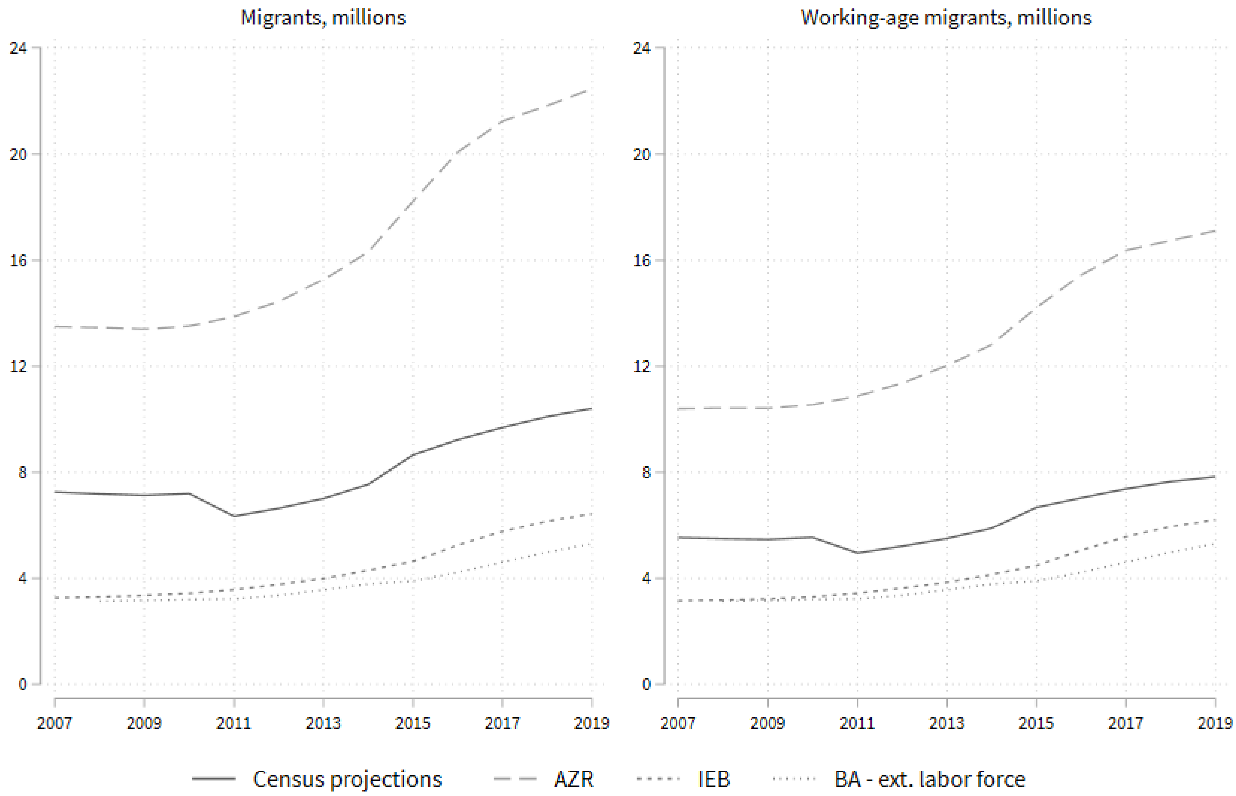
Source: IEB, INKAR, BA, own calculations. Panel on the left plots the monthly labor force, in millions. The full line is based on the "abhängige zivile Erwerbspersonen" definition of the Federal Employment Agency, the long-dotted line uses the extended definition "zivile Erwerbspersonen" which also includes self-employed. "IEB" is the estimate of the labor based on the social security records, "IEB adj." tries to correct for civil servants. The panel on the left provides estimates of the monthly inactive, in millions. "BA" approx. estimates the inactive as the difference between the 18-65 year old population and the official extended labor force number. "IEB" is a conservative estimate of the number of inactive from social security records, "IEB adj." also adds borderline inactive cases (e.g. individuals in training measures).

Figure 1..11: Transition rates over time



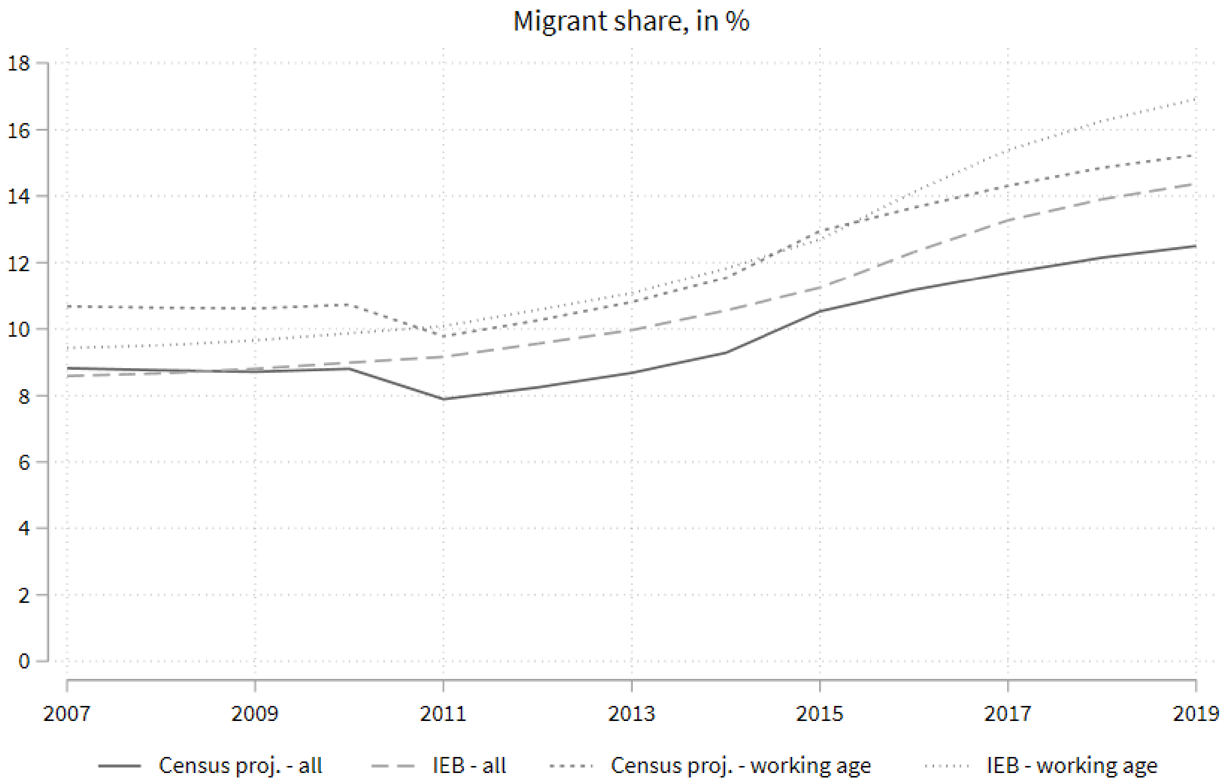
Source: IEB, BA, own calculations. Yearly averages. The job-finding rate is the number of unemployed individuals who find a job in the coming month over the total number of unemployed individuals in a month. Fraction is multiplied by 100 to give percentage points.

Figure 1..12: Migrant stocks over time



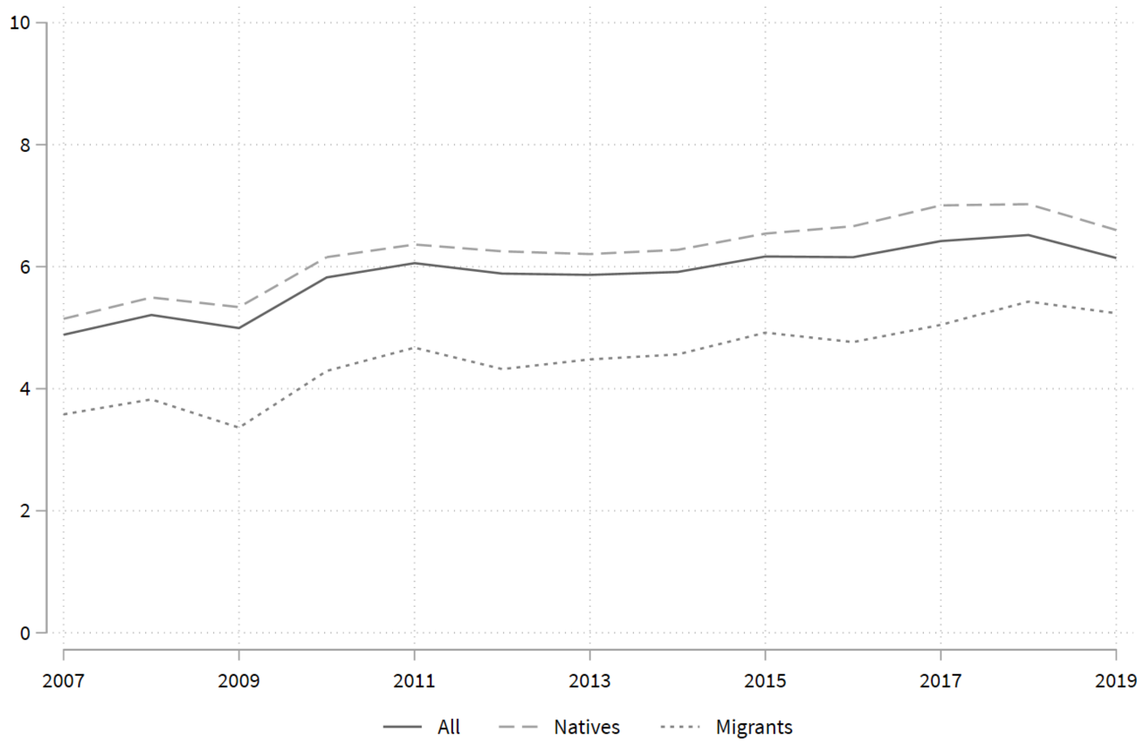
Source: IEB, Destatis, BA own calculations. Migrant share in percentage points of the population. Census projections are retrieved from Genesis Destatis Table 12411-0007. Central foreign register (AZR) data is retrieved from Genesis Destatis Table 12521-0003. IEB is from the Integrated Employment Biographies counting all individuals on the 30^{th} June of each year. BA is the extended labor force from a custom extract by the Federal Employment Agency. Working age is 18-64.

Figure 1..13: Migrant shares over time



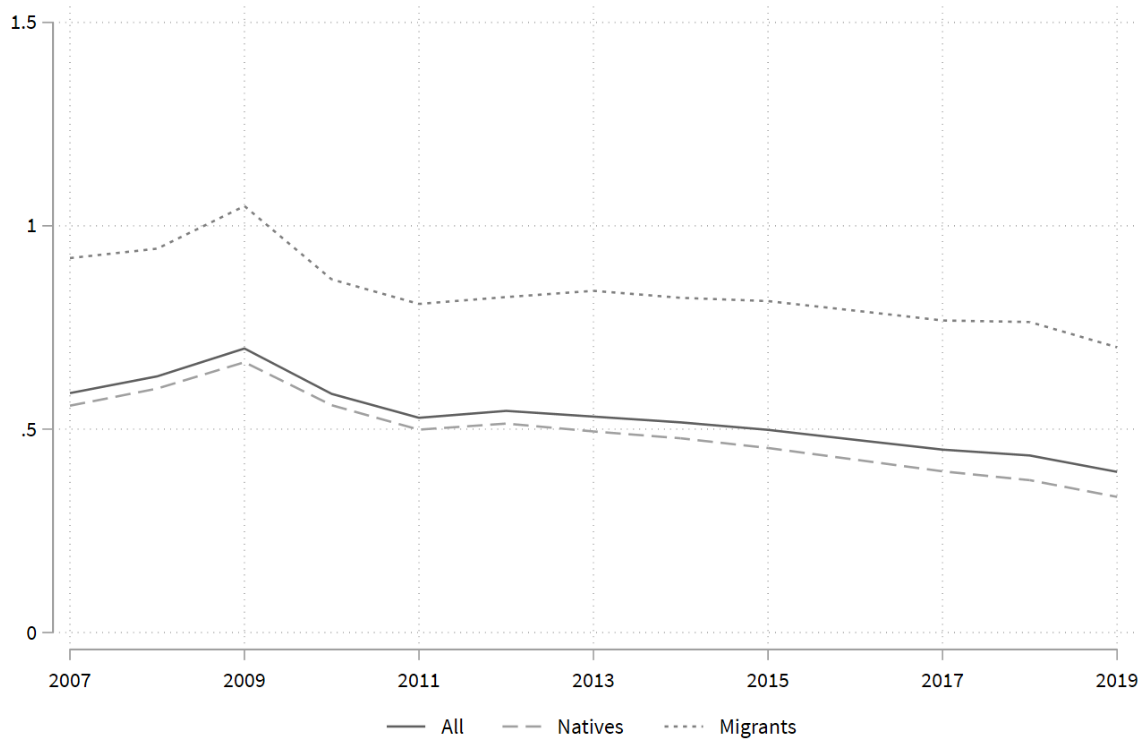
Source: IEB, Destatis, own calculations. Migrant share in percentage points of the population. Census projections are from Genesis Destatis Table 12411-0007. IEB is from the Integrated Employment Biographies counting all individuals on the 30th June of each year. Working age is 18-64.

Figure 1..14: Job-finding rates in %



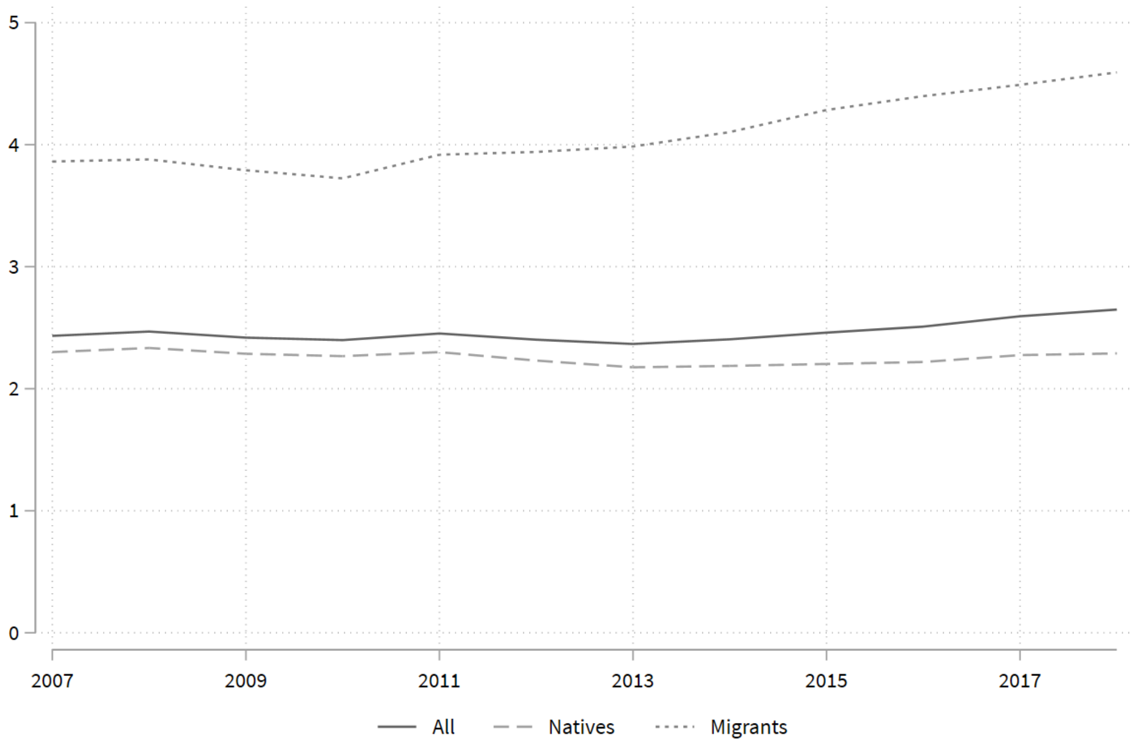
Source: Integrated Employment Biographies (IEB), own calculations. Yearly averages. The job-finding rate is the number of unemployed individuals who find a job in the coming month over the total number of unemployed individuals in a month. Fraction multiplied by 100 to give percentage points.

Figure 1..15: Job-separation rates in %



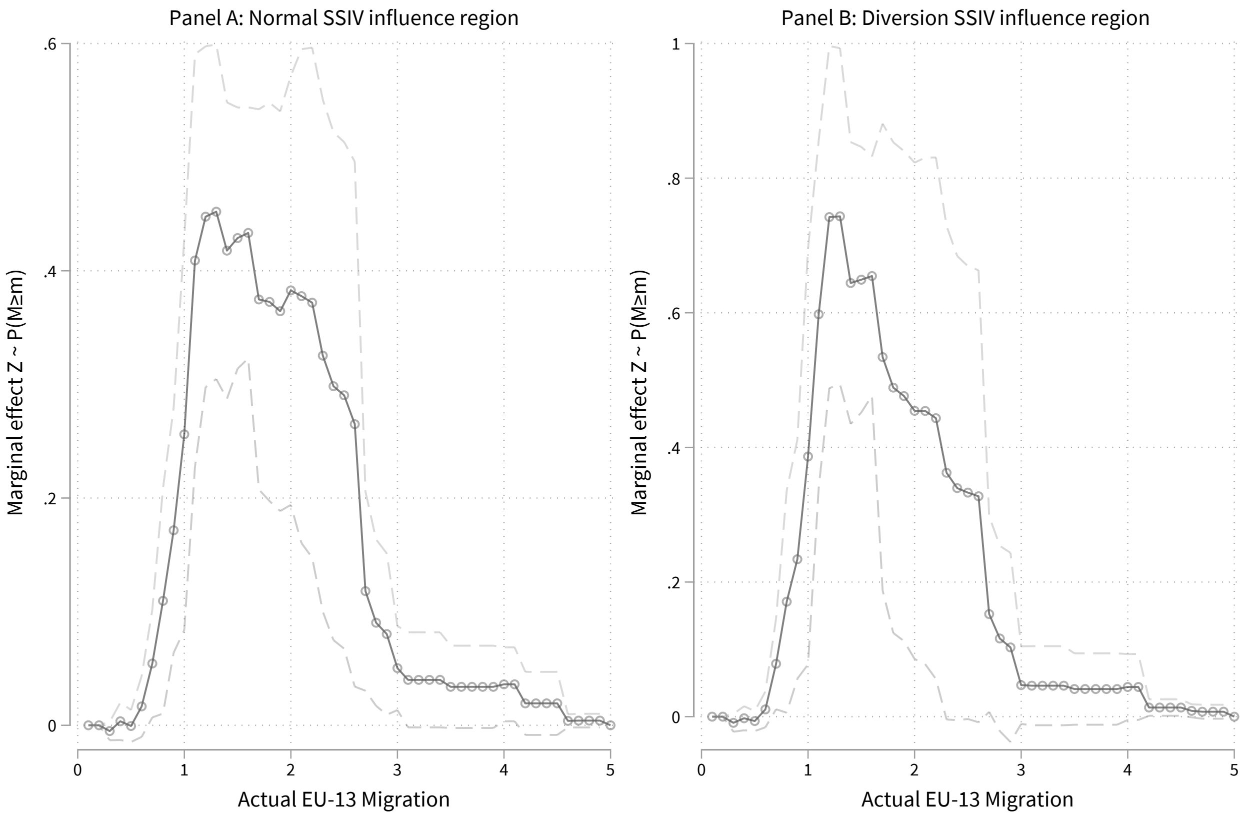
Source: Integrated Employment Biographies (IEB), own calculations. Yearly averages. The job-separation rate is the number of employed individuals who enter unemployment the coming month over the total number of employed individuals in a month. Fraction multiplied by 100 to give percentage points.

Figure 1..16: Job-switching rates in %



Source: Integrated Employment Biographies (IEB), own calculations. Yearly averages. The job-switching rate is the number of employed individuals who switch establishments the coming month over the total number of employed individuals in a month. Fraction multiplied by 100 to give percentage points.

Figure 1..17: Weighted regression weights



Source: Integrated Employment Biographies (IEB), own calculations. Figures plot regression coefficients and 95% confidence intervals from regressions of the instrument on indicators $\mathbb{I}[\Delta M \geq m]$ where m increases in steps of 0.1. The marginal effect of Z on $\mathbb{I}[\Delta M \geq m]$ is indicative on where the instrument marginally shifts the treatment variable. The left uses the standard shift-share instrument from eq. 1.2 and on the right uses the diversion instrument from eq. 1.3. Units are 223 commuting zones. Observations weighted by 2010 native population size. Standard errors clustered at the level of 50 broader commuting zones. The exercise is inspired by Rose and Shem-Tov, 2021 Figure 2 C, though the continuous instrument requires a different interpretation.

Table 1..11: End of transition phases for EU-13 workers in Germany

	EU-Accession	End of transition phase
EU-8	01.05.2004	01.05.2011
EU-2	01.01.2007	01.01.2014
Croatia	01.07.2013	01.07.2015

Source: Gallegos Torres et al., 2022. End of transition phase is the end date of the transition phase in Germany. EU-8 is Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, Slovenia, Cyprus, and Malta. EU-2 is Bulgaria and Romania.

Table 1..12: Origin groups

1. Estonia, Lithuania, Latvia
2. Slovakia
3. Czech Republic
4. Slovenia
5. Hungary
6. Poland
7. Bulgaria
8. Romania
9. Croatia
10. Malta, Cyprus

Source: Own calculations. The table gives the ten origin groups used to construct the shift-share instrument. Origins in one row are a group

Table 1..13: Top 20 origins of working-age migration changes by decade

1990 - 1999		2000 - 2009		2010 - 2019	
Origins	Share	Origins	Share	Origins	Share
Turkey	26.9	Turkey	24.6	Syria	13.0
Poland	6.7	Russia	10.7	Romania	12.3
Russia	4.9	Poland	10.5	Poland	12.2
Bosnia	4.8	Ukraine	5.9	Bulgaria	5.4
Italy	4.2	Iraq	4.2	Afghanistan	3.6
Croatia	4.1	Romania	3.9	Croatia	3.6
France	3.2	Kazakhstan	3.4	Turkey	3.4
Portugal	2.5	Serbia	2.9	Hungary	3.1
Iran	2.5	Morocco	2.9	Iraq	3.0
Ukraine	2.4	Afghanistan	2.7	Italy	2.8
Albania	2.3	Albania	2.6	Greece	2.3
Greece	2.3	Lebanon	2.5	Kosovo	2.1
Vietnam	2.3	China	2.1	Serbia	1.8
Romania	2.2	Bosnia	2.1	India	1.6
Morocco	1.8	Vietnam	2.1	Iran	1.5
Sri Lanka	1.5	Iran	1.9	Bosnia	1.5
Kazakhstan	1.5	Bulgaria	1.8	Czech Rep.	1.5
Afghanistan	1.5	Thailand	1.8	Eritrea	1.4
Lebanon	1.3	Serbia	1.6	Russia	1.3
Iraq	1.3	Syria	1.4	Spain	1.3

Source: IEB, own calculations. Migration flows are proxied using yearly differences in migrant stocks by origin. Stocks are based on working-age migrants, that is aged 18-65. The table uses the average for these differences for each decade. The share is the proportion to the total average inflows in that decade, multiplied by 100 to give percentage points.

Table 1..14: Summary statistics

	Full Sample	Influential CZ	Other CZ
Geography			
East	0.233	0.090	0.375
Rural	0.215	0.243	0.188
Rental prices	0.203	0.227	0.180
Population			
2010 Population, 1,000's	150.311	173.717	127.113
2010 Share high-skilled	10.065	9.665	10.462
1993 Migrant share	5.217	6.703	3.744
Economy			
2010 GDP pc, 1000's	29.127	30.521	27.745
2010 Income pc	2343.722	2415.430	2272.654
2010 Manufacturing share	33.280	33.858	32.707
N	223	111	112

Source: IEB, INKAR, RWI-GEO-REDX, own calculations. Full Sample are the 223 commuting used throughout the paper. Influential CZ are the commuting zones with EU-13 immigration between 1pp and 2.6pp, see discussion in text for more details. Other CZ are the remaining commuting zones which did not receive an EU-13 immigrants inflow between 1pp and 2.6pp.

Table 1..15: Normal SSIV Bartik decompositions for main outcomes

	β	α	γ	π	G
Panel A: Δ Ln Unemployment					
Romania	-3.795	0.358	-0.420	0.111	1.007
Hungary	-3.193	0.184	-0.281	0.088	0.475
Bulgaria	-0.605	0.184	-0.073	0.120	0.505
Czech Republic	-2.780	0.077	-0.455	0.164	0.169
Poland	41.031	0.075	0.984	0.024	1.750
Slovakia	-2.542	0.058	-0.387	0.152	0.124
Croatia	-7.326	0.033	-0.374	0.051	0.069
Estonia, Latvia, Lithuania	-0.578	0.021	-0.024	0.042	0.197
Slovenia	-9.873	0.010	-0.367	0.037	0.020
Cyprus, Malta	-79.375	-0.000	1.075	-0.014	0.003
Panel D: Δ Ln Vacancies					
Romania	-4.598	0.358	-0.509	0.111	1.007
Hungary	-4.749	0.184	-0.417	0.088	0.475
Bulgaria	-0.931	0.184	-0.112	0.120	0.505
Czech Republic	-4.732	0.077	-0.775	0.164	0.169
Poland	-80.350	0.075	-1.927	0.024	1.750
Slovakia	-4.831	0.058	-0.736	0.152	0.124
Croatia	-2.598	0.033	-0.133	0.051	0.069
Estonia, Latvia, Lithuania	-57.623	0.021	-2.433	0.042	0.197
Slovenia	8.433	0.010	0.313	0.037	0.020
Cyprus, Malta	86.823	-0.000	-1.176	-0.014	0.003
Panel B: Δ Job-finding rate					
Romania	36.322	0.358	4.021	0.111	1.007
Hungary	47.000	0.184	4.131	0.088	0.475
Bulgaria	42.864	0.184	5.155	0.120	0.505
Czech Republic	29.006	0.077	4.752	0.164	0.169
Poland	182.350	0.075	4.374	0.024	1.750
Slovakia	29.510	0.058	4.493	0.152	0.124
Croatia	54.107	0.033	2.761	0.051	0.069
Estonia, Latvia, Lithuania	87.954	0.021	3.713	0.042	0.197
Slovenia	63.557	0.010	2.362	0.037	0.020
Cyprus, Malta	-143.934	-0.000	1.950	-0.014	0.003
Panel E: Δ Wages (UE)					
Romania	4.152	0.358	0.460	0.111	1.007
Hungary	4.347	0.184	0.382	0.088	0.475
Bulgaria	4.907	0.184	0.590	0.120	0.505
Czech Republic	2.876	0.077	0.471	0.164	0.169
Poland	26.498	0.075	0.636	0.024	1.750
Slovakia	2.827	0.058	0.430	0.152	0.124
Croatia	5.444	0.033	0.278	0.051	0.069
Estonia, Latvia, Lithuania	12.541	0.021	0.529	0.042	0.197
Slovenia	5.874	0.010	0.218	0.037	0.020
Cyprus, Malta	-13.829	-0.000	0.187	-0.014	0.003

Source: IEB, own calculations. The table shows the estimates from the bartik decomposition described in Goldsmith-Pinkham et al., 2020 and using the authors package *bartik_weight*. β is the second-stage coefficient, α are the Rotemberg weights, γ are reduced-form estimates, π are first-stage estimates, and G is the Germany-wide growth rate between 2010 and 2014. Exercise controls for rural, east, and quintiles of the 1993 migrant share. Observations weighted by 2010 native population.

Table 1..16: Diversion SSIV Bartik decompositions for main outcomes

	β	α	γ	π	G
Panel A: Δ Ln Unemployment					
Romania	-3.795	0.857	-0.420	0.111	1.532
Bulgaria	-0.605	0.101	-0.073	0.120	0.177
Poland	41.031	0.012	0.984	0.024	0.182
Slovakia	-2.542	0.012	-0.387	0.152	0.016
Czech Republic	-2.780	0.007	-0.455	0.164	0.010
Hungary	-3.193	0.006	-0.281	0.088	0.010
Estonia, Latvia, Lithuania	-0.578	0.004	-0.024	0.042	0.025
Croatia	-7.326	0.001	-0.374	0.051	0.001
Cyprus, Malta	-79.375	-0.000	1.075	-0.014	0.002
Slovenia	-9.873	-0.000	-0.367	0.037	-0.000
Panel D: Δ Ln Vacancies					
Romania	-4.598	0.857	-0.509	0.111	1.532
Bulgaria	-0.931	0.101	-0.112	0.120	0.177
Poland	-80.350	0.012	-1.927	0.024	0.182
Slovakia	-4.831	0.012	-0.736	0.152	0.016
Czech Republic	-4.732	0.007	-0.775	0.164	0.010
Hungary	-4.749	0.006	-0.417	0.088	0.010
Estonia, Latvia, Lithuania	-57.623	0.004	-2.433	0.042	0.025
Croatia	-2.598	0.001	-0.133	0.051	0.001
Cyprus, Malta	86.823	-0.000	-1.176	-0.014	0.002
Slovenia	8.433	-0.000	0.313	0.037	-0.000
Panel B: Δ Job-finding rate					
Romania	36.322	0.857	4.021	0.111	1.532
Bulgaria	42.864	0.101	5.155	0.120	0.177
Poland	182.350	0.012	4.374	0.024	0.182
Slovakia	29.510	0.012	4.493	0.152	0.016
Czech Republic	29.006	0.007	4.752	0.164	0.010
Hungary	47.000	0.006	4.131	0.088	0.010
Estonia, Latvia, Lithuania	87.954	0.004	3.713	0.042	0.025
Croatia	54.107	0.001	2.761	0.051	0.001
Cyprus, Malta	- 143.934	-0.000	1.950	-0.014	0.002
Slovenia	63.557	-0.000	2.362	0.037	-0.000
Panel E: Δ Wages (UE)					
Romania	4.152	0.857	0.460	0.111	1.532
Bulgaria	4.907	0.101	0.590	0.120	0.177
Poland	26.498	0.012	0.636	0.024	0.182
Slovakia	2.827	0.012	0.430	0.152	0.016
Czech Republic	2.876	0.007	0.471	0.164	0.010
Hungary	4.347	0.006	0.382	0.088	0.010
Estonia, Latvia, Lithuania	12.541	0.004	0.529	0.042	0.025
Croatia	5.444	0.001	0.278	0.051	0.001
Cyprus, Malta	-13.829	-0.000	0.187	-0.014	0.002
Slovenia	5.874	-0.000	0.218	0.037	-0.000

Source: IEB, own calculations. The table shows the estimates from the bartik decomposition described in Goldsmith-Pinkham et al., 2020 and using the authors package *bartik_weight*. β is the second-stage coefficient, α are the Rotemberg weights, γ are reduced-form estimates, π are first-stage estimates, and G is the Germany-wide growth rate between 2010 and 2014. Exercise controls for rural, east, and quintiles of the 1993 migrant share. Observations weighted by 2010 native population.

Table 1..17: RESET Test for first stage

	Standard SSIV	Diversion SSIV
F-Stat	3.806	2.243
P-values	0.029	0.117

Source: IEB, own calculations. The table shows results from a RESET misspecification test. P-values are from a T-test that $H_0 : \gamma_2 = \gamma_3$ where $gamma_2$ and γ_3 are coefficients obtained from the regression $IV = \alpha + \theta X + \gamma_2 \widehat{Y}^2 + \gamma_3 \widehat{Y}^3 + \varepsilon$ and $\widehat{Y}$ were the predicted values from $IV = \alpha + \theta X + \varepsilon$. Covariates include dummies east, rural, and quintiles of the 1993 total migrant share. The code follows Torgovitsky, 2024 but clusters standard errors at the level of 50 broader commuting zones.

Table 1..18: Robustness for native unemployment and vacancies, normal shift-share

	(1) $\Delta\, ln\, U_{nat}$	(2) $\Delta\, Urate_{nat}$	(3) $\Delta\, ln\, V$	(4) $\Delta\, ln\, V_{fill}$	(5) $\Delta\, ln\, V_{new}$	(6) $\Delta\, V/U_{nat}$
Panel A. District-level						
ΔM_{EU13}	-0.057	-0.129	0.016	0.005	0.008	0.037
	(0.019)	(0.099)	(0.033)	(0.021)	(0.025)	(0.013)
ϵ	-0.010	-0.025	0.004	0.001	0.002	0.414
Panel B. Broader CZ-level						
ΔM_{EU13}	-0.079	-0.235	0.053	0.021	0.022	0.058
	(0.027)	(0.135)	(0.027)	(0.027)	(0.027)	(0.017)
ϵ	-0.012	-0.042	0.010	0.004	0.004	0.621
Panel C. No controls						
ΔM_{EU13}	-0.006	0.556	-0.032	0.038	0.022	0.033
	(0.021)	(0.196)	(0.023)	(0.038)	(0.034)	(0.021)
ϵ	-0.001	0.110	-0.007	0.009	0.005	0.386
Panel D. Additional controls						
ΔM_{EU13}	-0.065	-0.161	0.005	-0.001	0.000	0.045
	(0.026)	(0.119)	(0.027)	(0.022)	(0.021)	(0.016)
ϵ	-0.011	-0.032	0.001	-0.000	0.000	0.528
Panel E. Unweighted regression						
ΔM_{EU13}	-0.051	-0.088	-0.020	-0.035	-0.034	0.024
	(0.021)	(0.118)	(0.039)	(0.026)	(0.027)	(0.016)
ϵ	-0.008	-0.014	-0.004	-0.008	-0.008	0.215
Panel F. Exclude moderate immigration regions						
ΔM_{EU13}	-0.029	0.018	0.012	0.014	0.004	0.031
	(0.007)	(0.046)	(0.015)	(0.012)	(0.013)	(0.006)
ϵ	-0.005	0.003	0.003	0.003	0.001	0.420
Panel G. Doube-debiased machine learning						
ΔM_{EU13}	-0.054	-0.068	-0.026	-0.028	-0.029	0.029
	(0.023)	(0.123)	(0.035)	(0.023)	(0.024)	(0.019)
ϵ	-0.009	-0.013	-0.006	-0.007	-0.007	0.337

Cluster-robust standard errors in parentheses. Estimation uses the standard shift-share instrument. Outcomes, explanatory variables, and elasticity are described in Table 1.4 in the main text. *Panel A* runs the regression using 400 districts as the unit of observation. *Panel B* uses 105 commuting zones from Kropp and Schwengler, 2016 instead. *Panel C* includes no covariates. *Panel D* includes east, rural, and quintiles of the 1993 migrant share, 2010 natural logarithm of GDP per capita, 2010 price index from Klick and Schaffner, 2019, and the share of highly educated workers in 2010 from BHP. *Panel E* runs the regression without weighting observation by the baseline native population. *Panel F* excludes the 111 "influential CZ" from Table 1..14. *Panel G* runs a partially linear IV model using double-debiased machine learning with an elastic net, gradient boosting, random forest, support vector machine, and multilayer perceptron (each with three different parameter choices and K=5 cross-validation) and averages the results using non-negative least squares shortstacking with 10 resamples. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016. Double-debiased machine learning performed using *ddml* from A. Ahrens et al., 2024 and *pystacked* from A. Ahrens et al., 2023.

Table 1..19: Robustness for native unemployment and vacancies, diversion shift-share

	(1) $\Delta\, ln\, U_{nat}$	(2) $\Delta\, Urate_{nat}$	(3) $\Delta\, ln\, V$	(4) $\Delta\, ln\, V_{fill}$	(5) $\Delta\, ln\, V_{new}$	(6) $\Delta\, V/U_{nat}$
Panel A. District-level						
ΔM_{EU13}	-0.077	-0.078	0.073	0.042	0.038	0.078
	(0.028)	(0.094)	(0.055)	(0.040)	(0.040)	(0.023)
ϵ	-0.014	-0.015	0.017	0.011	0.010	0.865
Panel B. Broader CZ-level						
ΔM_{EU13}	-0.095	-0.211	0.084	0.027	0.027	0.084
	(0.035)	(0.129)	(0.038)	(0.033)	(0.032)	(0.027)
ϵ	-0.014	-0.038	0.015	0.005	0.005	0.895
Panel C. No controls						
ΔM_{EU13}	-0.016	0.650	-0.029	0.033	0.012	0.056
	(0.024)	(0.242)	(0.031)	(0.042)	(0.037)	(0.026)
ϵ	-0.003	0.129	-0.006	0.008	0.003	0.657
Panel D. Additional controls						
ΔM_{EU13}	-0.081	-0.141	0.062	0.035	0.036	0.079
	(0.035)	(0.124)	(0.048)	(0.039)	(0.037)	(0.029)
ϵ	-0.014	-0.028	0.013	0.008	0.008	0.916
Panel E. Unweighted regression						
ΔM_{EU13}	-0.063	0.024	0.039	0.007	-0.004	0.069
	(0.031)	(0.082)	(0.052)	(0.047)	(0.046)	(0.031)
ϵ	-0.010	0.004	0.008	0.001	-0.001	0.618
Panel F. Exclude moderate immigration regions						
ΔM_{EU13}	-0.029	0.048	0.031	0.023	0.011	0.040
	(0.006)	(0.029)	(0.016)	(0.014)	(0.015)	(0.005)
ϵ	-0.005	0.008	0.007	0.005	0.003	0.532
Panel G. Doube-debiased machine learning						
ΔM_{EU13}	-0.064	0.016	0.032	0.008	0.000	0.066
	(0.033)	(0.097)	(0.053)	(0.043)	(0.040)	(0.030)
ϵ	-0.011	0.003	0.007	0.002	0.000	0.773

Cluster-robust standard errors in parentheses. Estimation uses the standard shift-share instrument. Outcomes, explanatory variables, and elasticity are described in Table 1.4 in the main text. *Panel A* runs the regression using 400 districts as the unit of observation. *Panel B* uses 105 commuting zones from Kropp and Schwengler, 2016 instead. *Panel C* includes no covariates. *Panel D* includes east, rural, and quintiles of the 1993 migrant share, 2010 natural logarithm of GDP per capita, 2010 price index from Klick and Schaffner, 2019, and the share of highly educated workers in 2010 from BHP. *Panel E* runs the regression without weighting observation by the baseline native population. *Panel F* excludes the 111 "influential CZ" from Table 1..14. *Panel G* runs a partially linear IV model using double-debiased machine learning with an elastic net, gradient boosting, random forest, support vector machine, and multilayer perceptron (each with three different parameter choices and K=5 cross-validation) and averages the results using non-negative least squares shortstacking with 10 resamples. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016. Double-debiased machine learning performed using *ddml* from A. Ahrens et al., 2024 and *pystacked* from A. Ahrens et al., 2023.

Table 1..20: Relaxing exclusion restriction: Unemployment and Vacancies

	(1) $\Delta\, ln\, U_{nat}$	(2) $\Delta\, Urate_{nat}$	(3) $\Delta\, ln\, V$	(4) $\Delta\, ln\, V_{fill}$	(5) $\Delta\, ln\, V_{new}$	(6) $\Delta\, V/U_{nat}$
Panel A. Normal SSIV						
Require Violation	39 %	0	0	0	0	47 %
Reduced Form	-0.049	-0.119	0.010	0.002	0.002	0.032
	(0.013)	(0.073)	(0.019)	(0.013)	(0.014)	(0.009)
Panel B. Diversion SSIV						
Require Violation	32 %	0	0	0	0	51 %
Reduced Form	-0.084	-0.130	0.069	0.044	0.040	0.080
	(0.023)	(0.099)	(0.035)	(0.037)	(0.033)	(0.016)
N	223	223	223	223	223	223

Source: IEB, own calculations. Required violation is the percentage of the reduced form effect endogeneity required to get bounds which touch zero. Reduced form is the coefficient from a regression of the instrument on the outcome, including controls. Results are obtained using Conley et al., 2012 Union of Confidence Intervals method implemented through the *plausexog* command (Clarke, 2014) with γ set between 0 and the reduced form effect. I increase the upper or lower bound in steps of one-hundredth of the reduced form effect until the bounds stretch over zero. The step at which this is reached is the percentage violation required. Outcomes, regressions, standard errors, and weights are described in the main text.

Table 1..21: Robustness for native transition rates, normal shift-share

	(1) $\Delta\, UE_{nat}/U_{nat}$	(2) $\Delta\, EU_{nat}/E_{nat}$	(3) $\Delta\, JJ_{nat}/E_{nat}$	(4) $\Delta\, UI_{nat}/U_{nat}$	(5) $\Delta\, IU_{nat}/I_{nat}$
Panel A. District-level					
ΔM_{EU13}	0.136	-0.026	-0.054	-0.014	0.045
	(0.053)	(0.010)	(0.018)	(0.023)	(0.077)
ϵ	0.034	-0.061	-0.035	-0.019	0.018
Panel B. Broader CZ-level					
ΔM_{EU13}	0.165	-0.025	-0.043	-0.003	0.093
	(0.067)	(0.009)	(0.016)	(0.028)	(0.072)
ϵ	0.040	-0.057	-0.027	-0.004	0.036
Panel C. No controls					
ΔM_{EU13}	0.134	0.057	0.025	-0.023	0.058
	(0.053)	(0.021)	(0.028)	(0.017)	(0.073)
ϵ	0.035	0.140	0.017	-0.034	0.025
Panel D. Additional controls					
ΔM_{EU13}	0.162	-0.023	-0.048	-0.024	0.048
	(0.064)	(0.010)	(0.018)	(0.018)	(0.082)
ϵ	0.043	-0.057	-0.033	-0.035	0.020
Panel E. Unweighted regression					
ΔM_{EU13}	0.095	-0.039	-0.075	-0.038	-0.030
	(0.077)	(0.017)	(0.030)	(0.015)	(0.108)
ϵ	0.018	-0.071	-0.041	-0.046	-0.010
Panel F. Exclude moderate immigration regions					
ΔM_{EU13}	0.055	-0.016	-0.055	-0.043	0.073
	(0.042)	(0.007)	(0.019)	(0.013)	(0.034)
ϵ	0.014	-0.033	-0.035	-0.062	0.029
Panel G. Doube-debiased machine learning					
ΔM_{EU13}	0.083	-0.038	-0.069	-0.039	-0.022
	(0.072)	(0.016)	(0.029)	(0.014)	(0.098)
ϵ	0.022	-0.094	-0.047	-0.056	-0.010

Cluster-robust standard errors in parentheses. Estimation uses the standard shift-share instrument. Outcomes, explanatory variables, and elasticity are described in Table 1.4 in the main text. *Panel A* runs the regression using 400 districts as the unit of observation. *Panel B* uses 105 commuting zones from Kropp and Schwengler, 2016 instead. *Panel C* includes no covariates. *Panel D* includes east, rural, and quintiles of the 1993 migrant share, 2010 natural logarithm of GDP per capita, 2010 price index from Klick and Schaffner, 2019, and the share of highly educated workers in 2010 from BHP. *Panel E* runs the regression without weighting observation by the baseline native population. *Panel F* excludes the 111 "influential CZ" from Table 1..14. *Panel G* runs a partially linear IV model using double-debiased machine learning with an elastic net, gradient boosting, random forest, support vector machine, and multilayer perceptron (each with three different parameter choices and K=5 cross-validation) and averages the results using non-negative least squares shortstacking with 10 resamples. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016. Double-debiased machine learning performed using *ddml* from A. Ahrens et al., 2024 and *pystacked* from A. Ahrens et al., 2023.

Table 1..22: Robustness for native transition rates, diversion shift-share

	(1) $\Delta\, UE_{nat}/U_{nat}$	(2) $\Delta\, EU_{nat}/E_{nat}$	(3) $\Delta\, JJ_{nat}/E_{nat}$	(4) $\Delta\, UI_{nat}/U_{nat}$	(5) $\Delta\, IU_{nat}/I_{nat}$
Panel A. District-level					
ΔM_{EU13}	0.026	-0.017	-0.042	-0.021	0.108
	(0.074)	(0.008)	(0.025)	(0.028)	(0.068)
ϵ	0.007	-0.040	-0.027	-0.029	0.044
Panel B. Broader CZ-level					
ΔM_{EU13}	0.085	-0.020	-0.035	-0.012	0.144
	(0.045)	(0.006)	(0.018)	(0.029)	(0.075)
ϵ	0.020	-0.046	-0.021	-0.016	0.056
Panel C. No controls					
ΔM_{EU13}	0.059	0.066	0.038	-0.045	0.088
	(0.053)	(0.026)	(0.034)	(0.019)	(0.067)
ϵ	0.015	0.161	0.026	-0.066	0.038
Panel D. Additional controls					
ΔM_{EU13}	0.088	-0.016	-0.035	-0.032	0.104
	(0.071)	(0.008)	(0.020)	(0.023)	(0.082)
ϵ	0.023	-0.040	-0.024	-0.046	0.045
Panel E. Unweighted regression					
ΔM_{EU13}	-0.017	-0.019	-0.039	-0.039	0.139
	(0.077)	(0.006)	(0.021)	(0.019)	(0.074)
ϵ	-0.003	-0.035	-0.021	-0.047	0.046
Panel F. Exclude moderate immigration regions					
ΔM_{EU13}	0.036	-0.012	-0.055	-0.054	0.096
	(0.058)	(0.004)	(0.015)	(0.010)	(0.027)
ϵ	0.009	-0.026	-0.035	-0.078	0.038
Panel G. Doube-debiased machine learning					
ΔM_{EU13}	0.003	-0.021	-0.035	-0.037	0.101
	(0.072)	(0.007)	(0.022)	(0.019)	(0.079)
ϵ	0.001	-0.051	-0.024	-0.054	0.043

Cluster-robust standard errors in parentheses. Estimation uses the standard shift-share instrument. Outcomes, explanatory variables, and elasticity are described in Table 1.4 in the main text. *Panel A* runs the regression using 400 districts as the unit of observation. *Panel B* uses 105 commuting zones from Kropp and Schwengler, 2016 instead. *Panel C* includes no covariates. *Panel D* includes east, rural, and quintiles of the 1993 migrant share, 2010 natural logarithm of GDP per capita, 2010 price index from Klick and Schaffner, 2019, and the share of highly educated workers in 2010 from BHP. *Panel E* runs the regression without weighting observation by the baseline native population. *Panel F* excludes the 111 "influential CZ" from Table 1..14. *Panel G* runs a partially linear IV model using double-debiased machine learning with an elastic net, gradient boosting, random forest, support vector machine, and multilayer perceptron (each with three different parameter choices and K=5 cross-validation) and averages the results using non-negative least squares shortstacking with 10 resamples. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016. Double-debiased machine learning performed using *ddml* from A. Ahrens et al., 2024 and *pystacked* from A. Ahrens et al., 2023.

Table 1..23: Relaxing exclusion restriction: Transitions

	(1)	(2)	(3)	(4)	(5)
	$\Delta UE_{nat}/U_{nat}$	$\Delta EU_{nat}/E_{nat}$	$\Delta JJ_{nat}/E_{nat}$	$\Delta UI_{nat}/U_{nat}$	$\Delta IU_{nat}/I_{nat}$
Panel A. Normal SSIV					
Required Violation	36%	26%	39%	0	0
Reduced Form	0.120	-0.017	-0.040	-0.014	0.042
	(0.037)	(0.006)	(0.014)	(0.015)	(0.063)
Panel B. Diversion SSIV					
Required Violation	0	23%	0	0	0
Reduced Form	0.093	-0.016	-0.037	-0.027	0.105
	(0.073)	(0.006)	(0.020)	(0.027)	(0.084)
N	223	223	223	223	223

Source: IEB, own calculations. Required violation is the percentage of the reduced form effect endogeneity required to get bounds which touch zero. Reduced form is the coefficient from a regression of the instrument on the outcome, including controls. Results are obtained using Conley et al., 2012 Union of Confidence Intervals method implemented through the *plausexog* command (Clarke, 2014) with γ set between 0 and the reduced form effect. I increase the upper or lower bound in steps of one-hundredth of the reduced form effect until the bounds stretch over zero. The step at which this is reached is the percentage violation required. Outcomes, regressions, standard errors, and weights are described in the main text.

Table 1..24: Robustness for native starting wages, normal shift-share

	(1) $\Delta\, ln\, Wage$	(2) $\Delta\, ln\, Wage(UE)$	(3) $\Delta\, ln\, Wage(IE)$	(4) $\Delta\, ln\, Wage(JJ)$
Panel A. Districts-level				
ΔM_{EU13}	0.018	0.022	0.037	0.017
	(0.005)	(0.006)	(0.015)	(0.005)
ϵ	0.007	0.009	0.015	0.006
Panel B. Broader CZ-level				
ΔM_{EU13}	0.025	0.025	0.048	0.023
	(0.005)	(0.005)	(0.014)	(0.005)
ϵ	0.009	0.009	0.019	0.008
Panel C. No controls				
ΔM_{EU13}	0.008	0.020	0.024	0.008
	(0.007)	(0.006)	(0.011)	(0.006)
ϵ	0.003	0.008	0.010	0.003
Panel D. Additional controls				
ΔM_{EU13}	-0.081	-0.141	0.062	0.035
	(0.035)	(0.124)	(0.048)	(0.039)
ϵ	-0.032	-0.058	0.027	0.013
Panel E. Unweighted regression				
ΔM_{EU13}	0.017	0.018	0.046	0.017
	(0.006)	(0.004)	(0.020)	(0.007)
ϵ	0.005	0.006	0.016	0.005
Panel F. Exclude moderate immigration regions				
ΔM_{EU13}	0.018	0.023	0.028	0.020
	(0.005)	(0.002)	(0.010)	(0.005)
ϵ	0.007	0.009	0.011	0.007
Panel G. Doube-debiased machine learning				
ΔM_{EU13}	0.015	0.018	0.045	0.016
	(0.006)	(0.005)	(0.019)	(0.006)
ϵ	0.006	0.007	0.020	0.006

Cluster-robust standard errors in parentheses. Estimation uses the standard shift-share instrument. Outcomes, explanatory variables, and elasticity are described in Table 1.4 in the main text. *Panel A* runs the regression using 400 districts as the unit of observation. *Panel B* uses 105 commuting zones from Kropp and Schwengler, 2016 instead. *Panel C* includes no covariates. *Panel D* includes east, rural, and quintiles of the 1993 migrant share, 2010 natural logarithm of GDP per capita, 2010 price index from Klick and Schaffner, 2019, and the share of highly educated workers in 2010 from BHP. *Panel E* runs the regression without weighting observation by the baseline native population. *Panel F* excludes the 111 "influential CZ" from Table 1..14. *Panel G* runs a partially linear IV model using double-debiased machine learning with an elastic net, gradient boosting, random forest, support vector machine, and multilayer perceptron (each with three different parameter choices and K=5 cross-validation) and averages the results using non-negative least squares shortstacking with 10 resamples. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016. Double-debiased machine learning performed using *ddml* from A. Ahrens et al., 2024 and *pystacked* from A. Ahrens et al., 2023.

Table 1..25: Robustness for native starting wages, diversion shift-share

	(1) $\Delta\, ln\, Wage$	(2) $\Delta\, ln\, Wage(UE)$	(3) $\Delta\, ln\, Wage(IE)$	(4) $\Delta\, ln\, Wage(JJ)$
Panel A. Districts-level				
ΔM_{EU13}	0.026	0.028	0.054	0.024
	(0.007)	(0.009)	(0.022)	(0.007)
ϵ	0.010	0.011	0.022	0.009
Panel B. Broader CZ-level				
ΔM_{EU13}	0.029	0.026	0.052	0.026
	(0.006)	(0.006)	(0.018)	(0.005)
ϵ	0.010	0.010	0.020	0.009
Panel C. No controls				
ΔM_{EU13}	0.007	0.014	0.030	0.008
	(0.009)	(0.007)	(0.013)	(0.007)
ϵ	0.003	0.006	0.013	0.003
Panel D. Additional controls				
ΔM_{EU13}	-0.081	-0.141	0.062	0.035
	(0.035)	(0.124)	(0.048)	(0.039)
ϵ	-0.032	-0.058	0.027	0.013
Panel E. Unweighted regression				
ΔM_{EU13}	0.026	0.021	0.054	0.027
	(0.007)	(0.007)	(0.026)	(0.007)
ϵ	0.008	0.007	0.019	0.008
Panel F. Exclude moderate immigration regions				
ΔM_{EU13}	0.020	0.024	0.033	0.022
	(0.005)	(0.002)	(0.010)	(0.005)
ϵ	0.007	0.010	0.014	0.008
Panel G. Doube-debiased machine learning				
ΔM_{EU13}	0.025	0.021	0.053	0.022
	(0.008)	(0.007)	(0.023)	(0.007)
ϵ	0.010	0.009	0.023	0.009

Cluster-robust standard errors in parentheses. Estimation uses the standard shift-share instrument. Outcomes, explanatory variables, and elasticity are described in Table 1.4 in the main text. *Panel A* runs the regression using 400 districts as the unit of observation. *Panel B* uses 105 commuting zones from Kropp and Schwengler, 2016 instead. *Panel C* includes no covariates. *Panel D* includes east, rural, and quintiles of the 1993 migrant share, 2010 natural logarithm of GDP per capita, 2010 price index from Klick and Schaffner, 2019, and the share of highly educated workers in 2010 from BHP. *Panel E* runs the regression without weighting observation by the baseline native population. *Panel F* excludes the 111 "influential CZ" from Table 1..14. *Panel G* runs a partially linear IV model using double-debiased machine learning with an elastic net, gradient boosting, random forest, support vector machine, and multilayer perceptron (each with three different parameter choices and K=5 cross-validation) and averages the results using non-negative least squares shortstacking with 10 resamples. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016. Double-debiased machine learning performed using *ddml* from A. Ahrens et al., 2024 and *pystacked* from A. Ahrens et al., 2023.

Table 1..26: Relaxing exclusion restriction: Wages

	(1) $\Delta\, ln\, Wage$	(2) $\Delta\, ln\, Wage(UE)$	(3) $\Delta\, ln\, Wage(IE)$	(4) $\Delta\, ln\, Wage(JJ)$
Panel A. Normal SSIV				
Required violation	44 %	54 %	28 %	42 %
Reduced Form	0.014	0.017	0.029	0.014
	(0.005)	(0.005)	(0.009)	(0.005)
Panel B. Diversion SSIV				
Required violation	46 %	47 %	20 %	43 %
Reduced Form	0.025	0.026	0.051	0.023
	(0.009)	(0.009)	(0.014)	(0.009)
N	223	223	223	223

Source: IEB, own calculations. Required violation is the percentage of the reduced form effect endogeneity required to get bounds which touch zero. Reduced form is the coefficient from a regression of the instrument on the outcome, including controls. Results are obtained using Conley et al., 2012 Union of Confidence Intervals method implemented through the *plausexog* command (Clarke, 2014) with γ set between 0 and the reduced form effect. I increase the upper or lower bound in steps of one-hundredth of the reduced form effect until the bounds stretch over zero. The step at which this is reached is the percentage violation required. Outcomes, regressions, standard errors, and weights are described in the main text.

Table 1..27: Estimates job creation and firm entry

	Plants			Jobs per plant		
	(1)	(2)	(3)	(4)	(5)	(6)
	Δ Exits	Δ Entries	Δ Net	Δ Exits	Δ Entries	Δ Net
Panel A. OLS						
ΔM_{EU13}	6.044	66.690	60.646	0.016	0.070	0.053
	(7.457)	(20.851)	(17.992)	(0.015)	(0.021)	(0.017)
ϵ	10.350	114.199	103.849	0.028	0.119	0.092
Panel B. 2SLS, normal IV						
ΔM_{EU13}	10.306	104.244	93.937	0.031	0.109	0.077
	(7.344)	(29.072)	(27.152)	(0.015)	(0.029)	(0.028)
ϵ	17.648	178.504	160.855	0.053	0.186	0.133
Panel C. 2SLS, diversion IV						
ΔM_{EU13}	15.283	123.116	107.833	0.055	0.158	0.103
	(7.901)	(32.154)	(31.939)	(0.012)	(0.040)	(0.040)
ϵ	26.171	210.821	184.650	0.094	0.270	0.176
$\bar{Y}$	830.032	1080.233	250.201	2.140	2.708	0.567
$\Delta\bar{Y}$	17.232	-48.641	-65.873	0.066	0.007	-0.059
N	223	223	223	223	223	223

Cluster-robust standard errors in parentheses. Outcomes are the natural logarithm of the gross domestic product deflated to 2015 prices (1), the natural logarithm of sales tax revenue deflated to 2015 prices (2), the natural logarithm value added per employee in 2015 prices (3), the natural logarithm of the net new jobs created, that is the number of new inflows into employment minus the outflows from employment (4). Outcomes (1) to (3) are from INKAR, outcome (4) is from BHP. ΔM_{EU13} is the change in EU-13 migrant stock between 2014 and 2010 relative to the 2010 population in percentage points. Controls include the total migrant share in 1993 as well as indicators for rural region and eastern Germany. Regressions are weighted by 2010 native population. $\bar{Y}$ is the mean of the dependent variable. ϵ is the elasticity at mean values, calculated as $\epsilon \approx \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \cdot 100$ for logarithmic outcomes and $\epsilon = \beta \cdot \frac{\bar{M}_{2010}}{\bar{L}_{2010}} \frac{1}{\bar{Y}_{2010}} \cdot 100$ for outcomes in levels. Two-stage least square estimation performed using *ivreg2* from Baum et al., 2016.

Chapter 2

Why Immigration Need Not Harm Wages: A Behavioral Explanation

Citation: Reuben, E., Stips, F., & Verheyden, B. (2024). Why Immigration Need Not Harm Wages: A Behavioral Explanation. Unpublished mimeo.

2.1 Introduction

While testing the code for the lab experiment of this paper, we employed workers through UpWork, a popular website for hiring freelancers. The workers we hired for this project had an hourly wage of four dollars. We decided to pay them six dollars. The difference was little by Luxembourgian standards, and we felt the productivity gain from more diligent work outweighed the extra cost. In other words, we were paying efficiency wages (Akerlof, 1982).

An old idea in labor economics, efficiency wages became a prominent explanation for why wages don't fall during a recession (Bewley, 1999), yet have not received attention to rationalize the lack of wage effects many studies find concerning immigration (Foged et al., 2022; Nedoncelle et al., 2024). In this paper, we make an attempt at this. We study short-run wage effects of immigration in markets where efficiency wage considerations may be important.[32] In such markets, wages serve not only to attract workers, but also to induce effort on the job. The dual purpose of wages complicates the optimal response of employers to labor supply increases.[33] Introducing efficiency wages thus yields a more nuanced perspective on the wage effects of immigration.

We begin by showing a series of stylized facts to motivate the analysis. First, we show that a considerable fraction of respondents believe that reducing effort after a wage cut is socially acceptable. We run a survey experiment on respondents in Luxembourg and elicit social norms about effort responses to wage cuts using incentivized guessing exercises à la Krupka and Weber, 2013. We find that more than half of the respondents anticipate others to cut effort if presented with a wage cut of 7% and that the results are broadly similar whether the reason for the wage cut is framed in terms of immigration or a recession. Second, we show that the fair wage expectation of migrants depends on home country factors. We use data from the German Socio-Economic Panel (SOEP)

[32] It is worth highlighting that not all markets mirror the assumptions of the efficiency-wage model, or at least less so. For some jobs, such as, for example, warehouse clerks, we would expect that effort can be more easily monitored such that the motivational impact of wages is less important compared to, for example, software engineers working in home office.

[33] Similar arguments have been made related to monopsony. Efficiency wage considerations limit the ability of employers to use their market power as this would decrease productivity (Langella & Manning, 2021).

and show that migrants from higher income origins deem higher wages as "fair" as do migrants who have been in Germany for longer. The correlations remain intact even after conditioning on the current wage level. The different wage expectations between migrants and natives will be a key building block of the theoretical model. Third, using data from a recent meta-study (Foged et al., 2022), we show tentative evidence that the percentage point wage effect of immigration is more negative the larger the inflow. This potential nonlinearity is an intuition the efficiency-wage model helps to rationalize.

We build a model to rationalize these facts. Workers have reference-dependent utility over wages with loss aversion towards wage cuts and an idiosyncratic preference for specific employers. Firms post job offers to attract workers and produce goods using capital and labor in efficiency units. They enjoy monopsony power over workers, yet must consider that labor productivity depends on the effort workers choose. When immigrants enter the market, labor supply increases, and, in turn, marginal productivity decreases. Natives' reference wage is based on the pre-immigration equilibrium. Due to loss aversion, they will reduce their effort when the competitive wage cut is applied. Migrants will nonetheless offer high effort since we assume that the home-country wage is lower. This creates an interesting trade-off for firms: either keep the pre-immigration wage and retain high effort from all workers, or offer low wages and obtain high effort only from immigrants. Which strategy is more profitable depends, in expectation, on the share of migrants in the labor market. We show that there exists a unique threshold of immigrant share below which firms find it more profitable to maintain high wages. This means that for inflows below the cutoff, immigration will have no effect on wages. The results are robust to extensions introducing imperfect substitutability between natives and migrants or the possibility of endogenous reference wages, which respond to the magnitude of the immigration shock.

One difficulty of testing the predictions from the model is that, under plausible parameter assumptions, the discontinuity in wage effects occurs only at migration inflows above 10% or even 20% of the labor force. To the best of our knowledge, this is larger than any migration episodes for which wage data is available.[34] We, therefore, choose a different route and test the predictions using a laboratory labor market (Charness & Kuhn, 2011). Besides the ability to manipulate the size of the migration shock, the laboratory also offers exceptionally clean identification and control over important decision inputs, such as the reference wage. Naturally, concerns of external validity should caution direct policy conclusions from the findings of the laboratory labor market. Yet, for the purpose of testing theory, and in the sense that our design captures the relevant aspects of the model, it can be a powerful tool with high internal validity. It also appears to be the first paper on the wage effects of immigration to use this research design.

The experimental design is an adaptation of the classic gift-exchange experiment of Fehr et al., 1993. Participants are randomly assigned to one of three markets with six workers and five employers each. They interact over 16 rounds, each comprising a market and a work stage. During the market stage, employers post public job offers, which workers can accept. During the work stage, workers decide how much effort to provide. After the round, employers can send private job offers to workers they employed in the previous round. The design until here is similar to S. Altmann et al., 2014. The main innovation in this paper is the introduction of an immigration episode. We implement this by dissolving one of the three markets after eight rounds have been played. The employers leave the experiment, and the workers are distributed to the two other markets. We move one worker to one of the remaining markets and five to the other. This generates a small and a large immigration shock in the destination markets. Both shocks are unexpected and administered through a short set of instructions before the ninth round. We also introduce an income gap between the origin and the destination markets. This is an essential aspect of the theory, as the lower reference wage means firms expect migrants to exert high effort even at lower wages. We introduce this by parametrizing the origin market at lower productivity levels so that migrants will, on average, experience an income gain after moving.

The analysis compares the evolution of wages, effort, and profits across destination

[34] For example, the labor force of Miami increased by 7% following the Cuban refugee inflow in 1980. Perhaps the largest migration episodes in history were to the United States during the age of mass migration and to Germany after the end of the Second World War. Both periods lack data sources with information on wages.

markets before and after the immigration shock. We use the individual- and job-level data to test a number of pre-specified hypotheses on the adjustment of the experimental labor markets. The findings reveal that wages initially dropped in both markets after the immigration episode. They recovered quickly in the small inflow destination, but continued to fall in the large inflow destination. Incumbent effort decreased considerably in both markets. In the small inflow condition, that led employers to increase wages gain. In the destination receiving the large immigrant inflow, however, employers continued to offer lower since migrant workers were providing higher effort. Total effort provision did not decrease in the large shock destination despite the wage drop. This is partly due to origin parametrization as a low-income market, such that migrants experience an income even if average wages at destination fall. The income gains of migrants are particularly substantial in the small inflow destination and come at no expense since incumbent workers or firms do not lose. In the large inflow destination, on the other hand, migrant income gains are more modest, and incumbent workers lose considerably. The winners are firms, whose profits increase sharply after the large immigration inflow. This shows that, as the share of immigrants increases disproportionately, the threat of effort cuts loses its bite, and firms can accrue a larger part of the migration surplus by cutting wages.

Regarding mechanisms, we find that market tightness increased considerably more in the market receiving the large inflow, as evidenced by a decrease in the number of unfilled job offers or the number of seconds a job offer remained in the market. Wages and effort dropped most strongly in single-period contracts, which indicates that some long-term cooperation survived the increase in outside options. One of the advantages of the experimental labor market is that many confounders, such as demand effects, imperfect substitution, or native migration, are ruled out by design. We can thus focus the discussion of mechanisms on how choices reacted to changes in market conditions. Delving into the role of behavioral traits, we find that positive reciprocity, measured exactly as in Falk et al., 2018, is an important determinant of gift exchange among workers. Negative reciprocity seems less important in shaping the effort response to immigration. On the firm side, we find that firms with higher risk preferences were more willing to cut wages after immigration. We also provide direct evidence that effort drops discontinuously after workers receive a wage cut, which is quantitatively much more important than the magnitude of the wage cut. We thus confirm the important modeling assumption that workers will respond with a discrete effort drop for any wage cut. This assumption is critical for our theoretical model and similar efficiency-wage models, such as Kaur, 2019.

Results from this study feed into a large literature studying the wage effects of immigration (for reviews see Dustmann et al., 2016; Edo, 2019; National Academies of Sciences & Medicine, 2017). In the competitive partial equilibrium model with perfect substitution, immigration is viewed as an increase in the labor supply, consequently reducing the wage (Borjas, 2014). Several arguments have been made to rationalize the lack of negative wage empirical studies find, including that migrants induce natives to upgrade skills and occupations, or that they boost entrepreneurship (Azoulay et al., 2021; Kerr & Kerr, 2020; Llull, 2018; Peri & Sparber, 2009). More generally, immigrants and natives may be imperfect substitutes in production, which makes the price response across skill groups unclear (Ottaviano & Peri, 2012). We add to this literature by showing that even if migrants and natives are perfect substitutes and migrants have no positive externalities on the destinations' economy, we need not expect negative wage effects. Under efficiency wages, wages serve not only to attract labor, but also to induce effort at work. This complicates the firms' optimal response and creates a downward wage rigidity. In fact, under our behavioral assumptions, the labor demand curve is *not always* downward sloping (Borjas, 2003), but flat at low levels of immigration.

If the wage effect discontinuity occurs at migration inflows larger than those observed in the recent past, is the exercise in this paper solely an intellectual curiosity? Only partly so. It also has practical implications for the design of the optimal immigration policy. The call for *open borders* is prominent among left-wing intellectuals and politicians. In a counterfactual where a popular destination like the United States or Germany lifts all immigration restrictions, we could reasonably expect immigration inflows to reach double-digits. While we do not want to overemphasize the direct applicability of the thresholds predicted by our model (since it is purposefully kept simple), the broader point is that the presence of nonlinearities should warrant caution when extrapolating findings gathered at moderate levels of immigration toward very different counterfactuals. More research with comparable research designs across countries is needed to trace out the non-linear wage response.

Our theory contributes to the literature studying the adjustment to shocks in labor markets with incomplete contracts (S. Ahrens et al., 2017; Egger et al., 2013; Eliaz & Spiegler, 2014; Heidhues & Kőszegi, 2008; Kaur, 2019), and specifically to efficiency-wage models examining the effects of immigration (Carter, 1999; Malchow-Møller et al., 2012; Müller, 2003). Relative to the latter models, our theory extends the efficiency-wage model by including workers' reference-dependent preferences and firms' monopsony power. Both assumptions are motivated by ample empirical evidence, for example, that workers adjust their effort based on wage changes relative to their reference points (Abeler et al., 2011; Coviello et al., 2022; Kaur, 2019). The result is a modern efficiency-wage model where market power shapes the wage effect of immigration: workers' effort response limits firms' ability to impose markdowns, but the ability of native workers' ability to do so may erode as immigrants become a larger share of the labor force. This microfondation of how firms' market power changes with immigration extends previous models assuming that migrants accept lower wage offers due to worse outside options (Amior & Manning, 2020; Amior & Stuhler, 2024; Borjas & Edo, 2023). The effort-response mechanism adds a layer to this mechanism. It also delivers the novel prediction that changes in market power may be irrelevant at low levels of immigration. Our model identifies critical thresholds of immigration below which firms choose to maintain the natives' reference wage to ensure high effort, and beyond which it becomes more profitable to reduce wages despite potential decreases in worker effort. Furthermore, our analysis connects with studies involving imperfect substitutability between native and immigrant workers (Card, 2009; Ottaviano & Peri, 2012). We show that the impact of behavioral responses is amplified by imperfect substitutability, as firms have stronger incentives to maintain high wages to preserve native workers' high effort levels. Finally, we generalize wage expectations based on market conditions, accounting for pre-immigration and post-immigration market wages, the latter being inspired by rational expectations à la Kőszegi and Rabin (2006).

This paper also contributes to the experimental study of labor markets, particularly those studying the dynamics of incomplete contract formation.[35] In a related exercise, Brandts et al., 2010 find that a proportional increase in the number of buyers and sellers does not affect outcomes in a gift-exchange market. Unlike them, we increase only the number of workers. Closer to our setting, Brandts and Charness, 2004 vary the ratio of buyers to sellers and find little impact on market outcomes. Relatedly, Schram et al., 2010 vary whether there is an excess of firms or workers in markets with complete or incomplete markets and find that most the surplus accrues to the short side of the market. Our design offers several innovations over existing experiments. First, we introduce immigration as a within-session change, which allows to trace out the dynamic reaction to a shock. Secondly, we explicitly manipulate the reference wage of the new workers and inform employers about this. Heterogenous reference wages add a second source of competitive pressure. More broadly, and outside the lab setting, we contribute to the study of how social norms are affected by market competition. Starting with Falk and Szech, 2013, a vibrant experimental literature investigates whether markets erode social norms.[36] Recently, Ziegler et al., 2024 indirectly varied the market by manipulating how many goods traders can sell and found that erosion occurs most strongly when markets are large and, thus, the individual influence is low. Though our model does not rely on pro-social preferences, it is similar since the influence of individual decisions by workers diminishes as more workers are in the market.

The remainder of this paper proceeds as follows. The following section 2.2 provides a series of stylized facts to motivate the theoretical analysis in section 2.3. Section 2.4 describes the details of the laboratory experiment and section 2.5 presents the results. Lastly, section 2.6 concludes.

[35] A vibrant literature emerged following the seminal contribution of Fehr et al., 1993, see for example Hannan et al., 2002, Brown et al., 2004, Brandts and Charness, 2004, Brown et al., 2012, Falk et al., 2006, Owens and Kagel, 2010, S. Altmann et al., 2014, Falk et al., 2015, Işgın and Sopher, 2015, Schram et al., 2010, Charness et al., 2017, Hannan et al., 2002.

[36] See for example Bartling et al., 2015, Bartling et al., 2023, Irlenbusch and Saxler, 2019, Sutter et al., 2020, Kirchler et al., 2016, or Ziegler et al., 2024.

2.2 Stylized Facts

We begin to motivate our analysis of efficiency wages and the theoretical modeling decisions by presenting a series of stylized facts. First, workers' effort responds to wage changes. The response is similar, if not stronger, when the wage cut is due to immigration than when it is due to a recession. Second, migrants' reference wages are shaped by the home country, but assimilate towards the destination wage over time. Third, the wage effect of immigration seems to be nonlinear in the size of the immigration shock. The last fact is also a key intuition the model can rationalize.

2.2.1 Effort provision responds to wage changes

We begin by discussing the role of wage cuts in shaping effort provision at work. For this, we rely on survey data from 826 residents of Luxembourg[37], collected as part of a larger online survey covering various topics related to societal challenges. Inspired by Kahneman et al., 1986, our module consists of a hypothetical scenario in which wages decreased by 7% due to immigration or a recession. The framing in terms of immigration or recession was randomized in equal proportion across participants. The version with immigration would read as follows.[38]

> *Due to a massive inflow of immigrants to Luxembourg, job-seekers fiercely compete for positions. A company responds by reducing its starting pay for new hires by 7%. Hence, new hires earn significantly less than in their previous jobs, whereas incumbent employees' wages remain unaffected. Considering these circumstances, what do you believe is a fair expectation for the effort of the new hires?*
>
> *In light of the massive immigration, it seems fair that newly hired workers, whose wages decreased, would exert [1. significantly less, 2. slightly less, 1. the*

[37] The original sample is 912, from which we exclude anybody who failed at least one of the three attention tests.

[38] Attentive readers will notice that we varied only the wages of new hires. This was done after discussion with HR professionals who informed the research team that it would be illegal in Luxembourg to reduce incumbent wages for the reasons mentioned in the scenario.

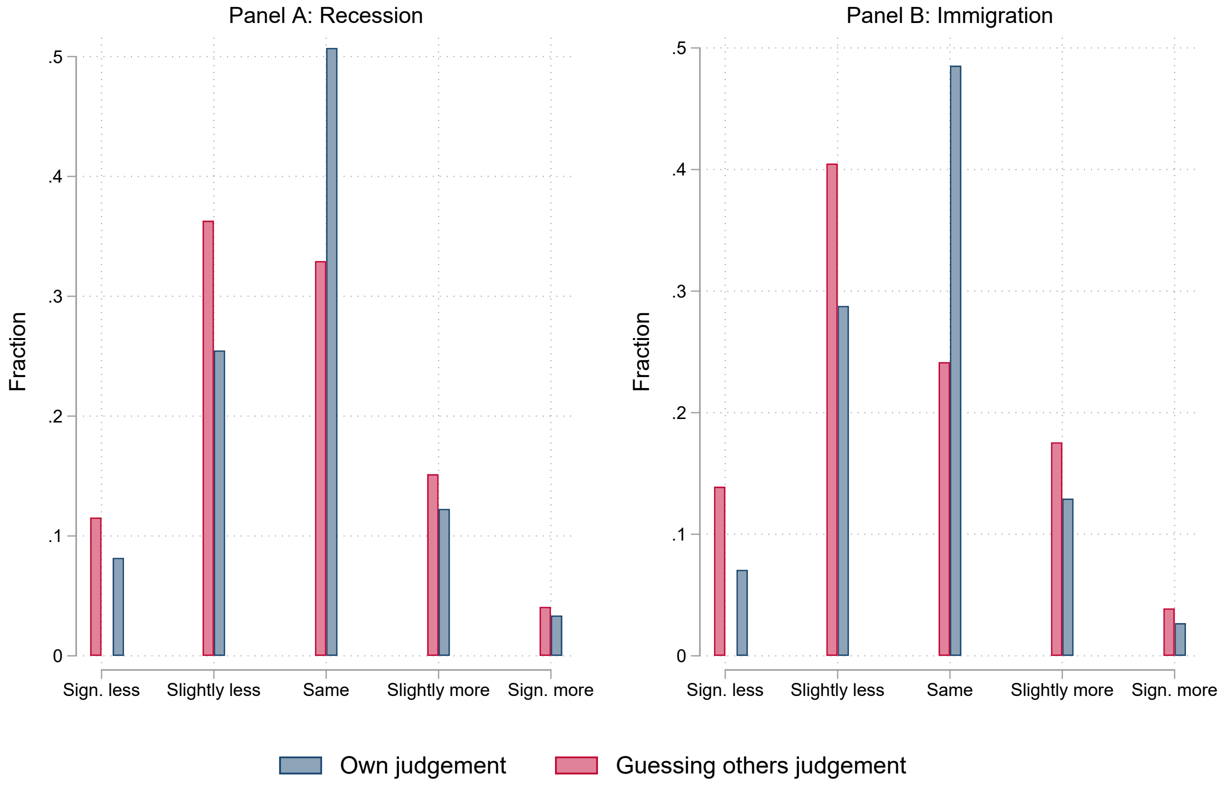
same amount of, 4. slightly more, 5. significantly more] effort than in their previous job.

Respondents first had to assess what effort response they found fair using the 5-point Likert scale shown in the question. Afterwards, they had to guess which answer was the most frequent among all other respondents. This exercise was incentivized.[39]

Figure 2.2.1 shows the distribution of effort responses on the five-point Likert scale from "significantly less effort" to "significantly more effort". The blue bars indicate the assessments of the own effort response, while the red bars are the incentivized responses to guess the most common answer among all other participants. We regard the latter as more informative about the social norm concerning the appropriate effort response to a wage cut of 7% due to a recession or an immigrant inflow.

[39] The incentivization scheme was based on a point contest. The survey included 10 norm-elicitation questions, which incentivized participants to guess other participants' most frequent response over various topics related to societal challenges, akin to the classic Krupka and Weber, 2013 method. Each norm-elicitation question would lead to a gain of 10 points if the provided answer corresponded to the actual most frequent answer among all study participants. The payment scheme was structured as follows: participants whose total scores ranked among the top 50 of all participants received a payment of 30€. Those whose scores ranked between the 51st and 150th positions received 20€, and participants ranked between the 151st and 350th positions received 10€.

Figure 2.2.1: Effort responses to wage changes



Source: Online survey with Luxembourgish respondents, N=826. The question is as described in the main text. The blue bars give the distribution of responses respondents gave to the questions. The red bars show the distribution of participants' guesses about the most common among other participants.

Two results emerge from the figure. First, a non-negligible share of respondents believes that effort may respond to wage cuts. In total, around 42% find it fair if workers reduce effort, of which 7-8% also if effort is reduced significantly. Interestingly, around 58% anticipate others doing so, and only a third to a quarter believe the most common answer would be the same effort. The gap between the assessment of the own response and the assessment of others' responses confirms the importance of incentivized norm elicitation methods.

The second finding is that the effort response to wage cuts due to immigration is very similar to the effort response to wage cuts due to a recession. The visual impression is confirmed in Appendix Table 2..1 where a Mann-Whitney U-test between treatment arms fails to reject the null of no difference between migration and recession framing. The result is also confirmed using a t-test on a binary indicator equal to one for any effort reduction in Appendix Table 2..2. The self-reported means are virtually identical, in the guessing exercise the effort response is even stronger when wages are cut due to immigration (p-value = 0.06).

Taken together, the results imply that potential effort responses to wage changes are important to assess the wage effects of immigration. The threat of effort reduction when cutting wages due to immigration is as relevant as it is in the case of recessions. Section 2.3 provides a model of firms competing for workers in the presence of migration shocks when incumbent workers' effort reacts to wages.

2.2.2 Migrants wage expectations assimilate from origin wages

In this section, we present descriptive evidence suggesting that the wage expectation of migrants differs from that of natives. We show that expectations towards a fair wage depend on the home country's income level, but that they assimilate towards the destination country wage over time spent at the destination.

We do so by making use of the Socio-Economic Panel (SOEP), the largest representative panel survey in Germany. The survey contains information about country of origin and the year of immigration. Between 2009 and 2019, some waves also included a module asking respondents what amount of gross income they would consider as fair.[40] We use data from all years for which this variable is available and drop non-migrants to yield a sample of 4600 observations from 3019 individuals.[41] The results are provided in Figure 2.2.2. Panel (a) depicts the binned mean and linear fit between fair wage and origin GDP per capita. Panel (b) depicts the relationship between the fair wage expectation and the years since migration. Both graphs use pooled, unweighted data without further controls. We find that the fair wage expectation is increasing in both the origin country's GDP per capita and the years since migration. The positive relationships are less pronounced but still present when controlling for age, education, and even the respondents' wage (Figure 2..1 and 2..2 in the Appendix). We view this as evidence that migrants reference wages differ from natives' because they are shaped by home country wages, an assumption that we make to develop the model in section 2.3. It also has the interesting practical implication that immigrants will affect wages less the longer they have been in the

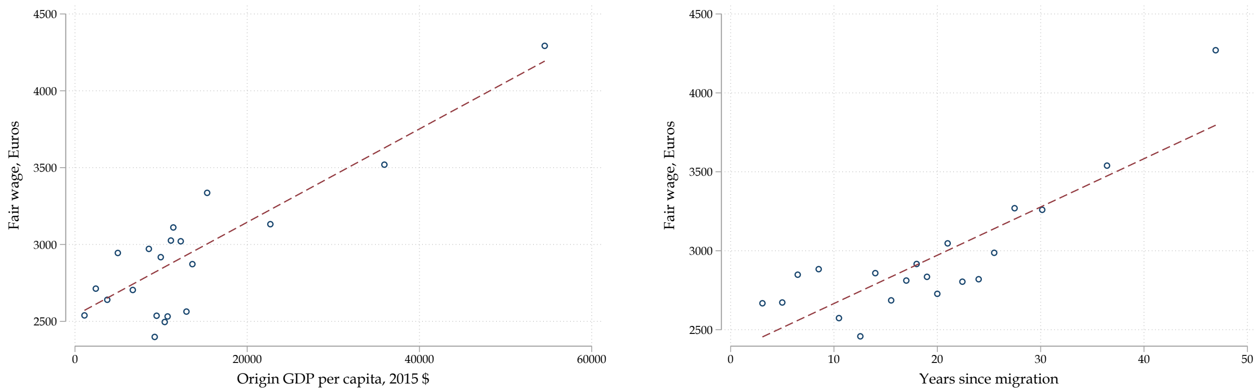
[40] The variable is plh0138. Since most people in the survey are wage laborers, we equate wage and total income.

[41] We dropped one observation who reported a fair wage of 400,000 Euros, thus six times higher than the second-highest value. The results without this exclusion are less pronounced but fall in the same direction.

Figure 2.2.2: Anchoring and assimilation of fair wage expectations

(a) Fair wages and origin GDP per capita

(b) Fair wages and years since migration



Source: Data from SOEP, 2021 and The World Bank, 2024a, own calculations. Sample restricted to migrants (N=4550). Native average is 3,621.5, migrant average is 2,936.1. Y-axis is the fair wage, proxied through fair income from the question plh0138 "*How high would your gross income have to be in order to be just?*" in gross nominal Euro. The X-axis is the origin-country nominal GDP per capita in Panel (a) and the years since migration in Panel (b). Estimation done using *binscatter2* from Droste, 2024.

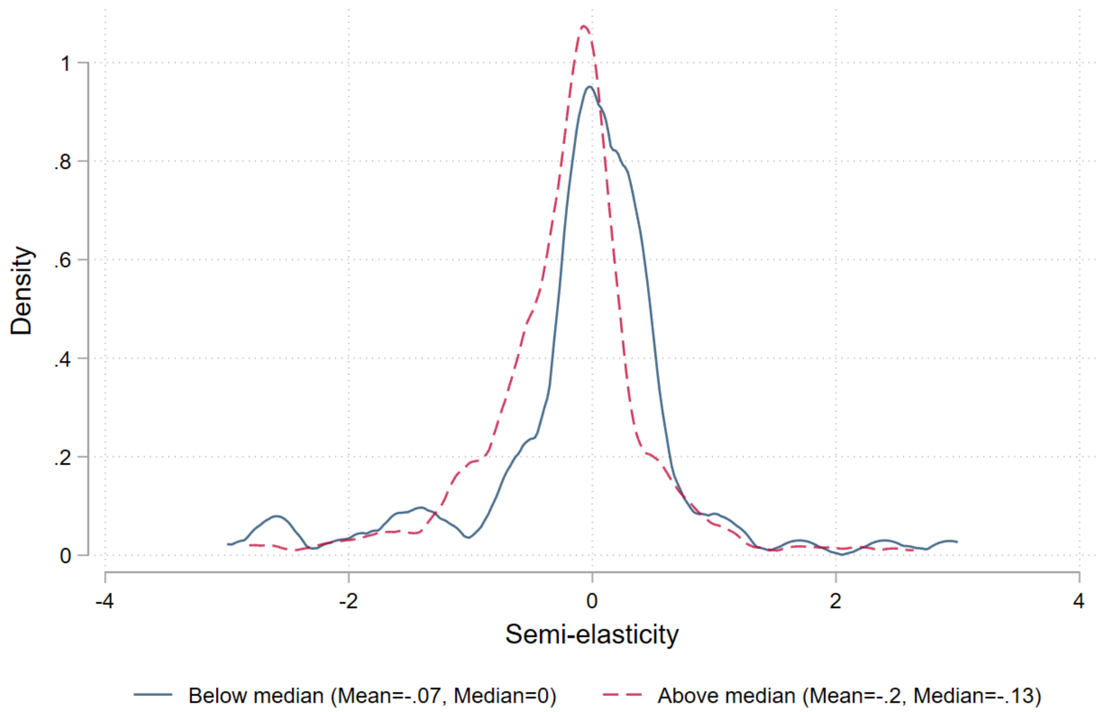
destination since the gap in wage expectations closes over time.

2.2.3 Wage effects vary with the size of the migration shock

We argue that a narrow focus on the average wage effect of immigration masks important heterogeneities. One important heterogeneity is the size of the immigration shock. To motivate this, we use data from the recent meta-analysis of Foged et al., 2022. We follow the original authors' procedure of dropping OLS estimates if IV estimates are available and dropping outliers above the author's cutoff (though the qualitative findings are robust to keep the outliers in the sample). Furthermore, we focus on studies with data at a yearly or lower frequency. This restriction is useful since we study the short-run effect of immigration on wages and want to exclude studies measuring changes across decades.[42] We combine the set of estimates with data on the inflow of migrants relative to the working-age population using data from Standaert and Rayp, 2022 and The World Bank, 2024b. We assign the average migrant inflow rate during the observation period and then split the sample of wage estimates above and below the median of migration inflows. The results are shown in Figure 2.2.3 below.

[42] The entire controversy about the effect immigration may have on native wages is only about the short-run. In the long run, when capital adjusts, the effect on wages is null even in the standard cometitive partial equilibrium model. It is unclear how long capital takes to adjust, yet the 10-year gap of studies using census data is very likely sufficient. We, therefore, exclude these studies.

Figure 2.2.3: Distribution of wage semi-elasticities



Source: Data from Foged et al., 2022, Standaert and Rayp, 2022, and The World Bank, 2024b. N=521. The X-axis gives the semi-elasticity of immigration on native wages, the Y-axis the density. The blue line is the density of wage estimates in studies where the average migration inflow during the sample period was below median. The red line is the density of wage estimates in studies where the average migration inflow during the sample period was above median.

If the effect of immigration on native wages is linear, we would expect the density of semi-elasticities to be identical for migration inflows above and below the median. This is not the case. The distribution of wage estimates below the median, depicted in the blue line, has considerably more density in the positive region. In contrast, the distribution above the median, depicted in the red line, has substantially more density in the negative area. The median wage estimate is 0 below and -0.13 above the median. In Appendix Table 2..3 column (1), we conduct a quantile regression for differences across medians in the two samples and can reject the null of no difference at 5% significance level. We repeat this exercise in Appendix Figure 2..3, keeping only the authors' preferred estimate for each study. As expected, the small sample size yields imprecise estimates of the distribution of estimates. Still, the density of negative estimates is larger for studies with migration inflows above the median, and consequently, both the median and mean elasticity are more negative. However, we cannot reject the null hypothesis of no difference in column (2) of Table 2..3.

This section showed indicative evidence that the wage effect of immigration is non-linear in the size of the immigration shock. The evidence showed some differences, albeit quantitatively not very pronounced. However, note that the immigration shocks used in the studies were all in the same rather small region. The lack of support raises difficulties in estimating non-linearities, which one might expect to occur only at high immigration levels. This is one of the advantages of the laboratory experiment: We can simulate immigration inflows larger than what has been observed in reality. But before doing so, in the following section, we present a theoretical model to rationalize why we might expect non-linear wage effects of immigration in the first place.

2.3 Theory

This model examines how an immigration shock affects a labor market where firms have monopsony power and hire loss-averse workers using incomplete contracts. Natives' reference wages are based on the pre-immigration equilibrium. When immigrants enter the market, the increased labor supply reduces marginal productivity, potentially leading to a decrease in natives' effort provision due to loss aversion. We analyze firms' profit-maximizing strategies and show that there exists a unique threshold of the immigration inflow below which it is more profitable for firms to pay natives' reference wages.

2.3.1 Baseline Model: before the immigration shock

We consider a labor market with n native workers and J firms. Firms compete for workers by setting wages w, and workers choose to work for the firm that provides them with the highest utility. The utility derived by individual i working in firm j is $U_{ij} = u\left(w_j, e_i, \epsilon_{ij}\right)$, which increases in firm j's wage w_j, decreases in worker i's effort e_i, and increases in worker i's preference for working specifically at firm j, ϵ_{ij}.[43] For the time being, effort is

[43] ϵ_{ij} captures on how i values j's characteristics, which is unobserved by j.

exogenous and normalized to 1.[44] Hence, we can write

$$U_{ij} = \alpha w_j + \epsilon_{ij},$$

where α captures the marginal utility of consumption.

Firm competition. Firms compete to attract workers by posting a wage offer w_j for $j \in \{1, ..., J\}$, and workers will opt for the firm which provides them the highest utility level. From the firms' perspective, the preference parameter ϵ_{ij} is unobserved and treated as random component. Following the canonical discrete choice model McFadden, 1973, we assume that ϵ_{ij} is distributed identically and independently across all workers and firms and has a Type I extreme value distribution. Under this assumption, for any wage w_j offered by firm j, and any vector of wages offered by competing firms $\mathbf{w}_{-j} = (w_1, \ldots, w_{j-1}, w_{j+1}, \ldots, w_J)$, the probability that a worker chooses firm j is:[45]

$$P_j\left(w_j, \mathbf{w}_{-j}\right) = \frac{\exp(\alpha w_j)}{\sum_{k=1}^{J} \exp(\alpha w_k)}.$$

Hence, firm j attracts in total L_j workers, where

$$L_j = n \cdot P_j.$$

The effective labor input of firm j is $L_j e$, where $e = 1$ for now. The firm's production function is

$$Y_j = A\left(L_j e\right)^{\theta} K^{\gamma},$$

where total factor productivity A and capital K are exogenously fixed and θ is smaller than 1. Defining the price of one unit of the firm's output as p, using $\Omega = AK^{\gamma}$ and setting $e = 1$, we can write the firm's profit Π_j as:

$$\Pi_j = p\Omega L_j^{\theta} - L_j w_j, \tag{2.1}$$

[44] In the next section, effort will be endogenized to reflect natives' response to possible wage changes induced by immigration. This section aims to determine natives' reference wage prior to such shocks, so for now, effort can remain exogenous without loss of generality.

[45] For further details on the derivation and properties of the multinomial logit model, see Train, 2009.

where $L_j \cdot w_j$ is the salary mass of the firm. When they compete for workers, firms set wages with the objective of maximizing their profits, taking other firms' wage offers as given. A marginal increase in the firm's wage offer has the following impact on profits:

$$\frac{\partial \Pi_j \left(w_j, \mathbf{w}_{-j}\right)}{\partial w_j} = \frac{\partial L_j}{\partial w_j} p\Omega\theta L_j^{\theta-1} - \frac{\partial L_j}{\partial w_j} w_j - L_j.$$

Increasing the wage brings $\frac{\partial L_j}{\partial w_j}$ more workers to the firm. These additional workers increase revenues by $p \cdot \Omega \cdot \theta L_j^{\theta-1}$ and must be paid w_j. In addition, the marginal wage increase has to apply to all workers, leading to an additional increase in the salary mass by L_j. Firm j's best response, $w_j\left(\mathbf{w}_{-j}\right)$, is implicitly represented by the following first-order condition $\frac{\partial \Pi_j\left(w_j, \mathbf{w}_{-j}\right)}{\partial w_j} = 0$. We focus on the symmetric equilibrium where all identical firms offer the same wage. In this case, each firm's probability of attracting a worker boils down to $P_{ij} = 1/J$. Proposition 1 characterizes this symmetric equilibrium.

Proposition 1. *The symmetric equilibrium wage is*

$$w^* = p\frac{\partial Y}{\partial L} - M. \tag{2.2}$$

where $\frac{\partial Y}{\partial L} = \theta p\Omega \left(\frac{n}{J}\right)^{\theta-1}$ *and* $M = \frac{1}{\alpha}\frac{J}{J-1}$.

The equilibrium profits are

$$\Pi^* = (1-\theta)\, pY^* + \frac{n}{\alpha}\frac{1}{J-1},$$

where $Y^* = \Omega\left(\frac{n}{J}\right)^{\theta}$.

Proof. See Appendix 2.6. □

The first term contained in w^* is the marginal revenue product of labor $p\frac{\partial Y}{\partial L}$, where the marginal productivity of labor is $\frac{\partial Y}{\partial L} = \theta p\Omega\left(\frac{n}{J}\right)^{\theta-1}$. The second term, $M = \frac{1}{\alpha}\frac{J}{J-1}$, is the markdown, that is, how much below marginal productivity firms can set wages. The markdown depends on the firms' market power, which is decreasing in the number of firms (which captures the intensity of competition) and in workers' sensitivity to wage differences (captured by α).

2.3.2 Immigration shock and reference wages

We now consider the impact of an immigration shock: n_I immigrants enter the labor market, which was previously composed of only n natives. The total size of the labor force becomes $n + n_I$, and firms cannot discriminate between native and immigrant workers.[46]

Our core assumption in this section is that native workers demonstrate loss aversion with respect to their wages. The wage obtained by natives before the migration shock has become their reference point. The reference wage w_r is the equilibrium wage w^* described in Proposition 1:

$$w_r = w^* = \theta p \Omega \left(\frac{n}{J}\right)^{\theta - 1} - \frac{1}{\alpha}\frac{J}{J-1}. \tag{2.3}$$

Due to loss aversion, wage changes relative to w_r are asymmetric. When wages are equal to or exceed w_r, utility is weighted by a factor of 1. However, if wages fall below w_r, utility is weighted by a factor of $\lambda < 1$. This asymmetry is represented by the reference-dependent function:

$$R\left(w_j, w_r\right) = 1 \cdot I\left[w_j \geq w_r\right] + \lambda\left(1 - I\left[w_j \geq w_r\right]\right),$$

which multiplies utility as follows:

$$U_{ij}\left(w_j, e_i; w_r\right) = R\left(w_j, w_r\right)\frac{\alpha w_j}{e_i} + \epsilon_{ij}.$$

Loss aversion becomes particularly relevant in a migration episode, as the equilibrium wage is expected to decline due to diminishing marginal productivity (see equation 2.2). Consequently, native workers may experience wage reductions. If contracts are incomplete (i.e., firms cannot enforce worker effort), workers may mitigate the negative impact of these wage cuts by reducing their effort. We assume that this effort adjustment fully compensates for the associated utility loss, which implies that effort behavior follows the reference-dependence function: $e_i^* = R\left(w_j, w_r\right)$ for all w_j. This implies the following effort

[46] The workforce composition is assumed to be identical across all firms, i.e. a fraction I of immigrants and a fraction $1 - I$ of natives.

levels:

$$e^* = 1 \text{ if } w_j \geq w_r, \tag{2.4}$$

$$e^* = \lambda < 1 \text{ if } w_j < w_r. \tag{2.5}$$

Under the effort function $e_i^* = R\left(w_j, w_r\right)$, the baseline utility function is preserved for all wage levels:

$$U_{ij}\left(w_j, e^*; w_r\right) = \alpha w_j + \epsilon_{ij}\ \forall w_j.$$[47]

While utility is continuous in w, firms' production and profits are not. Paying natives marginally less than w_r results in a discrete production drop since their effort decreases to $\lambda < 1$. Native workers have a reference wage of w_r, whereas immigrants, coming from a country with lower wage levels, have a lower reference wage, denoted as $w_I < w_r$. Immigrant workers exhibit the same productivity and preferences as native workers, apart from the difference in reference wages, such that $U_{ij}\left(w_j, e^*; w_I\right) = \alpha w_j + \epsilon_{ij}$.

Consequently, the firm's effective labor input, $E\left[e^*\right] L_j$, can take on three levels depending on the wage offered. When $w_j \geq w_r$, all workers are paid at least their reference wage and exert full effort, leading to an expected effort level of $E\left[e^*\left(w_j \geq w_r\right)\right] = 1$. When $w_j \in [w_I, w_r[$, only immigrant workers receive at least their reference wage and exert full effort, while native workers exert reduced effort. In this wage interval, the expected effort level is given by $E\left[e^*\left(w_j \in [w_I, w_r[\right)\right] = \bar{e} = I \cdot 1 + (1 - I) \cdot \lambda$, where $I = \frac{n_I}{n+n_I}$ represents the proportion of immigrant workers in the labor force. Finally, if $w_j < w_I$, all workers are paid less than their reference wage, and the expected effort level is $E\left[e^*\left(w_j < w_I\right)\right] = \lambda$.

Anticipating the effort response $E\left[e^*\left(w_j\right)\right]$, firms compete in wages for workers, leading to the firm's first-order condition:

$$\frac{\partial \Pi_j\left(w_j, \mathbf{w}_{-j}; E\left[e^*\right]\right)}{\partial w_j} = \frac{\partial L_j}{\partial w_j} p\Omega\theta E\left[e^*\right]^\theta L_j^{\theta-1} - \frac{\partial L_j}{\partial w_j} w_j - L_j = 0.$$

Using symmetry and simplifying, we can isolate the symmetric wage offer for a given

[47] Introducing loss aversion and endogenous effort in a way that preserves continuity in the utility function is critical for solving the discrete choice model.

expected effort, $\widehat{w}\left(E\left[e^{*}\right]\right)$:

$$\widehat{w}\left(E\left[e^{*}\right]\right)=\theta p\Omega E\left[e^{*}\left(w\right)\right]^{\theta}\left(\frac{n+n_{I}}{J}\right)^{\theta-1}-\frac{1}{\alpha}\frac{J}{J-1}. \tag{2.6}$$

It is important to note that since the arrival of immigrants reduces marginal productivity, $\widehat{w}\left(E\left[e^{*}\right]\right)$ remains below the reference wage even if firms were certain to receive high effort from all workers : $\widehat{w}\left(E\left[e^{*}\right]=1\right)<w_{r}$ (See Appendix 2.6). Hence, in equilibrium, the competitive wage $\widehat{w}\left(E\left[e^{*}\right]\right)$ is below natives' reference wage, leading natives to exert low effort. Firms thus receive an average effort level $\bar{e}=I+\left(1-I\right)\lambda$, leading to the competitive wage after immigration $\widehat{w}\left(E\left[e^{*}\right]=\bar{e}\right)\equiv w_{c}$, where:

$$w_{c}=\theta p\Omega\left(I+\left(1-I\right)\lambda\right)^{\theta}\left(\frac{n+n_{I}}{J}\right)^{\theta-1}-\frac{1}{\alpha}\frac{J}{J-1}. \tag{2.7}$$

However, offering this competitive wage w_{c} may not be the optimal strategy for firms. Indeed, it may be in their interest to offer the native reference wage to ensure high effort by native workers. This approach, however, is costly, as it involves paying workers more than their marginal productivity (adjusted for the monopsonistic markdown). The larger the workforce per firm $\frac{n+n_{I}}{J}$, the more expensive this strategy becomes. On the other hand, the benefits of paying w_{r} and obtaining normal effort with certainty (relative to paying w_{c} and obtaining $\bar{e}$) are greater when $\bar{e}$ is low, that is when λ is low and when the proportion of natives (who exert λ) is high. Therefore, in the case of a small immigration shock, paying the reference wage has relatively low costs and high benefits. Conversely, a large immigration shock increases the costs and reduces the benefits of the reference wage strategy.

There exists a unique threshold in the share of immigrant workers, I, such that firms obtain the same profit levels by either paying w_{r} and obtaining $e=1$, or paying w_{c} and obtaining $\bar{e}$. We denote this cut-off by $\bar{I}$, which is such that $\Pi\left(w_{r};1\right)=\Pi\left(w_{c};\bar{e}\right)$, where after some computations,

$$\begin{aligned}
\Pi\left(w_r;1\right) &= p\Omega\left[\left(\frac{n+n_I}{J}\right)^{\theta}-\frac{n+n_I}{n}\theta\left(\frac{n}{J}\right)^{\theta}\right]+\frac{n+n_I}{\alpha}\frac{1}{J-1},\\
\Pi\left(w_c;\bar{e}\right) &= \left(1-\theta\right)p\Omega\left(\left(I+\left(1-I\right)\lambda\right)\frac{n+n_I}{J}\right)^{\theta}+\frac{n+n_I}{\alpha}\frac{1}{J-1}.
\end{aligned}$$

After additional computations, we can show that $\Pi\left(w_r;1\right)=\Pi\left(w_c;\bar{e}\right)$ boils down to the following equality, which implicitly defines $\bar{I}$:

$$1-\theta\left(1-\bar{I}\right)^{-(1-\theta)}=\left(1-\theta\right)\left(\left(\bar{I}+\left(1-\bar{I}\right)\lambda\right)\right)^{\theta}.^{48} \tag{2.8}$$

We can now formally describe the equilibrium and the role played by the immigration shock magnitude.

Proposition 2. *After the immigration shock, the symmetric equilibrium wage and profits are:*

- *If $I<\bar{I}$ (small immigration shock),*

$$\begin{aligned}
w^* &= w_r, && (2.9)\\
E\left[e^*\right] &= 1, && (2.10)\\
\Pi^* &= \Pi\left(w_r;1\right)>\Pi\left(w_c;\bar{e}\right). && (2.11)
\end{aligned}$$

- *If $I>\bar{I}$ (large immigration shock),*

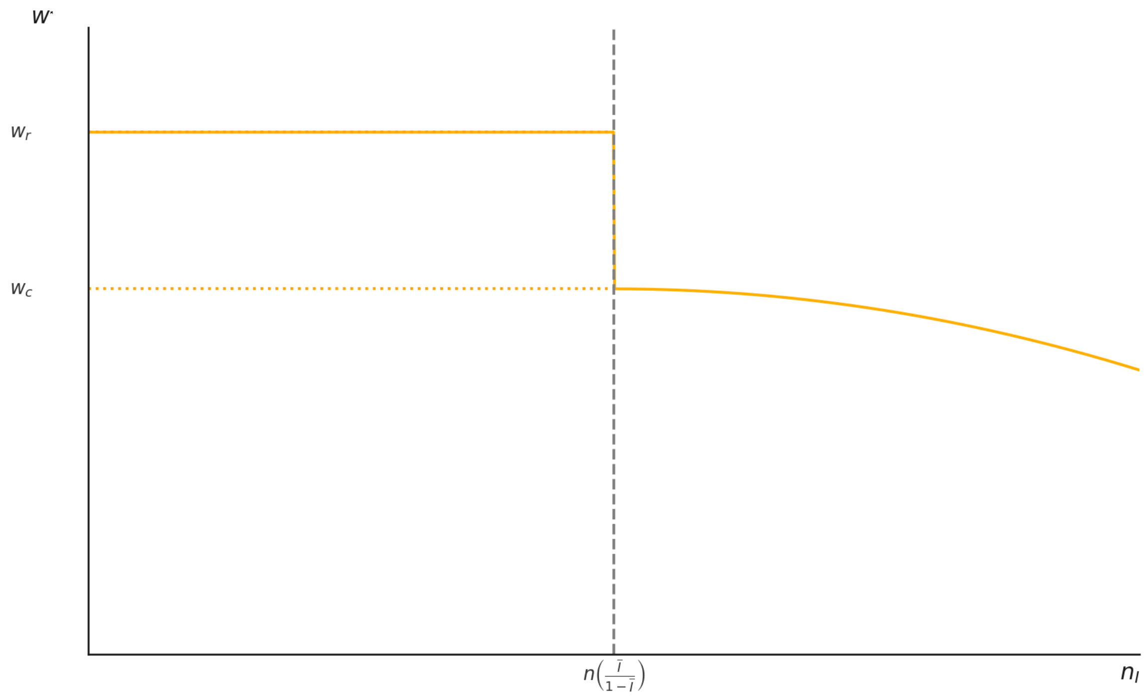
$$\begin{aligned}
w^* &= w_c<w_r, && (2.12)\\
E\left[e^*\right] &= \bar{e}=I\cdot 1+\left(1-I\right)\cdot\lambda<1, && (2.13)\\
\Pi^* &= \Pi\left(w_c;\bar{e}\right)>\Pi\left(w_r;1\right). && (2.14)
\end{aligned}$$

Proof. See Appendix 2.6. □

Proposition 2 is illustrated in Figure 2.3.1, which shows how wages respond to the number of immigrants n_I entering the labor force. Note that since $I=\frac{n_I}{n+n_I}$, the condition

[48] See Appendix 2.6.

Figure 2.3.1: Equilibrium wage w^* as a function of the number of immigrants n_I



The horizontal axis represents the magnitude of immigrant flow, while the vertical axis represents the equilibrium wage. When the immigrant flow is inferior to the threshold $n\left(\frac{\bar{I}}{1-\bar{I}}\right)$, the equilibrium wage is natives' reference wage. Beyond this threshold, the wage drops to the competitive level and is accompanied by natives providing low effort.

$I < \bar{I}$ is equivalent to $n_I < n\left(\frac{\bar{I}}{1-\bar{I}}\right)$. In other words, wages remain at the level of the reference wage if the number of immigrants does not exceed a fraction $\frac{\bar{I}}{1-\bar{I}}$ of the number of natives n:

$$\frac{\partial w^*}{\partial n_I} = \frac{\partial w_r}{\partial n_I} = 0 \text{ for } n_I < n\left(\frac{\bar{I}}{1-\bar{I}}\right).$$

Figure 2.3.1 shows that the wage is indeed constant at w_r until $n_I = n\left(\frac{\bar{I}}{1-\bar{I}}\right)$, and that a discrete wage drop to w_c occurs when n_I exceeds that threshold. Further increases in n_I then lead to a continuous decrease in wages, which results from decreasing marginal productivity:

$$\frac{\partial w^*}{\partial n_I} = \frac{\partial w_c}{\partial n_I} = -\frac{\theta(1-\theta)\,p\Omega\bar{e}^{\theta}}{J}\left(\frac{n+n_I}{J}\right)^{\theta-2} < 0 \text{ for } n_I > n\left(\frac{\bar{I}}{1-\bar{I}}\right).$$

To conclude, let us discuss the magnitude of $\bar{I}$ as this threshold plays a crucial role in the impact of immigration on wages. Analyzing equation (2.8), it is important to note that $\bar{I}$ is contained between 0 and 1, and that it is decreasing in λ. In other words, the threshold above which the equilibrium shifts from the reservation wage to the competitive wage has a relevant magnitude, and the larger λ, the lower this threshold. This means that countries with institutions that allow firms to enforce higher minimal levels of effort are less likely to maintain the reference wage for the same immigration shock magnitude (than countries with weaker effort enforcement). Table 2..9 in Appendix 2.6 provides a calibration of $\overline{I}$ for various values of low effort λ. When $\lambda = 0.7$ (i.e., low effort represents 70% of normal effort), firms offer competitive wages (rather than sticking to the reference wage) if the migration labor force exceeds 21% of the total labor force ($\overline{I} \simeq 0.21$). This threshold goes down to 15% for $\lambda = 0.8$. For $\lambda = 0.9$, a migration shock of 9% suffices to cut wages.[49] Note that in the trivial case where effort is entirely enforceable by firms, i.e., $\lambda = 1$, we have that $\overline{I} = 0$: since natives cannot respond to wage cuts, firms would apply the competitive wage irrespective of the magnitude of the immigration shock.

Note that we implicitly assumed that migrants always exert normal effort at the competitive wage, meaning that the competitive wage is always larger than migrants' reference wage.[50] From the perspective of the experiment, it is, however, worth discussing the implications of a case where immigrants and natives would have the same reference wage. Then, firms would anticipate that any wage offer below w_r would result in *all* workers offering low effort. In this case, profits under the competitive wage would be lower, implying a higher threshold $\overline{I}$. In other words, wage cuts are far less likely when immigrants have the same reference wage as natives.[51]

Finally, two extensions of the model are provided in the Appendix. Since the previous results were obtained under the assumption of perfect substitution between natives and immigrants, the first extension, in Appendix 2.6, introduces imperfect substitution. This generalization leads to the same qualitative result as Proposition 2, although imperfect substitutability increases the threshold $\overline{I}$. This is because, ceteris paribus, firms' incentives to maintain high effort by natives are reinforced by the fact that immigrants are imperfect substitutes. Hence, firms will offer the reference wage w_c under imperfect substitution in situations where they would have paid the competitive wage under perfect

[49] We obtain identical results for $\theta = 0.6$.

[50] This would require a migration shock that is so large that marginal productivity of labor would reach the level of the immigrant's origin country.

[51] For instance, for $\theta = 0.7$ and $\lambda = 0.7$, $\overline{I}$ would equal 0.35 (instead of 0.21 when the competitive wage is assumed always larger than migrants' reference wage).

substitution. The second extension generalizes natives' reference wage formation by considering the impact of immigrant competition on wage expectations. In Appendix 2.6, the reference wage is a convex combination of natives' past wage and of the (lower) competitive wage they would obtain after immigration. As a consequence, the larger the immigration inflow, the lower the reference wage and the stronger the incentive for firms to partially decrease the equilibrium wage while satisfying natives' adjusted wage expectations. This extension provides an explanation to the experiment's finding that the large immigration inflow may induce a wage cut without triggering an effort drop among natives.

2.4 Experimental Design

This section describes the design and procedures of the laboratory experiment. It is a modified version of the classic gift-exchange game Fehr et al., 1993. The key innovation is to introduce a "migration shock" after half of the rounds have been played, which we do by moving players across markets with income differences. The instructions and ex-post questionnaire are available in Appendix 2.6.

2.4.1 Baseline

Each session begins by randomly assigning participants to one of three markets and to a role, which is either worker or employer. Participants maintain this role throughout the session. There are six workers and five employers in each of the three markets. Workers and firms will interact over the course of 16 periods.

Each period consists of two stages: a market stage and a work stage. During the market stage, firms post job offers, and workers decide whether to accept them. A work contract lasts one period and entails an offered wage level between 0 and 100 and a desired effort level, which can be either "normal" or "low". Job offers sent during the market stage are public.[52] This means all participants can see the job offer, and all workers can accept

[52] Note that job offers do not contain employer identifiers, and employers are not shown identifiers of the workers who accept their offers. This means that employers and workers do not directly observe whom they interact with. Longer cooperation is only possible through private reemployment offers. This design reduces the importance of reputation effects as compared to previous gift-exchange experiments.

it. The first worker to accept a job offer secures it and enters a working relationship with the firm. Employers can hire up to two workers in one period. Workers can accept one job offer or remain unemployed. The market phase ends if all workers have been employed, all employers indicated they are done hiring, or if the maximum duration of 150 seconds is reached.

Workers who accept a job offer enter the work stage, during which they decide how much effort to provide. They know the effort level firms requested but can choose freely whether to comply. Workers can choose between two effort levels, which are labeled as "normal" and "low" effort.[53] Providing effort is costly to workers. To provide "normal" effort, they have to pay a cost of effort of $c(e) = 20$. To provide a "low" level of effort, they have to pay a cost of effort of $c(e) = 10$. Employers can observe the choice of effort but have no means to enforce it. After the work stage, payoffs are distributed to all parties, results are shown, the period concludes, and the following period begins.

Before the next market phase starts, employers can send private reemployment offers to one or both of the workers they employed in the previous round. Workers will see this offer during the next market phase. They can choose either the private offer or one of the public offers. Apart from that, private offers work just as public job offers. They contain a wage and a desired effort level. Workers can accept at most one offer, either private or public. Employers can employ at most two workers, including workers hired through private or public offers.

A worker's payoff in any period depends on whether they are employed and, if so, what wage they receive and what effort they provide. If a worker is employed, their payoff is the difference between the wage w and the cost of effort $c(e)$. If they are unemployed, they receive no payment. The per-round payoff function π_w is thus

$$\pi_W = \begin{cases} w - c(e) & \text{if employed} \\ 0 & \text{otherwise.} \end{cases} \tag{2.15}$$

[53] We choose the label normal as we want to emphasize that providing low effort is a negatively reciprocal act, rather than providing the higher effort choice being the positively reciprocal act. That is because we want to create a situation in which the market is in the higher effort equilibrium before the immigration shock happens. The case when the market is in the low effort equilibrium before the immigration shock happens is not of theoretical interest since it coincides with standard non-behavioral models.

An employer's payoff depends on the number of workers they employ, the wages they pay, and the effort workers choose. If an employer hires one worker, they will receive a revenue of 160 if the worker provides normal effort and 80 if the worker chooses low effort. If the employer hires two workers, they will receive a revenue of 140 for *each* worker choosing normal effort e_H and 70 for each worker choosing low effort e_L. The profit in both cases is the difference between the total revenue and the wages paid to all workers. If an employer hires no workers, they receive a payment of zero for that round. Note that the payoff function implies that workers are perfectly substitutable and that productivity decreases in the number of workers, an assumption akin to decreasing marginal returns in the theoretical model.[54] The employer's per-round payoff function π_E is given by

$$\pi_E = \begin{cases} 140e_H + 70e_L - w_1 - w_2 & \text{if two workers} \\ 160e_H + 80e_L - w_1 & \text{if one worker} \\ 0 & \text{otherwise.} \end{cases} \tag{2.16}$$

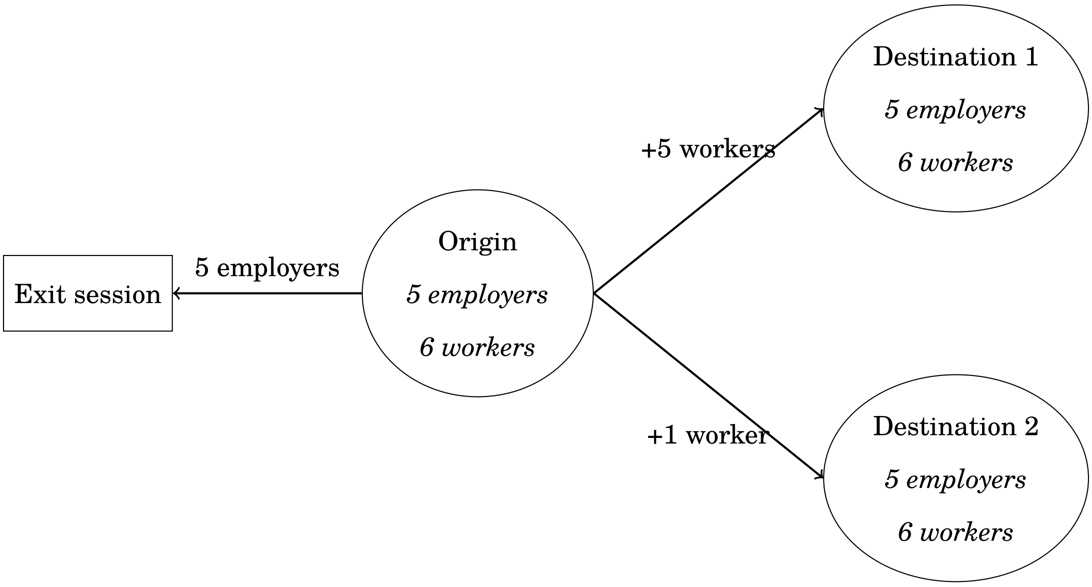
After a period is concluded, there is a short break showing the results of the market phase and a question for employers whether they want to offer a private contract to the workers they hired in the last round. The final payout is the sum of the profits from all rounds.

2.4.2 Migration

The baseline structure described above continues for the first eight rounds of each session. Participants are informed at the beginning that it consists of two parts, and they will receive a second short instruction to inform them about the changes in the second part of the experiment. In reality, the second part of the experiment is virtually identical to the first part. The only difference is that between part 1 and part 2, a migration episode occurs. Figure 2.4.1 illustrates the logic of the migration episode.

[54] Strictly speaking, average as opposed to marginal returns are decreasing, but since we have no clear sequential structure in hiring we cannot introduce marginal productivity. However, the trade-off of decreasing returns to hiring more workers is given in both cases.

Figure 2.4.1: Migration Episode



The migration episode entails that one of the three markets is dissolved after part 1. Employers from this market will skip the second part and move directly to the ex-post questionnaire. The six workers of the origin market will move to one of the other two destination markets. Five workers will move to destination 1, and one to destination 2. This means destination 1 will contain five employers and eleven workers, while destination 2 will contain five employers and seven workers for the remaining rounds in part 2. We do so to implement the "large" and "small" migration shocks.

Subjects receive short instructions about the migration episode. Employers are provided with more information about the newcomers than workers; in particular, the income level of the origin market is revealed to them. This was done to ensure that employers are informed about the reference wage of the newcomers. The instructions shown in the mid-break are provided in Appendix 2.6. The remaining markets continue with the same structure as the baseline game, but with more workers.

2.4.3 Treatments

Treatment is assigned at the market level. We use a two-by-two factorial design, varying the intensity of the immigration shock and whether there is an income difference between the origin and destination market. Table 2.4.1 below shows the four possible treatment states. Note that we omit a control state without migration. We do so to save power and instead use data from the rounds before the migration state as a comparison.

Table 2.4.1: Treatment matrix

	Small shock	Large shock
Same income	SN	LN
Income difference	SI	LI

We vary the immigration intensity by varying the number of workers that move from origin to destination. Specifically, a *small* migration shock means that the market receives one worker. A *large* migration shock entails receiving five new workers during the migration episode. As discussed previously, we implement this by distributing the six workers from one origin market to two destinations. This approach minimizes the number of "lost observations", that is, subjects who do not participate in the post-migration market. While we implement large and small migration shocks simultaneously, workers in both states are unaware of the other group and treatment state. These treatment conditions manipulate the degree of competition following the migration shock.

We also vary whether there is an income difference between the markets. We use two currencies: points are the currency in the destination market, and tokens are the currency in the origin. In the same income treatment condition, the exchange rate between the two currencies is one-to-one. The origin and destination are thus identical, and migrants experience no income difference after moving from one market to another. In the income difference condition, the exchange rate between points and tokens is 0.5, meaning that one point provides as many pounds as two tokens. This exchange rate is real in the sense that firms are also less productive in the origin market and receive only half the revenue for each unit of high or low effort from workers. This means that workers and employers in the origin market will initially profit less from similar offers. Migrants continue to see the wage displayed in tokens at destination but will see, on average, higher wage levels than before. These treatment conditions serve to manipulate the reference point of migrants. Under income differences, they will perceive wage offers in the new market as higher and are more likely to provide high effort. Employers in the destination market are informed about this in the mid-break instruction: They receive information about the effort migrants provided for different wage levels during a pilot. Workers in the destination market are only informed that there are more workers.

2.4.4 Procedures

The experimental sessions were conducted at the Centre for Decision Research and Experimental Economics (CeDEx) at the University of Nottingham. Participants were recruited through the Online Recruitment System for Economic Experiments (ORSEE) platform (Greiner, 2015) from the CeDEx subject pool. The software was programmed in Otree (D. L. Chen et al., 2016) and deployed using Heroku servers located in Europe with 1 XL web dyno and Postgres Standard 2 database.[55] Full replication codes for the analysis will be published as an OSF repository after implementation. The Institutional Review Boards at NYU Abu Dhabi and the Luxembourg Institute of Socio-Economic Research (LISER) approved the ethical standards of the experimental design.

We will conduct 20 experimental sessions with 33 participants each, ten of these have been conducted for the current draft version.[56] The total number of participants was 660, including students from all subject areas. Each session began with a set of instructions, which participants read from within the program. Participants had to successfully answer three comprehension questions after the instructions before the session. The comprehension quiz tested their understanding of the profits employers and workers make under different scenarios. After the sessions, students completed a 5-minute ex-post questionnaire covering basic socio-demographic information and behavioral traits following Falk et al., 2023.

Sessions lasted 60 to 70 minutes on average. Profits from all rounds count towards the final payoff. The average payout was 18.5£. Participants who went bankrupt received a minimum payment of 12£. Payments were administered through PayPal after the session was concluded. Personal details used for registration and payment were stored separately from the main dataset used for analysis.

[55] OTree programs are available at https://github.com/FelixStips/gefinal/.

[56] Initially, sessions were planned to include 36 participants, with six employers and six workers per market. However, the larger lab with 40 seats at NUBS had to close during our study period, and we had to rely on the smaller laboratory. We thus decided to remove one employer per market and run sessions of 33 participants.

2.5 Preliminary Results

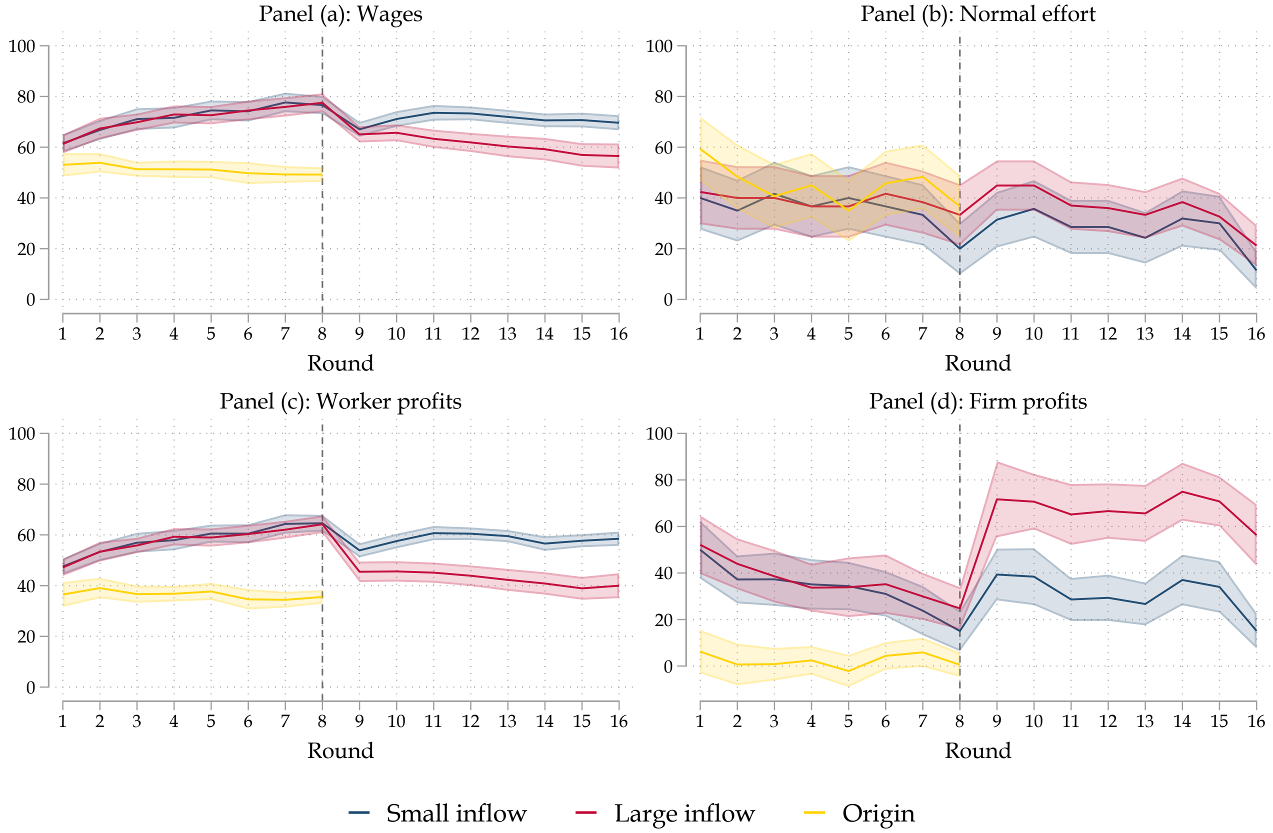
This section presents the results, which, in this draft, are based on data from ten experimental sessions. These sessions were implemented using the condition with an income difference between origin and destination. The main part of the analysis is based on pre-specified hypotheses, as laid out in AsPredicted entry #158104 (see also Appendix 2.6). However, we have included additional analysis that we find too interesting to dust in the file drawer. We also added another short ex-post questionnaire after the first five sessions (see Appendix 2.6). To communicate this transparently, we will mention whenever a hypothesis was prespecified. We understand readers may wish to place less weight on findings lacking this statement, but hope they agree that the additional results enrich the understanding overall. The fact that Clemens and Lewis, 2022 is the only other preregistered study on the wage effects of immigration shows, by revealed preference, that readers likely deem results from ex-post hypotheses informative as well.

2.5.1 Main Results

We begin the analysis with a visual inspection of the evolution of the main outcomes across markets over time. Below, as preregistered, Figure 2.5.1 plots the average wage, effort provision, and profits in each round and market. The blue and red lines indicate the two destination markets, while the gold line is the origin market, which dissolves after round eight.

Average wages depicted in panel (a) responded immediately and strongly in both destinations. In the round after the immigration shock, the average wage of employed workers fell by 12.5% in the destination receiving one worker and by 16.1% in the destination receiving five workers. This is a large magnitude and markedly different from the pre-immigration trends, where wages grew similarly in both markets. After the initial drop, wages recovered somewhat in the destination with the small inflow. They did not recover in the destination receiving the inflow, where the wage decline continued until

Figure 2.5.1: Average outcomes by market and round



Source: Own calculations. The figures plot the mean and 95% confidence intervals of different in outcomes for each round. The yellow line is the origin market, the blue line is the destination that received one worker, and the red line is the destination that received five workers. The outcomes are in panel (a) the average accepted wage, in panel (b) the share of workers providing normal effort, in panel (c) the average per-round profit of employers, and in panel (d) the average per-round profit of workers in the market.

the end of the experiment.[57] In Appendix Figure 2..4, we show that wage evolution is very similar when using offered wages, which includes wages of unfilled jobs, or when using only the wages of incumbent workers. The latter is the most commonly used measure in the wage effects of immigration literature.

Workers' effort provision in the destination responded less than wages to the inflow of workers. In Panel (b) of Figure 2.5.1, we illustrate this by plotting the share of workers choosing normal effort over time in each market. We do not find a substantial decrease, particularly if we exclude the last two rounds, when the willingness of workers to cooperate may drop due to other considerations. Also interesting is that effort provision was, if anything, slightly lower in the small market despite wages being cut less. In Appendix

[57] It is interesting to note that wages in the origin market stagnated. This is because lower productivity meant that firms could not offer high wages, as evidenced by the evolution of firm profits in panel (d). This implies that the experimental design successfully created a wage gap between origin and destination such that migrants experience a wage increase when moving. This was important because the gap in reference wages between migrants and natives is critical to the theoretical model.

Figure 2..5, we delve deeper into the effort response. In Panel (a), we see that firms initially tried to ask more often for high effort, and more so in the large shock, but over time, they stopped. Indeed, it seems as if firms asked increasingly less for high effort over time. In panel (b), we see that the share of incumbent workers providing normal effort decreased in both markets.

If wages decrease more than effort, firm profits should increase, and worker profits shrink. This is what we observe in panels (c) and (d) of Figure 2.5.1. The average per-round profit of workers was around 64 points before the shock. In the round immediately after the shock, it dropped by 16.6% in the market receiving one worker and by 29.2% in the market receiving five workers. Worker profits recover over time in small shock conditions but continue to fall in the large shock market. The reverse can be seen for firm profits, which increase massively and remain high in the large inflow destination while increasing less and only temporarily in the small shock. This, coupled with the effort response, shows that firms in the small market did not manage to exploit workers as they did in the large shock condition. In Appendix Figure 2..6 we distinguish between the profit of incumbent versus migrant workers. Migrants experienced an income gain relative to the origin market, while incumbent workers witnessed an income drop.

Next, we want to quantify the differences in outcomes before and after immigration and conduct statistical inference on them. In Table 2.5.1, we do so by comparing average outcomes before and after the shock. In columns (1) and (2), we run the following prespecified regression separately for each destination market

$$y_{i,t,s} = \alpha + \beta \cdot \mathbb{I}[t \geq 9] + \gamma_s + \varepsilon_{i,t,s} \tag{2.17}$$

where $y_{i,t,s}$ is the outcome of a job or an individual i in round t of session s, $\mathbb{I}[t \geq 9]$ is an indicator taking the value of one if the round is after the migration shock occurred, γ_s are session fixed effects, and $\varepsilon_{i,t,s}$ is an error-term clustered at the round-market-level. The parameter of interest is β, which estimates the difference in means in the outcome before and after the migration shock. Since the immigration episode is exogenous, the simple before-after comparison should be an unbiased estimate of the average treatment effect of the respective immigration shock. The only identification threat is the presence of round effects, for example, because endgame considerations in the last round decrease workers' effort provision. To mitigate these concerns, we will conduct a robustness check excluding the first and last three rounds from the analysis in section 2.5.3.

In column (3), we also explore differences across destinations using the following prespecified difference-in-difference regression

$$y_{i,t,m,s} = \alpha + \beta \cdot \mathbb{I}[t \geq 9] \cdot \mathbb{I}[m = large] + \eta \cdot \mathbb{I}[t \geq 9] + \zeta \cdot \mathbb{I}[m = large] + \gamma_s + \varepsilon_{i,t,m,s} \quad (2.18)$$

where the notation is as before except that $m = \{small, large\}$ denotes the two destination markets and $\mathbb{I}[m = large]$ is an indicator taking the value of one if a job or individual was in the market that received the large immigration shock of five workers. The parameter of interest is β, which estimates how much more the outcome reacted to the immigration episode in the large shock market vis-à-vis the small shock market. The assumption of parallel trends holds by definition since the markets are identical before the shock. Hence, the results yield unbiased estimates of the average treatment effect on the treated of receiving five instead of one new worker.

At the top of Panel A, we confirm the previous impression that the post-migration wage drop was significant only in the large shock. The results in column (2) show that wages after the large shock were, on average, 10.5 points lower than before. This corresponds to a 14.6% drop relative to the pre-immigration average of 71.5. For an increase in labor supply of 83.33%, the wage change corresponds to a wage elasticity of -0.176. The results are very similar when considering native wages and remain intact also when considering offered wages. The market receiving the small supply shock, instead, witnessed a small and insignificant wage drop, as shown in column (1). The corresponding wage elasticity is -0.06 relative to the 16.67% labor supply increase. The coefficient on native wages is even smaller, while the coefficient on wage offers also flips sign. In column (3), we assess the double difference between the small and large shock markets and find that the wage response in the large market was significantly stronger. This should be no surprise, since wages responded only in the large shock destination.

The remaining outcomes in Panel A of Table 2.5.1 show that the effort response to immigration was less strong than the response in wages, at least in terms of statistical

Table 2.5.1: Main regression results

	(1) Pre-post small	(2) Pre-post large	(3) Double-difference
Panel A. Job-level outcomes			
Wages	-0.715	-10.459***	-9.701***
	(1.974)	(2.136)	(2.861)
Native wages	-0.540	-11.002***	-10.482***
	(1.958)	(2.041)	(2.791)
Offered wages	1.929	-5.507***	-7.337***
	(1.510)	(1.826)	(2.291)
Normal effort	-0.077**	-0.028	0.050
	(0.035)	(0.028)	(0.044)
Native normal effort	-0.097**	-0.070***	0.028
	(0.034)	(0.021)	(0.039)
Request normal effort	-0.041	-0.032	-0.003
	(0.027)	(0.033)	(0.043)
Panel B. Individual outcomes			
Firm profit	-1.920	31.197***	33.118***
	(4.479)	(3.509)	(5.583)
Worker profit	3.258	-5.375*	-8.633**
	(2.351)	(2.989)	(3.219)
Native worker profit	0.271	-14.996***	-15.267***
	(2.086)	(1.974)	(2.819)
Migrant worker profit	21.181***	6.170***	-15.011***
	(1.713)	(1.602)	(2.339)
Session FE	Yes	Yes	Yes
Individuals	120	160	280
Accepted jobs	1039	1270	2309
Jobs	2077	1972	4049

Source: own calculation. Results in columns (1) and (2) display estimates from the regression described in eq. 2.17, while columns (3) displays estimates from the regression described in eq. 2.18. Standard errors clustered at the round-market level in parentheses. Significance: * 0.10 ** 0.05 *** 0.01. Outcomes: Wages is the pay in experimental currency of all accepted job offers. Native wages in experimental currency of all job offers that incumbent workers accepted. Offered wages are the pay of all job offers, including open and canceled. Normal effort is an indicator taking the value of one if a worker provided normal effort, native normal effort conditions on the workers being an incumbent. Shirking takes the value of one if a worker provided low effort conditional on the job offer requiring normal effort. Request normal effort takes the value of one if any job offer requested normal effort, zero if it requested low effort.

significance, and that this is mainly so because migrants provided higher effort. The probability that a worker provided normal effort decreased by almost -0.08 in the small shock, which is significant at the 5% level and relatively large in magnitude compared to the pre-immigration average of 0.386. The effect was much smaller and insignificant in the large shock condition. It seems workers were more able to retaliate against low wages with low effort in small immigration destination. The row below illustrates part of the answer. Effort provision of incumbent workers did fall after the immigration episode also in the large shock condition. The gap between total and incumbent effort is likely due to migrants providing normal effort more often. Since they represent a larger portion of the market, the wedge in the large shock condition is larger. Note that this mechanism is very similar to the model. Since the chance of getting an incumbent worker decreases as the immigration inflow increases, firms prefer low wages since they expect to hire migrants more frequently who provide normal effort nonetheless. Lastly, we test whether employers adjusted to the worker's effort response by requesting normal effort less often and find, again, that they did so only in the small shock market, if at all.

In Panel B of Table 2.5.1, we confirm that firm profits increased and workers decreased sharply only in the large immigration shock market. The magnitude of these effects is rather striking. For example, the effects of native worker profits in column (2) are larger than they were for the wage drop. This implies that there was also a small displacement effect, which is smaller than the intensive margin wage effect. Also, the differences between the destination markets are striking as none of the effects are significant in the small shock markets in column (1), and the double-difference specifications in column (3) yield strong, significant differences between the markets. Perhaps the most interesting aspect of the results is comparing native and migrant worker profits across markets. In the small shock condition in column (1), neither incumbent firms nor incumbent workers lose significantly due to immigration. Yet, migrant workers gain a lot by moving from a low-income to a high-income market (Clemens, 2011). Excluding origin firms for the moment, the small immigration episode was a clear Pareto improvement since nobody lost, yet some gained considerably. This is markedly different in the large shock condition in column (2). Migrant workers still gain somewhat by moving to the higher income market. However, the gain is much smaller than in the small immigration destination. Also, incumbent workers lose much profit in the large immigration destination. Instead, the winners of the large immigration episode are firms whose profits increase considerably. The increased competition allowed firms to grab most of the migration surplus.

This section aimed to analyze the response of wages, effort, and worker and firm profits in response to the immigration shocks. The findings revealed that wages declined significantly in the market that received the large immigration shock while only slightly in the market receiving the small immigration inflow. The effort response of natives was negative in both markets. The response was less relevant in the large inflow destination since there were more migrants to return high effort at low wages. It seems that incumbent workers in the small shock condition were better able to punish the initial wage drop such that firms restored previous wages. Worker profits declined, and firm profits increased significantly only in the large inflow market. The following section will delve more into the mechanisms behind these results.

2.5.2 Mechanisms

In this section, we delve into mechanisms that may explain the previous results. Relative to other papers studying the wage effects of immigration, we have an easier task at this because the market we study was tightly controlled. This means that migrants and natives were perfect substitutes, there were no effects on firm creation or product demand, and there was also no possibility that incumbents move to other regions. With these explanations out of the way, we can focus on the dynamics of worker-employer interactions and the role of behavioral traits.

Table 2.5.2 reports results from identical regressions to those in the previous section but uses other outcomes to understand how market conditions changed following immigration. The employment duration outcome was prespecified. In Panel A, we look at market-level outcomes; in the top row, we see that the number of unfilled jobs, that is, jobs that were canceled or remained open until the end of the round, decreased slightly in both markets post-immigration. The overall magnitude is not very large, but the small standard errors yield high statistical significance. Interestingly, the decrease was much stronger in the large shock condition. This aligns with the previous results, indicating that market tightness increased strongly only in the large shock condition. This reading appears to be confirmed in Panel B, where, at the top, we assess the effect of immigration on the duration in seconds a job offer stayed in the market. On average, job offers remained open for less time in part two across both markets. Partly, this may be explained by learning, in the sense that employers know from previous experience whether an offer will be filled. Yet, this cannot explain why vacancy duration would fall more in the large immigration shock. The increase in labor market tightness and decrease in vacancy duration must, therefore, also be linked to differences in market conditions.

Another relevant feature of our market is the possibility of employers extending private re-employment offers to workers they hired the previous round. This allows them to establish longer-term employment relationships with workers. A possibility is that workers do not cooperate with employers because of pro-sociality but rather profit from longer-term collaboration. In this view, and considering the strong wage changes we observe in the large shock market, we would expect that multi-period cooperation decreases. Yet, we do not observe this; if anything, the opposite: the number of re-employment offers and the average number of consecutive rounds employers and workers interact increases more in the large shock, though both differences are insignificant. In Appendix Table 2..6, we take a descriptive look at the distribution of employment relationships lasting one versus multiple periods. We see that multi-period contracts increased in the large immigration destination while they decreased in the small shock market.

Table 2.5.2: Other outcomes regression results

	(1) Pre-post small	(2) Pre-post large	(3) Double-difference
Panel A. Market outcomes			
Unfilled job offer	-0.061***	-0.288***	-0.226***
	(0.020)	(0.024)	(0.031)
Re-employment offer	0.015	0.115**	0.099
	(0.039)	(0.045)	(0.059)
Employment duration	0.111	0.416	0.308
	(0.177)	(0.264)	(0.313)
Panel B. Page times, in seconds			
Vacancy duration	-34.438***	-61.346***	-26.320*
	(7.933)	(11.911)	(13.823)
Effort decision	-3.050***	-2.685***	0.360
	(1.070)	(0.925)	(1.348)
Re-employment decision	-3.293**	-0.274	3.004*
	(1.297)	(1.050)	(1.617)
Session FE	Yes	Yes	Yes
Individuals	120	160	280
Jobs	2077	1972	4049

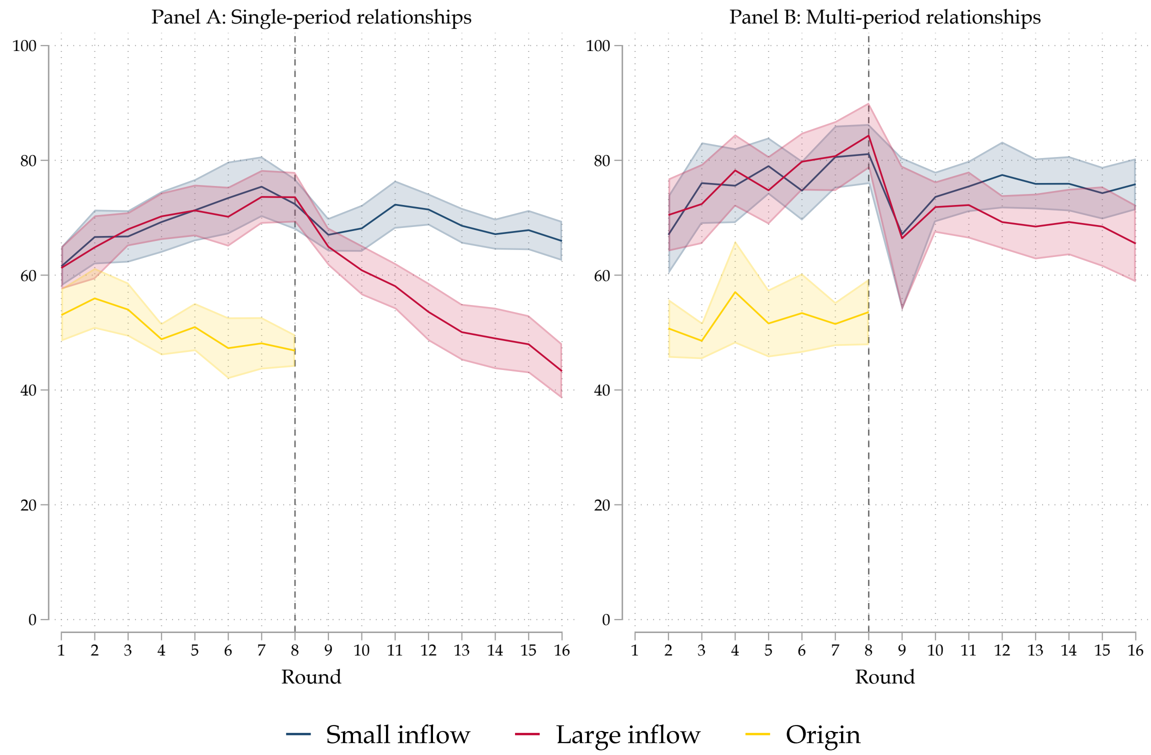
Source: own calculation. Results in columns (1) and (2) display estimates from the regression described in eq. 2.17, while columns (3) displays estimates from the regression described in eq. 2.18. Standard errors clustered at the round-market level in parentheses. Significance: * 0.10 ** 0.05 *** 0.01. Outcomes: Unfilled job offers are the number of job offers that were canceled or remained open until the end of the round. Re-employment offer is the number of additional job offers that employers send to workers they employed the previous round. Employment duration is, among all accepted jobs in a round, how many rounds the employer and workers have interacted. Vacancy duration is the average number of seconds a job offer was open. Effort decision is the average number of seconds the workers use to decide what effort to provide. Re-employment decision is the average number of seconds employers used to decide whether and what private job to offer to workers they employed in the previous round.

The previous analysis suggests that a potential adjustment channel is whether to engage in longer-term relationships. If multi-round employment relationships offer higher wages, the wage gap between destinations will be reduced, as we have seen that employment duration increased more in large immigration shock. As such, changes in longer-term cooperation alone cannot rationalize the main results on wages. Yet, it is interesting to assess whether wage adjustment occurred only in single-period contracts or whether employers reduced wages also in longer-term relationships. Therefore, in Figure 2.5.2, we distinguish the wage response of single-period and multi-period employment relationships.[58] Indeed, the wage response seems to be larger for single-period contracts. Longer-term relationships also experienced a wage drop, which was stronger in the large inflow market. However, the overall wage drop was smaller than for single-period

[58] Note that employers couldn't extend contracts spanning over the migration shock after round eight. This causes a large confidence interval in round nine of the multi-period relationships. If they formed one, it was by accident, since public contracts are anonymous.

contracts. The regression analysis in Appendix Table 2..5 confirms this view. The wage drop was significant for single- and multi-period relationships in the large shock market and insignificant for both after the small shock. The analysis also suggests an interesting insight into the effort response. In the small market, the negative effort response occurred more often in multi-period contracts, while in the large shock, the drop was primarily for single-period contracts. The fact that effort drops less in the large shock and particularly not in multi-period employment relationships indicates that large inflows induce job insecurity and a stronger demand for multi-period contracts, which prevents negative effort responses despite wage decreases. The extension of the theory dealing with endogenous reference points rationalizes such a mechanism within our theoretical framework.

Figure 2.5.2: Single- vs multi-period wage response



Source: Own calculations. The figures plot the mean and 95% confidence intervals of different in outcomes for each round. The yellow line is the origin market, the blue line is the destination that received one worker, and the red line is the destination that received five workers. The outcomes are in panel (a) the average accepted wage of a new working relationship, in panel (b) the average wage of an existing working relationship, that means the worker and employer already had an employment relationship in the previous round.

Next, we aim to understand individual decision-making better and relate choice differences across individuals to well-established behavioral traits. This allows us to link the results back to the theoretical model. A preregistered hypothesis was that the effort decision of workers depends on the wage they receive.[59] In Appendix Figure 2..7, we confirm this hypothesis visually, as the binscatter plots find a positive relationship between the wage the worker received and the share of workers choosing normal effort as well as a negative relationship between the wage and the share of workers shirking, that is providing low effort despite the job offer requiring normal effort. In columns (1) to (3) of Table 2.5.3, we probe whether this relationship holds conditional on an extensive set of fixed effects and basic worker's sociodemographic. We find that the positive relationship remains intact and, if anything, becomes stronger after introducing covariates. In the most demanding specification, we find that the probability of a worker choosing normal effort increases by 0.011 for each wage point, a large magnitude for which we can rule out null effects at high statistical precision.

Throughout the paper, we have assumed that the positive relationship between wages and effort is due to gift exchange (Akerlof, 1982). This means workers return effort for wages with a cooperative mindset rather than to gain utility from dynamic reputation considerations. Gift exchange is closely linked to the notion of positive reciprocity in the Rabin, 1993 sense that players choose kinder strategies against players who treat them well, or at least not wrong. We would, therefore, expect that workers who are more positively reciprocal will also return normal effort more often for higher wages. We test this, as prespecified, by including items from the Global Preference Survey (GPS) (Falk et al., 2018, 2023) in the ex-post questionnaire (see questionnaire in Appendix 2.6). We use their weights on the survey items to construct an index of positive reciprocity replicating the validated GPS measure. Later in this section, we will use similar indices for negative reciprocity and risk preference, also constructed to replicate the GPS. We then interact, as prespecified, with the measure of positive reciprocity with the wage and assess its impact on effort provision. As expected, the interaction is positive and statistically significant. This means that the share of workers providing normal effort increases more strongly among workers with high positive reciprocity scores. We confirm the hypothesis is also conditional on fixed effects and workers' controls.

[59] In reality, certainly, other factors will surely play a role as well. In the laboratory, however, workers knew almost nothing about their employers. Absent emotional relationships, the currency is the only social aspect left.

Table 2.5.3: Conditional wage-effort schedule

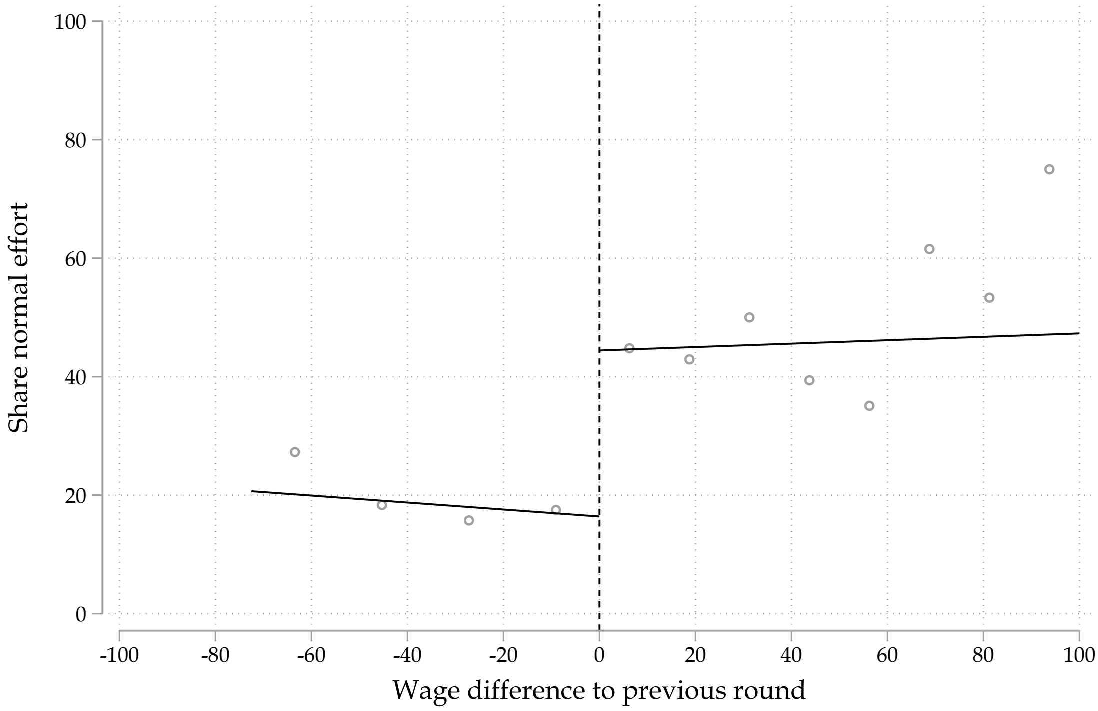
	(1)	(2)	(3)	(4)	(5)	(6)
Wage	0.007***	0.011***	0.011***	0.008***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Positive reciprocity				-0.178	-0.178	-0.186
				(0.133)	(0.133)	(0.130)
Wage × positive reciprocity				0.005**	0.005**	0.004**
				(0.002)	(0.002)	(0.002)
Observations	2786	2786	2786	2786	2786	2786
Session FE	Yes	Yes	Yes	Yes	Yes	Yes
Market FE	No	Yes	Yes	No	Yes	Yes
Round FE	No	Yes	Yes	No	Yes	Yes
Worker controls	No	No	Yes	No	No	Yes

Source: own calculations. Standard errors clustered at the market-round level in parentheses. Significance: * 0.10 ** 0.05 *** 0.01. The outcome is the effort the worker provided, taking the value of one for normal effort and zero for low effort. Worker controls include age, gender, and indicators for foreign-born, and completed university degree.

Readers may have noticed that we make a more specific functional form assumption about the wage-effort relationship in the theory section, namely that effort provision drops discretely when workers receive a wage cut relative to the reference point. We motivated this discrete drop with loss aversion, though, in principle, reciprocity could also motivate a discrete drop if the wage cut is seen as defecting behavior. However, one may wonder whether effort provision will not decrease proportionally to the size of the wage cut. Therefore, in Figure 2.5.3, we investigate how effort responds to wage changes between rounds. Specifically, we plot the share of workers providing normal effort against the wage change relative to the previous rounds' wage. The plot is rather remarkable in that it looks exactly like a regression discontinuity design with a threshold at zero. Below that threshold is a 29% drop in the share of workers providing normal effort. Also, in line with our theory, the slope of the effort-wage-change relationship below zero is rather flat. This means that most of the negative effort adjustment occurs as a discrete drop for any wage cut, and not as a continuous response depending on the wage loss. This confirms a central functional form assumption used in the theory section.[60]

[60] Longterm followers of the project will know that the theoretical assumption predated the empirical confirmation.

Figure 2.5.3: Effort discontinuity in wage changes



Source: own calculations. Accepted jobs from all rounds and markets used (N = 2,786). The X-axis is the difference between the current and the previous rounds' wage. The Y-axis is the share of workers choosing normal effort. If a worker was not employed in the previous round, they had a wage of zero. The results, available upon request, are very similar if these cases are excluded from the analysis. Linear fits are with polynomials of degree one. Automated bandwidth selection using polynomial regression with the IMSE-optimal evenly-spaced method. Implemented using the package *rdplot* from Calonico et al., 2017.

To link the discussion about the effort response to immigration-induced wage cuts back to behavioral traits, we would expect to find among the workers whose effort provision drops after the large migration shock an overrepresentation of negatively reciprocal workers, that is, workers who punish unacceptable behavior, even if they incur a material loss from it. In this case, the material loss is reneging on the possibility of receiving a re-employment offer, though the cost might be negligible if workers were unhappy with wages. Therefore, in Table 2..4, we conduct the prespecified test linking the negative reciprocity score of a worker, as measured by the GPS items, to their effort response to immigration. We implement this by interacting the continuous measure for negative reciprocity with the post-treatment dummy in each destination. We find some evidence that workers with a higher negative reciprocity score provide less effort, at least in the small shock market. Yet, we do not find that they responded differently to the immigration shock. This is tentative evidence that positive reciprocity is the more relevant behavioral trait.

Lastly, we want to explore the firm side of the market. Clearly, the firm side is less realistic than the worker side in our experimental design. Yet, it has a feature that may be important in practice: employers must decide whether to adjust their wage-setting strategy when market conditions change. They may offer lower wages to try to exploit workers. Reducing wages may increase profits but may prevent them from filling all jobs or induce recruited workers to deliver low effort. The decision thus entails an element of risk-taking. In Table 2.5.4, we explore this further and conduct a prespecified test to inquire whether employers with higher risk preference responded more to the immigration shock. For that, we use the GPS survey item for risk preference and interact it with the post-treatment dummy, similar to the previous analysis. The results show that employers with higher risk preferences offered high wages before immigration. This result makes sense since offering high wages in the hope that workers will reciprocate is somewhat risky. More interesting, though, is that these employers also reacted more strongly to the immigration shock. The coefficients on the interaction between risk preference and post-treatment are large and significant throughout, meaning that more risk-taking employers offered lower wages after the immigration shock. It seems it takes courage to respond to changes in market conditions.

Table 2.5.4: Employer risk preference and wage offer response to immigration

	Small shock		Large shock	
	(1)	(2)	(3)	(4)
Post	7.271***	7.386***	1.299	0.616
	(1.963)	(1.948)	(2.786)	(2.871)
Risk preference	12.259***	16.997***	7.353**	4.868
	(2.075)	(3.566)	(3.318)	(3.359)
Post × risk preference	-10.298***	-10.614***	-13.807***	-12.600***
	(2.860)	(2.787)	(3.793)	(3.922)
Observations	2077	2077	1972	1972
Session FE	Yes	Yes	Yes	Yes
Employer controls	No	Yes	No	Yes

Source: own calculations. Standard errors clustered at the market-round level in parentheses. Significance: * 0.10 ** 0.05 *** 0.01. The outcome is the wage offered by the employer. Employer controls include age, gender, and indicators for foreign-born, and completed university degree.

In summary, the findings demonstrate that immigration impacts market conditions by increasing market tightness, affecting vacancy duration, and influencing long-term employment relationships. Behavioral traits, such as positive reciprocity and risk aversion, are important determinants of individual responses. Furthermore, we could provide evidence for a key assumption in our model, namely that effort provision exhibits a discrete drop below the reference wage. This evidence is also useful for other models using this assumption, for example, Kaur, 2019.

2.5.3 Robustness

In this section, we conduct two types of robustness checks. The first is to exclude decisions from workers and employers with low cognitive ability scores. The second is to exclude the data from the first and last three rounds. The qualitative findings of the main text are not affected by either choice.

First, we conduct the prespecified robustness check of dropping observations from individuals with a cognitive inability score in the top tenth percentile. The idea is to reduce the noise by excluding the choices of individuals who are less likely to have understood the decision structure. As preregistered, the cognitive inability score is calculated using the time required to read the instructions and the number of tries to answer all three questions of the comprehension quiz correctly:

$$Cogn = \mathbb{I}[time < 20p] + \mathbb{I}[time > 80p] + Tries(Q1) + Tries(Q2) + Tries(Q3) \tag{2.19}$$

where $\mathbb{I}[time < 20p]$ and $\mathbb{I}[time > 80p]$ are indicators for individuals who were in the lowest or highest 20^{th} percentiles of time taken to read the instructions and $Tries(Q1)$ through $Tries(Q3)$ is the number of tries the user needed to provide a correct answer to comprehension quiz question one to three, respectively. Note that the number of tries is discrete and will impact the final score more than the page times. This was intended since incorrect answers to the comprehension quiz are more informative than differences in reading times. Individuals in the top tenth percentile of cognitive inability will have accumulated ten or more points in this measure. Dropping them means removing 33 out of the 330 players, 26 workers and seven employers. For job-level analyses, it means reducing the total number of job offers from 4,967 to 4,271.

Appendix Figure 2..8 and Appendix Table 2..7 replicate the results from Figure 2.5.1 and Table 2.5.1 in the main results section, but using the sample which excludes decisions from individuals with high cognitive inability score. The results in the figure seem almost identical to those in the main text. Perhaps the only difference is in panel (b) concerning effort provision, where the gap between the small and large shock destination appears smaller. The results in Table 2..7 are very similar to the main results overall. It seems the response on wages and worker profits is in the large market is slightly smaller, and that the wage gain of migrants is even larger with this sample definition. Yet all coefficients remain in the same ballpark and retain statistical significance. Perhaps the only difference is that the native effort reduction in the large shock is less pronounced and only significant at the 90% level. Yet, even here, the coefficient is similar to the main results. Overall, the robustness check shows that our findings are not driven by the decisions of individuals who did not understand the experiment well.

Another important robustness check is to exclude the first and last rounds from the analysis. These rounds may be different in particular ways. During the first round, players are still learning the game and experimenting with different strategies. In other words, the market is trying to reach equilibrium. The last round may be affected by endgame considerations, leading workers to shirk more often. This should not concern the double-difference regressions, as we would not expect these patterns to differ across markets. Yet, we rely also on simple pre-post comparisons within each market, and here the first and last rounds might introduce bias if they distort the outcomes in meaningfully different ways.[61] Therefore, in Appendix Table 2..8 we replicate the results of Table 2.5.1 but excluding the first and last three rounds. The findings, available upon request, are qualitatively the same when excluding only one or two rounds at each tail. However, we prefer to show the most demanding version of this robustness check, which excludes almost half of our sample. In practice, this means keeping all individuals but excluding

[61] For example, if we expect that initially, workers shirk too much because they don't yet know that workers and firms can both benefit by achieving the high-wage-high-effort equilibrium, and they will also shirk too often in the last round because it's over afterward, this would not be a substantial threat to our identification strategy since the distortion moves in the same direction and will be canceled out. However, suppose workers initially provide too often higher effort and, in the last round, too often shirk. In that case, both distortions reinforce each other, which will cause the pre-post comparisons to be biased.

1,802 of the available 4,967 job offers.

The results in Appendix Table 2..8 are reassuring because the magnitude and significance of the main coefficients remain very close to the results presented in the main text. It seems that the wage response to immigration is more pronounced in the small shock destination with this sample. However, the wage drop in the large shock destination is also stronger, so the double difference remains large and significant. The increase in firm' and decrease in workers' profits seem more pronounced in the large shock market under the new sample definition. However, we doubt this difference would be significant in hypothesis tests. Overall, the results remain very stable when excluding the first and last rounds from the analysis. This confirms our hypothesis that round effects are likely not a driver of our findings.

2.6 Conclusion

This paper investigates how immigration impacts wages, effort, and profits in markets where efficiency wage considerations influence wage-setting. Efficiency-wage considerations complicate firms' optimal responses since wages serve not only to attract labor but also to induce effort at work. We develop a model to describe this process and test its implications in a controlled laboratory experiment. By constructing conditions with a high and low inflow of migrant workers, we isolate the behavioral and economic responses to increased labor supply in markets with reference-dependent wage expectations.

The findings show that wages dropped in both markets following the immigration shock, recovering only in the small immigration destination. Migrants, coming from a low-income market, continued to provide high effort despite wage declines, whereas incumbent workers, facing wage cuts, reduced their effort substantially. This contrast led firms in the large inflow market to maintain lower wages, gaining from the differential effort response between incumbents and migrants. Firm profits increased notably only in the large inflow market, where the influx of migrants weakened incumbents' wage bargaining power, illustrating the nuanced efficiency wage effect on market dynamics in high-immigration settings.

Behavioral mechanisms play a critical role in these adjustments. Positive reciprocity among workers influenced effort decisions, while risk preferences among firms impacted wage offers following the immigration shock. Markets that received large inflows saw higher tightness and shorter vacancy duration as employers leveraged increased competition to maintain lower wages.

It is important to note that we found these results despite setting up a market in which many of the theoretical arguments for why immigration might benefit natives were not present. Migrants and natives were perfect substitutes in production, and there were no other spillovers on demand, innovation, and firm creation that the previous literature found. The theory section showed that the general results also hold when introducing imperfect substitutability. These mechanisms are undoubtedly relevant in practice. We have shown that even without them, it is possible to develop a theoretical model and a form of real market interaction where the wage effects of immigration are negligible at moderate levels of immigration.

We have shown partial wage stickiness may arise endogenously due to the principal-agent relationship inherent to labor contracts. This is a plausible explanation based on prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), a well-established and empirically validated theory of human choice. Yet, there may well be other mechanisms that also favor such a response. More research into these mechanisms, also discerning between different behavioral microfondations, would be interesting to complement our research. More broadly, we think that the analysis of how labor markets adjust to immigration can benefit considerably from the insights of behavioral economics. This applies to wage effects on natives and other important topics, such as immigrant integration or attitudes toward immigration.

Appendix

Table 2..1: Nonparametric test between immigration and recession framing

	Observations	Z-value	Exact P-value
Self-reported	826	0.383	0.7023
Guessing others	826	1.205	0.2284

Source: own calculations. Results from a Mann–Whitney U-test for differences between immigration and recession framing. Question is described in the main text. Self-reported are the answers respondents gave, guessing others are the values from the incentivized belief elicitation. Exact p-value is the p-value from a Fisher randomization inference procedure.

Table 2..2: Parametric test between immigration and recession framing

	Observations	Difference	P-value
Self-reported	826	-0.022	0.507
Guessing others	826	-0.066	0.060

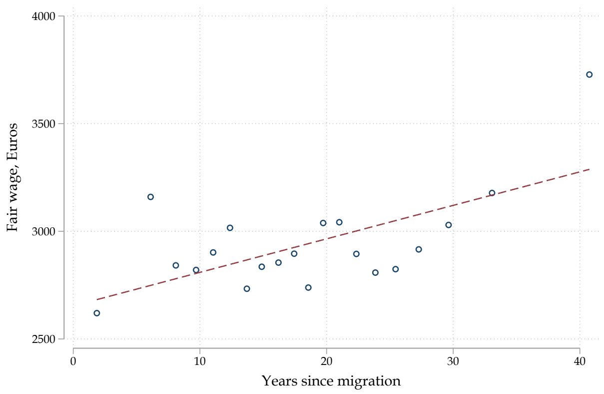
Source: own calculations. The outcome takes the value of one for any reduction in effort. Difference = mean(recession) - mean(immigration). P-value is from a two-sided T-test for differences between immigration and recession framing. Question is described in the main text. Self-reported are the answers respondents gave, guessing others are the values from the incentivized belief elicitation.

Table 2..3: Quantile regression results

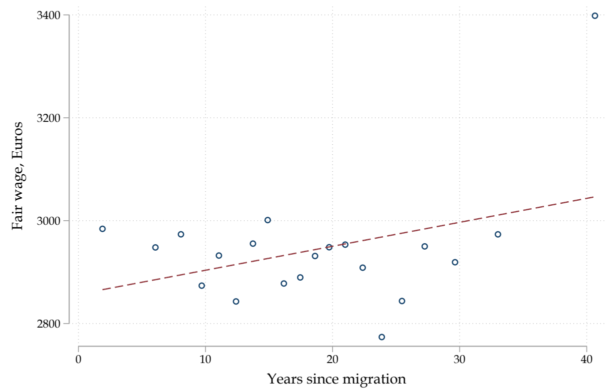
	Full sample	Preferred spec.
Above median	-0.136**	-0.094
	(0.044)	(0.327)
Constant	0.000	0.090
	(0.034)	(0.141)
N	521	28

Heteroscedasticity-robust standard errors in parentheses. Significance: * 0.10 ** 0.05 *** 0.01. The dependent variable is the estimate of semi-elasticity of immigration on wages. Above median is an indicator taking the value of one if the average migrant inflow as share of population was above the median during the study period.

Figure 2..1: Fair wage expectations and years since migration



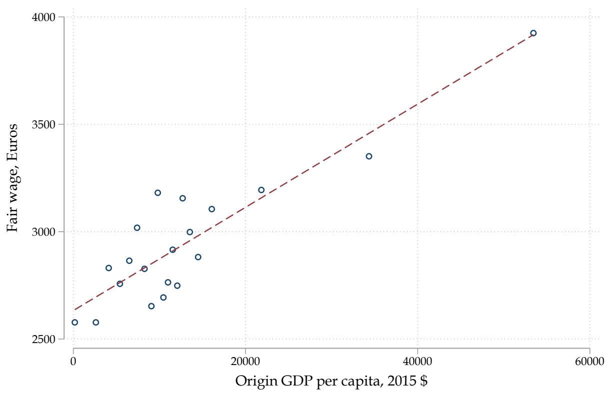
(a) Controls: age, age squared, education



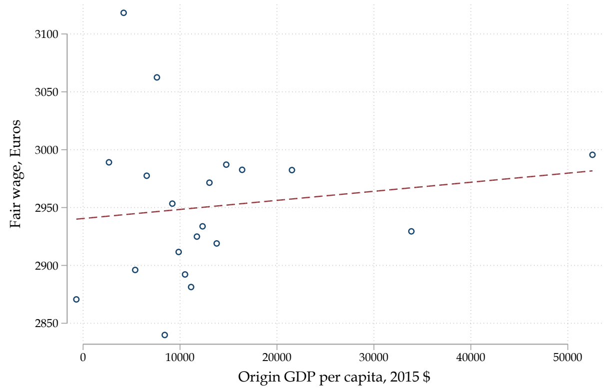
(b) Controls: age, age squared, education, own wage

Source: Source: Data from SOEP, 2021 and The World Bank, 2024a, own calculations. Sample restricted to migrants. Y-axis is the fair wage, proxied through fair income from question plh0138 "*How high would your gross income have to be in order to be just?*" in gross nominal Euro. The X-axis is the origin-country nominal GDP per capita in Panel (a) and the years since migration in Panel (b). Estimation done using *binscatter2* from Droste, 2024.

Figure 2..2: Fair wage expectations and origin GDP per capita



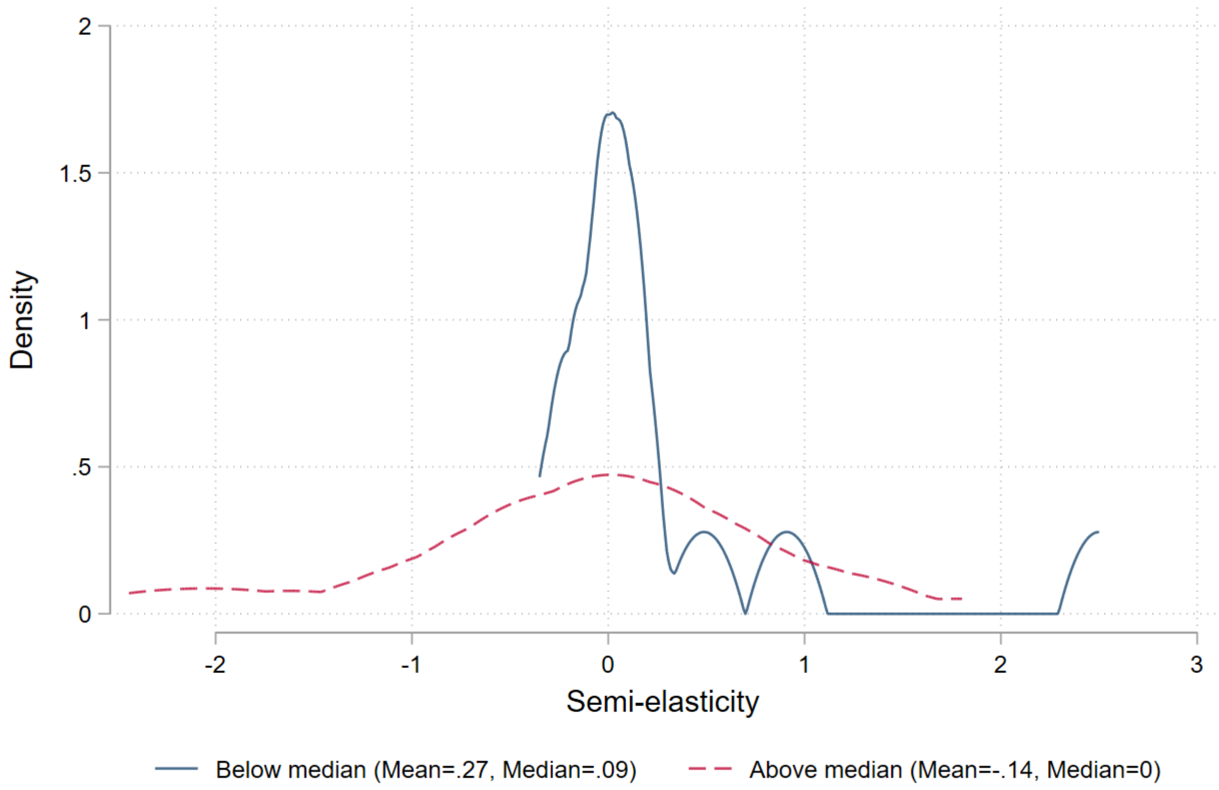
(a) Controls: age, age squared, own wage



(b) Controls: age, age squared, education, own wage

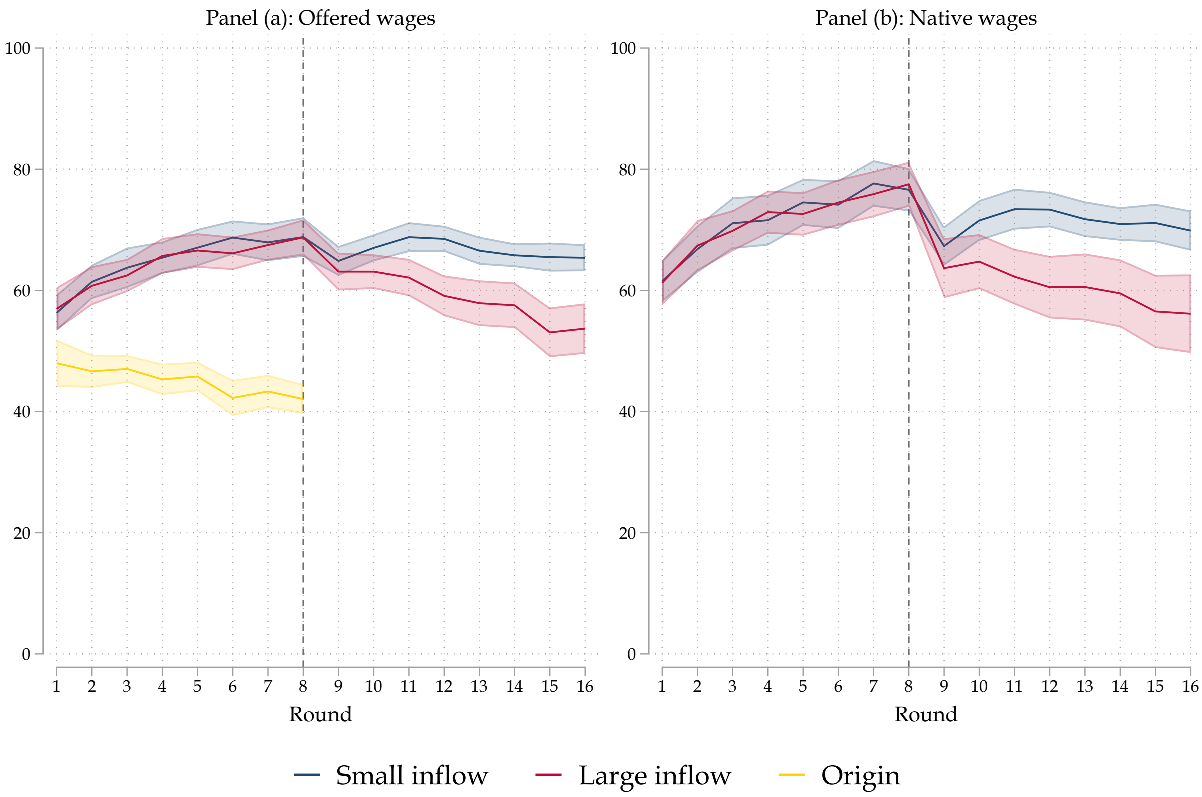
Source: Source: Data from SOEP, 2021 and The World Bank, 2024a, own calculations. Sample restricted to migrants. Y-axis is the fair wage, proxied through fair income from question plh0138 "*How high would your gross income have to be in order to be just?*" in gross nominal Euro. The X-axis is the origin-country nominal GDP per capita in Panel (a) and the years since migration in Panel (b). Estimation done using *binscatter2* from Droste, 2024.

Figure 2..3: Distribution of wage semi-elasticities



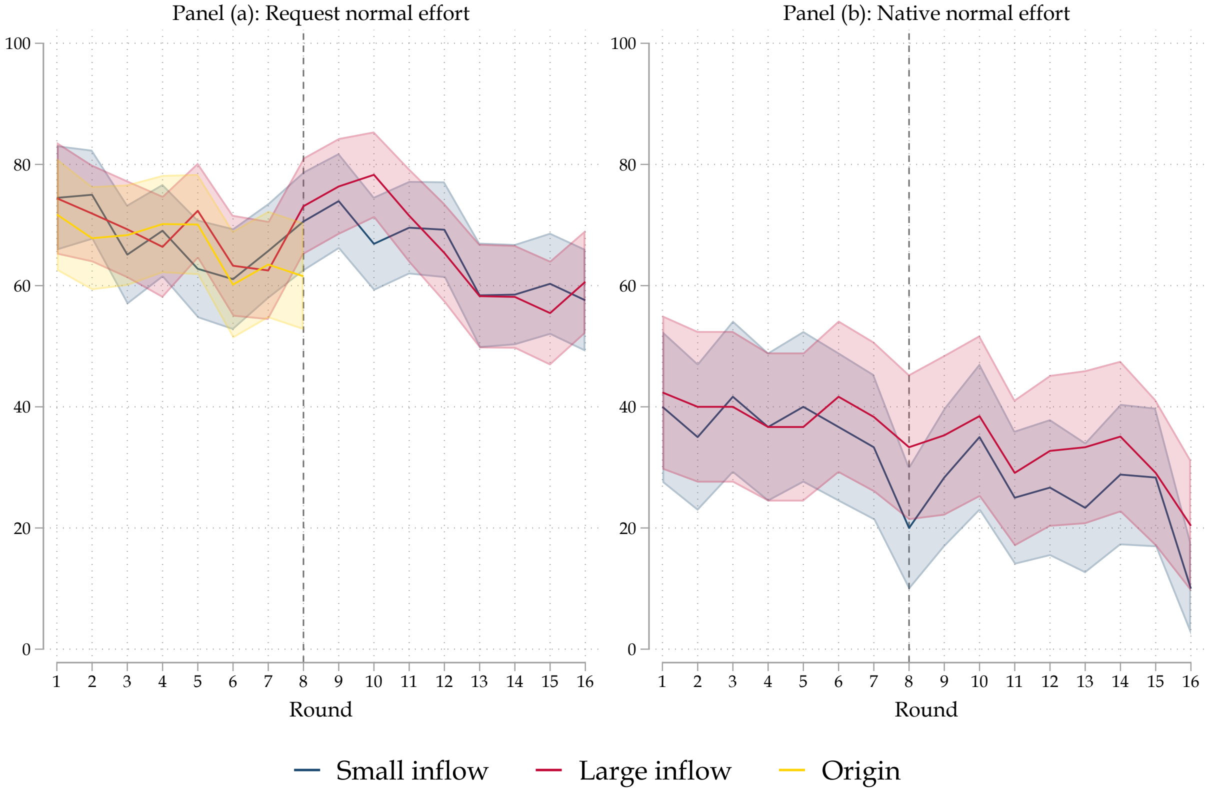
Source: Data from Foged et al., 2022 and Standaert and Rayp, 2022. N = 28. This figure keeps only the authors' preferred estimate for each study in our sample. The X-asis gives the semi-elasticity of immigration on natives wages, the Y-axis the density. The blue line is the density of wage estimates in studies where the average migration inflow during the sample period was below median. The red line is the density of wage estimates in studies where the average migration inflow during the sample period was above median.

Figure 2..4: Further wage outcomes by market and round



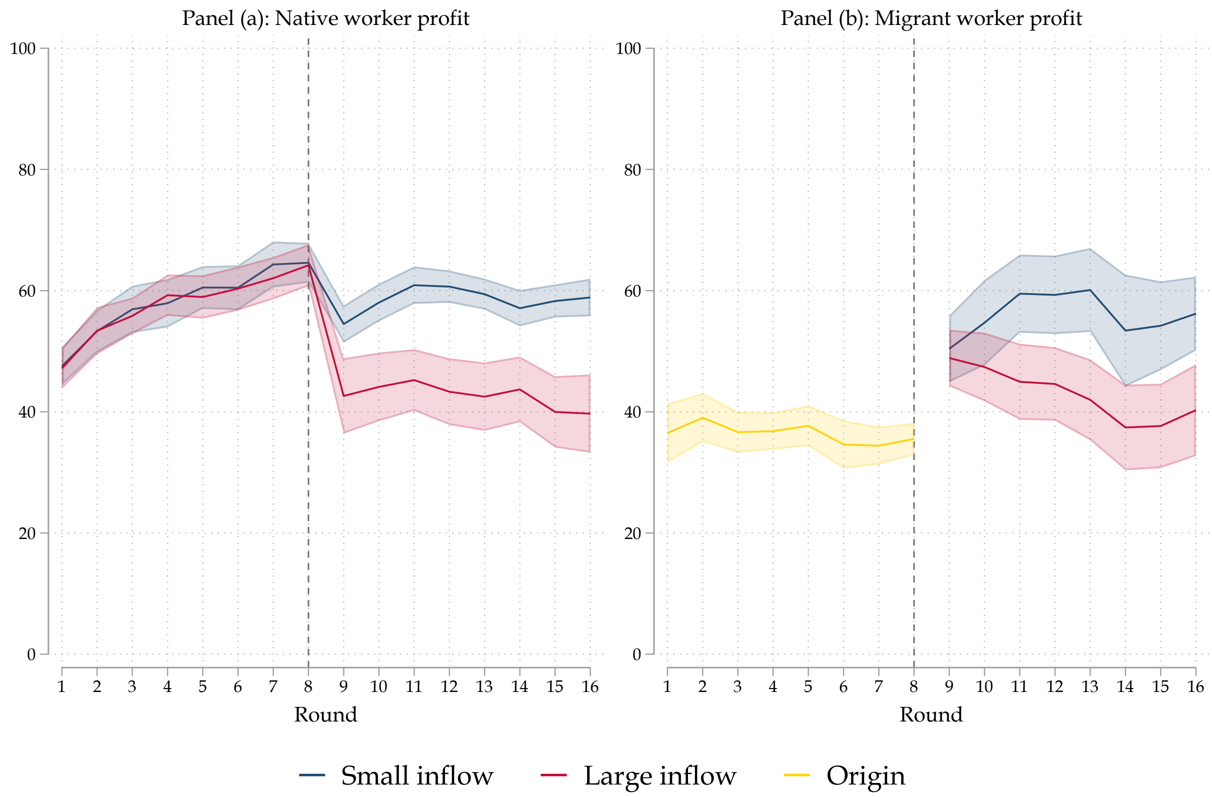
Source: Own calculations. The figures plot the mean and 95% confidence intervals of different in outcomes for each round. The yellow line is the origin market, the blue line is the destination that received one worker, and the red line is the destination that received five workers. The outcomes are in panel (a) the average offered wage and, in panel (b), the average accepted wage of incumbent workers.

Figure 2..5: Further effort outcomes by market and round



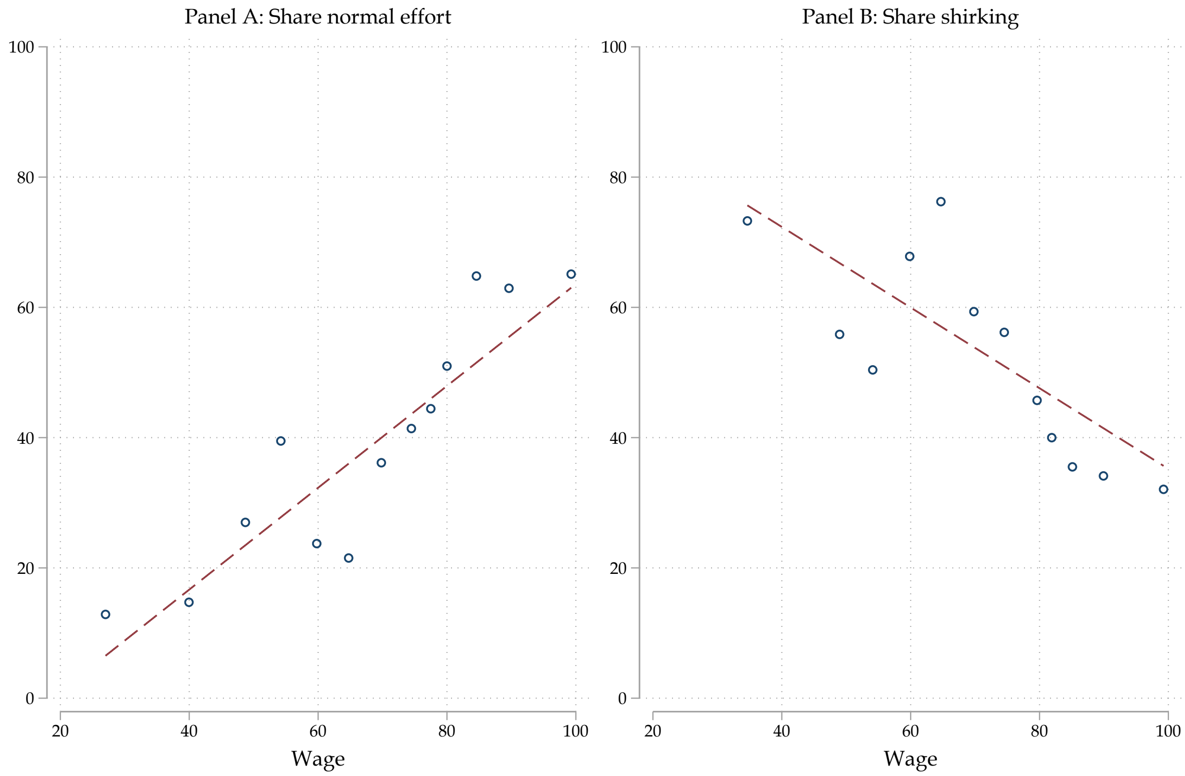
Source: Own calculations. The figures plot the mean and 95% confidence intervals of different in outcomes for each round. The yellow line is the origin market, the blue line is the destination that received one worker, and the red line is the destination that received five workers. The outcomes are in panel (a) the share of job requesting normal effort and, in panel (b), the share of workers shirking, that is provide low effort despite the job requesting normal effort.

Figure 2..6: Further worker profit outcomes by market and round



Source: Own calculations. The figures plot the mean and 95% confidence intervals of different in outcomes for each round. The yellow line is the origin market, the blue line is the destination that received one worker, and the red line is the destination that received five workers. The outcomes are in panel (a) the average per-round profit of native workers and, in panel (b), the average per-round profit of migrant workers.

Figure 2..7: Wage-effort schedule



Source: own calculations. Accepted jobs from all rounds and markets used (N = 2,786). The X-axis is the wage workers received. The Y-axis is in panel A the share of workers choosing normal effort, and in panel B the share of worker who shirk, that is they choose low effort despite the job offer requesting for normal effort. Estimation done using *binscatter2* from Droste, 2024.

Table 2..4: Negative reciprocity and effort response

	Small shock		Large shock	
	(1)	(2)	(3)	(4)
Post	-0.139*	-0.137*	-0.037	-0.030
	(0.072)	(0.072)	(0.051)	(0.050)
Negative reciprocity	-0.197*	-0.238**	0.018	0.060
	(0.100)	(0.098)	(0.061)	(0.064)
Post $\times$ neg. reciprocity	0.144	0.145	0.022	-0.008
	(0.121)	(0.121)	(0.080)	(0.080)
Observations	1039	1039	1270	1270
Session FE	Yes	Yes	Yes	Yes
Worker controls	No	Yes	No	Yes

Source: own calculations. Standard errors clustered at the market-round level in parentheses. Significance: * 0.10 ** 0.05 *** 0.01. The outcome is the effort the worker provided, taking the value of one for normal effort and zero for low effort. Worker controls include age, gender, and indicators for foreign-born, and completed university degree.

Table 2..5: Main outcome regression for single- vs. multi-period contract

	(1) Pre-post small	(2) Pre-post large	(3) Double-difference
Panel A. Single-period employment relationship			
Wages	0.174	-13.901***	-12.950***
	(2.162)	(3.249)	(3.700)
Native wages	0.305	-13.898***	-13.351***
	(2.147)	(2.911)	(3.417)
Normal effort	-5.105	-10.829	-4.592
	(4.998)	(7.012)	(8.581)
Native normal effort	-7.471	-13.575**	-4.802
	(4.753)	(5.477)	(7.191)
Panel B. Multi-period employment relationship			
Wages	-2.291	-7.456***	-7.124**
	(1.537)	(2.087)	(2.627)
Native wages	-2.011	-8.453***	-8.014***
	(1.513)	(2.150)	(2.554)
Normal effort	-10.757*	3.961	12.063
	(5.802)	(6.176)	(8.366)
Native normal effort	-12.108*	-1.622	8.345
	(5.870)	(5.914)	(8.148)
Session FE	Yes	Yes	Yes

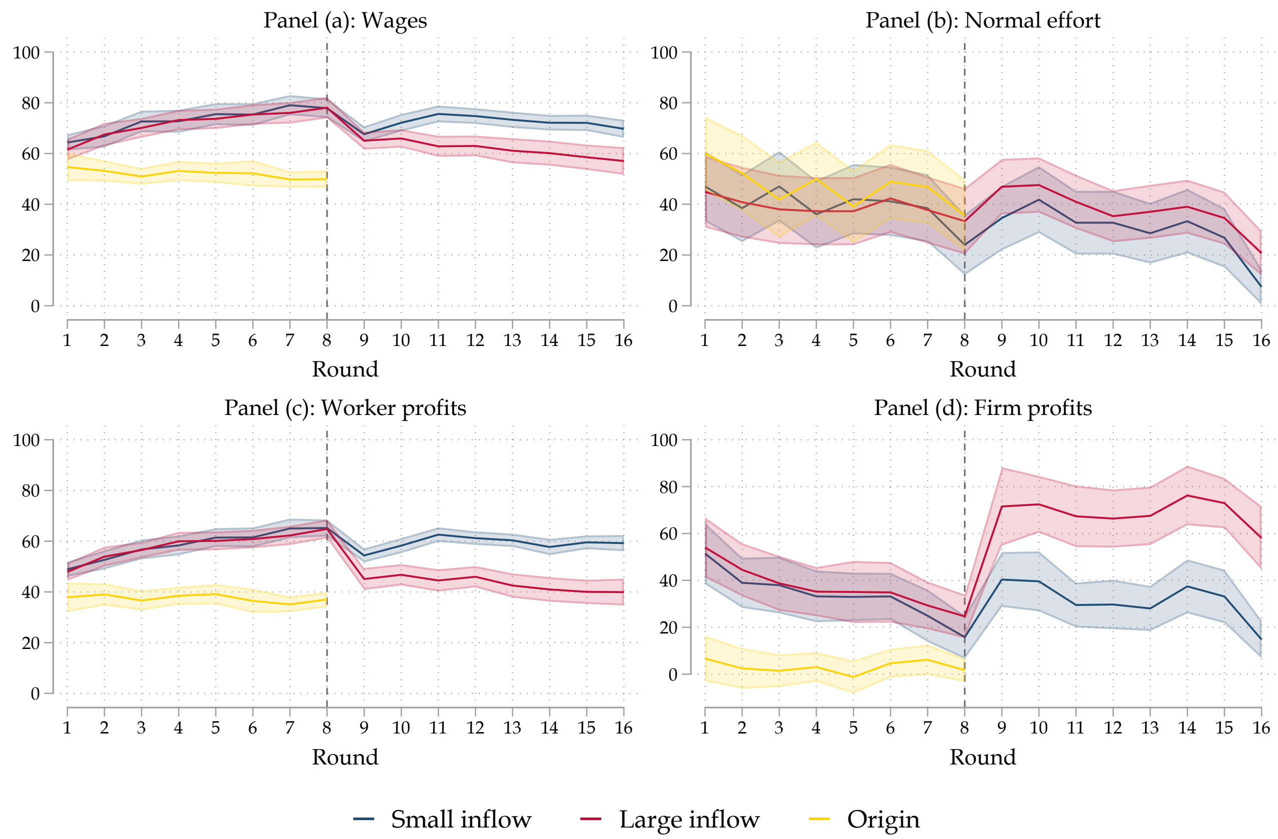
Source: own calculation. Results in columns (1) and (2) display estimates from the regression described in eq. 2.17, while columns (3) displays estimates from the regression described in eq. 2.18. Standard errors clustered at the round-market level in parentheses. Significance: * 0.10 ** 0.05 *** 0.01. Outcomes: Wages is the pay in experimental currency of all accepted job offers. Native wages in experimental currency of all job offers that incumbent workers accepted. Normal effort is an indicator taking the value of one if a worker provided normal effort, native normal effort conditions on the workers being an incumbent, native normal effort is the same measure, but considering only the jobs that were filled by incumbents.

Table 2..6: Frequency of single- vs. multi-period employment

	Small shock		Large shock	
	Pre	Post	Pre	Post
Single-period	61.3%	62.1%	65.8%	50.5%
Multi-period	12.8%	11.1%	13.7%	39.4%
No employment	0.0%	0.2%	0.2%	10.1%

Source: own calculation. A job is a multi-period if the employer and worker also had an employment relationship is the previous round. Pre is the first eight rounds, post is round nine and later when the immigration episode has already occurred. The shares are calculated relative to the total available workers. In the pre-shock that is 6 workers $\times 8$ rounds $\times 10 = 480$. By the same argument, the post is 560 in the small shock and 880. The calculation only use the two destination markets, and not the origin market.

Figure 2..8: Robustness I of average outcomes by market and round



Source: Own calculations. The figures plot the mean and 95% confidence intervals of different in outcomes for each round. The yellow line is the origin market, the blue line is the destination that received one worker, and the red line is the destination that received five workers. The outcomes are in panel (a) the average accepted wage, in panel (b) the share of workers providing normal effort, in panel (c) the average per-round profit of employers, and in panel (d) the average per-round profit of workers in the market.

Table 2..7: Robustness check I of main regression results

	(1) Pre-post small	(2) Pre-post large	(3) Double-difference
Panel A. Job-level outcomes			
Wages	-1.125	-10.282***	-9.256***
	(1.988)	(2.081)	(2.830)
Native wages	-1.335	-10.255***	-8.979***
	(1.947)	(1.935)	(2.717)
Offered wages	1.508	-5.217**	-6.953***
	(1.432)	(1.774)	(2.211)
Normal effort	-0.095**	-0.013	0.085
	(0.042)	(0.032)	(0.052)
Native normal effort	-0.112**	-0.050*	0.063
	(0.039)	(0.024)	(0.046)
Request normal effort	-0.036	-0.013	0.006
	(0.028)	(0.038)	(0.046)
Panel B. Individual outcomes			
Firm profit	-1.976	31.997***	33.973***
	(4.577)	(3.655)	(5.746)
Worker profit	2.367	-5.942*	-8.309**
	(2.227)	(2.944)	(3.299)
Native worker profit	0.255	-14.632***	-14.887***
	(2.146)	(1.967)	(2.857)
Migrant worker profit	25.175***	5.024**	-20.151***
	(2.414)	(1.967)	(3.041)
Session FE	Yes	Yes	Yes
Individuals	106	143	249
Accepted jobs	847	1063	1910
Jobs	1801	1707	3508

Source: own calculation. Sample excludes 33 individuals which were in the top tenth percentile of cognitive inability score. Results in columns (1) and (2) display estimates from the regression described in eq. 2.17, while columns (3) displays estimates from the regression described in eq. 2.18. Standard errors clustered at the round-market level in parentheses. Significance: * 0.10 ** 0.05 *** 0.01. Outcomes: Wages is the pay in experimental currency of all accepted job offers. Native wages in experimental currency of all job offers that incumbent workers accepted. Offered wages is the pay of all job offers, including open and cancelled. Normal effort is an indicator taking the value of one if a worker provided normal effort, native normal effort conditions on the workers being an incumbent. Shirking takes the value of one if a worker provided low effort conditional on the job offer requiring normal effort. Request normal effort take the value of one if any job offer requested normal effort, zero if it requested low effort.

Proof of Proposition 1

We start by studying the impact of w_j on the firm's probability of attracting a worker, $P_j\left(w_j, \mathbf{w}_{-j}\right) = \frac{\exp(\alpha w_j)}{\sum_{k=1}^{J}\exp(\alpha w_k)}$. Using the quotient rule, this derivative of P_j is:

$$\frac{\partial P_j}{\partial w_j} = \frac{\alpha \exp(\alpha w_j) D - \exp(\alpha w_j)\alpha \exp(\alpha w_j)}{D^2},$$

Table 2..8: Robustness check II of main regression results

Source: own calculation. Sample excludes 33 individuals which were in the top tenth percentile of cognitive inability score. Results in columns (1) and (2) display estimates from the regression described in eq. 2.17, while columns (3) displays estimates from the regression described in eq. 2.18. Standard errors clustered at the round-market level in parentheses. Significance: * 0.10 ** 0.05 *** 0.01. Outcomes: Wages is the pay in experimental currency of all accepted job offers. Native wages in experimental currency of all job offers that incumbent workers accepted. Offered wages is the pay of all job offers, including open and cancelled. Normal effort is an indicator taking the value of one if a worker provided normal effort, native normal effort conditions on the workers being an incumbent. Shirking takes the value of one if a worker provided low effort conditional on the job offer requiring normal effort. Request normal effort take the value of one if any job offer requested normal effort, zero if it requested low effort.

where $D = \sum_{k=1}^{J} \exp(\alpha w_k)$. This simplifies to:

$$\frac{\partial P_j}{\partial w_j} = \frac{\alpha(D - \exp(\alpha w_j))\exp(\alpha w_j)}{D^2} = \alpha \frac{(D - \exp(\alpha w_j))}{D} \frac{\exp(\alpha w_j)}{D}.$$

Recognizing that $P_{ij} = \frac{\exp(\alpha w_j)}{D}$, we obtain:

$$\frac{\partial P_{ij}}{\partial w_j} = \alpha(1 - P_j)P_j.$$

Now, remember that the firm's first-order condition is

$$\frac{\partial \Pi_j\left(w_j, \mathbf{w}_{-j}\right)}{\partial w_j} = n\frac{\partial P_j}{\partial w_j} p\Omega\theta L_j^{\theta-1} - n\frac{\partial P_j}{\partial w_j} w_j - nP_j = 0.$$

Substituting for $\frac{\partial P_{ij}}{\partial w_j}$ and L_j, and after some simplifications, we can write

$$\frac{\partial \Pi_j\left(w_j, \mathbf{w}_{-j}\right)}{\partial w_j} = 0 \iff \alpha(1 - P_j)\left[p\Omega\theta n^{\theta-1}P_j^{\theta-1} - w_j\right] - 1 = 0.$$

Using symmetry, $w_j = w^*$ for all j and $P_j = \frac{1}{J}$, the first-order condition further simplifies to the symmetric equilibrium condition:

$$\alpha(1 - \frac{1}{J})\left[p\Omega\theta\left(\frac{n}{J}\right)^{\theta-1} - w^*\right] - 1 = 0.$$

Solving for w^*, we obtain

$$w^* = \theta p\Omega\left(\frac{n}{J}\right)^{\theta-1} - \frac{1}{\alpha}\frac{J}{J-1}.$$

Equilibrium firm profits are thus:

$$
\begin{aligned}
\Pi^* &= p\Omega\left(\frac{n}{J}\right)^{\theta}-\left(\frac{n}{J}\right)w^* \\
&= p\Omega\left(\frac{n}{J}\right)^{\theta}-\left(\frac{n}{J}\right)\theta p\Omega\left(\frac{n}{J}\right)^{\theta-1}+\left(\frac{n}{J}\right)\frac{1}{\alpha}\frac{J}{J-1} \\
&= (1-\theta)\,p\Omega\left(\frac{n}{J}\right)^{\theta}+\frac{n}{\alpha}\frac{1}{J-1}.
\end{aligned}
$$

Proof of Proposition 2

We have shown that there are two symmetric equilibrium candidates, $(w_c, \bar{e})$ and $(w_r, 1)$. We need to prove under which conditions firms obtain higher profits from each candidate. The comparison of firms' profits under the two strategies is $\Pi^*(w_r, 1) \lesseqgtr \Pi^*(w_c, \bar{e})$, which, using $L_j = \frac{n+n_I}{J}$ (2.1), (2.3) and (2.7), can be rewritten as:

$$
p\Omega\left[\left(\frac{n+n_I}{J}\right)^{\theta}-\frac{n+n_I}{n}\theta\left(\frac{n}{J}\right)^{\theta}\right]+\frac{n+n_I}{\alpha}\frac{1}{J-1}\lesseqgtr(1-\theta)\,p\Omega\left((I+(1-I)\,\lambda)\,\frac{n+n_I}{J}\right)^{\theta}+\frac{n+n_I}{\alpha}\frac{1}{J-}
$$

The markdown component of profits cancels out, and after factoring out $p\Omega$ on each side of the equality sign, one obtains

$$
\left[\left(\frac{n+n_I}{J}\right)^{\theta}-\frac{n+n_I}{n}\theta\left(\frac{n}{J}\right)^{\theta}\right]\lesseqgtr(1-\theta)\left((I+(1-I)\,\lambda)\,\frac{n+n_I}{J}\right)^{\theta}.
$$

After factoring out $\left(\frac{1}{J}\right)^{\theta}$ on each side of the equality sign and isolating $\frac{n+n_I}{n}$, we obtain

$$
\left[\left(\frac{n+n_I}{n}\right)^{\theta}-\frac{n+n_I}{n}\theta\right]\lesseqgtr(1-\theta)\left((I+(1-I)\,\lambda)\,\frac{n+n_I}{n}\right)^{\theta}.
$$

Dividing both sides by $\left(\frac{n+n_I}{n}\right)^{\theta}$, and using $\frac{n+n_I}{n}=(1-I)^{-1}$:

$$
1-\theta\,(1-I)^{-(1-\theta)}\lesseqgtr(1-\theta)\,(I+(1-I)\,\lambda)^{\theta}.
$$

Note that the left-hand side (LHS), which relates to $\Pi^*(w_r, 1)$, is strictly decreasing in I and is equal to $1-\theta$ for $I=0$ and tends to $-\infty$ for $I=1$. As per the right-hand side

(RHS), which relates to $\Pi^*(w_r, 1)$, it is strictly increasing in I and equal to $(1-\theta)\lambda^\theta$ for $I = 0$ and to $1-\theta$ for $I = 1$. Hence, when $I = 0$, the LHS is greater than the RHS ; as I increases the LHS decreases and the RHS increases ; and when $I = 1$, the LHS is smaller than the RHS. Hence, there is a unique immigration level $\bar{I}$ such that

$$1 - \theta (1 - I)^{-(1-\theta)} = (1-\theta)\left(\left(\bar{I} + \left(1 - \bar{I}\right)\lambda\right)\right)^\theta \iff \Pi^*(w_r, 1) = \Pi^*(w_c, \bar{e}).$$

Finally, when $I < \bar{I}$, the LHS is larger than the RHS, which implies that $\Pi^*(w_r, 1) > \Pi^*(w_c, \bar{e})$. Conversely, when $I > \bar{I}$, the LHS is smaller than the RHS, which implies that $\Pi^*(w_r, 1) < \Pi^*(w_c, \bar{e})$.

Extension: Imperfect Substitutability between Migrants and Natives

In this extension, we use a Constant Elasticity of Substitution (CES) production function to allow for imperfect substitutability between immigrants and natives. We show that when the high effort of native workers cannot be easily replaced by immigrant labor, firms are more inclined to maintain higher wages to preserve that high effort level. Hence, compared to the baseline model, a larger influx of immigrants is needed before it becomes advantageous to switch to competitive wages.

The CES production function is given by:

$$Y_j = \Omega\left[(L_{Nj}e_N)^\rho + (L_{Ij}e_I)^\rho\right]^{\frac{\theta}{\rho}},$$

where $L_{Nj} = (1-I)L_j$ and $L_{Ij} = IL_j$ are the labor inputs from natives and immigrants respectively, e_N and e_I are their respective efforts, and $\rho = \frac{\sigma - 1}{\sigma} \in]-\infty, 1]$ and $\sigma \geq 0$ represents the elasticity of substitution between native and immigrant labor. The total labor input is therefore $L_j\left[((1-I)e_N)^\rho + (Ie_I)^\rho\right]^{\frac{1}{\rho}}$, which is greater than the baseline labor input $L_j E[e^*]$ as soon as immigrants and natives are imperfect substitutes (i.e., when ρ is smaller than 1). The firm's profit function is:

$$\Pi_j = p\Omega L_j^\theta\left[((1-I)e_N)^\rho + (Ie_I)^\rho\right]^{\frac{\theta}{\rho}} - L_j w_j.$$

The first-order condition with respect to w_j is:

$$\frac{\partial \Pi_j}{\partial w_j} = \frac{\partial L_j}{\partial w_j}\left[p\Omega\theta L_j^{\theta-1}\left[\left((1-I)\,e_N\right)^{\rho}+\left(Ie_I\right)^{\rho}\right]^{\frac{\theta}{\rho}}-w_j\right]-L_j=0.$$

Unlike the baseline, the inflow of immigrants has an ambiguous effect on the marginal productivity of labor. Indeed, while marginal productivity is decreasing via $L_j^{\theta-1}$, it is positively impacted by immigrant-native complementarity. If migrants and natives exhibit a high level of complementarity (i.e., ρ is low) and marginal productivity does not decrease too quickly in the number of workers (i.e. if θ is close to 1), the net effect of the inflow of immigrants on the competitive wage might be positive.[62] In such a case, immigration raises the competitive wage above the pre-immigration reference wage, leading to a trivial solution. Indeed, the competitive wage is then compatible with natives exerting high effort, and paying the reference wage can never be an equilibrium, as in the complete contract case. Our focus is thus on the more interesting scenario where complementarity is not too strong and marginal productivity decreases sufficiently to cause the competitive wage to become lower than the reference wage. In this case, the competitive wage is lower than the reference wage w_r and implies that natives exert reduced effort λ, as in the baseline model. Using $L_j=\frac{n+n_I}{J}$, this competitive wage is:

$$w_c=\theta\left(\frac{n+n_I}{J}\right)^{\theta-1}p\Omega\left[\left((1-I)\,\lambda\right)^{\rho}+I^{\rho}\right]^{\frac{\theta}{\rho}}-\frac{1}{\alpha}\frac{J}{J-1}. \tag{2.20}$$

Profits under the competitive wage w_c and reduced native effort λ are thus:

$$\Pi(w_c,e_N=\lambda,e_I=1)=(1-\theta)\,p\Omega\left(\left[\left((1-I)\,\lambda\right)^{\rho}+I^{\rho}\right]^{\frac{1}{\rho}}\frac{n+n_I}{J}\right)^{\theta}+\frac{n+n_I}{\alpha}\frac{1}{J-1}.$$

In contrast, profits under the reference wage w_r and normal effort by all workers are:

$$\Pi(w_r,e_N=1,e_I=1)=p\Omega\left[\left(\frac{n+n_I}{J}\right)^{\theta}\left[(1-I)^{\rho}+I^{\rho}\right]^{\frac{\theta}{\rho}}-\frac{n+n_I}{n}\theta\left(\frac{n}{J}\right)^{\theta}\right]+\frac{n+n_I}{\alpha}\frac{1}{J-1}.$$

[62] Formally, the competitive wage compatible with full effort is larger than the reference wage if $\left[((1-I))^{\rho}+I^{\rho}\right]^{\frac{1}{\rho}}>(1-I)^{-\frac{1-\theta}{\theta}}$. For a given I, the left-hand side is decreasing in ρ, whereas the righ-hand side is decreasing in θ. Hence the condition is satisfied for low ρ and high θ.

Table 2..9: Values of $\overline{I}$ for various calibrations of λ and ρ

$\overline{I}$	$\rho = 1$	$\rho = 0.95$	$\rho = 0.9$	$\rho = 0.667$
$\lambda = 0.7$	0.21	0.25	0.31	0.60
$\lambda = 0.8$	0.15	0.2	0.26	0.58
$\lambda = 0.9$	0.09	0.12	0.19	0.57

The comparison of firms' profits under the two strategies is $\Pi^* (w_r, 1) \lesseqgtr \Pi^* (w_c, \overline{e})$:

$$p\Omega \left(\frac{n+n_I}{J}\right)^{\theta} \left[((1-I))^{\rho} + (I)^{\rho}\right]^{\frac{\theta}{\rho}} - \frac{n+n_I}{n}\theta p\Omega \left(\frac{n}{J}\right)^{\theta} \lesseqgtr (1-\theta)\, p\Omega \left(\left[((1-I)\,\lambda)^{\rho} + I^{\rho}\right]^{\frac{1}{\rho}} \frac{n+n_I}{J}\right)^{\theta}.$$

After after factoring out $p\Omega$, and dividing both sides by $\left(\frac{n+n_I}{J}\right)^{\theta}$, one obtains

$$\left[((1-I))^{\rho} + (I)^{\rho}\right]^{\frac{\theta}{\rho}} - \left(\frac{n+n_I}{n}\right)^{1-\theta} \theta \lesseqgtr (1-\theta) \left(\left[((1-I)\,\lambda)^{\rho} + I^{\rho}\right]^{\frac{1}{\rho}}\right)^{\theta}.$$

Using $\frac{n+n_I}{n} = (1-I)^{-1}$ on the left-hand side, one obtains:

$$\left[(1-I)^{\rho} + I^{\rho}\right]^{\frac{\theta}{\rho}} - \theta\,(1-I)^{-(1-\theta)} \lesseqgtr (1-\theta) \left[((1-I)\,\lambda)^{\rho} + I^{\rho}\right]^{\frac{\theta}{\rho}}.$$

Hence, the equation characterizing $\overline{I}$ is here:

$$\left[(1-\overline{I})^{\rho} + \overline{I}^{\rho}\right]^{\frac{\theta}{\rho}} - \theta\left(1-\overline{I}\right)^{-(1-\theta)} = (1-\theta) \left[\left((1-\overline{I})\,\lambda\right)^{\rho} + \overline{I}^{\rho}\right]^{\frac{\theta}{\rho}}.$$ [63]

As in the baseline, firms will obtain higher profits by paying the reference wage (i.e., wages are unaffected by the immigration shock) if and only if the migration inflow I is lower than $\overline{I}$. We conclude this extension by calibrating $\overline{I}$ using relevant values of ρ, θ and λ. Ottaviano and Peri, 2012 find an elasticity of substitution between native and immigrant workers of about 20, which implies $\rho = \frac{20-1}{20} = 0.95$. Card, 2009 reports estimates suggesting an elasticity in the range of 3 to 10. For $\sigma = 10$, $\rho = \frac{10-1}{10} = 0.9$, and for $\sigma = 3$, $\rho = \frac{3-1}{3} \simeq 0.667$. Assuming $\theta = 0.7$, we report the different values of $\overline{I}$ for these three calibrations of ρ and for low effort values of $\lambda = 0.7, 0.8,$ and 0.9 in Table 2..9.

The key insight from Table 2..9 is that, ceteris paribus, a lower elasticity of substitution

[63] Note that for $\rho = 1$, we obtain equation (2.8).

between immigrants and natives encourages firms to keep paying the reference wage. Specifically, as the substitutability between immigrant and native workers decreases (i.e., as σ decreases), the critical migration threshold $\overline{I}$ increases.

For instance, columns 1 and 2 allow us to compare $\overline{I}$ between the cases of perfect substitutes ($\sigma \to \infty$, which corresponds to the baseline model) and $\sigma = 20$. When the low effort level is 0.9, firms will pay the reference wage to ensure high effort if the migration shock is lower than 12

The intuition behind this result is as follows: When immigrants are *imperfect substitutes* for natives, firms have stronger incentives to maintain natives' high effort and thus to continue paying the reference wage. This is because the high effort exerted by native workers is more valuable when it cannot be easily replaced by immigrant labor.

To sum up, the qualitative result of the baseline model remains—firms switch to competitive wages when the migration shock exceeds $\overline{I}$—but it takes a larger immigration shock for this switch to become profitable when natives' labor is less substitutable by immigrants.

Extension: Expectation-based reference wages

In this extension, we acknowledge that native workers may adjust their wage expectations based on evolving market conditions. In particular, in the face of a large immigration shock and increased competition, native workers may indeed consider that the pre-immigration wage w^* (see equation 2.3) is nolonger a valid reference point. Workers who would fully adapt their wage expectations to new market conditions would provide high effort under the post-migration competitive wage $\widehat{w}(1)$ (see equation 2.6). Such an adaptation of the reference wage based on rational expectations is in line with the approach of Kőszegi and Rabin, 2006.

We generalize the reference wage by allowing it to depend on both the pre-immigration wage w^* and the post-immigration market wage compatible under full effort $\widehat{w}(1)$. Specifically, we define the reference wage w_r as a convex combination of w^* and $\widehat{w}(1)$:

$$w_r = (1-\tau)w^* + \tau\widehat{w}(1),$$

where $\tau \in [0,1]$ captures the degree to which workers adapt their reference wage to new market expectations. When $\tau = 0$, workers fully anchor their reference wage to the pre-immigration wage w^*, as in the baseline model. When $\tau = 1$, they instead fully adjust their reference wage to the new competitive wage $\widehat{w}(1)$. Substituting the expressions for w^* and $\widehat{w}(1)$ and simplifying, we obtain:

$$w_r = \theta p \Omega \left[(1-\tau) \left(\frac{n}{J} \right)^{\theta-1} + \tau \left(\frac{n+n_I}{J} \right)^{\theta-1} \right] - \frac{1}{\alpha} \frac{J}{J-1}. \tag{2.21}$$

Unlike the baseline model, the generalized reference wage now depends on the magnitude of the immigrant inflow:

$$\frac{\partial w_r}{\partial n_I} = -\tau (1-\theta) \theta p \Omega J^{1-\theta} (n+n_I)^{\theta-2} < 0.$$

The reference wage decreases with the magnitude of the immigration shock due to decreasing marginal productivity, and this effect is more pronounced if natives are more tolerant to wage adjustments (if τ is large).

As in the baseline, firms face two possible strategies: either pay the reference wage w_r and obtain full effort from all workers ($E[e^*(w_r)] = 1$), or pay the competitive wage knowing that, as soon as $\tau < 1$, this wage will be below w_r and will lead to low effort by native workers.[64] This competitive wage w_c is thus based on expected effort $\bar{e} = I + (1-I)\lambda$):

$$w_c = \widehat{w}(\bar{e}) = \theta p \Omega (I + (1-I)\lambda)^{\theta} \left(\frac{n+n_I}{J} \right)^{\theta-1} - \frac{1}{\alpha} \frac{J}{J-1}.$$

The decision of firms' wage-setting strategy is based on the comparison of profits between $\Pi^*(w_r, 1)$ and $\Pi^*(w_c, \bar{e})$.

First, using $L_j = \frac{n+n_I}{J}$, (2.1), (2.21), $\Pi^*(w_r,1) = p\Omega \left(\frac{n+n_I}{J}\right)^{\theta} - \left(\frac{n+n_I}{J}\right) w_r$ can be expressed after substitution for w_r as:

$$\Pi^*(w_r, 1) = p\Omega \left[(1-\theta\tau) \left(\frac{n+n_I}{J} \right)^{\theta} - \frac{n+n_I}{n} (1-\tau) \theta \left(\frac{n}{J} \right)^{\theta} \right] + \frac{n+n_I}{\alpha} \frac{1}{J-1}.$$

[64] The only case where the competitive wage is compatible with high effort provision by natives is when $\tau = 1$, as it implies that $w_r = \widehat{w}(1)$.

Table 2..10: Values of $\overline{I}$ for various calibrations of λ and ρ

$\overline{I}$	$\tau = 0$	$\tau = 0.2$	$\tau = 0.4$	$\tau = 0.6$
$\lambda = 0.7$	0.21	0.25	0.29	0.37
$\lambda = 0.8$	0.15	0.18	0.22	0.29
$\lambda = 0.9$	0.09	0.11	0.13	0.18

Note that when $\tau = 0$, this expression is the same as in the baseline, whereas when $\tau = 1$, it corresponds to profits when firms pay the competitive wage and receive high effort (which is equivalent to the case of complete contracts).

Second, when firms offer the competitive wage $w_c = \widehat{w}\left(\overline{e}\right) < w_r$, their profits are, as in the baseline:

$$\Pi\left(w_c;\overline{e}\right) = \left(1-\theta\right) p\Omega\left(\left(I + \left(1-I\right)\lambda\right)\frac{n+n_I}{J}\right)^{\theta} + \frac{n+n_I}{\alpha}\frac{1}{J-1}.$$

Firms thus choose to offer the reference wage if $\Pi^*\left(w_r,1\right) > \Pi^*\left(w_c,\overline{e}\right)$, which boils down to:

$$\left(1-\theta\tau\right) - \theta\left(1-\tau\right)\left(1-I\right)^{\theta-1} > \left(1-\theta\right)\left(\left(I+\left(1-I\right)\lambda\right)\right)^{\theta}.$$

This inequality is equivalent to $I < \overline{I}$, where $\overline{I}$ is implicitly defined by:

$$\left(1-\theta\tau\right) - \theta\left(1-\tau\right)\left(1-I\right)^{\theta-1} - \left(1-\theta\right)\left(\left(I+\left(1-I\right)\lambda\right)\right)^{\theta} = 0.$$

Again, note that $\overline{I}$ is identical to the baseline when $\tau = 0$. As in the first extension, we report different values of $\overline{I}$ for low effort values of $\lambda = 0.7$, 0.8 and 0.9, and for wage-cut tolerance levels of $\tau = 0$, 0.2, 0.4 and 0.6.

The threshold $\overline{I}$ increases with natives' adaptation parameter τ. For a given immigration inflow I, the larger τ, the more likely I is smaller than $\overline{I}$. In other words, the more natives adapt their wage expectations to new market conditions, the more firms are to prefer the strategy of offering the reference wage w_r in order to obtain full effort. When $\tau = 0.4$, i.e. when 40% of the reference wage is based on the past market wage and 60% on the decreased marginal productivity after immigration, and for low effort $\lambda = 0.9$, firms will pay high wages to maintain high native effort as long as the immigration shock does not exceed 13% of the total workforce. It is important to note however that this "high" wage is now inferior to the previous market wage w^*.

This extension brings the novel prediction that under incomplete contracts combined with workers' adaptability, firms are encouraged reduce wages in a controlled way so as to maintain high effort. While this wage decrease must be weaker than the actual marginal productivity decrease, it shows that relatively large immigration shocks need not trigger a negative effort response.

Consent Form

DESCRIPTION:

You are invited to participate in a research study on economic decisions in labor markets. The experiment is split into two parts. In the first part, you will read the instructions and answer some questions to check that you understood them. On another day, you will come to the laboratory for economic experiments at the Nottingham University Business School (NUBS). You will interact with other participants in a market situation. The choices you and other participants make will determine the payout you receive. At the end, there will be a short questionnaire asking you basic information about your background and preferences.

TIME INVOLVEMENT:

Your participation will take up to 90 minutes in total.

PAYMENT:

Your participation will depend on your decisions and the decisions of other participants. The average payout will be in line with wage regulations of the university.

PROTECTION OF CONFIDENTIALITY:

The information that you give in the study will be handled confidentially. Your information will be assigned a code number. The list connecting your name to this code will be kept in a locked file. When the study is completed and the data have been analyzed, this list will be destroyed. Your name will not be used in any report.

PARTICIPANT'S RIGHTS:

If you have read this form and have decided to participate in this project, please understand your participation is voluntary, and you have the right to withdraw your consent or discontinue participation at any time. If you discontinue before finishing the online part or do not show up for the in-person experiment, you will not earn any money. If you discontinue during the in-person experiment, you will earn a show-up fee of £5. You have the right to refuse to answer particular questions. The results of this research study may be presented at scientific or professional meetings or published in scientific journals.

CONTACT INFORMATION:

Last Name: ___________________

First Name: ___________________

If you have any questions, concerns, or complaints about this research, its procedures, risks, and benefits, contact the Researcher (Felix Stips `felix.stips@liser.lu`).

AGREEMENT:

I agree to participate in the research study described above.

Signature: __________________________________

Date: ____________

Instructions

In the following, we present representative instructions for an employer in the high-income market. They are based on Altmann et al. (2014) and adapted for our needs. The instructions for employers in the low-income market are the same, except that some dynamic fields are adjusted to match the parameters of the market. The instructions for workers are very similar and provided upon request. The second set of instructions shown during the mid-break are provided in Appendix 2.6. The instructions were shown to participants through the software. Dashed lines indicated page breaks in the software version.

Welcome

You are now participating in an economic experiment. This is a research project which is designed to study economic decision-making.

The experiment consists of two parts. After you finish both parts, there will be a short questionnaire followed by the payment. In total, the experiment should last up 90 minutes.

During the game, you will interact **anonymously** with other participants through the computer. This means you will not know the identity of those you interact with, and they will not know yours. The participants you interact with are not necessarily sitting close to you.

Please **do not talk or communicate** with the other participants during the experiment. This might disturb other people and invalidate the session. Also, please turn off or put away your mobile phone.

Payment

At the beginning of the experiment, you will receive an **initial endowment of 12 Pounds**. During the experiment, you can earn more money by accumulating points.

All points you earn will be converted to Pounds at the end of the experiment. The exchange rate will be:

1 point = 0.01 Pounds

The amount earned will depend on your decisions and those of the other participants. You will be paid via PayPal after the experiment is completed.

Instructions – Part 1

This set of instructions describes the first part of the experiment. After you finish the first part, there will be another set of instructions for the second part. The second part is very similar to the first part.

Please read the following instructions carefully. After the instructions, you will be asked a **set of questions to ensure you understand** the instructions. The experiment will begin after every participant has answered all questions.

If you have any **questions, please raise your hand**. We will answer your questions at your cubicle.

General Information

The experiment is divided into **16 rounds**, 8 rounds in the first part and 8 rounds in the second part.

You have been randomly assigned to one of two roles: either an employer or a worker.

> **You will take the role of employer.** You will keep this role for the entire experiment.

In the first part, you will interact in a group of 11 randomly assigned participants. Each group has 5 employers and 6 workers.

Short Overview of the Experimental Procedures

In each round, employers choose to employ zero, one, or two workers. To employ a worker, an employer needs to pay a wage. In return, employed workers provide effort. Workers choose how much effort to provide after they accept a job offer. Effort is costly to workers, so they make profits only if their effort costs are lower than the wage they receive.

On the other end, your profits will depend on the

1. the number of workers you employ
2. the wages you pay
3. the effort workers provide

You make profits if the effort you receive is worth more than wage you pay.

The first part lasts for 8 rounds. Each round has two phases:

1. First, a **market phase** lasting for 150 seconds. During this phase, employers can submit job offers, which workers can accept. When submitting a job offer, an employer specifies two things

- The wage they will pay
- The effort they want the worker to provide

Once employers submit offers, workers get to see them and decide which to accept. The first worker to accept an offer will get it. **Each worker can accept at most one job offer in each round. Each employer can employ at most two workers**. Unaccepted offers are deleted at the end of the market phase.

You can also send a **private re-employment offer to each worker you employed in the previous round**. Only the worker to whom the offer is directed can accept a private offer.

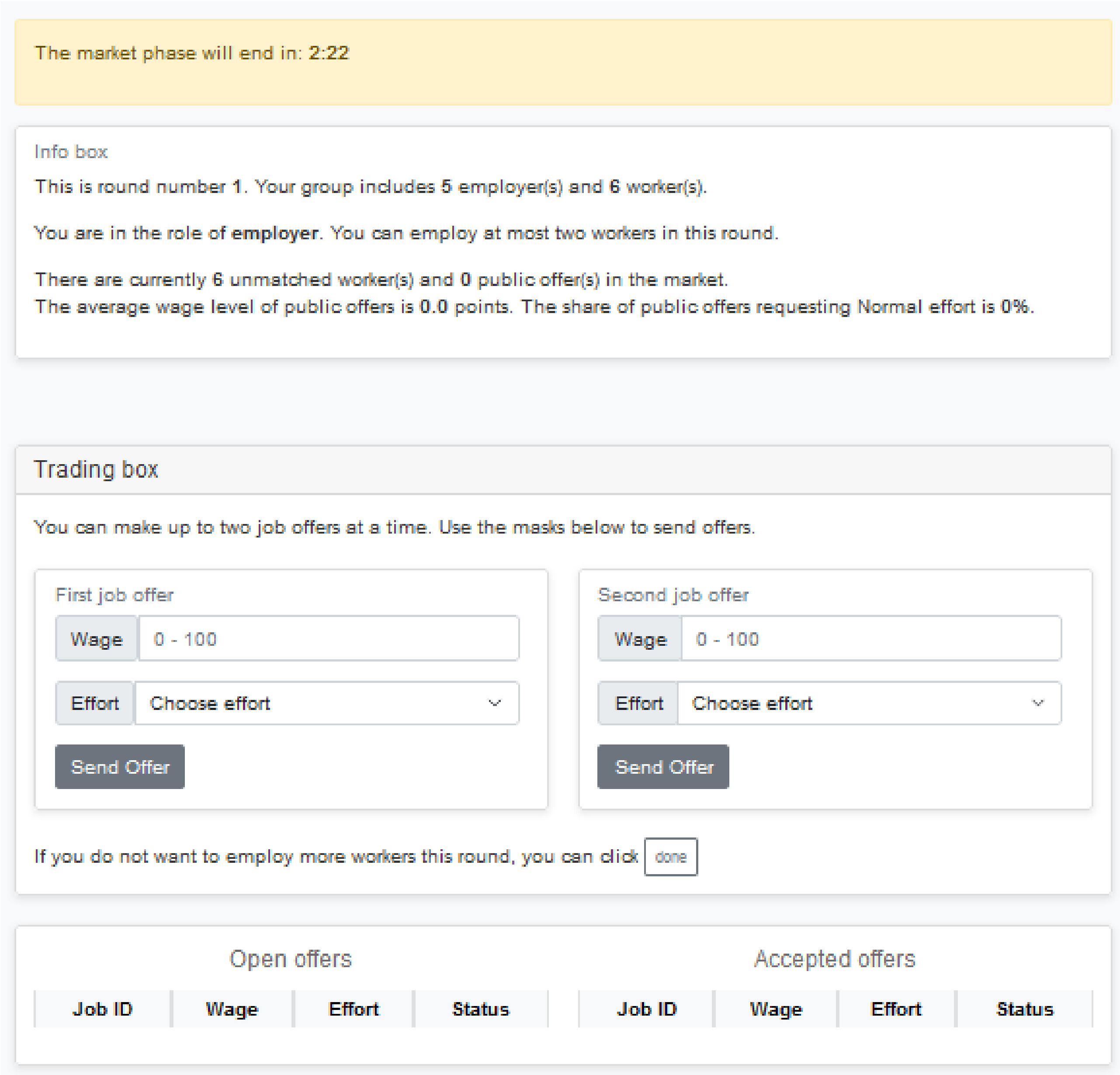
2. After the market phase, workers who accepted a job offer enter the production phase. In this phase, workers decide their effort. **Workers are not forced to provide the effort desired by the employer. They are free to choose any effort**.

Once all workers have chosen their effort, everyone's profits are calculated and the next round starts. The sum of your profits across all rounds will determine your earnings for the experiment.

The Market Phase in Detail

Each round begins with a **market phase**.

The figure below shows a preview of the market phase screen. The white box below the timer gives **information** about the state of the market.

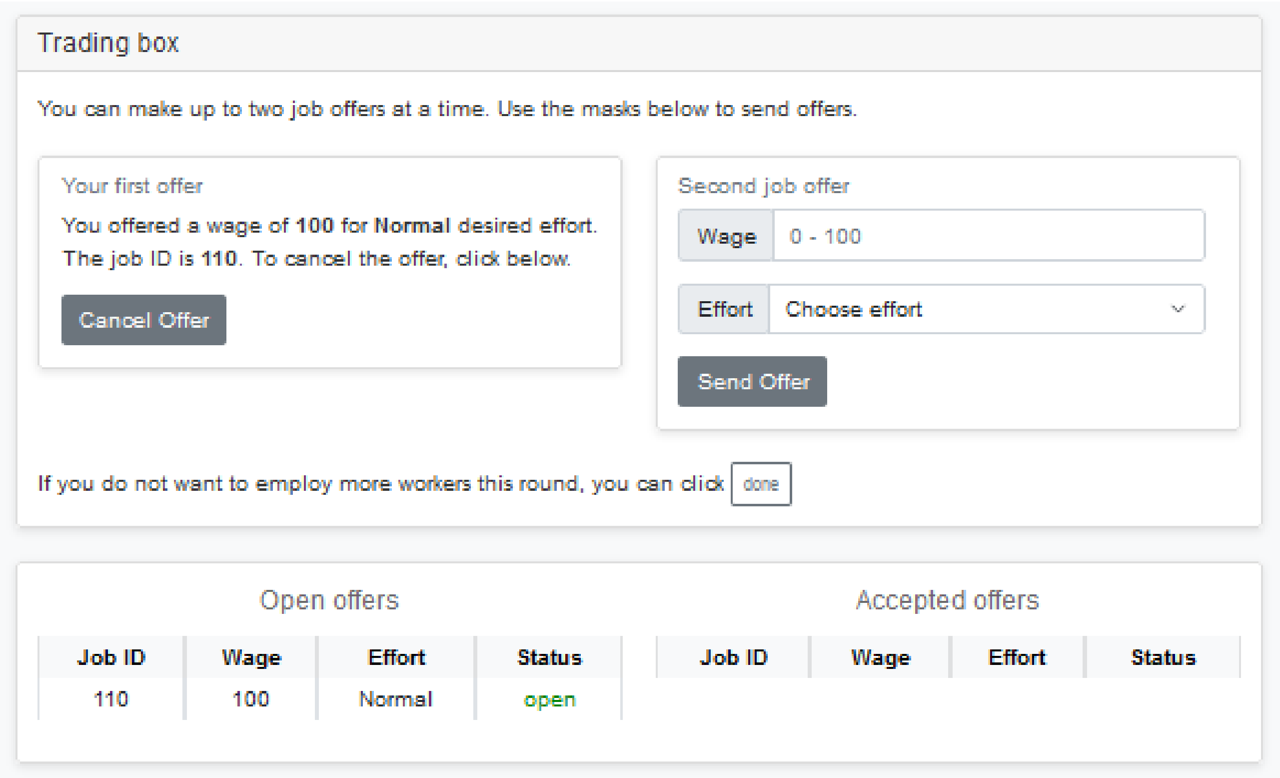


To **make a job offer**, use "Trading box" in the middle of the page. You can send up to two job offers at the same time.

To send a job offer, you must:

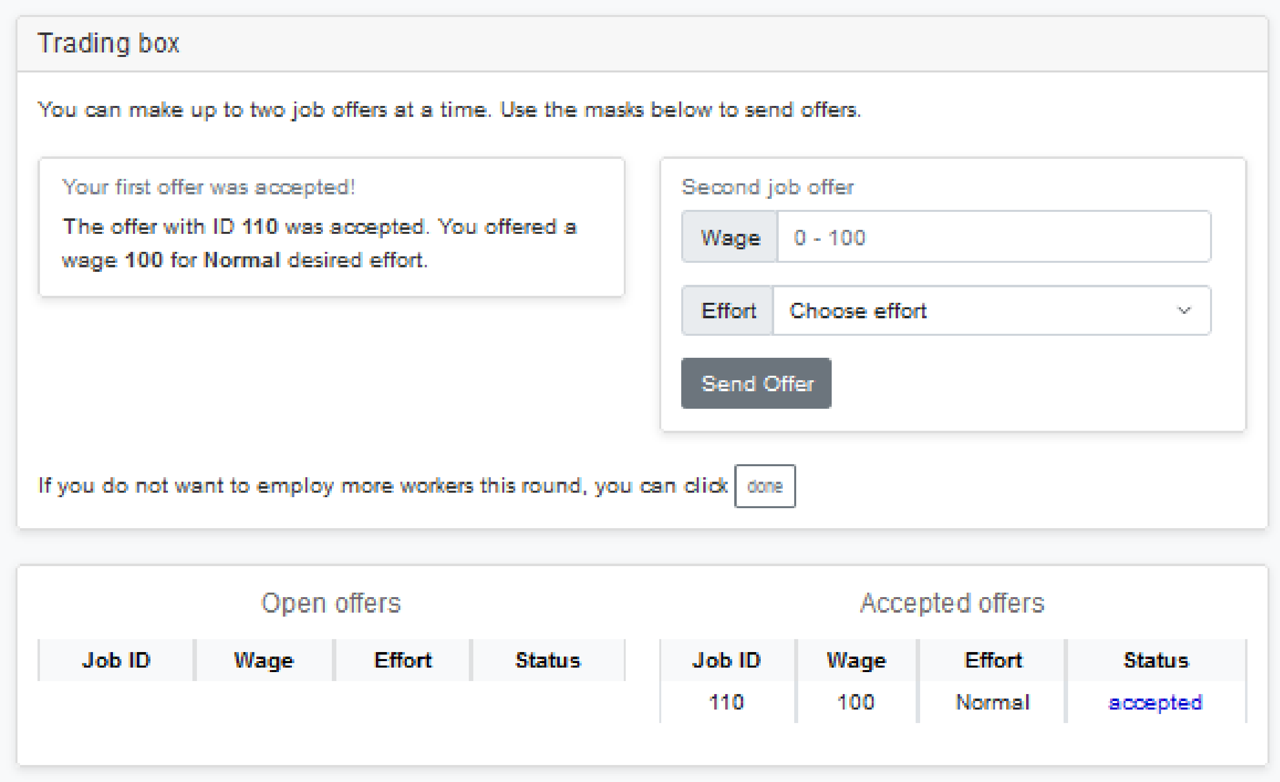
1. choose a wage between 0 and 100 points.
2. indicate whether you want the worker to provide **"normal"** or **"low"** effort.
3. click on "Send Offer".

Once you **submit a job offer**, it will appear in the table "Open offers" at the bottom of the page. See the figure below for an example of how the screen will look.



Workers can accept an offer as long as it is "open". You can have up two open job offers at a time. You can also submit no offer at all. To withdraw or change a job offer, click on "Cancel Offer".

Open offers from all employers are visible to everyone in the "Open Offers" table. **Workers can accept at most one offer during each round**, and each job offer can only be accepted by one worker. If your offer is accepted, you will see a notification and the offer will appear in the "Accepted Offers" table. See the figure below for an example of how the screen will look after a worker accepts your first job offer.



If you do not want to send any more offers in this round, you can click the "done" button in the bottom of the Trading Box.

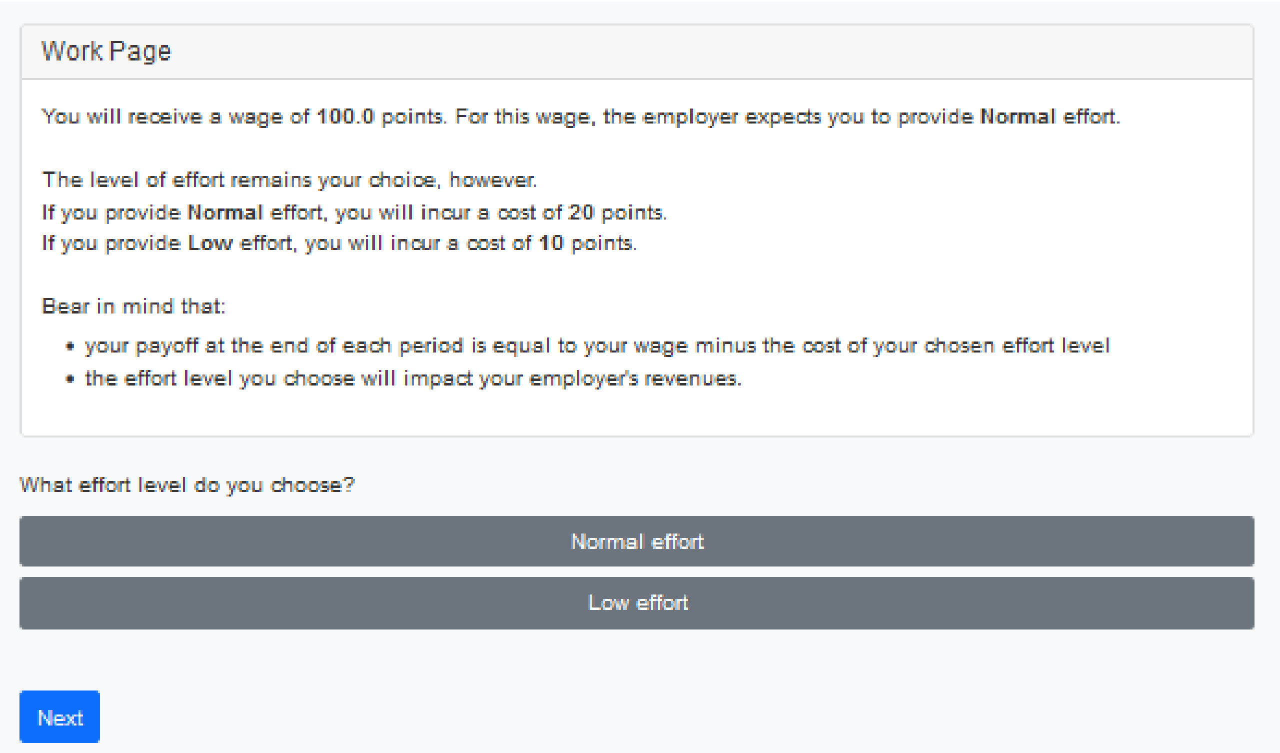
The market phase **ends** when all workers have accepted an offer, all employers have either filled their jobs or clicked the "done" button, or else when the time limit of 150 seconds is reached. The time left in the current round will be displayed in a yellow box at the top of the page. When the market phase is over, no offers cannot be submitted or accepted.

The Production Phase in Detail

Following the market phase, all workers who have accepted a job offer enter the production phase. In this phase, workers choose their effort. **The effort requested by employers is not binding**. Hence, workers can provide **normal** or **low** effort irrespective of the effort indicated in the job offer.

The page workers see looks like the one shown below. At the top of the page, workers see the wage they accepted and the effort requested by the employer. Below, they see the cost of each effort choice.

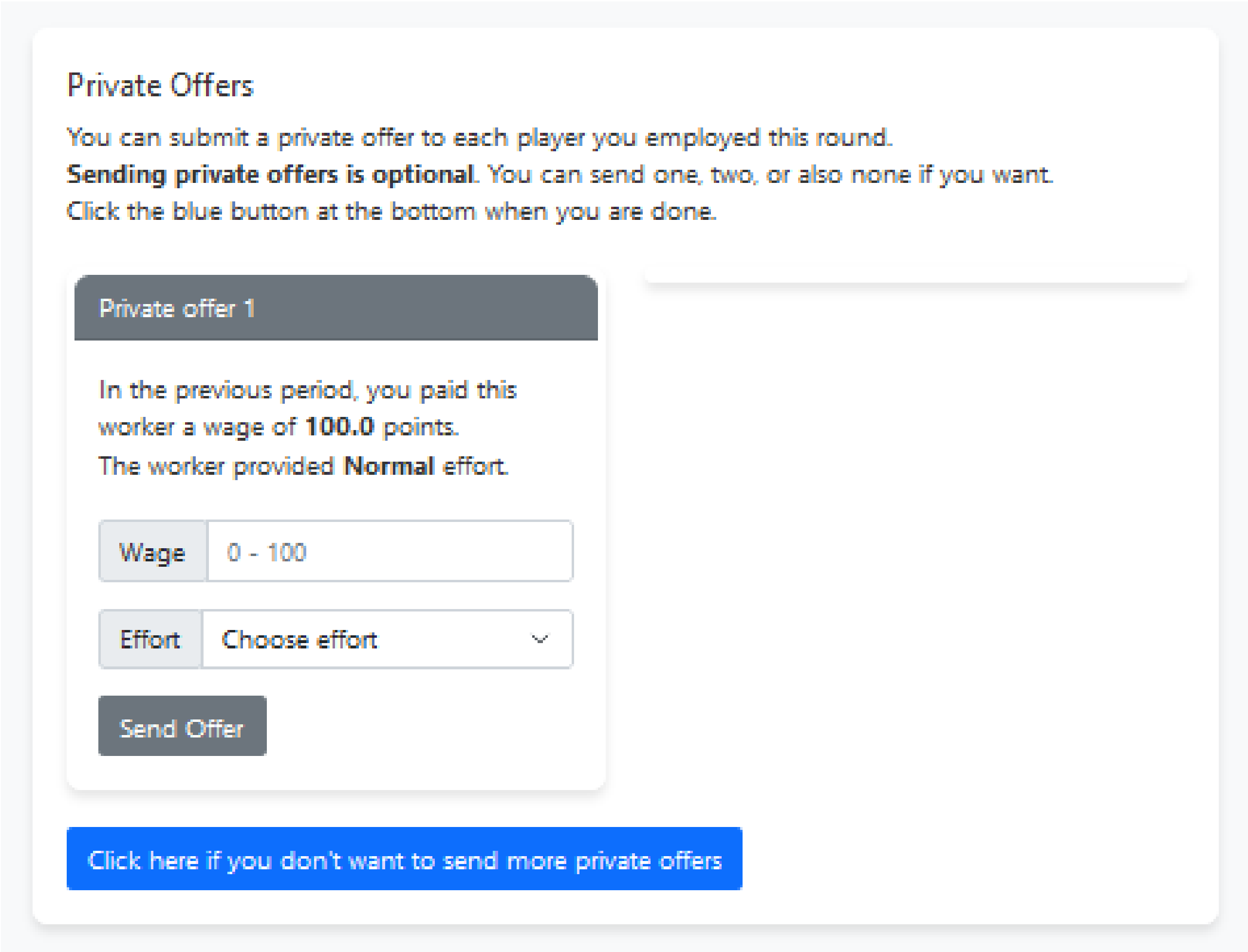
- **The cost of normal effort is 20 points.**
- **The cost of low effort is 10 points.**



Private Re-Employment Offers

You can send a **private offer to each worker you employed in the previous round**. Before the next round starts, you can decided whether and what private offers to submit.

Private offers work just as public offers: You need to specify a wage and an effort, and then click on "Send Offer". Click "next" once you are done sending private offers. Note that you cannot cancel the offer later on.



A private offer can only be seen by the worker you've sent it to. As such, private offers **allow you to establish a longer-term relationship** with workers.

How are the workers' profits calculated?

In each round, the workers' profits depend on whether they are employed, their wage, and their effort level.

- If a worker does not accept a job offer, they earn **0 points**.
- If a worker accepts a job offer and chooses Low effort, their profits equal the wage minus the cost of low effort: **wage – 10 points**.
- If a worker accepts a job offer and chooses Normal effort, their profits equal the wage minus the cost of normal effort: **wage – 20 points**.

Example: if a worker receives a wage of 50 points and chooses normal effort, the worker's profits in this round are 50 – 20 = 30 points. If, instead, the worker chooses Low effort, their profits are 50 – 10 = 40 points.

As you can see, a worker's profits are larger the higher their wage. Their profit is also higher if they choose Low effort. Note that a worker can make a loss if the effort costs are higher than their wage.

How are your profits calculated?

Your profits in each round are determined by the wages you pay to your workers, the effort they choose, and the number of workers you employ.

If you **do not employ any workers** in a round, you earn **0 points** in that round.

If you employ one or two workers, your profit is the difference between your total revenue and the wages you pay:

Profit = Total revenue – total wage payments.

The total revenue depends on the number of workers you employ and their effort, according to the table below.

Case	**Total revenue**
Employ two workers. Both workers chose Normal effort	280
Employ two workers. One worker chooses Normal effort and the other chooses Low effort	210
Employ two workers. Both workers chose Low effort	140
Employ one worker. The worker chooses Normal effort	160
Employ one worker. The worker chooses Low effort	80

Example: Suppose you employ two workers. You pay one worker a wage of 60 points and the other worker a wage of 40 points.

- If both workers choose Normal effort, then your profits equal to 280 – 60 – 40 = 180 points.
- Alternatively, if one worker chooses Normal effort and the other Low effort, then your profits equal 210 – 60 – 40 = 110 points.

Note that you **make a loss** if your total wage payments are larger than the total revenue. Losses will be paid from your initial endowment.

Are you ready?

The experiment will not start until all participants are familiar with the instructions. To ensure this is the case, we ask you to **answer** three questions on the following pages.

Click on the next button below when you are ready to answer these questions.

Mid-break instructions

In the following, the mid-break instructions. The first page is the same for all participants, afterwards there are four cases: (i) employers in the low-income market, (ii) workers from the low-income market, (iii) employers in the high-income market, and (iv) workers in the high-income market. For the latter two, the number of workers joining the market would depend on whether the market is large or small shock. Here we use the version with the large shock. The instructions were shown to participants through the software. Dashed lines indicated page breaks in the software version. Note on case iii: X and Y adjust based on the average in part 1. Z is then the corresponding predicted effort level using a linear regression $effort = \alpha + \beta \times wage + \epsilon$ *on pre-test data. Wage is defined in 5-points brackets.*

Case i: Employer from low-income market

Welcome to Part 2

The first part of the experiment is now concluded.

The following page will explain the differences between part 1 and part 2.

Overview

You have been randomly selected to exit the market. The second part of the experiment will comprise only of a short questionnaire.

You will finish before the other participants. Please remain leave the room quietly.

Case ii: Worker from low-income market

Welcome to Part 2

The first part of the experiment is now concluded.

The following page will explain the differences between part 1 and part 2.

Overview

The second part of the experiment is very similar to the first part. There will be 8 more rounds to play. Each round will comprise of a market and production phase just like before.

Differences from Part 1

You have been moved to a different market. This market is different to the market you were in before. In addition to you, there are 5 employers and 11 workers in this market.

Case iii: Employer from high-income market

Welcome to Part 2

The first part of the experiment is now concluded.

The following page will explain the differences between part 1 and part 2.

Overview

The second part of the experiment is very similar to the first part. There will be 8 more rounds to play. Each round will comprise of a market and production phase just like before.

Differences from Part 1

The difference is that 5 workers from another market have joined this market. The new workers played in a different market, where firms were less productive and thus paid lower wages. For this reason, a high share of workers provide high effort even at low wages. For example, the average wage in your market was X and Y% of workers provided high effort. In a previous session, Z% of workers from the other market were for providing high effort for this wage.

Case iv: Worker from high-income market

Welcome to Part 2

The first part of the experiment is now concluded.

The following page will explain the differences between part 1 and part 2.

Overview

The second part of the experiment is very similar to the first part. There will be 8 more rounds to play. Each round will comprise of a market and production phase just like before.

Differences from Part 1

The difference is that 5 workers from another market have joined this market. The number of employers is the same as in the first part of the experiment. This means that there are 5 employers and 11 workers in this market.

Ex-post survey items

- Age:

  What is your age?

- Gender:

  What is your gender? (Choices: Female, Male, Other, Prefer not to say)

- Education:

  What is the highest level of education you have completed? (Choices: Less than high school, High school, Some college, Bachelor's degree, Master's degree, Doctorate, Other, Prefer not to say)

- Field of study:

  What is your field of study? (Choices: Arts, Business, Economics, Finance, Law, Mathematics, Medicine, Psychology, Natural Sciences, Engineering, Not applicable, Other, Prefer not to say)

- Country of birth:

  What is your country of birth? (Choices: United Kingdom, Other European, Other non-European, Prefer not to say)

- Self-assessment: willingness to take risks in general:

  Please tell us, in general, how willing or unwilling you are to take risks. Please use a scale from 0 to 10, where 0 means "completely unwilling to take risks" and a 10 means you are "very willing to take risks". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. (1-10 Likert scale)

- Gift in exchange for help:

  Please think about what you would do in the following situation. You are in an area you are not familiar with, and you realize you lost your way. You ask a stranger for directions. The stranger offers to take you to your destination. Helping you costs the stranger about 20 £ in total. However, the stranger says he or she does not want any money from you. You have six presents with you. The cheapest present costs 5£, the most expensive one costs 30£. Do you give one of the presents to the stranger as a “thank-you”-gift? If so, which present do you give to the stranger? (Choices: No present’, ’The present worth 5£’, ’The present worth 10£’, ’The present worth 15£’, ’The present worth 20£’, ’The present worth 25£’, ’The present worth 30£’)

- Self-assessment: willingness to return a favor :

  When someone does me a favor I am willing to return it. (Agree-disagree 1-10 Likert scale)

- Self-assessment: willingness to take revenge:

  If I am treated very unjustly, I will take revenge at the first occasion, even if there is a cost to do so. (Agree-disagree 1-10 Likert scale)

- Self-assessment: willingness to punish unfair behavior toward self:

  How willing are you to punish someone who treats you unfairly, even if there may be costs for you? (1-10 Likert scale)

- Self-assessment: willingness to punish unfair behavior toward others:

  How willing are you to punish someone who treats others unfairly, even if there may be costs for you? (1-10 Likert scale)

- Lottery choice sequence using staircase method:

  Please imagine the following situation: You can choose between a sure payment and a lottery. The lottery gives you a 50 percent chance of receiving 300 Euro. With an equally high chance you receive nothing. Now imagine you had to choose between the lottery and a sure payment. We will present to you five different situations. The lottery is the same in all situations. The sure payment is different in every situation. What would you prefer: a 50 percent chance of winning 300 Euro when at the same time there is 50 percent chance of winning nothing, or would you rather have the amount of 160 Euro as a sure payment? (a) lottery → go to question 17 (b) sure payment → go to question 2 ...

  See Falk et al., 2023 Appendix E for a full description of the staircase tree.

Additional ex-post survey items

Note: We added these questions after five sessions were completed. They were shown after the market rounds were completed and before the main ex-post survey items. Each player type (employers in destination, incumbent workers in destination, migrants) were shown different questions. Employers in the sending market were not shown any questions. Below we report the questionnaires for each type.

Destination Employer Questionnaire

We would like to know more about your experience in part 2 (after the mid-break). Please answer all questions considering your experiences during this second part of the experiment.

1. How did the workers' effort in part 2 compare to what you expected?

   I think the workers provided...

   - a lot more effort than I expected.
   - somewhat more effort than I expected.
   - the same level of effort as I expected.
   - somewhat less effort than I expected.
   - a lot less effort than I expected.
   - I don't know.

2. How important was retaining your existing worker(s)?

   - Not important
   - Slightly important
   - Moderately important
   - Very important
   - Extremely important
   - I don't know

3. How confident were you in your ability to hire workers in each round?

   - Not confident
   - Slightly confident
   - Moderately confident

- Very confident
- Extremely confident
- I don’t know

4. Which of the following best describes your strategy when the number of workers increased in part 2?
    - I prioritized cost-saving by lowering wages.
    - I maintained wages to preserve worker effort.
    - I followed other employers’ wage offers.
    - Other
    - I did not have a clear strategy.

Incumbent Worker Questionnaire

We would like to know more about your experience in part 2 (after the mid-break). Please answer all questions considering your experiences during this second part of the experiment.

1. Did the increased number of workers on the market change your motivation to provide effort?
    - It highly demotivated me
    - It slightly demotivated me
    - It did not impact my motivation
    - It slightly motivated me
    - It highly motivated me
    - I don’t know
2. Did you perceive that wages were different in part 2?
    Compared to part 1 I think that on average wage offers in part 2 were...
    - higher
    - about the same
    - lower

    (a) If "higher": How much higher were the wage offers? Enter a number in percentage points or write "NA" if you don’t know.

I think that on average the wage increased by ... %.

(b) If "lower": How much lower were the wage offers? Enter a number in percentage points or write "NA" if you don't know.

I think that on average the wage decreased by ... %.

3. How did you personally feel about these new wage offers?
   - I found the new wage very unfair
   - I found the new wage somewhat unfair
   - Neutral
   - I found the new wage somewhat fair
   - I found the new wage very fair
   - I don't know

4. Do you think it is reasonable for employers to adjust wages when market conditions change, such as an increase in the number of workers?
   - Yes definitely reasonable
   - Yes somewhat reasonable
   - Neither reasonable nor unreasonable
   - No somewhat unreasonable
   - No definitely unreasonable
   - I don't know

5. Did you consider that there was a higher risk of becoming unemployed in part 2?
   - Yes definitely
   - Yes somewhat
   - Unsure
   - No unlikely
   - I don't know

6. Which of the following emotions did you experience in response to the arrival of new workers in the market? (Please select all that apply)
   - Threatened
   - Confident

- Resentful toward employers
- Resentful toward new workers
- Grateful toward employers
- Grateful toward new workers
- Indifferent
- Other
- I did not experience any particular emotions

Newcomer Worker Questionnaire

We would like to know more about your experience in part 2 (after the mid-break). Please answer all questions considering your experiences during this second part of the experiment.

1. Did you perceive that wages were different in the new market?

   Compared to part 1 I think that on average wage offers in part 2 were...

   - higher
   - about the same
   - lower

   (a) If "higher": How much higher were the wage offers? Enter a number in percentage points or write "NA" if you don't know.

   I think that on average the wage increased by ... %.

   (b) If "lower": How much lower were the wage offers? Enter a number in percentage points or write "NA" if you don't know.

   I think that on average the wage decreased by ... %.

2. Did the wages in the market motivate you to provide more effort?

   - Yes I was significantly more motivated
   - Yes I was somewhat more motivated
   - My motivation remained the same as before
   - No I was less motivated
   - No I was significantly less motivated

3. As the rounds progressed in part 2 how did your motivation evolve in the new market?

- I became more motivated over time
- My motivation remained the same
- I became less motivated over time
- I don't know

4. Did you feel any pressure to provide more effort than incumbent workers?

    - Yes absolutely
    - Yes a bit
    - No not so much
    - No not at all
    - I don't know

Pre-analysis plan

1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

The aim of this study is to investigate the impact of migration shocks on wages, employment, and the efficiency of local labor markets under incomplete contracts and gift exchange. The main question is whether, as predicted by our model, the size of the migration shock and the relative wages difference between origin and destination determine the wage and employment response in such a labor market.

Main Hypotheses:

- H1. Wages and effort drop significantly only in the large shock condition, not in the small shock condition.
- H2. Effect of immigration in the large shock condition is stronger when there is an income difference.

Exploratory Hypotheses:

- H3. Workers choose high effort more often when offered high wages.
- H4. The correlation between wages and effort is stronger for workers with higher positive reciprocity score.
- H5. Effort in large shock condition drops more strongly for workers with higher negative reciprocity score.
- H6. Firms with higher risk preference offer lower wages in large shock condition.

3) Describe the key dependent variable(s) specifying how they will be measured.

The outcomes of interest will be wages, effort, profits, and (un-)employment.

Wages:

- What we will commonly refer to as "wage" is the wage of accepted work contracts. The wage can be between 0 and 100 points.

- "Offered wages" will be wage in all contracts, including those that were not accepted (but excluding cancelled offers).
- We will distinguish between "wage" and "native wages". The latter will be the accepted wage of all workers excluding the wages of migrants.

Effort:

- What we will commonly refer to as "effort" is the effort, which workers choose to give. Effort is binary and can be either "normal" or "low".
- "Desired effort" will be the effort which employers asked in the job offer.
- We will distinguish between "effort" and "native effort". The latter will be the given effort of all workers excluding the effort choices of migrants.

Profits:

- Firm profits will be the total effort gain minus the total wage paid and zero if no worker is employed.
- Worker profits will be the received wages minus the cost of effort and zero is no contract is accepted.

(Un-)employment:

- Unemployment is the number of workers who did not accepted a job offer during a round.
- We will distinguish between "unemployment" and "native unemployment". The latter will be the number of unemployed excluding migrants.
- We will also calculate the average length on consecutive working relationship for worker-firm pairs measured in rounds.

For H4 the outcome we will from the correlation between wages and effort and use that as outcome.

4) How many and which conditions will participants be assigned to?

The baseline design is a gift-exchange market game with six workers and six employers who interact over sixteen rounds with two stages. During the market stage, employers offer wage contracts. Workers can accept an offer or remain unemployed. Employers can employ up to two workers. During the work stage, workers who entered an employment relationship have to decide what effort level to provide. The parametrization is such that the gift exchange equilibrium can be mutually beneficial for employers and workers in the baseline condition. Wage offers are anonymous, but employers have to possibility to extend re-employment offers for workers from the previous round.

We conduct a 2x2 design:

- First arm is either large or small immigration shock. The small immigration shock means that the market receives one new worker. The large immigration shock means that the market receives five new workers. Both occur after eight of the sixteen rounds have been played and after a second short instruction.
- Second treatment arm is whether there is an income difference between the origin market and the destination markets. The no income difference condition implies that the exchange rate for wages between the origin and destination is 1:1. The income difference condition implies that the exchange rate is 2:1. This means that employers and workers in the origin country make less profit and that workers moved from origin to destination will experience an income increase (and presumably accept lower wage offers than natives).

Randomization for the first treatment arm occurs at the group level within sessions. In each session, we run three markets, one origin and two destination markets. One of the destination markets receives a large immigration shock and one a small immigration shock. Randomization for the second treatment arm occurs between sessions, with half of the sessions having an income difference and half not.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

We will conduct descriptive, causal, and exploratory analyses.

Descriptive analysis may include a summary table of main outcomes and covariates as well as a graphical representation of outcomes by round and treatment status.

Causal:

- Hypotheses H1 and H2 will be tested using experimental variation between treatment states. For H1 we compare post-migration outcomes between sessions with a small migration shock to outcomes in sessions with a large migration shock. For H2, we are interested in the double difference between large and small shock and no income difference vs income difference. Estimation will be conducted using regression analysis and, possibly, also nonparametric tests.

- Hypotheses H5 and H6 will be tested as heterogeneity analyses on treatment effects either parametrically as interaction effects in regression analysis or using nonparametric tests for treatment effect heterogeneity.

Exploratory:

- For H3 we will use regression analysis with effort choice as outcome. Candidate explanatory variables include round, treatment, wage, wage drop compared to previous round dummy, contract renewed, positive reciprocity, negative reciprocity, risk aversion. Controls might be extended with justification.
- For H4 we may extend the previous regression to include an interaction between wages and the positive reciprocity score.

No adjustment for multiple hypothesis testing will be made.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

Exclusions:

- Main analysis based on the full sample with no restrictions.
- Robustness check excluding decisions of participants top 10th percentile of cognitive inability score.

Outliers:

- No adjustments for outliers will be made.

7) How many observations will be collected or what will determine sample size?

Eligibility:

- Students from all study areas at the University of Nottingham who registered in the ORSEE system.

Sample size:

- 720 participants participating in 20 sessions of 36 participants each. Only full sessions count. If participants leave or there are technical problems during a session, we will conduct a new session.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

We will collect the following additional data:

- Socio-demographics: Age, gender, highest education, field of studies, area of birth
- Preferences: Risk-taking, positive reciprocity, negative reciprocity (following Table 1 in Falk et al., 2018
- Cognitive inability from passive instruction data: Cogn = 1(time < 20p) + 1(time > 80p) + Tries(Q1) + Tries(Q2) + Tries(Q3)

Chapter 3

Cross-border Mobility Responses to COVID-19 in Europe: New Evidence from Facebook Data

Citation: Docquier, F., Golenvaux, N., Nijssen, S., Schaus, P. & Stips, F. (2022). Cross-border mobility responses to COVID-19 in Europe: new evidence from Facebook data. *Globalization and Health*, 18(41). https://doi.org/10.1186/s12992-022-00832-6.

3.1 Background

Covid-19 is a disease induced by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first known cases of Covid-19 were registered in December 2019 in Wuhan (China), but the virus rapidly spread into many countries, leading to a global pandemic that has held the world captive for many months. As until recently no effective vaccine or medication was available, government policies were mainly targeted towards tracing and disrupting infection chains. Many countries introduced coercive measures and disincentives to limit within- and cross-border mobility of people in the hope of reducing the virus propagation. Whether these policies were effective or not remains an open question, and there is much to be gained from better understanding the evolution and determinants of people's mobility during Covid times.[65]

There is strong evidence that *within-border* (or internal) people's mobility declined during the Covid-19 crisis. Existing literature relies on big data provided by private cellular phone companies,[66] and documenting spatial movements in real time. Caselli et al. (2020) uses Vodafone data for Italy, Portugal and Spain, and finds that women and younger people show the largest drop in mobility. Chetty et al. (2024) combines telco data with household surveys to highlight a sharp decline in short-distance mobility, as proxied by daily time spent at parks, retail and recreation, grocery, transit locations, and workplaces. Using mobility data from the analytics company SafeGraph, Goolsbee and Syverson (2021) finds that the decline in mobility in New York and in four other U.S. cities is mostly driven by the fear of infection, rather than by legal restrictions. Although it also finds a significant impact of non-pharmaceutical interventions (NPIs) such as closing nonessential businesses, sheltering in place, and school closures, the dominant role of infection threats is confirmed by Maloney and Taskin (2020), who relies on Google mobility data.

By contrast, there is scant evidence of the impact of Covid-19 on *cross-border* (or

[65] We understand mobility as "the movement of human beings in space and time" Zhang and Wang (2021). This broad definition is in line with the mobility measure that we use in our quantitative analysis, as that includes all types of movement.

[66] By big data, we mean extremely large data sets that may be analyzed computationally to reveal patterns, trends and associations, especially those related to human behaviors and interactions, according to the *Oxford Languages* (https://languages.oup.com/google-dictionary-en/).

international) mobility, which is due to the absence of high-frequency data on border crossings.[67] In this paper we aim to fill this research gap by addressing the following research questions: (i) How has the Covid-19 impacted cross-border movements of people? (ii) Are these changes due to coercive measures (such as containment policies or international travel bans) or by the fear of contracting the virus?

Understanding the determinants of cross-border mobility responses to Covid-19 is important for economic and epidemiological reasons. Cross-border movements of people predominantly consists of labor commuting flows and business travels. Economically speaking, labor mobility is a key ingredient for growth and competitiveness in normal times. And in a pandemic context, restrictions placed on how workers move around can slow down economic recovery prospects, by making it more difficult for businesses to hire productive workers. They can also induce severe economic impacts on cross-border workers and their families. Epidemiologically speaking, the role that mobility is playing in the spread of the disease is still unclear. Using SafeGraph data for New York city and for other U.S. cities, Glaeser et al. (2022) find that (internal) mobility increased the spread of the disease in the early stage of the pandemic. Basellini et al. (2021) also finds a strong association between internal movements and mortality using Google mobility data for the UK. In the same vein, S. Chen et al. (2020) shows that (internal and international) travel bans enacted during the Chinese Lunar New Year holiday helped reduce the spread of the virus, and Ruktanonchai et al. (2020) argue that an appropriate coordination would considerably improve the likelihood of eliminating community transmission throughout Europe. By contrast, others expressed skepticism about the epidemiological consequences of travel bans, arguing that the impacts of these restrictions are not well understood Lee et al. (2020) or poorly effective Askitas et al. (2021), Chinazzi et al. (2020), Errett et al. (2020), Kraemer et al. (2020) and Tian et al. (2020) once patient zero has already spread the virus across regions.

Without taking any position on the fact that cross-border mobility should be limited

[67] The recent OECD migration outlook reveals that issuance of new visas and permits in OECD countries plummeted by 46% in the first half of 2020 (by 72% in the second quarter), as compared with the same period in 2019. However, international migrants and refugees account for a tiny proportion, not a say a negligible proportion, of daily cross-border movements of people between European countries. Daily flows predominantly consist of commuting workers and business travels.

or encouraged, we use a unique database on daily mobility of European Facebook users to shed light on the evolution of cross-border movements of people during an entire pandemic year, and to compare the effects of coercive policies with those related to the fear of infection. Our results contrast with those obtained for internal mobility. The following sections successively describe our data sources, methods and findings.

3.2 Data

3.2.1 Border crossings

Data on daily cross-border mobility is obtained from Facebook (denoted by FB, henceforth) for the period from the 29th of February 2020 to the 28th of February 2021 Facebook (2021). The database documents cross-border flows of FB users with *location services enabled*, who travel from an origin to a destination country by any means of transportation (car, train, air, etc.) during each 24-hour time period[68]. Only daily flows with a minimum of 1,000 movers are reported in the dataset. All flows below 1,000 are set to zero in order to minimize re-identification risk of FB users. This means that our outcome variable is left-censored. To limit the impact of censoring, we focus on 45 country pairs (involving 30 contiguous European countries) characterized by at least 25% of uncensored values of daily traffic during the period of observation. This selection limits the ability to generalize our results but is necessary to limit the impact of censoring, and allow smooth estimation with Machine Learning methods. Further limiting the impact of censoring, we use the 7-day centered rolling average of daily flows.

Although FB data has high coverage, FB users are not a random sample of the population. This raises concerns about representativeness. Figure 3..1 in the Appendix mitigates these concerns by showing a strong association between daily movements of FB users and the (estimated) number of daily border crossings in the pre-Covid-19 period, which are

[68] A number of recent studies such as Alexander et al. (2019) and Rampazzo et al. (2021) use data extracted from FB advertising API to identify migrants and model migration patterns. Note that our data source is different as we only have access to aggregate bilateral counts, without further details on movers' characteristics. The advantage of our data source is, however, that it is available on a daily basis.

presented in Appendix Table 3..1.[69] However Ribeiro et al. (2020) shows that FB users are over-represented in the population aged 20-40, with a high level of education and an above-average income level. It means that groups of individuals under 20 or over 65 and those with lower income/education levels are under-represented in FB data. Although people under 20 and over 65 form a tiny minority of the population under investigation (i.e., commuters and business travelers), this is a limitation of our work. FB users are more likely to belong to the richest and healthiest parts of the movers' population.

Let us denote by $M^F_{i\to jt}$ the count of FB movers from country i to country j at day t. When focusing on contiguous countries (i.e., the pairs of countries that exhibit the largest numbers of daily cross-border movements by far, and that are the least affected by censoring rules), the number of movers from i to j is almost identical to the number of movers from j to i at each day t (i.e., $M^F_{i\to jt} \simeq M^F_{j\to it}$). The reason is that border crossings predominantly consist of back-and-forth movements of commuting workers and business travelers, who move for short periods and for economic reasons[70]. This is also the case in the summer vacation period when considering a 7-day centered rolling average of daily flows. This means that $M^F_{i\to jt}$ and $M^F_{j\to it}$ are reflecting the same reality, and say nothing about the primary direction of the flows. For this reason, we define the level of *bilateral traffic* of FB users between countries i and j as:

$$T^F_{ijt} \equiv \textbf{Max}\left[M^F_{i\to jt}, M^F_{j\to it}\right], \tag{3.1}$$

and see it as a proxy for the scaled sum of the two unobserved unidirectional flows between the two countries, $\phi(M_{i\to jt} + M_{j\to it})$, where the scale factor ϕ denotes the fraction of FB users in the actual number of movers (denoted by $M_{i\to jt}$ and $M_{j\to it}$). Using the maximum as weighting scheme has the advantage to limit the impact of censoring in the case movers in one of the two directions are below the threshold. As $T^F_{ijt} = T^F_{jit}$, we can get rid of the dyadic dimension of the data, treat each country pair as a one-dimensional observation,

[69] We also find a strong association between the number of FB users and population size at the regional level. Results are available upon request.

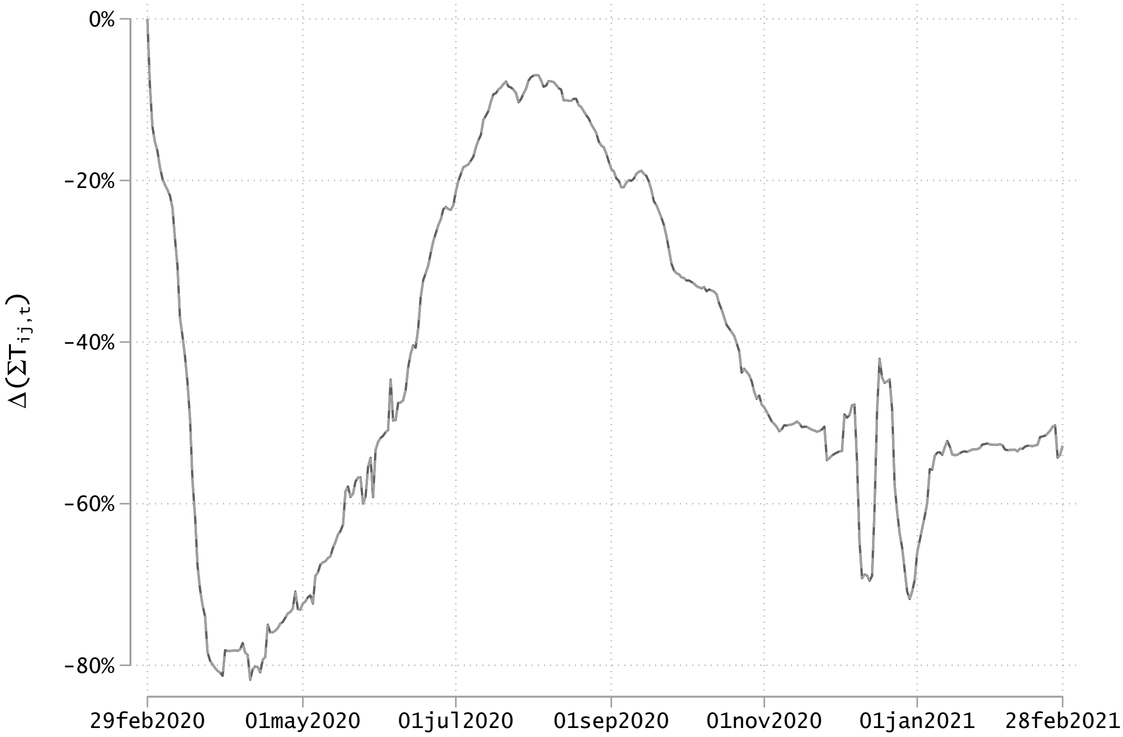
[70] Note that we only observe the total number of movers within a day, and neither the duration of their stay nor the direction of the movement. The vast majority of movers cross the border twice as they return to their origin country within a day or so. This is why we model cross-border traffic and use a trick below to account for the presumed direction of the flows.

and divide the size of the sample by two. In the methodological section, however, we explain how priors about the primary direction of the flows can be used to improve the quality of fit of our models.

To avoid dealing with re-scaling issues, we express traffic counts as relative deviations from their initial or pre-Covid-19 levels – in our case, the levels observed at the outset of the pandemic (denoted by day 0). We thus use the relative deviation in bilateral traffic between day 0 and day t, $\tau_{ijt} \equiv \frac{T_{ijt}-T_{ij0}}{T_{ij0}}$, as a variable of interest instead of focusing on the level of traffic T_{ijt} itself. Modelling relative deviations is also helpful to avoid over-fitting large corridors at the expense of small corridors, and mitigates representativeness issues even if the scale factor (ϕ) varies across country pairs.

Figure 3.2.1 portrays these relative deviations in the aggregate level of traffic between all country pairs included in our sample. The curve largely mirrors the three phases of the pandemic, depicting a stark drop in traffic in March 2020, a recovery during the spring and summer periods, and a new contraction in the post-summer period. Between end of February and early April 2020, the aggregate traffic level decreased from 720,000 to 130,000, implying a 82% drop. Aggregate traffic never fully recovered to the February levels in our period of observation. This also holds true during the summer vacation period when international travels were largely liberalized. The pace and strength of these changes vary across the three phases of the pandemic. The drop in March 2020 was strong and sudden, while the summer peak and the post-summer contraction were more gradual.

Figure 3.2.1: Aggregate Traffic Deviations from Pre-Covid Levels

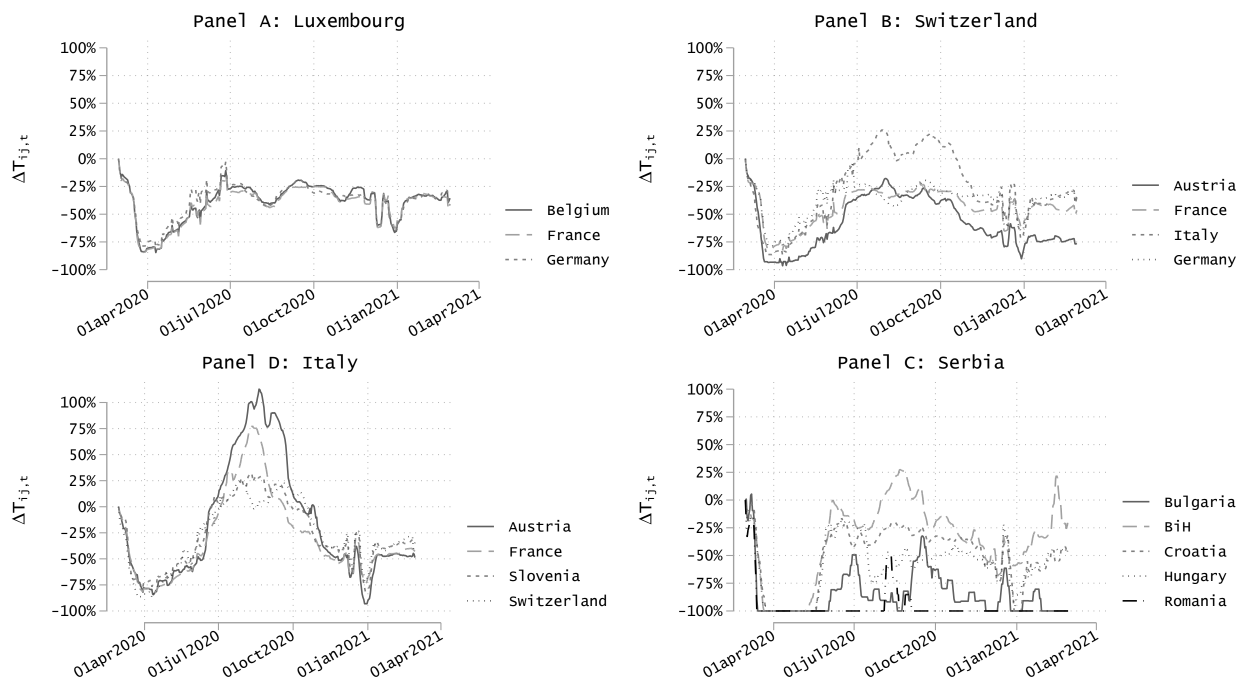


Source: Facebook data on daily border crossings. Notes: Y-axis represents the average of τ_{ijt}, the percentage change (times 100) in the 7-day moving average traffic compared to $t = 0$ over all destinations. The weights are the traffic levels observed in pre-Covid-19 period (i.e., $t = 0$).

Aggregate fluctuations mask large differences across country pairs. Bilateral traffic returned to its pre-Covid level in a minority of cases. For the majority of corridors, however, the traffic level has not fully recovered. This is illustrated in Figure 3.2.2, which depicts the evolution of people's traffic in corridors involving four open countries, namely Luxembourg, Switzerland, Italy, and Serbia. Luxembourg is the country with the highest share of cross-border workers in Europe. Given the economy's high reliance on cross-border workers, the Luxembourg government has never implemented international travel restrictions during the pandemic. Luxembourg experienced a significant drop in traffic in March 2020, whatever the partner country. After one month of lockdown, traffic levels recovered pretty quickly until reaching a plateau at about -25% since June 2020. Switzerland is the country with the largest number of cross-border commuters in Europe. This country experienced a larger drop during the first lockdown, and a slower recovery. Furthermore, the variability across corridors is considerably greater than in Luxembourg.

Italy has been severely impacted by the pandemic, and responded with national and international travel bans. We observe similar patterns of contraction and recovery during the first two quarters of 2020, followed by a substantial increase in traffic during the holiday summer period, and a second lockdown-type contraction in the post-Summer period. Finally, the patterns observed in Serbia are less conclusive as they are more severely affected by censoring rules. Serbia is an important origin and transit country for migrants and refugees entering the EU. Overall these patterns illustrate the need to account for corridor-specific heterogeneity when analyzing the determinants of bilateral traffic. Variations are likely to be influenced by seasonal effects, epidemiological risks, and policy measures implemented in the countries. We now turn to the description of the data sources used to proxy epidemiological conditions and the stringency of national policies.

Figure 3.2.2: Traffic Deviations from Pre-Covid Levels for Selected Corridors



Source: Facebook data on daily border crossings. Notes: Y-axis represents τ_{ijt}, i.e. percentage change (times 100) in corridor 7-day moving average traffic compared to $t = 0$ in corridor ij.

3.2.2 Explanatory Features

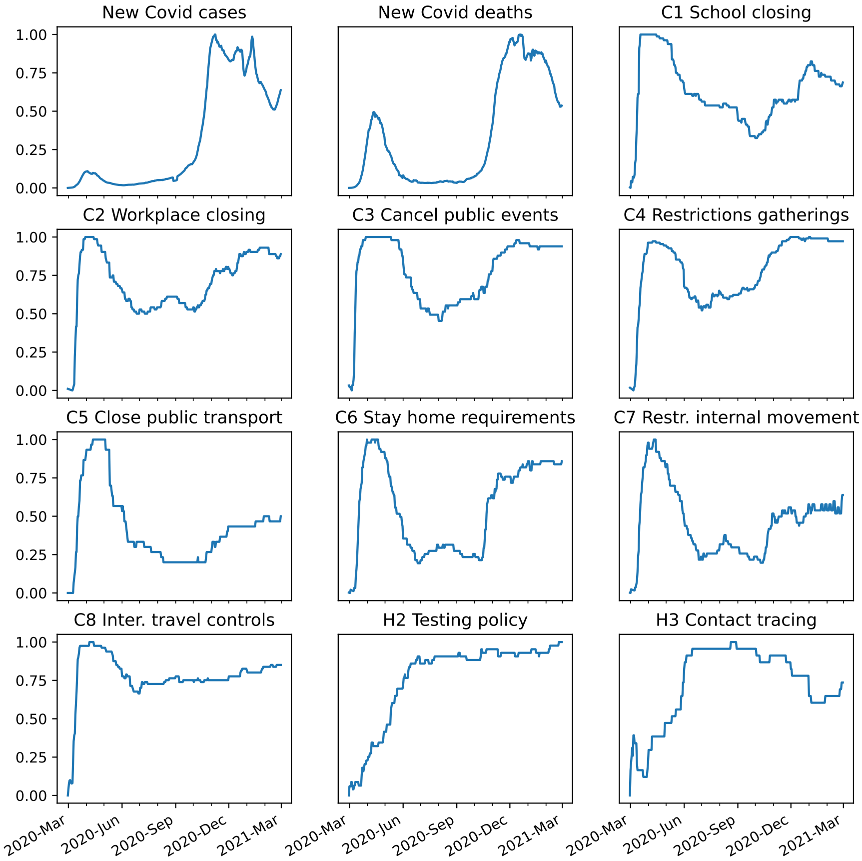
We link variations in cross-border traffic during the pandemic to daily changes in epidemiological conditions and containment policies in the countries involved. We proxy the severity of the pandemic with the daily numbers of new Covid-19 cases and new Covid-19 related deaths in each country using data from Dong et al. (2020). With regard to containment measures, we use data on daily policy responses from the Oxford Covid-19 Government Response Tracker (OxCGRT) Hale et al. (2020). The latter database consists of 18 ordinal indicators capturing the levels of nonpharmaceutical interventions (NPIs). Based on our priors as to which policies likely affect mobility, we choose to include the eight mobility-related measures that form the "containment and closure policies" block (denoted by C1-C8 in the database) as well as proxies for the intensity of testing and contact tracing (denoted by H2 and H3). We rescale all sets of predictors between 0 and 1, and align them with the definition of the outcome variable using the centered 7-day rolling average at each day.

For all features and days, Figure 3.2.3 displays the cross-country mean level of each containment index over time. Values close to one represent higher Covid-19 cases/deaths or more stringent responses. As containment policies were implemented in most countries during the second half of March 2020, maximum values are observed during this period. Testing and contract tracing were implemented more heterogeneously across countries and peaked in the summer of 2020. Reported number of new Covid-19 related cases/deaths were much greater during the second wave and peaked at the end of the year 2020.[71]

Figure 3.2.4 shows the cross-country correlations between each of the explanatory variables and the relative deviations in average traffic (i.e. deviation of the country-specific mean level of traffic with all potential partner countries in the sample). Containment policies are positively correlated with each other, and moderately correlated with epidemiological conditions, which allows us to include both sets of variables jointly in our regression and machine learning models. However, the fact that containment policies are correlated with each other raises concerns of multicollinearity, and motivates the usage of a limited number of synthetic policy indices. These indices are obtained by conducting a Principal Component Analysis (PCA) of all policy measures (C1-C8, and H2-H3) over the entire sample of observations, and by extracting the first two components. The first component mainly represents the C1-C8 measures which are strongly correlated with

[71] Epidemiological conditions are likely to subject to measurement errors. For example, testing and tracing practices played an important role in determining the number of detected cases.

Figure 3.2.3: Evolution of Covid cases and government policy measures

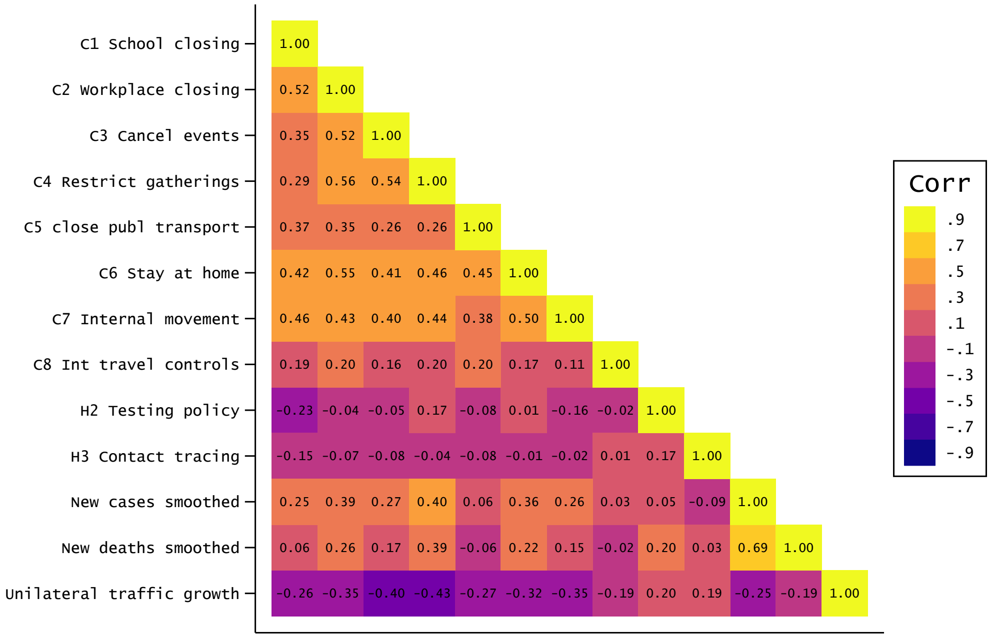


Source: Oxford Covid-19 Government Response Tracker (OxCGRT). Note: The values of each indicator is scaled between 0 and 1, and the average is computed using the 30 countries included in our sample.

each other while the second component corresponds to a higher variance for the H2-H3 measures. We will compare the results obtained when using a comprehensive specification, including all (collinear) explanatory variables, with those obtained when using a parsimonious specification, including the two synthetic PCA components as predictors.

Turning to the main correlations of interest (i.e., correlations between the relative deviations in cross-border traffic and each predictor), the figure displays a moderate negative correlation with containment policies, as well as low and positive correlations with testing and tracing policies. The strongest associations are obtained for cancellations of public events, restrictions on public gatherings, requirements to stay at home, and restrictions on internal movements. Two interesting observations arise from these partial

Figure 3.2.4: Correlation Matrix of the different variables



Source: Own computations. Notes: Unilateral traffic growth for each country i is the relative deviation in aggregate traffic involving country i, $\sum_{j=1}^{n} T_{ijt}$, as compared to the pre-Covid-19 period ($t = 0$).

correlations. First, the negative correlation between measures of epidemiological intensity (Covid-19 cases and deaths) and changes in total traffic are rather low. This might suggest that, in contrast with studies on within-border mobility, the fear of being infected might play a less important role in explaining changes in cross-border movements. Second, among containment policies, the implementation of international travel bans is far from being the most strongly correlated covariate. This suggests that travel bans might be effective in limiting non-essential travels, but less effective to limit labor commuting flows and business travels, which represent the overwhelming majority of daily border crossings between European countries (see Table 3..1). By contrast, cancellations of events and constraints on internal movements and gatherings are highly correlated with traffic variations, possibly because such constraints better proxy changes in economic activity and incentives to move. In the following section, we describe the regression and machine learning methods used to test these hypotheses using bilateral traffic growth as a dependent variable.

3.3 Methods

Mobility patterns identified in the previous section might result from various factors such as travel bans, sanitary measures influencing economic costs and incentives to move for work and business (e.g. sectoral lockdown, work-from-home practices) or for leisure (e.g. shops, restaurant and bar closings), or the fear of the virus itself. Our goal here is to identify the determinants of the relative deviation in daily traffic of people between country i and country j (τ_{ijt}), considering all NPIs and epidemiological daily indicators (x_{it} and x_{jt}) during the Covid-19 crisis. Our models are also used to predict the effects of epidemiological restrictions, NPIs and mobility restrictions on traffic counts.

We combine two analytical methods, Econometric Modelling (EM) and Machine Learning (ML). EM and ML techniques are generally used for different purposes. In EM, gravity models are used to explain human mobility flows between two countries.[72] EM models require imposing one analytical specification for the response function, which governs the derivatives of the dependent with respect to covariates. In the gravity specification, a large set of fixed effects are included to capture relevant confounders, and allow for identification of potential causal effects based on the so-called within variation. ML techniques are at the other extreme of the bias-variance tradeoff. They do not require strong analytical assumptions and allow, by design, to explore a larger set of regression functions including linear or polynomial combinations of the covariates. This increase of the so-called model *capacity* comes with two drawbacks. First, the models are more complex and are usually not easy to interpret. Contrary to EM, ML techniques might computationally suffer from the inclusion of large sets of control fixed effects. Second, there is always a risk of overfitting the training data and the identification of causation links is usually not an objective per se. Despite these difficulties, we use both EM and ML to identify converging messages, and the triangulation of results can serve to strengthen our evidence base.

We use four models exploring a broad range of learning techniques: (i) A gravity model based on the linear regression method Letouzé et al. (2009); (ii) A K-nearest neighbors

[72] Gravity models aim to predict bilateral flows between two entities based on their size, on their economic characteristics, and on the distance between them Beine et al. (2016).

method (KNN), which predicts the dependent variable by interpolation of its nearest observation neighbors in the training set Peterson (2009); (iii) A Gradient Boosting method (GBoost), whose predictions are based on a set of decision tree models Friedman (2001); (iv) A Multi-layer Perceptron (MLP), which is a classic neural network approach Glorot and Bengio (2010) and Hinton (1990). The last three models rely on different ML regressors, each based on a distinct type of technique. We assess the predictive performance of each model using the very same (and standard) cross-validation ML methodology. The goal of the study is not to design a forecasting model, but rather to identify the main determinants of mobility, and to investigate whether these different approaches generate converging findings. Therefore, instead of the validating our model on a particular sub-period (as is usually done to evaluate a time-series model), the observations composing the cross-validation folds are randomly chosen within the full sample. All models are implemented in Python via the Skicit-learn library Pedregosa et al. (2011).

3.3.1 Approaches with or without directional priors

Ideally, mobility models aim to characterize the evolution of the unidirectional flow of people ($M_{i\to jt}$) from an origin country i to a destination country j at day t, or of their relative deviation from the initial reference period ($\mu_{i\to jt} \equiv \frac{M_{i\to jt}-M_{i\to j0}}{M_{i\to j0}}$), based on a set of features available for the same time period. Without loss of generality, the general functional form f_M of such a model can be written as:

$$\mu_{i\to jt} = f_M(x_{it}, x_{jt}, d_{ij}, d_t) + \eta_{i\to jt} \tag{3.2}$$

where x_{it} represents the set of origin-specific determinant, x_{jt} is a set of destination-specific determinants, d_{ij} is a set of bilateral dummies capturing time-invariant bilateral resistance (including initial $M_{i\to j0}$, distance, language proximity, cultural proximity, etc.), d_t a set of day dummies capturing weekdays and seasonal trends (e.g. holiday season, general feeling of risk when traveling, etc.), and $\eta_{i\to jt}$ an error term. In our case, the vectors of explanatory variables x_{it} and x_{jt} capture the set of NPIs and epidemiological variables, and should also be interpreted as variations from period 0 since x_{i0} and x_{j0} are equal to zero in the pre-Covid-19 period.

With FB data, the primary direction of the cross-border flows is unknown, which implies that $M_{i\to jt}$ and $M_{j\to it}$ cannot be distinguished *a priori*. Instead, we observe the relative deviation in bilateral traffic, τ_{ijt}, and we have to estimate the function f_T linking bilateral traffic to the set of explanatory features without being able to distinguish between origin- and destination-specific determinants. A learning approach without directional priors writes as:

$$\tau_{ijt} = f_T(x_{it}, x_{jt}, d_{ij}, d_t) + \eta_{ijt} \tag{3.3}$$

It is possible, however, to discipline the model with priors about the direction of the flows. As bilateral traffic is a proxy for the sum of unidirectional flows ($T_{ijt} \simeq \phi(M_{i\to jt} + M_{j\to it})$), relative deviations in T_{ijt} can be expressed as a weighted sum of the relative deviations in unidirectional flows: $\tau_{ijt} = \omega_{i\to j,0} \times \mu_{i\to jt} + \omega_{j\to i,0} \times \mu_{j\to it}$, where $\omega_{i\to j,0} = 1 - \omega_{j\to i,0}$ is the pre-Covid-19 share of unidirectional cross-border flows from country i to country j in total traffic between the two countries. Estimates for $\omega_{i\to j,0}$ are constructed using pre-Covid-19 data on commuters, air travels and international migration, and then used as priors to discipline the model (these shares are depicted in Figure 3..2 in Appendix).

We can thus create two sets of weighted features, namely X^o_{ijt} for origin-specific effects, and X^d_{ijt} for destination-specific effects, defined as follows:

$$X^o_{ijt} = \omega_{i\to j0} x_{it} + \omega_{j\to i0} x_{jt}$$
$$X^d_{ijt} = \omega_{i\to j0} x_{jt} + \omega_{j\to i0} x_{it}.$$

The model with directional priors is obtained after replacing (x_{it}, x_{jt}) in Eq. (3.3) by (X^o_{ijt}, X^d_{ijt}). It writes as:

$$\tau_{ijt} = f_T(X^o_{ijt}, X^d_{ijt}, d_{ij}, d_t) + \eta^{od}_{ijt}. \tag{3.4}$$

If the true model for unidirectional flow ($f_M(.)$) was linear, plugging weighted covariates in the estimated model for bilateral traffic ($f_T(.)$) would allow retrieving the true origin- and destination-specific coefficients of interest accurately. Although this is not the case when $f_M(.)$ is non linear, using weighted covariates might improve the quality of fit or facilitate the interpretation of the results. The rationale is that the effect of policies depends on where they are implemented and on the primary direction of the flows. Suppose $\omega_{i \to j0} \simeq 1$ (i.e., flows mostly go from i to j), then an increase in restrictions/stringency at destination (resp. at origin) makes X^d_{ijt} positive (resp. X^o_{ijt} positive) and is more (resp. less) likely to reduce the flow of cross-border movements. Using directional priors allows approximating origin- and destination-specific effects without observing the direction of the flows during the pandemic year.

3.3.2 Permutation Feature Importance

To identify the main causes of daily mobility variations, the importance of each feature is computed for the different regression approaches with *permutation feature importance* A. Altmann et al. (2010). It has a the advantage of working similarly for all regression models considering them as a black-box models Molnar (2018). It is defined to be the decrease in the regression score when a single feature value is randomly shuffled across observations. More exactly, for a given model f, it first calculates a baseline score S_f provided by f when it is fitted, and then evaluated with a certain metric on the whole sample. Then for each possible feature x the modified score $S^*_{f,x}$ is computed by evaluating f on the transformed data set where the values of feature x are randomly permuted across all observations. The mean importance of the feature x for the model f is computed as:

$$I_{f,x} = \frac{1}{K} \sum_{k=1}^{K} \frac{S_f - S^*_{f,x,k}}{S_f} \tag{3.5}$$

where K is the number of random permutations realized for each feature, and $S^*_{f,x,k}$ is the $S^*_{f,x}$ score for the k^{th} permutation. In order to compare the importance values of the different models in an equivalent manner, the values $I_{f,x}$ are scaled between 0% and 100% separately for each model f. The mean features importance $I_{f,x}$ are computed over 10 permutations using the negative mean absolute error (MAE) and and the Root Mean Squared Error (RMSE). This means that the bigger $I_{f,x}$, the more permutations of the feature x degrades the quality of predictions for the model f and the feature is considered as more importantly associated with the target variable.

3.4 Results

We present our results in two steps. First, we assess the predictive power of the various models. This implies comparing learning methods with or without directional priors and with or without day/corridor control dummies. We compare their predictive performance by using out-of-sample predictions and computing the MAE and RMSE. Second, we use the estimated models to rank the importance of different features relying on permutation techniques. Using multiple models allows assessing the robustness of our findings.

3.4.1 Validation of models

We first investigate whether adding priors about the direction of the flows and/or adding a full bunch of day and corridor dummies improves the performance of our learning models. Models without directional priors are described in Eq. (3.3), while models with priors use weighted regressors, as described in Eq. (3.4). A 10-fold cross-validation over the whole data set (from the 29$^{\text{th}}$ February 2020 until the 28$^{\text{th}}$ February 2021) is realized for each model to assess its performance. Table 3.4.1 reports the MAE, RMSE and their standard error across cross-validation folds obtained with different learning methods.

It shows that directional priors ($\omega_{i\to j,0}$ and $\omega_{j\to i,0}$) do not bring significant additional predictive power under most learning approaches when day and corridor dummies are not factored in (see Panels A and B). The only exception is the G-Boost method. On the contrary, when day and corridor dummies are included (Panels C and D), adding directional priors slightly improve the quality of fit with virtually all learning techniques (except with KNN). In addition, we show below that distinguishing between origin- and destination-specific effects as in Eq. (3.4) makes the interpretation of the results much easier. Therefore, the model with directional priors will be prioritized in the rest of the analysis.

Second, we investigate whether the inclusion of day-specific effects – i.e., 366 time dummies, d_t, that are common to all corridors and capture unobserved variations such as seasonal changes, synchronized fears of infection, etc. – and corridor-specific effects – i.e., 45 corridor dummies, d_{ij}, that are time invariant and capture unobserved variations

Table 3.4.1: MAE comparison of the different models with or without directional priors and dummies

	Panel A: No dummies - No prior				**Panel B: No dummies - Priors**			
	Linear	KNN	G-Boost	MLP	Linear	KNN	G-Boost	MLP
avg MAE	0.194	0.018	0.073	0.051	0.201	0.019	0.042	0.057
std MAE	(0.009)	(0.001)	(0.001)	(0.005)	(0.005)	(0.001)	(0.001)	(0.003)
avg RMSE	0.285	0.042	0.106	0.083	0.287	0.043	0.064	0.089
std RMSE	(0.017)	(0.009)	(0.003)	(0.011)	(0.009)	(0.005)	(0.002)	(0.008)
	Panel C: Dummies - No prior				**Panel D: Dummies - Priors**			
	Linear	KNN	G-Boost	MLP	Linear	KNN	G-Boost	MLP
avg MAE	0.135	0.020	0.050	0.041	0.134	0.020	0.049	0.038
std MAE	(0.005)	(0.001)	(0.002)	(0.003)	(0.005)	(0.001)	(0.001)	(0.003)
avg RMSE	0.210	0.045	0.081	0.068	0.203	0.047	0.077	0.064
std RMSE	(0.010)	(0.006)	(0.006)	(0.006)	(0.009)	(0.005)	(0.002)	(0.005)

Note: The table compares the performances of the 4 different approaches (Linear, KNN, G-Boost and MLP) with and without directional priors ($\omega_{i\to j,0}$), and with or without day/corridor dummies (d_t and d_{ij}). Errors are computed from a 10-fold cross-validation on the whole sample.

such as the skill level of the cross-border workforce, linguistic and cultural proximity between countries, etc. – improves the predictive power of our models. Again, we perform another 10-fold cross-validation on different versions of each model. Panels C and D in Table 3.4.1 includes both sets of dummies jointly, with or without directional priors. In the absence of directional priors, the inclusion of day and corridor dummies reduces the MAE and RMSE by 20 to 30% whatever the learning technique used. When directional priors are factored in, the dummies improve the performance of the linear and MLP models, whereas they deteriorate the quality of fit under the KNN and G-Boost models. This is because adding day and corridors dummies drastically increases the number of parameters to be estimated, and some ML approaches (like KNN) are known to suffer from the curse of dimensionality.

To further explore this issue, Table 3..2 in Appendix considers the model with directional priors and adds one set of dummies at a time. The inclusion of 45 corridor-specific dummies always improves the quality of fit. On the contrary, the inclusion of 366 day-specific dummies deteriorates the performance of KNN and G-Boost methods. This confirms that the gains from adding information about unobserved common time trends, which might already be captured by the relatively well synchronized trends in observed epidemiological conditions and containment measures, is outbalanced by the costs linked to the inflated dimensionality of the computation problem.

Third, ML techniques always outperform the linear EM model. This result was also expected given that ML is based on more complex prediction methods that allow for non-linear relationships between variables, and account for non-stationary variations contained in the matrices of X^o_{ijt} and X^d_{ijt}. The KNN always produces the best quality of fit. The error of this approach is minimal when the number of neighbors k used to estimate the relative deviations in traffic is low (say, 2 or 3). Its impressive performance in 10-fold cross-validation can be explained by the fact that the model finds a small number of observations for which the relative deviations in traffic are similar to those that must be predicted. In general, the closest neighbors are observations of the days preceding or following the daily level of traffic observed in the same corridor.

3.4.2 Main sources of variations in cross-border mobility in Covid times

In order to identify the government policy indicators having the greatest impact on relative deviations in traffic, the importance of each feature is computed for the different approaches involving directional priors and dummies. Directional priors allow us to distinguish the effects of origin-specific features from those of destination-specific features. Table 3.4.2 presents the results from this exercise. Features are ranked by decreasing order on the basis of the average predictive power across the four learning techniques. The column 'Avg.' gives the mean value of error metric averaged over the four models. Panel A provides the results obtained with the saturated models including the large set of corridor and day dummies, which is the first-best model when using linear and MLP learning techniques. Panel B gives the results obtained with corridor dummies only, which is the first-best model when using the KNN and G-Boost techniques. Results of Panel B will be discussed in the next section. In the top part of the table, the models use all individual features depicted in Figure 3.2.3, despite the high level of correlation between some of them. In the bottom part of the table, the models use synthetic containment features derived from a PCA analysis.

When considering all individual features, the ranking based on their predictive power varies across models. We identify, however, several common and interesting findings.

Table 3.4.2: Feature ranking by origin and destination

Features	Panel A Corridor & Day dummies Linear	KNN	G-Boost	MLP	**Avg.**	Panel B Corr. dum. **Avg.**
Indiv. features						
Origin - C1 School closures	100	69	100	59	82	85
Destin - C1 School closures	94	60	36	85	68	68
Destin - C3 Cancel public events	19	64	77	68	57	57
Origin - C3 Cancel public events	10	100	33	65	52	46
Origin - C7 Restr. Internal movement	0	93	16	100	52	51
Origin - H2 Testing policy	14	7	87	92	50	43
Origin - C4 Restrictions gatherings	50	51	40	54	49	44
Origin - C6 Stay home requirements	92	16	12	59	45	21
Destin - C6 Stay home requirements	17	15	96	54	45	45
Destin - C4 Restrictions gatherings	6	48	70	25	37	42
Destin - C7 Restr. Internal movement	13	66	12	58	37	29
Origin - H3 Contact tracing	5	12	27	44	22	22
Origin - C8 International travel bans	1	1	45	40	22	18
Origin - New Covid deaths	0	7	73	5	21	46
Origin - New Covid cases	11	0	21	52	21	75
Destin - New Covid deaths	6	10	64	0	20	49
Destin - C5 Close public transport	34	6	1	39	20	9
Destin - H3 Contact tracing	31	21	2	14	17	7
Destin - C2 Workplace closing	5	33	7	23	17	21
Destin - C8 International travel bans	0	12	20	32	16	24
Destin - New Covid cases	12	0	43	6	15	38
Origin - C5 Close public transport	18	7	0	31	14	12
Origin - C2 Workplace closing	1	23	10	14	12	23
Destin - H2 Testing policy	9	0	2	10	4	6
Synthetic features						
Destin - Component 1	100	100	100	100	100	100
Origin - Component 1	13	75	20	49	39	48
Origin - Component 2	1	55	17	19	23	26
Destin - Component 2	9	50	0	0	15	17
Origin - New Covid cases	0	0	18	30	12	75
Destin - New Covid deaths	5	14	6	11	9	49
Origin - New Covid deaths	1	13	3	9	6	46
Destin - New Covid cases	0	2	1	6	2	38

Notes: The different features are ranked following the permutation importance method. For each approach, we provide results obtained with the model including day/corridor dummies (cols. 1-5) and the version including corridors dummies only (col. 6). Directional priors are always used to identify the effects of origin- and destination-specific features. The importance values of each feature is computed over 10 permutations using the negative mean absolute error (MAE). The resulted values are scaled between 0% and 100% separately for each model. The col. 'Avg.' averages the results obtained with the four learning techniques. The features are ranked according to the average importance of the models including the day/corridor dummies (Panel A). In Panel B, we only report the 'Avg.' score without reporting the model-specific results.

First, school closures in the origin country has the largest average impact on the variation in daily traffic. Remember that cross-border traffic predominantly consists of labor commuting flows and business travels. School closures at origin imply that many parent workers are forced to take parental leave and cannot commute to work. In the same vein, school closures in the destination country are also paralyzing economic activity in the destination country and reduce incentives to move. This result is in line with Vannoni et al. (2020), who finds that school closures were among the most important predictors of internal mobility in March 2020. *Second*, variations in traffic are mostly impacted by containment measures. Based on the average predictive power (col. 'Avg.'), ten out of the twelve most predictive features involve C-type containment measures implemented in the origin or destination countries. *Third*, the fear of being infected in the destination country, as proxied by the destination-specific number of Covid-19 deaths and cases (appearing in bold characters in Table 3.4.2) are among the least predictive features. *Fourth*, international travel bans in origin and destination countries (appearing in italics in Table 3.4.2) also have a low predictive power. We thus conclude that cross-border daily flows are marginally influenced by the fear of infection and international travel bans. They are mostly governed by the stringency of internal containment policies and by family constraints. It is worth noticing that models without directional priors deliver very similar results, as illustrated in Table 3..3 in the Appendix.

In addition to school closures, the top panel of of Table 3.4.2 suggests that the most important containment measures are the cancellation of public events, restrictions on gatherings, restrictions on internal movements and stay-home requirements. In addition, the twelve most predictive features include 5 destination-specific and 7 origin-specific measures. However, as illustrated in Figure 3.2.4, these measures are highly correlated at the national level. Hence, instead of feeding the model with correlated features, the bottom part of the table uses synthetic indices of containment and sanitary measures. We use a PCA analysis to reduce the dimensionality of the origin- and destination-specific containment measures and we extract the first two components of the PCA.

Remember that the first PCA component can be interpreted as an average index of stringency of containment measures (i.e. C1-C8 indices); the second component captures testing and tracing policies (i.e. H2 and H3). The results clearly reveal that the stringency of containment measures in the destination country has, by far, the greatest predictive power. The average stringency of containment measures at origin is the second most predictive power, with an average importance equal to 40% of that of the destination country. This comforts the idea that cross-border daily flows of people mostly involve economic/essential movements which can only be influenced by changes in incentives to move or coercive mobility constraints. In line with the top part of the table, variables influencing the fear of infection have negligible impacts on border crossings.

3.5 Discussion

The EM and ML techniques used in this paper allow highlighting a strong association between the evolution of bilateral traffic between contiguous countries and containment policies in the destination country as well as school closures in the origin country. Association does not imply causation. It could be argued that this statistical association is governed by the influence of unobserved characteristics affecting both policy changes and mobility simultaneously, or that a reverse causation mechanism operates (i.e. cross-border mobility influences policy reforms). The fact that our results are robust to the inclusion of a large set of day and corridor dummies capturing unobserved time- and corridor-specific characteristics strongly mitigates the first misspecification concern.

With regard to reverse causation, concerns are mitigated by the use of high-frequency data. We cannot reject the possibility of a mobility-driven propagation of the virus requiring new containment measures. However, such a mechanism takes time to operate. Mobility shocks at day t do not generate immediate and visible epidemiological consequences, and policy responses are also implemented with a certain delay. By contrast, our estimates suggest that containment policies are immediately associated with changes in cross-border mobility. This prudently supports the existence of a causation effect of containment measures on mobility.

An opposite argument that goes against the reverse causation issue is that it also takes time for information about epidemiological conditions to be assimilated by potential movers. Hence, the fear of infection could be better proxied by the lagged numbers of Covid-19 cases and deaths. Our results are preserved and even reinforced when using lagged proxies for the fear of infection. More precisely, traffic at day t is very badly predicted by the number Covid-19 cases and deaths observed one or two weeks before the date (see Table 3..4 in the Appendix). Hence, we can reasonably rule out that the low impact of infection fears at destination is driven by a misspecification problem.

In the same vein, it could be argued that fears are strongly synchronized across countries and captured by the day dummies. These concerns are mitigated by the fact that removing the day-specific dummies does not alter our conclusions. In Panel B of Table

3.4.2, day dummies are excluded. The only significant change is that the number of Covid cases in the origin country has a greater predictive power, possibly implying that people in sick leave self-isolate and stop moving. However, the predictive power of international travel bans (in italics) and fears to be infected at destination (epidemiological conditions at destination are in bold characters) remain low when using both individual and synthetic features. Again, the stringency of containment measures in the destination country has the greatest predictive power by far.

Finally, one may fear that our results can be affected by the quality of data as well as their comparability across countries (see Stoto et al. (2022) for a discussion of the strengths and weaknesses of data sources on Covid-19). Data on Covid-19 threats and NPIs (such as the number of new cases and deaths) are likely to be subject to country and time biases due to multiple reasons (e.g. changes in testing strategies, seasonal effects, variations in reporting delays or in the classification of Covid-related deaths, etc.). It is worth emphasizing, however, that the *Permutation Feature Importance* technique that we use is, by construction, robust to noise.[73] In addition, the systematic use of bilateral and time dummies (d_{ij} and d_t) allows us to mitigate the biases caused by unobserved variations in the quality of data across country pairs and periods.

3.6 Conclusion

Existing literature shows that people within-border mobility has drastically declined in times of Covid-19, primarily because of the fear to be infected in parts of the population. To the best of our knowledge, our study is the first to analyze the effect of Covid-19 and related containment measures on people's cross-border movements. In line with the findings above, we also document a sharp decline in cross-border mobility in general, especially during the first lockdown and in the second and third waves of the pandemic. However, these variations in cross-border mobility are mostly induced by local containment policies in the destination country, and school closures in both countries. The fear of infection and international travel bans have little influence on cross-border movements.

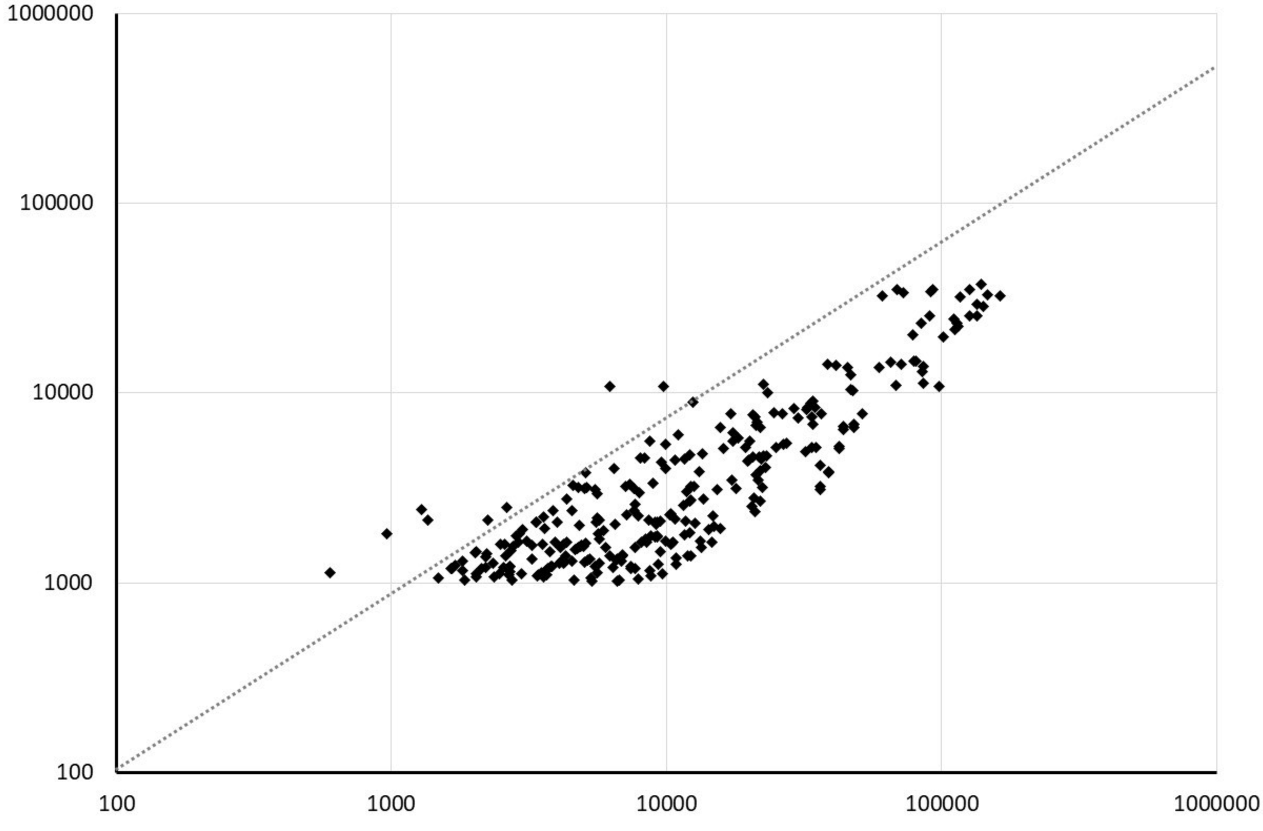
The likely reason is that cross-border daily flows of people are predominantly made

[73] It could be further supplemented with a leave-one-out sensitivity analysis on countries Stoto et al. (2022).

of commuting workers and business travelers who move for economic/essential reasons. These economic flows are observed between contiguous countries, and account for 99% of international movements of people when compared with the flows of migrants and refugees. Their magnitude varies with the economic costs and incentives of moving, which depend on lockdown measures and on the stringency of internal containment policies. In addition, international travel bans do not apply to commuters and businessmen. Although there is no consensus on the fact that these flows contribute to the propagation of the virus, policy-makers must be aware that economic movers hardly adapt their mobility decisions to epidemiological threats. Border crossings can only be controlled with internal coercive policies.

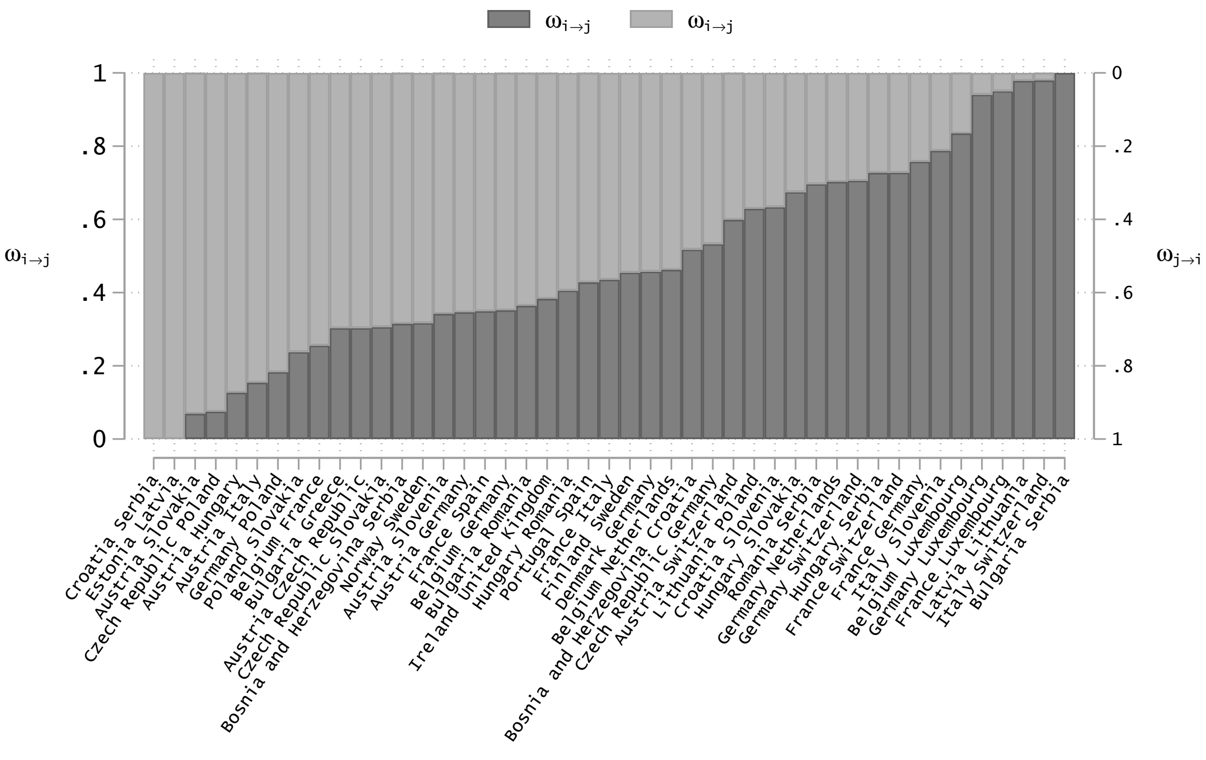
Appendix

Figure 3..1: Daily traffic estimates and FB data by corridor (1 March 2020)



Sources: Numbers of daily commuters are extracted from Eurostat data by Nuts2 region in 2019; Numbers of air passengers are extracted from Eurostat monthly statistics on air passenger transport on March 1, 2020; Data on international migrants are extrapolated from Abel and Cohen, 2019 for the year 2015, assuming a conservative 50% growth in the flows between 2015 and 2020. FB data are the Facebook data on daily border crossings on March 1, 2020. Note: All variables are expressed in logs.

Figure 3..2: Distribution of Weights



Sources: Numbers of daily commuters are extracted from Eurostat data by Nuts2 region in 2019; Numbers of air passengers are extracted from Eurostat monthly statistics on air passenger transport on March 1, 2020; Data on international migrants are extrapolated from Abel and Cohen, 2019 for the year 2015, assuming a conservative 50% growth in the flows between 2015 and 2020. Notes: Countries on X-Axis are ordered as $country_i$ $country_j$, thus linked to $\omega_{i\to j}$. The opposite weight $\omega_{j\to i} = 1 - \omega_{i\to j}$ refers to the weight of $country_j$ $country_i$.

Table 3..1: Estimates of the average count of daily cross-border outflows in 2019

	Daily count estimates				**As percent of Total**		
	Commuters	Air pass.	Migrants	Total	Commuters	Air pass.	Migrants
AUT	159678	169395	441	329514	0.484	0.514	0.001
BEL	133742	241103	729	375574	0.356	0.641	0.001
DNK	44732	45719	77	90528	0.494	0.505	0.000
FIN	22532	21311	195	44038	0.511	0.483	0.004
FRA	396457	417902	670	815029	0.486	0.512	0.000
DEU	472160	361659	2770	836589	0.564	0.432	0.003
GRC	1714	24700	4	26418	0.064	0.934	0.000
IRL	23892	108689	61	132642	0.18	0.819	0.000
ITA	139103	115605	952	255660	0.544	0.452	0.003
LUX	152407	4829	76	157312	0.968	0.030	0.000
NLD	102000	267917	296	370213	0.275	0.723	0.000
PRT	55785	87144	157	143086	0.389	0.609	0.001
ESP	79857	346395	36	426288	0.187	0.812	0.000
SWE	51000	87565	91	138656	0.367	0.631	0.000
GBR	76300	465893	873	543066	0.140	0.857	0.001
EU15	1911359	2765826	7428	4684613	0.408	0.590	0.001
BIH	17017	21296	0	38313	0.444	0.555	0.000
BGR	52571	33019	42	85632	0.613	0.385	0.000
HRV	36557	28639	33	65229	0.560	0.439	0.000
CZE	122160	62294	156	184610	0.661	0.337	0.000
EST	10642	18674	6	29322	0.362	0.636	0.000
HUN	100196	62174	217	162587	0.616	0.382	0.001
LTU	2414	14435	0	16849	0.143	0.856	0.000
NOR	8232	50767	178	59177	0.139	0.857	0.003
POL	205314	135350	109	340773	0.602	0.397	0.000
ROU	65550	81892	207	147649	0.443	0.554	0.001

Sources: Numbers of daily commuters are extracted from Eurostat data by Nuts2 region in 2019; Numbers of air passengers are extracted from Eurostat monthly statistics on air passenger transport in February 2019; Data on international migrants are extrapolated from Abel and Cohen (2019) for the year 2015, assuming a conservative 50% growth in the flows between 2015 and 2020.

Table 3..2: MAE comparison of the different models with and without day/corridor dummies

	Panel A: No dummies				**Panel B: Day dummies**			
	Linear	KNN	G-Boost	MLP	Linear	KNN	G-Boost	MLP
avg MAE	0.201	0.019	0.042	0.057	0.179	0.181	0.056	0.059
std MAE	(0.005)	(0.001)	(0.001)	(0.003)	(0.004)	(0.005)	(0.001)	(0.003)
avg RMSE	0.287	0.043	0.064	0.089	0.259	0.269	0.084	0.090
std RMSE	(0.009)	(0.005)	(0.002)	(0.008)	(0.008)	(0.008)	(0.002)	(0.004)
	Panel C: Corridor dummies				**Panel D: Corridor & Day dummies**			
	Linear	KNN	G-Boost	MLP	Linear	KNN	G-Boost	MLP
avg MAE	0.156	0.018	0.037	0.048	0.134	0.020	0.049	0.038
std MAE	(0.006)	(0.001)	(0.001)	(0.004)	(0.005)	(0.001)	(0.001)	(0.003)
avg RMSE	0.231	0.043	0.059	0.086	0.203	0.047	0.077	0.064
std RMSE	(0.010)	(0.005)	(0.002)	(0.009)	(0.009)	(0.005)	(0.002)	(0.005)

Notes: The table compares the performances of the 4 different approaches with and without the day- and corridor-specific dummies. All models are estimated with directional priors. Errors are computed from a 10-fold cross-validation on the whole data set.

Table 3..3: Averaged feature ranking across models and specifications without directional priors

Features	Panel A: Corridor & Day dummies					Panel B: Corridor dummies				
	Linear	KNN	G-Boost	MLP	**Avg.**	Linear	KNN	G-Boost	MLP	**Avg.**
C1 School closures	100%	78%	98%	91%	93%	100%	100%	86%	86%	93%
C3 Cancel public events	14%	100%	79%	84%	69%	21%	100%	60%	76%	64%
C6 Stay home requirement	56%	18%	78%	71%	56%	25%	23%	64%	38%	38%
C7 Restr. Internal movement	6%	97%	20%	100%	56%	1%	84%	27%	100%	53%
C4 Restrict gatherings	28%	60%	80%	49%	54%	9%	76%	65%	64%	53%
H2 Testing policy	11%	4%	64%	64%	36%	26%	2%	49%	30%	36%
New Covid deaths	3%	10%	100%	0%	28%	0%	77%	100%	0%	44%
C8 Inter. travel controls	0%	8%	47%	44%	25%	3%	17%	44%	32%	24%
H3 Contact tracing	18%	20%	20%	35%	23%	12%	1%	19%	32%	16%
New Covid cases	11%	0%	46%	35%	23%	11%	76%	63%	87%	59%
C5 Close public transport	26%	7%	0%	43%	19%	15%	0%	0%	31%	11%
C2 Workplace closing	2%	34%	12%	21%	17%	2%	46%	23%	33%	26%

Notes: The different features are ranked following the permutation importance method. For each approach, we provide results obtained with the model including day/corridor dummies (cols. 1-5) and the version including corridors dummies only (cols. 6-10). Directional priors are not included. The importance values of each feature is computed over 10 permutations using the negative mean absolute error (MAE). The origin- and destination-specific features importance are aggregated by taking the mean between the 2. Finally, the resulted values are scaled between 0% and 100% separately for each model. The last column in each panel presents the mean value of importance averaged over the four models. The features are ranked according to the average importance of the models including the corridor and day dummies.

Table 3..4: Feature ranking including lagged epidemiological conditions

Features	Panel A: Corridor & Day dummies					Panel B: Corridor dummies				
	Linear	KNN	G-Boost	MLP	**Avg.**	Linear	KNN	G-Boost	MLP	**Avg.**
C1 School closing	100%	69%	81%	68%	79%	100%	83%	79%	97%	90%
C3 Cancel public events	10%	100%	56%	78%	61%	19%	100%	58%	77%	63%
C7 Restr. Internal movement	5%	89%	22%	100%	54%	4%	74%	20%	100%	49%
C6 Stay home requirements	60%	20%	79%	55%	54%	27%	24%	74%	31%	39%
C4 Restrictions gatherings	12%	57%	100%	33%	51%	5%	74%	100%	32%	53%
H2 Testing policy	12%	6%	69%	54%	35%	20%	2%	67%	29%	30%
H3 Contact tracing	23%	15%	37%	29%	26%	21%	0%	34%	34%	22%
C5 Close public transport	25%	11%	7%	32%	19%	14%	5%	4%	33%	14%
C8 Inter. travel controls	0%	7%	30%	28%	16%	11%	16%	28%	21%	19%
C2 Workplace closing	0%	32%	8%	19%	15%	2%	41%	5%	12%	15%
New Covid cases $t-7$	12%	0%	8%	12%	8%	14%	34%	6%	46%	25%
New Covid deaths t	4%	12%	12%	0%	7%	10%	49%	13%	12%	21%
New Covid cases t	1%	1%	0%	24%	7%	4%	41%	0%	68%	28%
New Covid cases $t-14$	4%	1%	9%	8%	6%	6%	33%	8%	0%	12%
New Covid deaths $t-14$	0%	12%	6%	5%	6%	0%	53%	7%	21%	20%
New Covid deaths $t-7$	1%	10%	4%	3%	5%	8%	48%	4%	44%	26%

Notes: The different features are ranked following the permutation importance method. The importance values of each feature is computed over 10 permutations using the negative mean absolute error (MAE). Four new variables are inserted in addition to the ones includes in the regression: 7- and 14-days of new Covid cases and deaths are included. For each approach, we provide results obtained with the model including day/corridor dummies (cols. 1-5) and the version including corridors dummies only (cols. 6-10). Directional priors are not included. The origin- and destination-specific features importance are aggregated by taking the mean of the 2. Finally, the resulted values are scaled between 0% and 100% separately for each model. The last column in each panel presents the mean value of importance averaged over the four models. The features are ranked according to the average importance of the models including the corridor and day dummies.

Chapter 4

A Replication of Jones & Marinescu (2022)

Citation: Bacher, E., Herrera-Rodriguez, M., Marino Fages, D., & Stips, F. (2023). A replication of Jones & Marinescu (2022). I4R Discussion Paper Series, No 80. https://hdl.handle.net/10419/278842.

4.1 Introduction

This paper replicates the paper by Jones and Marinescu (2022), henceforth JM, who study the effects of a universal cash transfer introduced in Alaska in 1982. The authors explore whether this policy led to a negative employment response. The main dataset is individual-level data from the Current Population Surveys (CPS), provided either by IPUMS or MORG, which the authors aggregate to the state level.

To analyze the effects of this policy, they use the synthetic control method. This method consists in the creation of a synthetic state using a weighted combination of several untreated states. This synthetic state is built so that it matches the trend of the outcome of the treated unit, Alaska, during the pre-treatment period. If the pre-treatment fit is good, then the synthetic state can be considered as a credible counterfactual for Alaska in the post-treatment period, and the treatment effect can be estimated as the difference between the two trends in the post-treatment period. The authors used data from 1977 to 2014, although data for one outcome (the number of hours worked in the previous week) were only available since 1979. The main result of the paper is that the authors do not find a negative employment response to the cash transfer but an increase in part-time work.

We conduct a series of exercises to study the reproducibility and robustness of the conclusions of JM. First, we check whether the replication package allows us to reproduce the main results of the paper. We also reproduce key parts of the synthetic control estimation in R to check whether the results are sensitive to the choice of software. Next, we repeat the empirical estimation using different covariates. We also vary the post-treatment period to assess how it affects the magnitude of the treatment effect. Moreover, we used two different estimation methods derived from the synthetic control framework to compare their results with those in the paper. Finally, we separate the placebos used in the paper between time and unit placebos.

4.2 Reproducibility

As a first step, we checked whether the Stata code could run as-is and whether its outputs correspond to the results reported in the paper. We found three typos in the code, which were easy to fix and did not compromise the reproducibility of the paper. Two input datasets required to recreate the main datasets were missing from the replication package. The data were provided by the authors upon request. With these changes, we were able to reproduce the results of the paper.

Table 4.2.1: Replication of Table 1 using the R package `tidysynth`.

	Alaska	Employment rate	Labor-force participation	Part-time rate	Hours worked last week
IPUMS data					
employed	0.639	0.639			
activelf	0.712		0.707		
parttime	0.103			0.104	
age1	0.108	0.102	0.098	0.099	
age2	0.154	0.135	0.125	0.124	
age3	0.691	0.642	0.672	0.672	
female	0.503	0.51	0.504	0.506	
ind1	0.361	0.361	0.337	0.359	
ind2	0.097	0.099	0.087	0.092	
ind3	0.035	0.063	0.064	0.034	
ind4	0.191	0.185	0.18	0.18	
ind5	0.078	0.119	0.161	0.152	
educ1	0.229	0.242	0.263	0.281	
educ2	0.396	0.388	0.42	0.397	
MORG data					
hourslw	37.98				37.93
age1	0.074				0.078
age2	0.155				0.139
age3	0.759				0.746
female	0.435				0.411
ind1	0.148				0.176
ind2	0.051				0.129
ind3	0.292				0.278
ind4	0.123				0.134
ind5	0.385				0.283
educ1	0.11				0.191
educ2	0.387				0.384

Note: Column 2 shows the average value of the covariates in the pre-treatment period for Alaska (treated unit). Columns 3-6 show these values for the synthetic control.

Next, we tried to reproduce the main findings using R (Team, 2022). Two packages were used for this: `Synth` (v. 1.1-6, Abadie et al., 2011) and `tidysynth` (v. 0.2.0, Dunford, 2023). The objective was to reproduce the results from the balance checks in Table 1 and the main results in Table 2 of the original paper. Table 4.2.1 shows the summary statistics for Alaska and the synthetic control obtained with R. The first panel gives the averages for Alaska, which are identical to those of the original paper. The next rows show the summary statistics for the placebo states used for each of the four outcomes. We can observe minor differences in the averages for employment, labor force participation, and part-time employment as compared to the averages in the paper. These differences are due to the fact that the estimation in R assigns slightly different weights. However, the differences are rather minimal, occurring at the second or third decimal place. The differences for the hours worked last week outcome are slightly larger, indicating larger differences in the matching for that outcome.

Table 4.2.2 shows the results from Table 2 of the original paper and, below, from our replication using R. For the first two outcomes, employment rate and part-time rate, we find similar results: the coefficients are very close, the statistical significance is the same as in the original paper, and the pre-period RMSE are almost identical. However, we find different results for the effect of the unconditional transfer on the participation in the labor force and on the number of hours worked in the previous week. For labor force participation, we find an average post-treatment effect twice larger than JM (0.027 in our replication against 0.012), and most importantly we find that this effect is statistically significant at the 10% level. On the other hand, for the number of hours worked in the previous week, we find a smaller and statistically insignificant effect (-0.442 against -0.796). The results for `tidysynth` and `Synth` are very similar.

Table 4.2.2: Replication of Table 2 using the R package `tidysynth`.

	Employment rate	Part-time rate	Labor-force participation	Hours worked last week
Original				
Average effect	0.001	0.018	0.012	-0.796
p-value	0.942	0.020	0.331	0.084
RMSE	0.005	0.003	0.013	0.394
Replication with R				
Average effect	0.004	0.017	0.027	-0.442
p-value	0.593	0.008	0.052	0.142
RMSE	0.004	0.001	0.009	0.058

The small differences in coefficients probably come from different optimization routines between R and Stata and are not specific to this paper. Regarding the differences in p-values, they probably result from differences in the implementation of the placebo computations. `Tidysynth` computes the weights of each variable only once and then applies these weights to all synthetic controls with placebo states, while the Stata code runs the synthetic control separately for each placebo state. In the latter case, the variable weights differ from one synthetic control to the other, which leads to differences in placebo results and hence in p-values. While we don't favor one approach over another, it is important to point out the differences in results that are due to different software implementations.

4.3 Replication

We tested the robustness of the main findings (Table 2), by performing several robustness checks:

1. <u>Including other covariates trying six easy-to-implement specifications (Table 4.3.1):</u>

Panel B replicates the results from Table 2 of the original paper. This controls for education composition, age structure, share of female, and industry composition of the workforce (for definitions see p. 323 of the original paper). In Panel A, we remove industry controls from the set of covariates. In the remaining panels, we subsequently add other covariates that were mentioned in the paper. These include GDP per capita in Panel C, oil revenues as a share of GDP in Panel D, net migration rates in Panel E, and government expenditure per capita in Panel F. Details on the exact definition of the variables can be found in the original paper, which we follow exactly by using the original variables from the replication package. In column (II) we look at the part-time employment rate and find that adding more covariates decreases the size of the coefficient slightly without deteriorating the pre-period fit. For example, when including GDP per capita, oil revenues as share of GDP and net migration as matching variables in Panel E, we find an average effect of 0.011, that is, one third smaller of that in the paper, while achieving a smaller root mean squared error in the pre-treatment period. Again, these results are quantitatively different, but, if anything, more in line with the main story of the paper. In column (III) we look at labor-force participation and find that the coefficient remains virtually unchanged when including further covariates. The results for hours worked in column (IV) are also robust to changing the covariates. The coefficient remains in the same ballpark. The results become more statistically significant, but also the pre-treatment fit diminishes even more than in the original specification.

Taken together, the results from Table 4.3.1 indicate that the main results of the paper are quite robust to a change of covariates. The effect on employment would positive and marginally significant with more controls, the effect on part-time employment would be smaller and barely significant. Both changes would not affect the main finding of no negative employment response. A caveat is that it was not possible to collect additional covariates from the original data within the time frame of this project, and the analysis is thus restricted to covariates mentioned in the paper.

2. Varying the post-treatment sample period length:

Perhaps the main weakness of the paper is the use of a short pre-treatment and long post-treatment period, which may lead to spurious pre-treatment fit (Abadie, 2021). While this issue is of conceptual nature and not within the realm of this replication, it does highlight another researcher degree of freedom: when to stop the post-treatment period. The authors used the maximum number of years available. Here, we assess if the results

Table 4.3.1: Replication of Table 2 using different covariates

	(I) Employment rate	(II) Part-time rate	(III) Labor-force participation	(IV) Hours worked last week
Panel A: age, education, and gender groups				
Average effect	0.015	0.020	0.023	-0.866
RMSE	0.009	0.003	0.011	0.656
P-value	0.353	0.006	0.120	0.070
Panel B: age, education, gender, and industry groups (original)				
Average effect	0.001	0.018	0.012	-0.796
RMSE	0.005	0.003	0.013	0.394
P-value	0.942	0.020	0.331	0.084
Panel C: age, education, gender, industry groups and gdp pc				
Average effect	0.025	0.009	0.018	-0.824
RMSE	0.006	0.003	0.014	0.538
P-value	0.097	0.141	0.169	0.082
Panel D: age, education, gender, industry groups, gdp pc and oil gdp				
Average effect	0.025	0.010	0.014	-0.899
RMSE	0.008	0.003	0.014	0.613
P-value	0.089	0.103	0.262	0.064
Panel E: age, education, gender, industry groups, gdp pc, oil gdp, and net migration				
Average effect	0.026	0.011	0.014	-0.866
RMSE	0.008	0.002	0.014	0.638
P-value	0.080	0.103	0.258	0.072
Panel F: age, education, gender, industry groups, gdp pc, oil gdp, net migration, and govt. exp.				
Average effect	0.023	0.003	0.015	-0.863
RMSE	0.008	0.006	0.014	0.631
P-value	0.102	0.544	0.245	0.069

Source: Jones and Marinescu (2022), own calculations. The table replicates the results of Table 2 of Jones and Marinescu (2022) using different covariates as described in the panels.

are robust to the use of a shorter post-treatment period. To test this, we re-estimated the main specification from Table 2 of the original paper, varying the number of years post-treatment included in the estimation sample. The results of this exercise are provided in the Appendix in Figures 4..1 - 4..4. The Y-Axis gives the average effect (equivalent to $\hat{\alpha}_1$ in the original table) and the X-Axis gives the number of years after 1982 which are included in the sample. Unfortunately, we omitted inference for time limitations.

Figure 4..1 provides the results for employment and shows that the effect decreases strongly with time up until ten years. While we don't know whether the coefficient would be significant, the short-run coefficient of 0.03 seems rather large. The mirror image is shown in Figure 4..3 for part-time employment, which initially is just below zero, but starts to increase over time. The increase here is stronger in the first post-treatment years and flattens later. While these differences may or may not be statistically significant,

they are, in any case, not qualitatively significant since they do not alter the paper's main result of no negative employment response. If anything, a shorter post-treatment period would have led the authors to find stronger demand effects for these outcomes. The results for labor force activity in Figure 4..2 have a two-humped shape, with maximums around 5 and 15 years after the treatment period length. They remain in the same range and are nowhere near a negative coefficient. Hours worked during the last week is the only outcome where the length seems to matter for the sign of the coefficient. The curve in Figure 4..4 is initially positive and crosses the origin only if we include at least ten years post-treatment.

In sum, the choice of post-treatment length matters qualitatively for the results on hours worked (though we don't know if statistically), quantitatively for employment and part-time employment (though not qualitatively in changing the message of the paper), and not at all for labor force activity.

<u>3. Using augmented synthetic control and synthetic DiD estimation (Table 4.3.2)</u>:

To check the robustness of the results to the estimation procedure used, we implement two recent extensions of the canonical synthetic control method. First, following Eli Ben-Michael and Rothstein (2021), we estimate an "augmented" synthetic control model that corrects the treatment effects for inexact pre-treatment fit. Additionally, we use the method proposed by Arkhangelsky et al. (2021), which performs a synthetic differences-in-differences estimation. In essence, this introduces time and unit fixed effects as well as unit and time weights in the spirit of the synthetic control framework.

Table 4.3.2 shows the replication of the treatment effects of the paper, as well as the parameters using the two alternative methods and the p-values. The bias-corrected synthetic control leads to effects closer to zero for the part-time rate. On the other hand, it leads to larger coefficients for employment rate, labor force participation, and the number of hours worked. None of the coefficients are statistically different from zero at the 10% level. The results from the synthetic differences-in-differences are also not significantly different from zero, except for labor force participation, which turns slightly negative and is significant at the 10 percent confidence level. The effect of hours worked decreases in size and loses its significance.

Overall, the results from this exercise exhibit some variability. The differences are strongest in the case of the synthetic differences-in-differences estimator. However, the results do not contradict the main message of the paper, as the employment effect remains null and even the effect on labor force participation is rather small.

Table 4.3.2: Replication of Table 2 using other estimation methods

	Employment rate	Part time rate	Labor force part.	Hours worked
ATE (Replication)	0.001	0.018	0.012	-0.796
ATE (Bias corrected)	0.015	0.000	0.018	-1.331
P-values (Bias corrected)	0.579	0.717	0.493	0.345
ATE (Synth. DiD)	0.002	0.002	-0.008	-0.08
SE (Synth. DiD)	0.004	0.003	0.004	0.070
P-values (Synth. DiD)	0.488	0.430	0.063	0.252

Notes: The table replicates the results of Table 2 of Jones and Marinescu (2022). It adds two other econometric methods: Line 2 show the estimates using the bias correction proposed by Eli Ben-Michael and Rothstein (2021). Lines 5 and 6 implement the synthetic control differences-in-differences estimator described by Arkhangelsky et al. (2021).

4. Using either time or state placebos:

Panel B of Figures 2 and 3 in the original article builds confidence bands around the synthetic control estimate using all possible combinations of years and states as placebos showing wide and dense confidence intervals. We reproduce these figures in the Appendix in Figures 4..5 - 4..8.

Although having all possible combinations of year and state placebos can help in computing the p-values, it might obscure a clear visual comparison if we want to understand the rank relative to specific kinds of placebos. Therefore, we replicated the results in Stata distinguishing between state and year placebos. Figures 4..9 - 4..12 show the results using states and Figures 4..13 - 4..16 show the results using different years. The results are similar, with the exception of a short-lived but positive effect on employment.

4.4 Conclusion

We explore the reproducibility and replicability of the results of Jones and Marinescu (2022). The results of the paper were reproducible after fixing a few minor issues with the replication package. The authors fixed the issues in their code so that the current version of the replication package should be reproducible. We could also reproduce some of their results using R, but found some small differences in effects. The positive effect on participation in the labor force became marginally significant, and the negative effect on hours worked became insignificant. This difference highlights the importance of cross-software reproducibility.

We also assess the robustness of the findings to changes in covariates, post-treatment period length and estimator choice. Using more covariates increases effect sizes on employment and reduces effect sizes on part-time employment. Changing the post-treatment period may lead to positive employment effects, smaller effects on part-time employment, and non-negative effects on hours worked (though we did not assess the statistical significance of these estimates). Using the augmented synthetic control method by Eli Ben-Michael and Rothstein (2021) we observe smaller effects on part-time employment. Using the synthetic differences-in-differences estimator by Arkhangelsky et al. (2021) also gives a very small negative effect on labor force participation, but turns the effect on hours worked insignificant. Lastly, using only state or year placebos may show a positive employment effect in the initial years after the cash transfer came into place.

Overall, we consider the robustness replication to be successful. The main result of no negative employment response holds throughout all replication attempts. Although the results for other outcomes are less robust, we did not find a consistently different set of results that would lead us to reject the claims of the paper. The excellent replication package and clear documentation helped us in our work considerably.

Appendix

Figure 4..1: Treatment period length - Employment

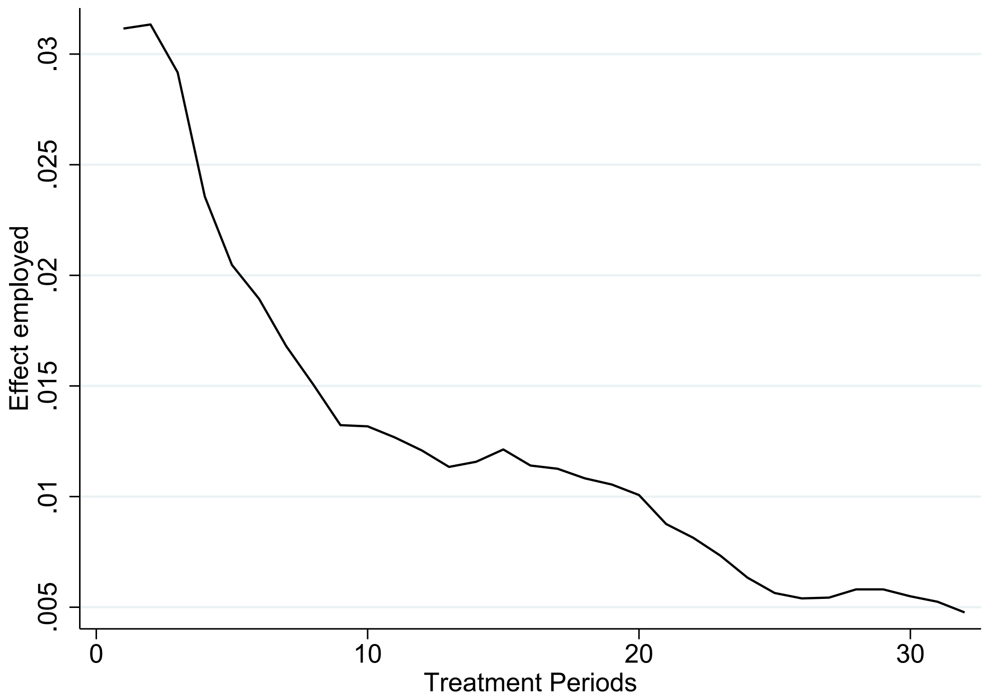


Figure 4..2: Treatment period length - Labor force activity

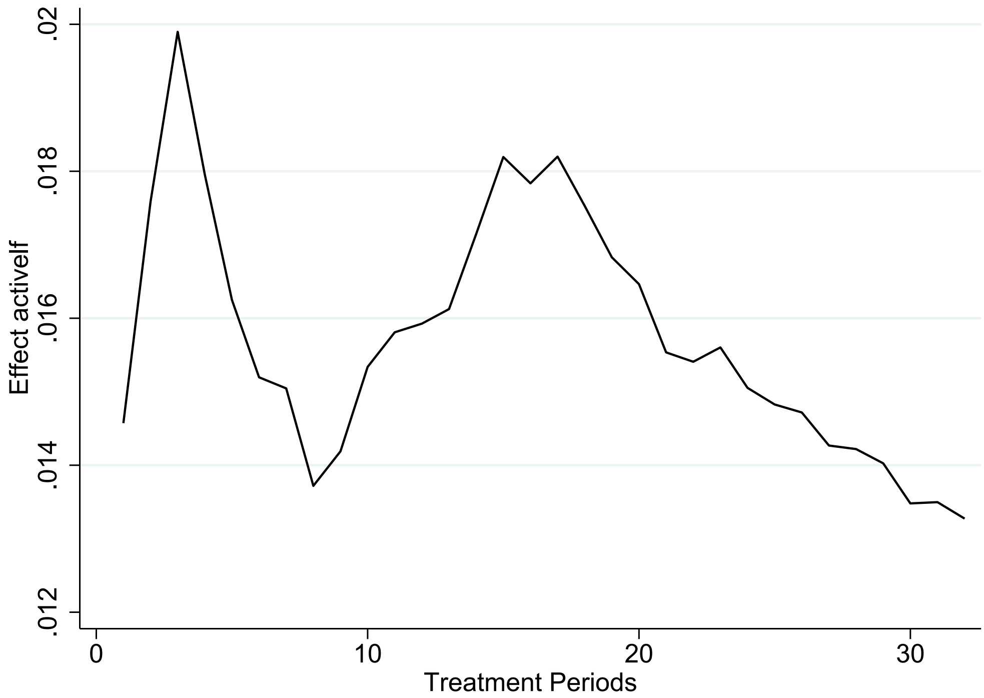


Figure 4..3: Treatment period length - Part-time employment

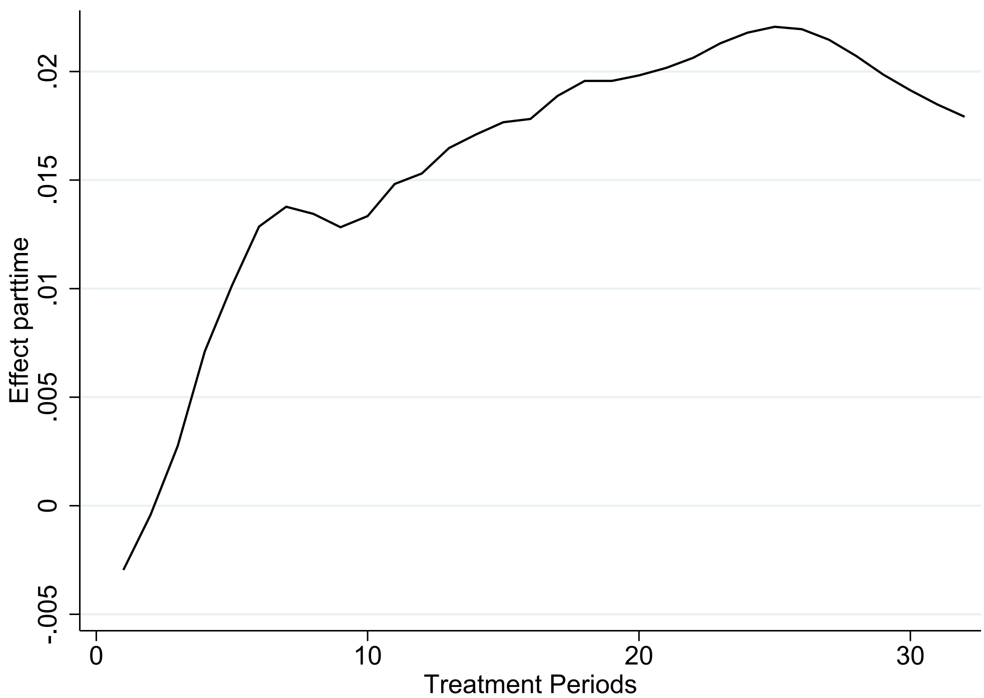


Figure 4..4: Treatment period length - Hours worked

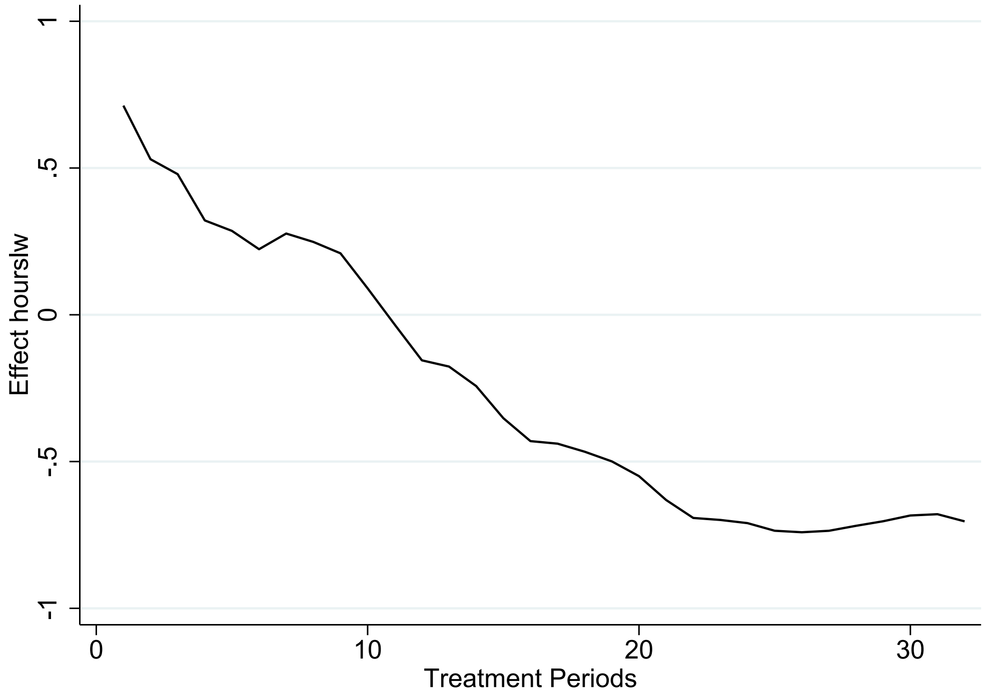


Figure 4..5: Placebo states and years - Employment

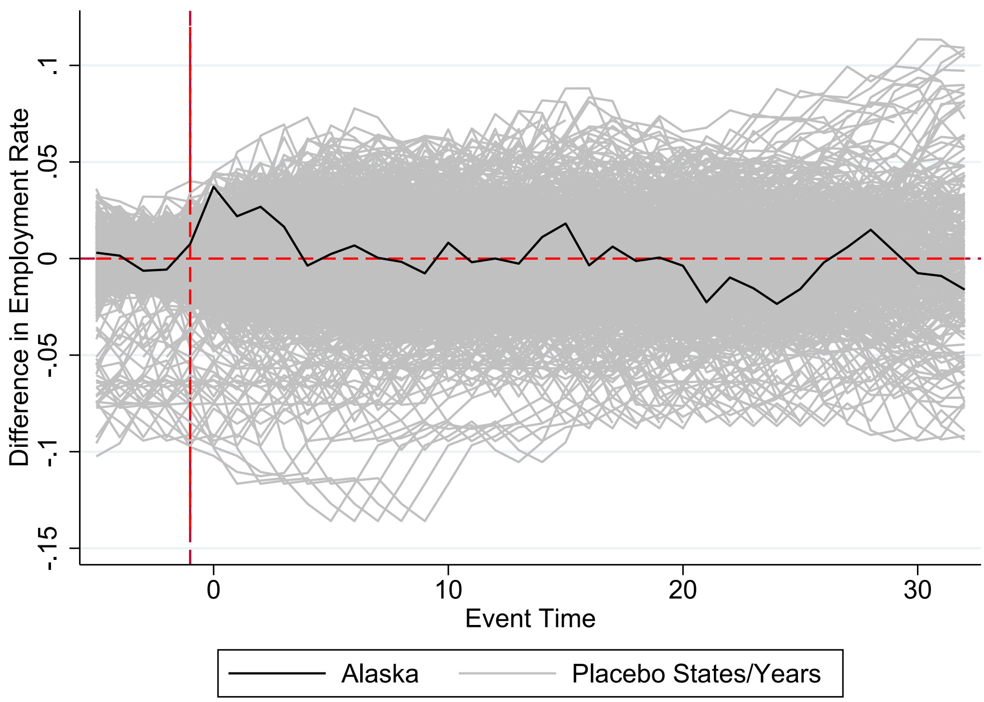


Figure 4..6: Placebo states and years - Labor force activity

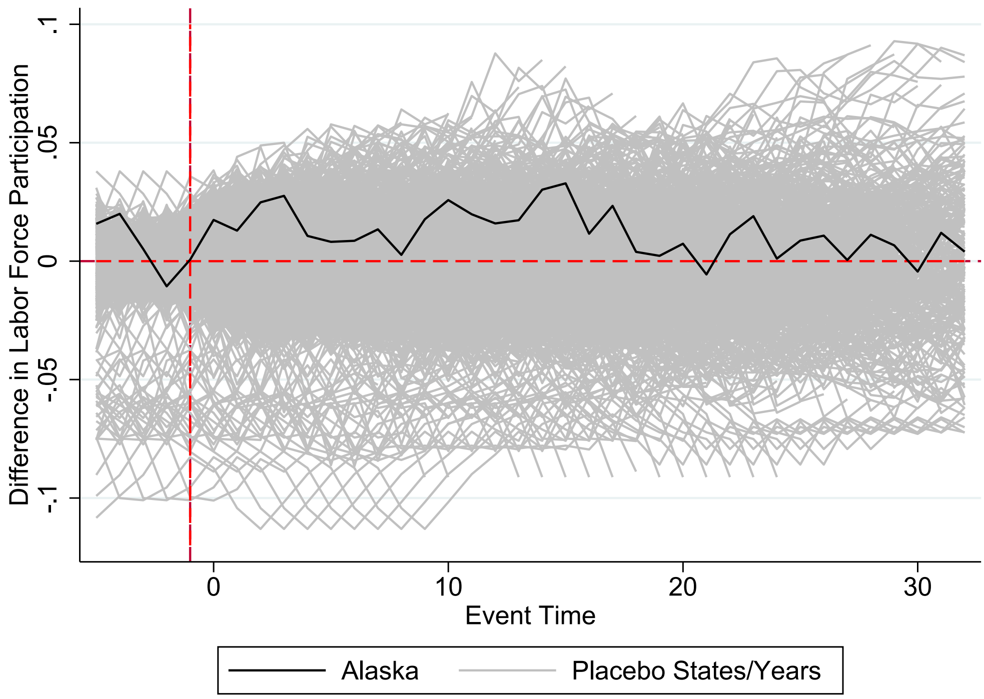


Figure 4..7: Placebo states and years - Part-time employment

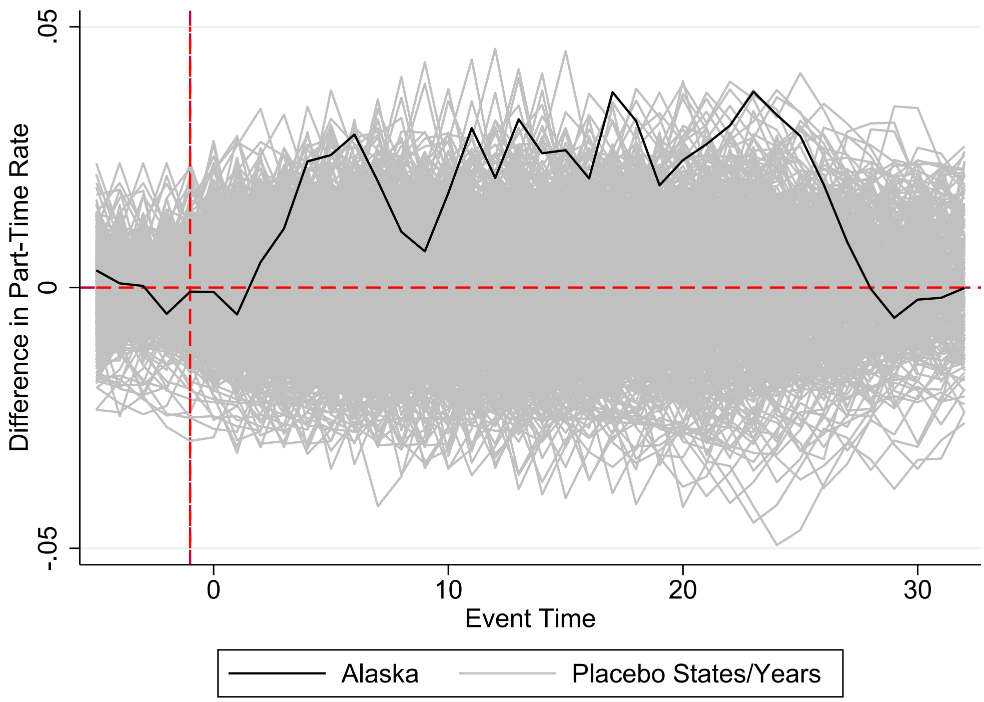


Figure 4..8: Placebo states and years - Hours worked

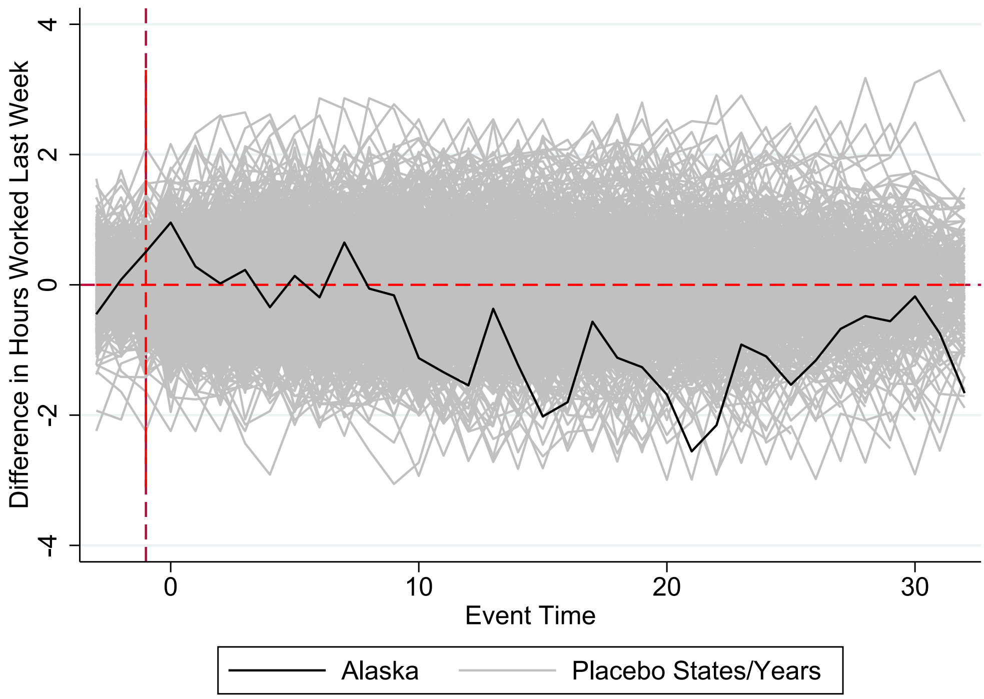


Figure 4..9: Placebo states - Employment

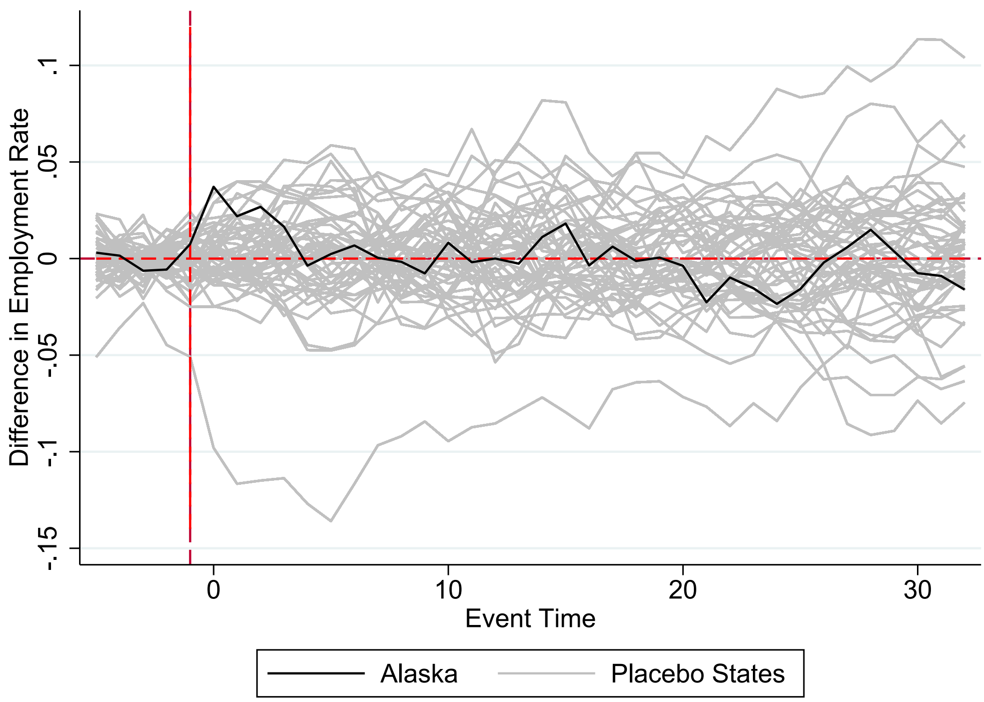


Figure 4..10: Placebo states - Part-time employment

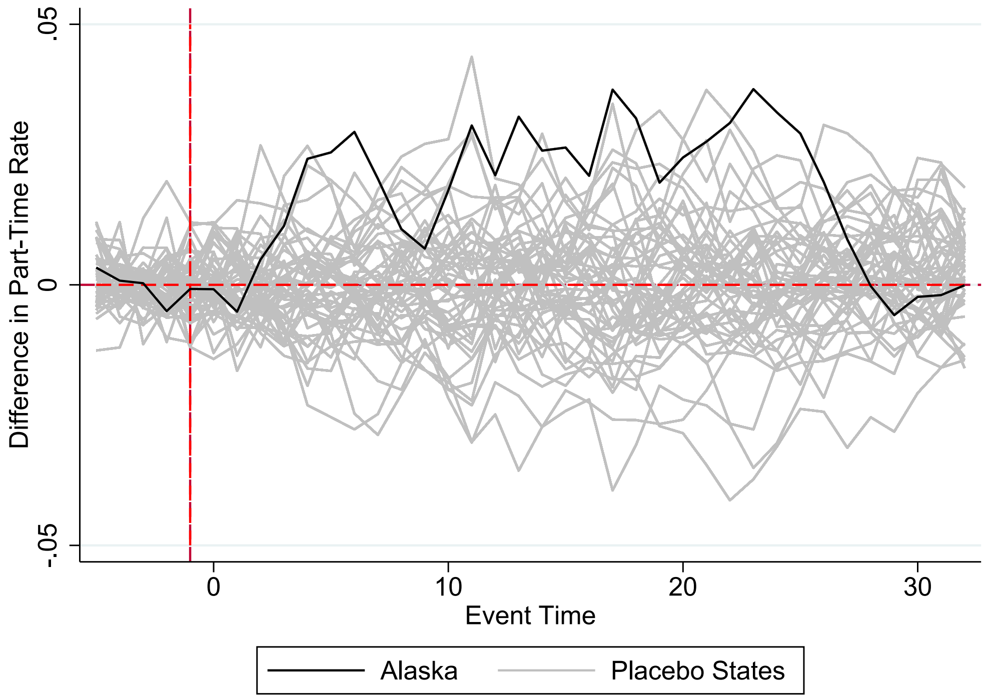


Figure 4..11: Placebo states - Labor force activity

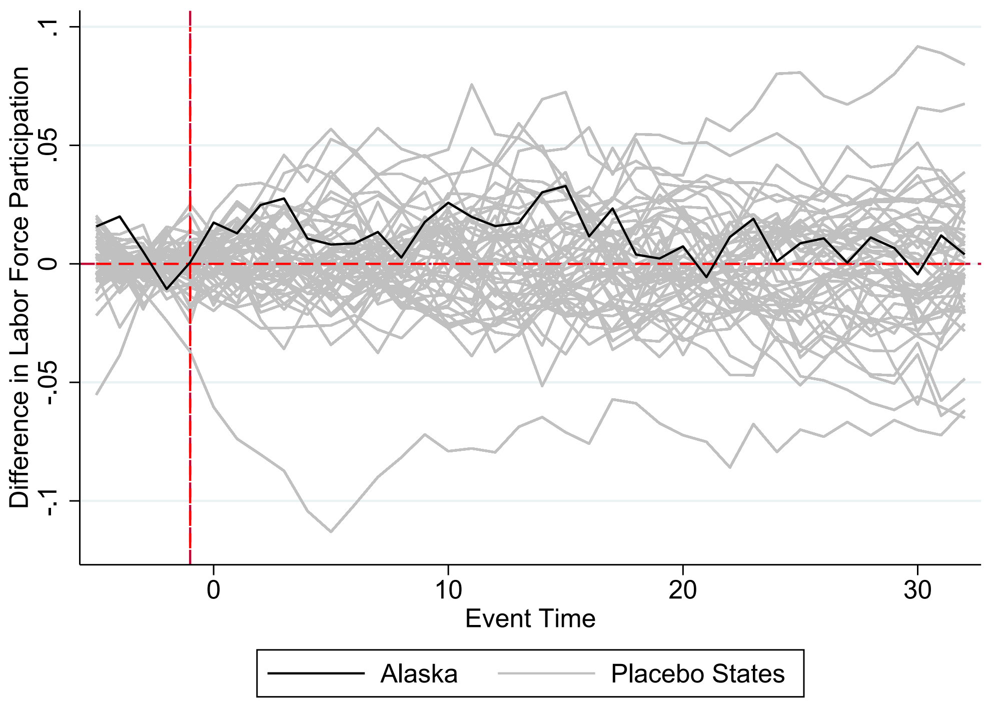


Figure 4..12: Placebo states - Hours worked

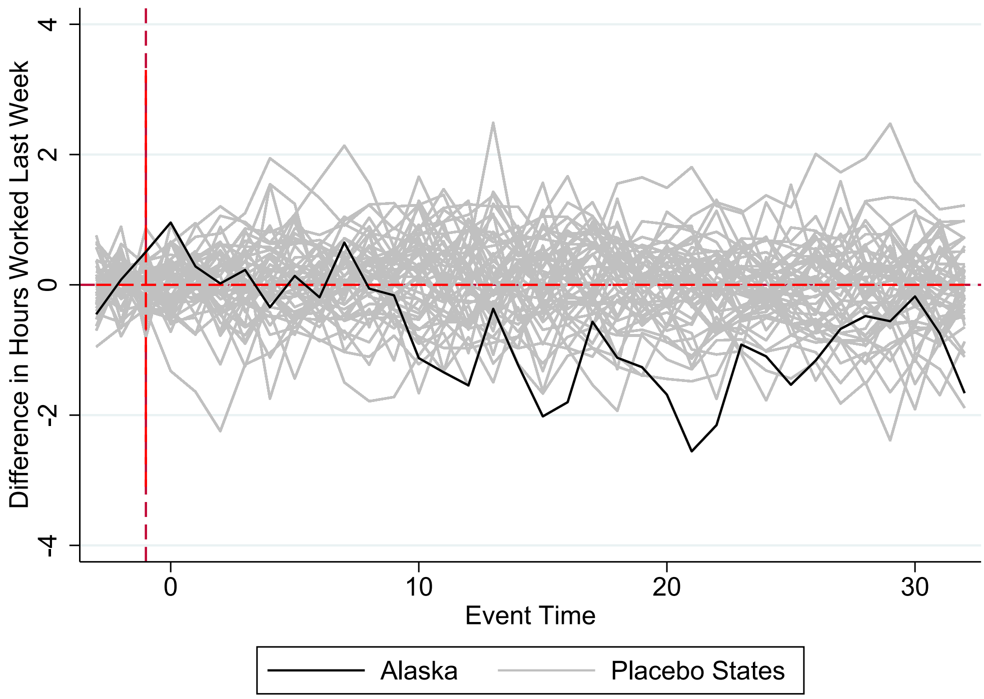


Figure 4..13: Placebo years - Employment

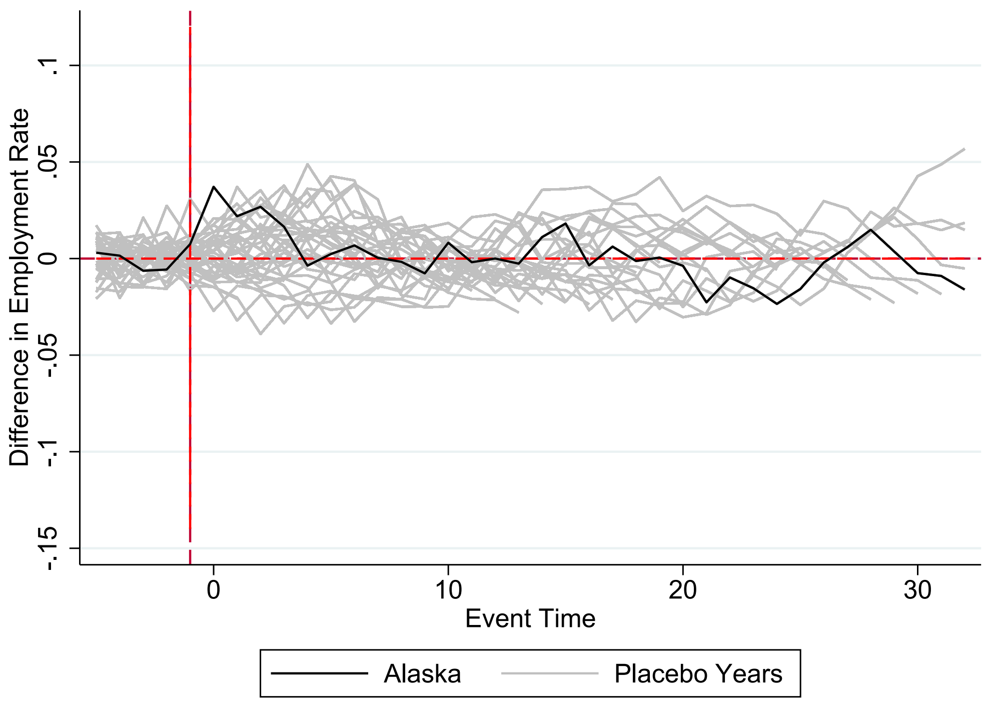


Figure 4..14: Placebo years - Labor force activity

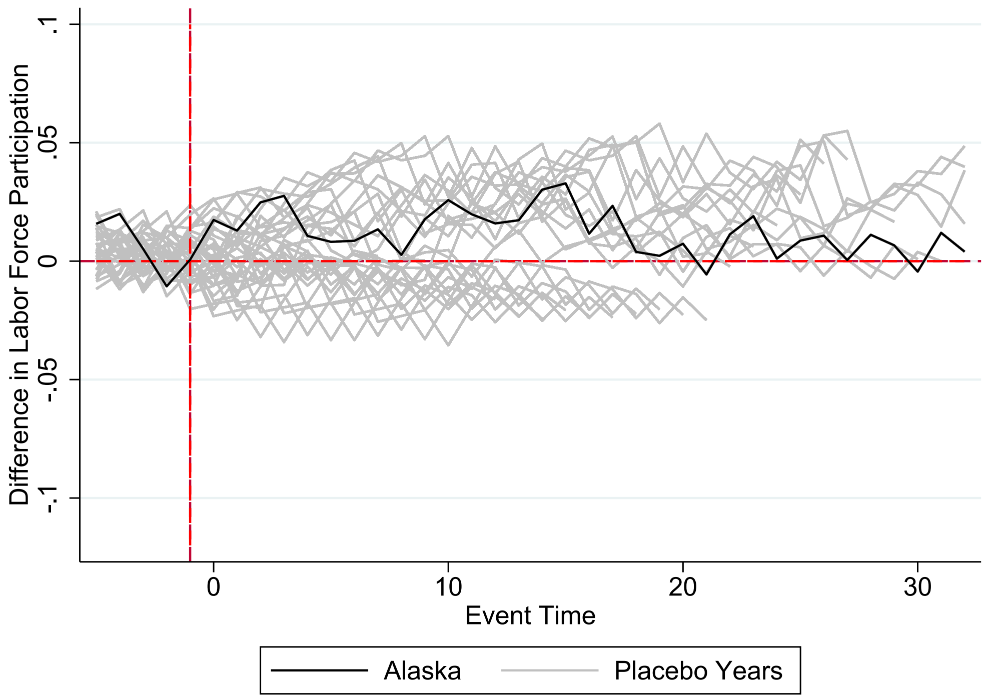


Figure 4..15: Placebo years - Part-time employment

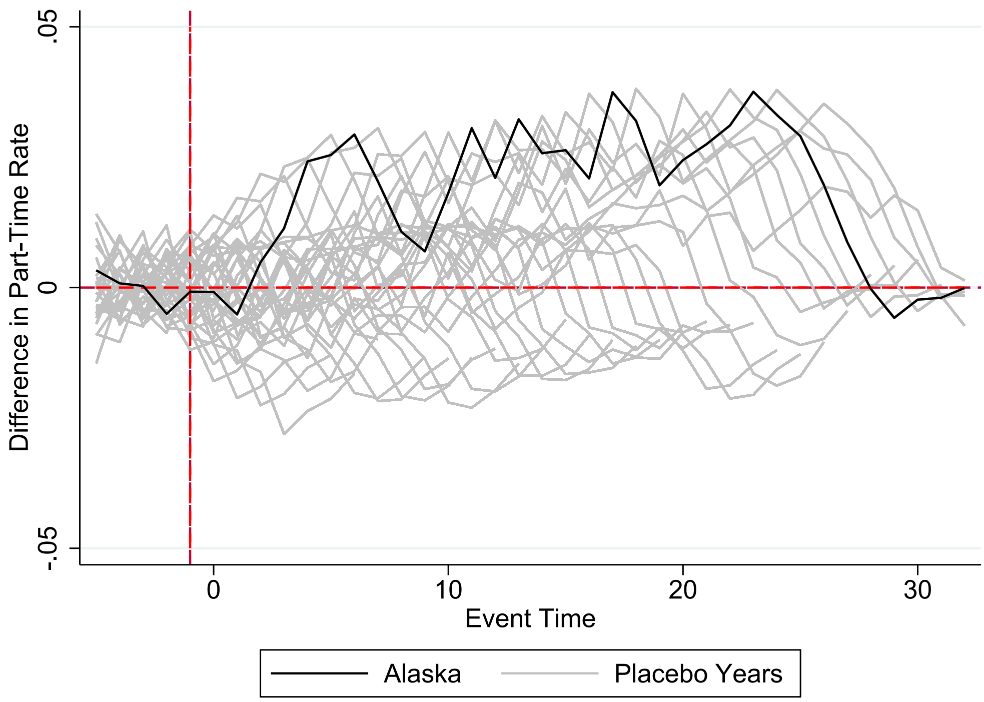
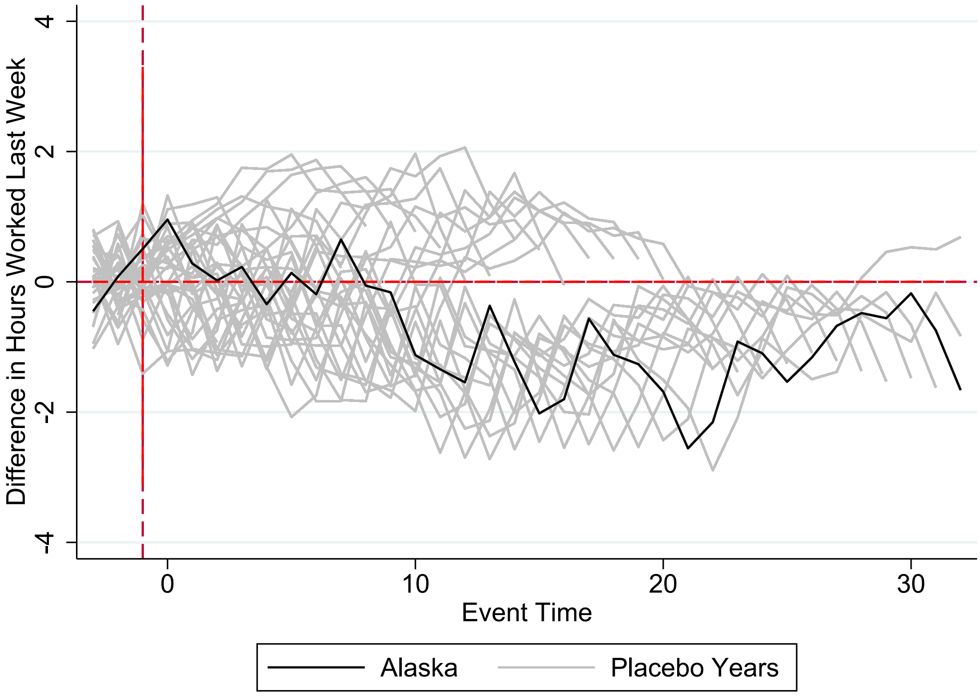


Figure 4..16: Placebo years - Hours worked



Bibliography

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, *59*(2), 391–425.

Abadie, A., Diamond, A., & Hainmueller, J. (2011). Synth: An R package for synthetic control methods in comparative case studies. *Journal of Statistical Software*, *42*(13), 1–17.

Abel, G. J., & Cohen, J. E. (2019). Bilateral international migration flow estimates for 200 countries. *Scientific data*, *6*(1), 1–13.

Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference points and effort provision. *American Economic Review*, *101*(2), 470–492.

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

Adunts, D., Brücker, H., Fendel, T., Hauptmann, A., Keita, S., & Konle-Seidl, R. (2022). Gesteuerte erwerbsmigration nach deutschland. *IAB-Forschungsbericht*, *23/2022*.

Ahrens, A., Hansen, C. B., & Schaffer, M. E. (2023). Pystacked: Stacking generalization and machine learning in stata. *The Stata Journal*, *23*(4), 909–931.

Ahrens, A., Hansen, C. B., Schaffer, M. E., & Wiemann, T. (2024). Ddml: Double/debiased machine learning in stata. *The Stata Journal*, *24*(1), 3–45.

Ahrens, S., Pirschel, I., & Snower, D. J. (2017). A theory of price adjustment under loss aversion. *Journal of Economic Behavior & Organization*, *134*, 78–95.

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *The Quarterly Journal of Economics*, *97*(4), 543–569.

Albert, C. (2021). The labor market impact of immigration: Job creation versus job competition. *American Economic Journal: Macroeconomics*, *13*(1), 35–78.

Alexander, M., Zagheni, E., & Polimis, K. (2019). The impact of hurricane maria on out-migration from puerto rico: Evidence from facebook data. *Population and Development Review*, *45*(3), 617–630.

Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. *Bioinformatics*, *26*(10), 1340–1347.

Altmann, S., Falk, A., Grunewald, A., & Huffman, D. (2014). Contractual incompleteness, unemployment, and labour market segmentation. *Review of Economic Studies*, *81*(1), 30–56.

Altonji, J. G., & Card, D. (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market* (pp. 201–234). University of Chicago Press.

Amior, M., & Manning, A. (2020). Monopsony and the wage effects of migration. *CEP Discussion Papers*, *1690, May 2020*.

Amior, M., & Stuhler, J. (2024). Immigration, monopsony and the distribution of firm pay. *CEP Discussion Papers*, *1971, January 2024*.

Anastasopoulos, L. J., Borjas, G. J., Cook, G. G., & Lachanski, M. (2021). Job vacancies and immigration: Evidence from the mariel supply shock. *Journal of Human Capital*, *15*(1), 1–33.

Anelli, M., Basso, G., Ippedico, G., & Peri, G. (2023). Emigration and entrepreneurial drain. *American Economic Journal: Applied Economics*, *15*(2), 218–252.

Angrist, J. D., & Imbens, G. W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American statistical Association*, *90*(430), 431–442.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, *111*(12), 4088–4118.

Askitas, N., Tatsiramos, K., & Verheyden, B. (2021). Estimating worldwide effects of non-pharmaceutical interventions on covid-19 incidence and population mobility patterns using a multiple-event study. *Nature Scientific Reports*, *11*(1972).

Azoulay, P., Jones, B. F., Kim, J. D., & Miranda, J. (2021). Immigration and entrepreneurship in the United States. *American Economic Review: Insights*.

Bartling, B., Fehr, E., & Özdemir, Y. (2023). Does market interaction erode moral values? *Review of Economics and Statistics*, *105*(1), 226–235.

Bartling, B., Weber, R. A., & Yao, L. (2015). Do markets erode social responsibility? *The Quarterly Journal of Economics*, *130*(1), 219–266.

Basellini, U., Alburez-Gutierrez, D., Del Fava, E., Perrotta, D., Bonetti, M., Camarda, C. G., & Zagheni, E. (2021). Linking excess mortality to mobility data during the first wave of covid-19 in england and wales. *SSM-Population Health*, *14*, 100799.

Battisti, M., Felbermayr, G., Peri, G., & Poutvaara, P. (2018). Immigration, search and redistribution: A quantitative assessment of native welfare. *Journal of the European Economic Association*, *16*(4), 1137–1188.

Baum, C. F., Schaffner, M. E., & Stillman, S. (2016). *Ivreg2: Stata module for extended instrumental variables/2sls, gmm and ac/hac, liml and k-class regression, v 4.1.10.* Retrieved June 15, 2020, from https://ideas.repec.org/c/boc/bocode/s425401.html

Beine, M., Bertoli, S., & Fernández-Huertas Moraga, J. (2016). A practitioners' guide to gravity models of international migration. *The World Economy*, *39*(4), 496–512.

Bellmann, L., Lochner, B., Seth, S., & Wolter, S. (2020). Akm effects for german labour market data. *FDZ-Methodenreport*, *01/2010*.

Berbée, P., Brücker, H., Garloff, A., & Sommerfeld, K. (2022). The labor demand effects of refugee immigration: Evidence from a natural experiment. *ZEW-Centre for European Economic Research Discussion Paper*, (22-069).

Bertoli, S., Brücker, H., & Moraga, J. F.-H. (2016). The european crisis and migration to germany. *Regional Science and Urban Economics*, *60*, 61–72.

Bewley, T. F. (1999). *Why wages don't fall during a recession*. Harvard university press.

Bilal, A. (2023). The geography of unemployment. *The Quarterly Journal of Economics*, *138*(3), 1507–1576.

Blandhol, C., Bonney, J., Mogstad, M., & Torgovitsky, A. (2022). *When is tsls actually late?* (Tech. rep.). National Bureau of Economic Research.

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

Borjas, G. J. (2006). Native internal migration and the labor market impact of immigration. *Journal of Human resources*, *41*(2), 221–258.

Borjas, G. J. (2014). *Immigration economics*. Harvard University Press.

Borjas, G. J., & Edo, A. (2021). Gender, selection into employment, and the wage impact of immigration. *IZA Discussion Paper*, *14261*.

Borjas, G. J., & Edo, A. (2023). Monopsony, efficiency, and the regularization of undocumented immigrants. *NBER Working Paper Series*, *31457*.

Borjas, G. J., Freeman, R. B., & Katz, L. F. (1996). Searching for the effect of immigration on the labor market. *The American Economic Review Papers and Proceedings*, *86*(2), 246–251.

Börschlein, E.-B., Bossler, M., Gürtzgen, N., Kubis, A., Küfner, B., Pohlan, L., & Popp, M. (2023). The iab job vacancy survey: Establishment survey on labour demand and recruitment processes, waves 2000 to 2020 and subsequent quarters 2006 to 2021. *FDZ-Methodenreport*, *17/2023*.

Borusyak, K., Hull, P., & Jaravel, X. (2022). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, *89*(1), 181–213.
Brandts, J., & Charness, G. (2004). Do labour market conditions affect gift exchange? some experimental evidence. *The Economic Journal*, *114*(497), 684–708.
Brandts, J., Gërxhani, K., Schram, A., & Ygosse-Battisti, J. (2010). Size doesn't matter! gift exchange in experimental labor markets. *Journal of Economic Behavior & Organization*, *76*(3), 544–548.
Brenzel, H., Gartner, H., & Schnabel, C. (2014). Wage bargaining or wage posting? evidence from the employers' side. *Labour Economics*, *29*, 41–48.
Brown, M., Falk, A., & Fehr, E. (2004). Relational contracts and the nature of market interactions. *Econometrica*, *72*(3), 747–780.
Brown, M., Fehr, E., & Falk, A. (2012). Competition and relational contract: The role of unemployment as a disciplinary device. *Journal of the European Economic Association*, *10*(4), 887–907.
Bundesagentur für Arbeit. (2024a). *Abgang und verbleib von arbeitslosen in beschäftigung – deutschland, länder, kreise, regionaldirektionen und agenturen*. Retrieved July 17, 2024, from https://statistik.arbeitsagentur.de/SiteGlobals/Forms/Suche/Einzelheftsuche_Formular.html;jsessionid=75F0FB7A443C61F70BAEF4E296C2E298?nn=1183188&topic_f=verbleib-alo-verbleib
Bundesagentur für Arbeit. (2024b). *Methodische hinweise zur arbeitslosigkeit und unterbeschäftigung*. Retrieved July 26, 2024, from https://statistik.arbeitsagentur.de/DE/Navigation/Grundlagen/Methodik-Qualitaet/Methodische-Hinweise/AST-Meth-Hinweise/AST-Meth-Hinweise-Nav.html
Bundesagentur für Arbeit. (2024c). *Minijob* [https://www.arbeitsagentur.de/lexikon/minijob].
Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR). (2022). *Inkar – indikatoren und karten zur raum- und stadtentwicklung*. Retrieved July 30, 2023, from https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deutschland/regionen/arbeitsmarktregionen/Arbeitsmarktregionen.html
Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR). (2024). Arbeitsmarktregionen in deutschland [https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deutschland/regionen/arbeitsmarktregionen/Arbeitsmarktregionen.html. Accessed: 2024-08-05].
Burda, M. C. (2000). Mehr arbeitslose - der preis für die osterweiterung? zur auswirkung der eu-osterweiterung auf die europäischen arbeitsmärkte im osten und westen. In L. Hoffmann (Ed.), *Erweiterung der eu: Jahrestagung des vereins für socialpolitik, gesellschaft für wirtschafts- und sozialwissenschaften, in mainz 1999* (pp. 79–103). Duncker & Humblot.
Caldwell, S., Haegele, I., & Heining, J. (2023). Bargaining and inequality in the labor market. *Working Paper*, *July 3, 2024*.
Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2017). Rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, *17*(2), 372–404.
Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, *19*(1), 22–64.
Card, D. (2009). Immigration and inequality. *American Economic Review Papers & Proceedings*, *99*(2), 1–21.
Card, D., Heining, J., & Kline, P. (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics*, *128*(3), 967–1015.
Carrillo-Tudela, C., Launov, A., & Robin, J.-M. (2021). The fall in german unemployment: A flow analysis. *European Economic Review*, *132*, 103658.
Carter, T. J. (1999). Illegal immigration in an efficiency wage model. *Journal of International Economics*, *49*(2), 385–401.
Caselli, F., Grigoli, F., Sandri, D., & Spilimbergo, A. (2020). Mobility under the covid-19 pandemic: Asymmetric effects across gender and age. *IMF Economic Review*, *70*(1), 105–138.

Castellanos, M. A. (2023). Immigration, parenthood and child penalties. *Working Paper*, *November 20, 2023*.
Charness, G., Cobo-Reyes, R., Jiménez, N., Lacomba, J. A., & Lagos, F. (2017). Job security and long-term investment: An experimental analysis. *European Economic Review*, *95*, 195–214.
Charness, G., & Kuhn, P. (2011). Lab labor: What can labor economists learn from the lab? *Handbook of labor economics*, *4*, 229–330.
Chassamboulli, A., Fontaine, I., Gálvez-Iniesta, I., & Gomes, P. (2024). Immigration and labour market flows. *Labour Economics*, *86*, 102454.
Chassamboulli, A., & Palivos, T. (2014). A search-equilibrium approach to the effects of immigration on labor market outcomes. *International Economic Review*, *55*(1), 111–129.
Chassamboulli, A., & Peri, G. (2020). The economic effect of immigration policies: Analyzing and simulating the us case. *Journal of Economic Dynamics and Control*, *114*, 103898.
Chen, D. L., Schonger, M., & Wickens, C. (2016). Otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.
Chen, S., Yang, J., Yang, W., Wang, C., & Bärnighausen, T. (2020). Covid-19 control in china during mass population movements at new year. *Lancet*, *395*, 764–766.
Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters [Published: 16 January 2018]. *The Econometrics Journal*, *21*(1), C1–C68.
Chetty, R., Friedman, J. N., & Stepner, M. (2024). The economic impacts of covid-19: Evidence from a new public database built using private sector data. *The Quarterly Journal of Economics*, *139*(2), 829–889.
Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Piontti, A. P. Y., Mu, K., Ross, L., Sun, K., Viboud, C., Xiong, X., Yu, H., Halloran, M. E., Longini Jr, I. M., & Vespignani, A. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak. *Science*, *368*, 395–400.
Christian, P., & Barrett, C. B. (2022). Spurious regressions and panel iv estimation: Revisiting the causes of conflict. *I4R Discussion Paper Series*.
Clarke, D. (2014). Plausexog: Stata module to implement conley et al's plausibly exogenous bounds [Statistical Software Components S457832, revised 08 Jul 2020, Accessed: 19 Sep 2024].
Clemens, M. A. (2011). Economics and emigration: Trillion-dollar bills on the sidewalk? *Journal of Economic perspectives*, *25*(3), 83–106.
Clemens, M. A., & Hunt, J. (2019). The labor market effects of refugee waves: Reconciling conflicting results. *ILR Review*, *72*(4), 818–857.
Clemens, M. A., & Lewis, E. G. (2022). *The effect of low-skill immigration restrictions on us firms and workers: Evidence from a randomized lottery* (tech. rep.). National Bureau of Economic Research.
Conley, T. G., Hansen, C. B., & Rossi, P. E. (2012). Plausibly exogenous. *Review of Economics and Statistics*, *94*(1), 260–272.
Coviello, D., Deserranno, E., & Persico, N. (2022). Counterproductive worker behavior after a pay cut. *Journal of the European Economic Association*, *20*(1), 222–263.
Dauth, W., Findeisen, S., Moretti, E., & Suedekum, J. (2022). Matching in cities. *Journal of the European Economic Association*, *20*(4), 1478–1521.
Delgado-Prieto, L. (2024). Immigration and worker responses across firms: Evidence from administrative records in colombia. *Working Paper*, *May, 2024*.
Destatis. (2023). *Genesis - wanderungen zwischen deutschland und dem ausland: Deutschland, jahre, staatsangehörigkeit, tabelle 12711-0007*. Retrieved June 15, 2024, from https://www-genesis.destatis.de/genesis//online?operation=table&code=12711-0007&bypass=true&levelindex=0&levelid=1716813437931#abreadcrumb
Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, *20*(5), 533–534.
Drechsler, J., Ludsteck, J., & Moczall, A. (2023). Imputation der rechtszensierten tagesentgelte für die beh. *FDZ-Methodenreport*, *05/2023*.

Droste, M. (2024). Stata-binscatter2: Really fast binned scatterplots in stata [Available at: https://github.com/mdroste/stata-binscatter2 (Accessed: 13 July 2024)].

Dunford, E. (2023). *Tidysynth: A tidy implementation of the synthetic control method, v 0.2.0*.

Dustmann, C., Fitzenberger, B., Schönberg, U., & Spitz-Oener, A. (2014). From sick man of europe to economic superstar: Germany's resurgent economy. *Journal of economic perspectives*, *28*(1), 167–188.

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

Dustmann, C., Otten, S., Schönberg, U., & Stuhler, J. (2023). Labor market effects of immigration – identification and interpretation. *Unpublisched mimeo*.

Dustmann, C., Schönberg, U., & Stuhler, J. (2016). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives*, *30*(4), 31–56.

Dustmann, C., Schönberg, U., & Stuhler, J. (2017). Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics*, *132*(1), 435–483.

Edo, A. (2019). The impact of immigration on the labor market. *Journal of Economic Surveys*, *33*(3), 922–948.

Edo, A. (2020). The impact of immigration on wage dynamics: Evidence from the algerian independence war. *Journal of the European Economic Association*, *18*(6), 3210–3260.

Egger, H., Egger, P., & Kreickemeier, U. (2013). Trade, wages, and profits. *European Economic Review*, *64*, 332–350.

Eli Ben-Michael, A. F., & Rothstein, J. (2021). The augmented synthetic control method. *Journal of the American Statistical Association*, *116*(536), 1789–1803.

Eliaz, K., & Spiegler, R. (2014). Reference dependence and labor market fluctuations. *NBER macroeconomics annual*, *28*(1), 159–200.

Errett, N. A., Sauer, L. M., & Rutkow, L. (2020). An integrative review of the limited evidence of international travel bans as an emerging infectious disease disaster control measure. *Journal of Emergency Management*, *18*, 7–14.

Eurostat. (2023). *Population on 1 january by age group, sex and citizenship (migr_pop1ctz)*.

Facebook. (2021). *Facebook data for good*. Retrieved February 19, 2021, from https://dataforgood.fb.com/

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, *133*(4), 1645–1692.

Falk, A., Becker, A., Dohmen, T., Huffman, D., & Sunde, U. (2023). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *Management Science*, *69*(4), 1935–1950.

Falk, A., Fehr, E., & Zehnder, C. (2006). Fairness perceptions and reservation wages—the behavioral effects of minimum wage laws. *The Quarterly Journal of Economics*, *121*(4), 1347–1381.

Falk, A., Huffman, D., & Macleod, W. B. (2015). Institutions and contract enforcement. *Journal of Labor Economics*, *33*(3), 571–590.

Falk, A., & Szech, N. (2013). Morals and markets. *Science*, *340*(6133), 707–711.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1993). Does fairness prevent market clearing? an experimental investigation. *The Quarterly Journal of Economics*, *108*(2), 437–459.

Foged, M., Hasager, L., & Yasenov, V. (2022). The role of labor market institutions in the impact of immigration on wages and employment. *The Scandinavian Journal of Economics*, *124*(1), 164–213.

Foged, M., & Peri, G. (2016). Immigrants' effect on native workers: New analysis on longitudinal data. *American Economic Journal: Applied Economics*, *8*(2), 1–34.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189–1232.

Gallegos Torres, K., Sommerfeld, K., & Bartel, J. (2022). 18 jahre eu-osterweiterung: Wo osteuropäer/innen in deutschland arbeiten. *ZEW-Kurzexpertise*, *3/22*.

Ganzer, A., Schmucker, A., Stegmaier, J., & Wolter, S. (2022). Establishment history panel 1975–2021 (bhp 7521 v1). *FDZ-Datenreport*, *15/2023*.

Glaeser, E. L., Gorback, C., & Redding, S. J. (2022). Jue insight: How much does covid-19 increase with mobility? evidence from new york and four other u.s. cities. *Journal of Urban Economics*, *127*, 103292.

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 249–256.

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, *110*(8), 2586–2624.

Goolsbee, A., & Syverson, C. (2021). Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *Journal of Public Economics*, *193*, 104311.

Grasso, G., & Tatsiramos, K. (2023). The impact of restricting fixed-term contracts on labor and skill demand. *IZA Discussion Paper*, *No. 16496*.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with orsee. *Journal of the Economic Science Association*, *1*(1), 114–125.

Gyetvay, S., & Keita, S. (2023). *Competition, mobility and immigration* (tech. rep.). Mimeo.

Hale, T., Webster, S., Petherick, A., Phillips, T., & Kira, B. (2020). *Oxford covid-19 government response tracker*. Blavatnik School of Government. Retrieved June 1, 2020, from https://github.com/OxCGRT/covid-policy-tracker

Hall, R. E., & Krueger, A. B. (2012). Evidence on the incidence of wage posting, wage bargaining, and on-the-job search. *American Economic Journal: Macroeconomics*, *4*(4), 56–67.

Hammer, L., & Hertweck, M. S. (2022). Eu enlargement and (temporary) migration: Effects on labour market outcomes in germany. *Deutsche Bundesbank Discussion Paper*, *02/2022*.

Hannan, R. L., Kagel, J. H., & Moser, D. V. (2002). Partial gift exchange in an experimental labor market: Impact of subject population differences, productivity differences, and effort requests on behavior. *Journal of Labor Economics*, *20*(4), 923–951.

Hart, J., & Clemens, M. (2019). A search and matching approach to business-cycle migration in the euro area. *30 Jahre Mauerfall - Demokratie und Marktwirtschaft - Session: Macroeconomics - Labor Markets*, *B06-V3*.

Hartung, B., Jung, P., & Kuhn, M. (2022). Unemployment insurance reforms and labor market dynamics. *Working Paper*, *July 2022*.

Heidhues, P., & Kőszegi, B. (2008). Competition and price variation when consumers are loss averse. *American Economic Review*, *98*(4), 1245–1268.

Hershbein, B., & Kahn, L. B. (2018). Do recessions accelerate routine-biased technological change? evidence from vacancy postings. *American Economic Review*, *108*(7), 1737–1772.

Hethey, T., & Schmieder, J. F. (2010). Using worker flows in the analysis of establishment turnover: Evidence from german administrative data. fdz methodenreport 201006 en. *FDZ-Methodenreport*, *06/2010*.

Hinton, G. E. (1990). Connectionist learning procedures. In Y. Kodratoff & R. S. Michalski (Eds.), *Machine Learning* (pp. 555–610). Morgan Kaufmann.

Illing, H. (2023). Crossing borders: Labor market effects of european integration. *IZA Discussion Paper*, *No. 15930*.

Irlenbusch, B., & Saxler, D. J. (2019). The role of social information, market framing, and diffusion of responsibility as determinants of socially responsible behavior. *Journal of Behavioral and Experimental Economics*, *80*, 141–161.

Işgın, E., & Sopher, B. (2015). Information transparency, fairness and labor market efficiency. *Journal of Behavioral and Experimental Economics*, *58*, 33–39.

Jaeger, D. A., Ruist, J., & Stuhler, J. (2018). Shift-share instruments and the impact of immigration. *NBER Working Paper Series*, *24285*.

Jones, D., & Marinescu, I. (2022). The labor market impacts of universal and permanent cash transfers: Evidence from the alaska permanent fund. *American Economic Journal: Economic Policy*, *14*(2), 315–40.

Jung, P., & Kuhn, M. (2014). Labour market institutions and worker flows: Comparing germany and the us. *The Economic Journal*, *124*(581), 1317–1342.
Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American Economic Review*, 728–741.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.
Kaur, S. (2019). Nominal wage rigidity in village labor markets. *American Economic Review*, *109*(10), 3585–3616.
Kerr, S. P., & Kerr, W. (2020). Immigrant entrepreneurship in america: Evidence from the survey of business owners 2007 & 2012. *Research Policy*, *49*(3), 103918.
Kirchler, M., Huber, J., Stefan, M., & Sutter, M. (2016). Market design and moral behavior. *Management Science*, *62*(9), 2615–2625.
Klick, L., & Schaffner, S. (2019). Fdz data description: Regional real estate price indices for germany -version 2. *RWI Projektberichte*.
Kline, P., & Moretti, E. (2013). Place based policies with unemployment. *American Economic Review*, *103*(3), 238–243.
Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, *121*(4), 1133–1165.
Kraemer, M. U. G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D. M., Group, O. C.-1. D. W., du Plessis, L., Faria, N. R., Li, R., Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., & Scarpino, S. V. (2020). The effect of human mobility and control measures on the covid-19 epidemic in china. *Science*, *497*, 493–497.
Kropp, P., & Schwengler, B. (2016). Three-step method for delineating functional labour market regions. *Regional Studies*, *50*(3), 429–445.
Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, *11*(3), 495–524.
Kuhn, M., Manovskii, I., & Qiu, X. (2021). The geography of job creation and job destruction. *NBER Working Paper Series*, *29399*.
Lachowska, M., Mas, A., Saggio, R., & Woodbury, S. A. (2022). Wage posting or wage bargaining? a test using dual jobholders. *Journal of Labor Economics*, *40*(S1), S469–S493.
Langella, M., & Manning, A. (2021). Marshall lecture 2020: The measure of monopsony. *Journal of the European Economic Association*, *19*(6), 2929–2957.
Lee, K., Worsnop, C. Z., Grépin, K. A., & Kamradt-Scott, A. (2020). A global coordination on cross-border travel and trade measures crucial to covid-19 response. *Lancet*, *395*, 1593–1595.
Letouzé, E., Purser, M., Rodríguez, F., & Cummins, M. (2009). Revisiting the migration-development nexus: A gravity model approach. *Human Development Research Paper*, *44/2009*.
Llull, J. (2018). Immigration, wages, and education: A labour market equilibrium structural model. *The Review of Economic Studies*, *85*(3), 1852–1896.
Malchow-Møller, N., Munch, J. R., & Skaksen, J. R. (2012). Do immigrants affect firm-specific wages? *The Scandinavian Journal of Economics*, *114*(4), 1267–1295.
Maloney, W., & Taskin, T. (2020). Determinants of social distancing and economic activity during covid-19: A global view. *Policy Research Working Paper Series*, *9242*.
Manacorda, M., Manning, A., & Wadsworth, J. (2012). The impact of immigration on the structure of wages: Theory and evidence from britain. *Journal of the European economic association*, *10*(1), 120–151.
McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In *Frontiers in econometrics* (pp. 105–142). Academic Press New York.
Michaillat, P. (2012). Do matching frictions explain unemployment? not in bad times. *American Economic Review*, *102*(4), 1721–1750.

Michaillat, P. (2023). Modeling the displacement of native workers by immigrants. *CEPR Discussion Paper*, *18601, November 2023*.
Michaillat, P., & Saez, E. (2015). Aggregate demand, idle time, and unemployment. *The Quarterly Journal of Economics*, *130*(2), 507–569.
Mitaritonna, C., Orefice, G., & Peri, G. (2017). Immigrants and firms' outcomes: Evidence from france. *European Economic Review*, *96*, 62–82.
Mogstad, M., & Torgovitsky, A. (2024). Instrumental variables with unobserved heterogeneity in treatment effects. *NBER Working Paper Series*, *32927*.
Molnar, C. (2018). *Interpretable machine learning, 2nd edition*. Leanpub.
Monras, J. (2020). Immigration and wage dynamics: Evidence from the mexican peso crisis. *Journal of Political Economy*, *128*(8), 3017–3089.
Montiel Olea, J. L., & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, *31*(3), 358–369.
Moretti, E., & Yi, M. (2024). Size matters: Matching externalities and the advantages of large labor markets. *NBER Working Paper Series*, *32250*.
Müller, T. (2003). Migration, unemployment and discrimination. *European Economic Review*, *47*(3), 409–427.
National Academies of Sciences, E., & Medicine. (2017). *The economic and fiscal consequences of immigration* (F. D. Blau & C. Mackie, Eds.). The National Academies Press.
Nedoncelle, C., Marchal, L., Aubry, A., & Hericourt, J. (2024). Does immigration affect wages? a meta-analysis. *Kiel Centre for Globalization (KCG) Working Paper*, *31*.
OECD. (2024a). *Annual labor statistics: Employment rate total: From 15 to 64 years for germany*. Retrieved May 15, 2024, from https://fred.stlouisfed.org/series/LREM64TTDEQ156S
OECD. (2024b). *Annual labor statistics: Inactivity rate total: From 15 to 64 years for germany*. Retrieved May 15, 2024, from https://fred.stlouisfed.org/series/LRIN64TTDEQ156S
OECD. (2024c). *Annual registered unemployment and job vacancies: Total economy: Registered unemployment for germany*. Retrieved May 15, 2024, from https://fred.stlouisfed.org/series/LMUNRRTTDEM156S
Orefice, G., & Peri, G. (2020). Immigration and worker-firm matching. *NBER Working Paper Series*, *26860*.
Ottaviano, G. I., & Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association*, *10*(1), 152–197.
Ottaviano, G. I., Peri, G., & Wright, G. C. (2018). Immigration, trade and productivity in services: Evidence from uk firms. *Journal of International Economics*, *112*, 88–108.
Owens, M. F., & Kagel, J. H. (2010). Minimum wage restrictions and employee effort in incomplete labor markets: An experimental investigation. *Journal of Economic Behavior & Organization*, *73*(3), 317–326.
Paserman, M. D. (2013). Do high-skill immigrants raise productivity? evidence from israeli manufacturing firms, 1990-1999. *IZA Journal of Migration*, *2*, 1–31.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.
Peri, G. (2012). The effect of immigration on productivity: Evidence from us states. *Review of Economics and Statistics*, *94*(1), 348–358.
Peri, G. (2016). Immigrants, productivity, and labor markets. *Journal of economic perspectives*, *30*(4), 3–30.
Peri, G., & Sparber, C. (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, *1*(3), 135–169.
Peri, G., & Yasenov, V. (2019). The labor market effects of a refugee wave: Synthetic control method meets the mariel boatlift. *Journal of Human Resources*, *54*(2), 267–309.
Peterson, L. E. (2009). *K-nearest neighbor*. Scholarpedia. Retrieved March 24, 2021, from http://www.scholarpedia.org/article/K-nearest_neighbor

Pflueger, C. E., & Wang, S. (2015). A robust test for weak instruments in stata. *The Stata Journal*, *15*(1), 216–225.
Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American Economic Review*, *83*(5), 1281–1302.
Rampazzo, F., Bijak, J., Vitali, A., Weber, I., & Zagheni, E. (2021). A Framework for Estimating Migrant Stocks Using Digital Traces and Survey Data: An Application in the United Kingdom. *Demography*, *58*(6), 2193–2218.
Ribeiro, F. N., Benevenuto, F., & Zagheni, E. (2020). How biased is the population of facebook users? comparing the demographics of facebook users with census data to generate correction factors. *12th ACM Conference on Web Science*, 325–334.
Roback, J. (1982). Wages, rents, and the quality of life. *Journal of political Economy*, *90*(6), 1257–1278.
Rose, E. K., & Shem-Tov, Y. (2021). How does incarceration affect reoffending? estimating the dose-response function. *Journal of Political Economy*, *129*(12), 3302–3356.
Rosen, S. (1979). Wage-based indexes of urban quality of life. In P. Mieszkowski & M. Straszheim (Eds.), *Current issues in urban economics* (pp. 74–104). Johns Hopkins University Press.
Ruktanonchai, N. W., Floyd, J. R., Lai, S., Ruktanonchai, C. W., Sadilek, A., Rente-Lourenco, P., Ben, X., Carioli, A., Gwinn, J., Steele, J. E., Prosper, O., Schneider, A., Oplinger, A., Eastham, P., & Tatem, A. J. (2020). Assessing the impact of coordinated covid-19 exit strategies across europe. *Science*, *369*(6510), 1465–1470.
Schram, A., Brandts, J., & Gërxhani, K. (2010). Information, bilateral negotiations, and worker recruitment. *European Economic Review*, *54*(8), 1035–1058.
Seibert, H., & Wapler, R. (2020). Einwanderung nach deutschland: Viele hochqualifizierte, aber auch viele ungelernte. *IAB-Kurzbericht*, *8/2020*.
Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, *95*(1), 25–49.
Sinn, H.-W. (2000). Eröffnungsrede. In L. Hoffmann (Ed.), *Erweiterung der eu: Jahrestagung des vereins für socialpolitik, gesellschaft für wirtschafts- und sozialwissenschaften, in mainz 1999* (pp. 9–16). Duncker & Humblot.
Sinn, H.-.-W. (2002). Eu enlargement and the future of the welfare state. *Scottish Journal of Political Economy*, *49*(1), 104–115.
SOEP. (2021). Soep-core v37eu: Socio-economic panel (soep), data from waves 1984-2020, version 37, eu edition [https://doi.org/10.5684/soep.core.v37eu].
Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, *65*(3), 557–586.
Standaert, S., & Rayp, G. (2022). Where did they come from, where did they go?: Bridging the gaps in migration data. *UNU-CRIS Working Papers*, *4*.
Stoto, M. A., Woolverton, A., Kraemer, J., Barlow, P., & Clarke, M. (2022). COVID-19 data are messy: Analytic methods for rigorous impact analyses with imperfect data. *Globalization and Health*, *18*(1), 2.
Sutter, M., Huber, J., Kirchler, M., Stefan, M., & Walzl, M. (2020). Where to look for the morals in markets? *Experimental Economics*, *23*(1), 30–52.
Team, R. C. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria.
The World Bank. (2024a). World bank open data: Gdp per capita (current us$) [Accessed: 2024-09-13].
The World Bank. (2024b). World bank open data: Population, total [Accessed: 2024-09-13].
Thomsen, U., Ludsteck, J., Schmucker, A., et al. (2018). Skilled or unskilled-improving the information on qualification for employee data in the iab employee biography. *FDZ-Methodenreport*, *9/2018*.
Tian, H., Liu, Y., Li, Y., Wu2, C.-H., Chen, B., Kraemer, M. U. G., Li, B., Cai, J., Xu, B., Yang, Q., Wang, B., Yang, P., Cui, Y., Song, Y., Zheng, P., Wang, Q., Bjornstad, O. N., Yang, R.,

Grenfell, B. T., … Dye, C. (2020). An investigation of transmission control measures during the first 50 days of the covid-19 epidemic in china. *Science*, *368*, 638–642.

Torgovitsky, A. (2024). Instrumental variables with unobserved heterogeneity in treatment effects: A guide for implementation in r and stata [Accessed: 2024-09-19].

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, *5*, 297–323.

Vannoni, M., McKee, M., Semenza, J. C., Bonell, C., & Stuckler, D. (2020). Using volunteered geographic information to assess mobility in the early phases of the covid-19 pandemic: A cross-city time series analysis of 41 cities in 22 countries from march 2nd to 26th 2020. *Globalization and health*, *16*(1), 1–9.

Zhang, A., & Wang, A. (2021). *Human mobility and smart city*. Scholarly Community Encyclopedia, Geography. Retrieved April 12, 2022, from https://encyclopedia.pub/entry/6764

Ziegler, A. G., Romagnoli, G., & Offerman, T. (2024). Morals in multi-unit markets. *Journal of the European Economic Association*, *22*(5), 2225.