Labor Market Disparities in Germany: Comparing First-Generation Refugees with First-Generation Migrants between 2016 and 2019

Anna Katharina Auer

Dissertation written under the supervision of professor Hugo Reis

Dissertation submitted in partial fulfilment of requirements for the MSc in Economics, at the Universidade Católica Portuguesa, March 20, 2025.

Abstract

This thesis provides descriptive evidence on labor market disparities between refugees and other migrants in Germany between 2016 and 2019, using the SOEP survey. Applying a probit model for the employment and an OLS wage regression for the wage gap, we find an overall significant disadvantage for refugees in the labor market. Yet, the gaps decrease after controlling for observable characteristics (e.g. education, language proficiency, and length of residence) and year, state and region of origin fixed effects whereby male refugees face a 11 p.p. lower probability of employment. Expanding this specification to the wage framework with sector/occupation fixed effects, we find that male refugees earn 14 percent less per hour than their migrant counterpart, while for female refugees this gap becomes statistically not significant. This implies, occupational and sectoral sorting plays an important role for the wage gap for female refugees. Disentangling the effect for each of the 16 states, our findings suggest strong heterogeneity within Germany.
Using an Oaxaca decomposition, we find that 63 percent of the employment gap and 45 percent of the wage gap can be attributed to differences in socioeconomic and demographic differences. Regarding the former, we find that age is the main driver. Whereas for the latter, full-time work experience and high education are the main variables contributing to this gap. This suggests the existence of labor market barriers either through discrimination or other unobserved factors for refugees with respect to wages, especially for males.

Author: Anna Katharina Auer

Title: Labor market disparities in Germany: Comparing First-generation refugees with first-generation migrants.

Keywords: Labor Economics, Refugee Migration, SOEP, Applied Economics, Oaxaca Decomposition, State Analysis

Resumo

Esta tese fornece evidências descritivas sobre as disparidades no mercado de trabalho entre refugiados e outros migrantes na Alemanha entre 2016 e 2019, usando o inquérito SOEP. Aplicando um modelo probit para o emprego e uma regressão salarial OLS para a diferença salarial, encontramos uma desvantagem global significativa para os refugiados no mercado de trabalho. No entanto, as disparidades diminuem após o controlo das caraterísticas observáveis, como a educação, a proficiência linguística e o tempo de residência, pelo que os refugiados do sexo masculino têm uma probabilidade de emprego 11 pontos percentuais inferior e ganham menos 14 por cento por hora do que os seus homólogos migrantes. No entanto, depois de considerarmos os factores observáveis e a seleção setorial e profissional, não encontramos uma penalização salarial significativa para as mulheres refugiadas. Separando o efeito para cada um dos 16 estados, os nossos resultados sugerem uma forte heterogeneidade dentro da Alemanha. Utilizando uma decomposição de Oaxaca, concluímos que 63 por cento da diferença de emprego e 45 por cento da diferença salarial podem ser atribuídos a diferenças socioeconómicas e demográficas. Relativamente à primeira, verificamos que a idade é o principal fator. Quanto à segunda, a experiência de trabalho a tempo inteiro e o ensino superior são as principais variáveis que contribuem para esta diferença. Isto sugere a existência de barreiras no mercado de trabalho, quer através de discriminação, quer através de outros factores não observados, para os refugiados no que diz respeito aos salários, especialmente para os homens.

Autor: Anna Katharina Auer

Título: Disparidades no mercado de trabalho na Alemanha:Comparação entre refugiados de primeira geração e migrantes de primeira geração.

Palavras-chave: Economia do Trabalho, Migração de Refugiados, SOEP, Economia Aplicada, Decomposição de Oaxaca, Análise do Estado

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Professor Hugo Reis, for his invaluable guidance, and encouragement throughout this process. His insightful feedback and expertise have been instrumental in shaping this thesis, and I am incredibly grateful for his patience and dedication.

I am also deeply thankful to Professor Pedro Raposo for his open-door policy, constructive criticism, and for inspiring me to strive for excellence in my academic journey.

A special thanks to my mother, whose unconditional love, and encouragement made it possible for me to achieve this goal in the first place.

I am also deeply grateful to my friends, whose unconditional support, and belief in me have been my greatest source of strength. Their companionship, words of encouragement, and countless moments of laughter have made this journey much more enjoyable.

Finally, I would like to acknowledge everyone who has, in one way or another, contributed to this work. Your support has been invaluable, and I sincerely appreciate it.

Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Definition
AME	Average marginal effects
BAMF	Bundesamt für Migration und Flüchtlinge (Federal Office for Migration and Refugees)
BAMF-FDZ	Forschungsdatenzentrum des Bundesamtes für Migration und Flüchtlinge (Research Data Centre of the Federal Office for Migration and Refugees)
Destatis	Statistisches Bundesamt (Federal Statistical Office of Germany)
DIW	Deutsches Institut für Wirtschaft (German Institute for Economic Research)
EU	European Union
Eurostat	European Statistical System
FE	Fixed Effects
FT	Full-time
IAB	Institut für Arbeitsmarkt - und Berufsforschung (Institute for Employment Research)
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
IV	Instrumental Variable
NACE	Nomenclature statistique des activités économiques dans la Communauté européenne (The Statistical Classification of Economic Activities in the European Community)
OLS	Ordinary Least Squares
p.p	percentage points
PT	Part-time
SOEP	Socio Economic Panel
UK	United Kingdom
WWII	World War Two

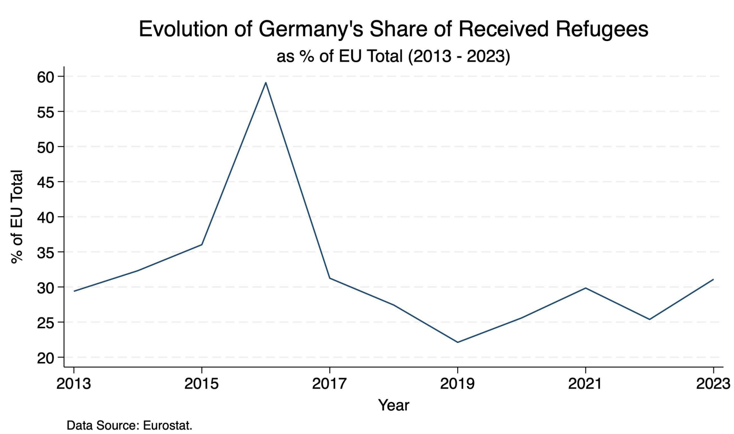
1 Introduction

Refugee flows have been increasing across the globe over the last decade. This trend started in 2013 marking unprecedented levels of refugee flows since 1994 (UNHCR, 2014). To put this into context, between 2013 and 2023 the number of refugees increased by 160 percent, reaching an all-time high at the end of 2023 with 43.4 million refugees. The majority (75 percent) has been hosted by low - and middle-income countries while only 25 percent has been hosted by high-income countries (UNHCR, 2024). Among others, war belongs to the main reason why people seek security in other countries.

The term *refugee* is commonly used to describe people seeking protection abroad. According to the 1951 Geneva Refugee Convention, refugees are individuals who are forcibly displaced and pursue protection in other countries (UNHCR, 1951). This convention provides an internationally recognized legal basis for refugees.

The German case is particularly interesting due to the large refugee inflows since 2015. In fact, among European member states, Germany has been receiving the largest number of asylum applicants. As shown in Figure 1, Germany received 60 percent of the refugees entering the European Union (EU) between 2015 and 2016, corresponding to an absolute value of 745 thousand asylum seekers. For Germany, this represents an increase in applications for asylum of 55 percent, compared to the previous year. Although this share has been decreasing since then, Germany remains the most popular host country in the EU for refugees. (Eurostat, 2025).

Figure 1: Evolution of Germany's Share of Received Refugees as Share of EU total



Notes: The figure reports the annual shares of refugees received by Germany to EU total (2013 - 2023).

To put into context, between 2013 and 2023 the refugee share relative to the total population increased by 3 percentage points (p.p.), reaching an unprecedented value of 4 percent. In absolute terms, this corresponds to 3 million refugees in 2023. Moreover, a large share of refugees are males belonging to the working-age group. For almost a decade, the main source country of refugees has been Syria (BAMF, 2024b).

At this point, it is important to distinguish between refugees and migrants. Typically, refugees are a subgroup of the migrant population, yet for the purpose of this thesis, we make a distinction between the two. Throughout the rest of this paper we refer to individuals who are not refugees as *migrants* or *other migrants*, while the term *refugee* is used exclusively for those who meet the definition of a refugee. This distinction can be made from a legal and economic perspective. Regarding the former, migrants are considered to leave on a voluntarily basis and have the ability to return to their home country. In contrast, refugees are forced to leave and are usually unable to return to their home country (Cortes, 2004). As already mentioned, refugees are internationally legally protected by the 1951 Geneva Convention, while migrants are not subject to such a protection scheme (UNHCR, 2024). Next, differentiating between refugees and migrants is important from an economic perspective, since migrants in general, chose to relocate as they expect to have better opportunities in the receiving country than in their home country (Borjas, 1987; Brell et al., 2020). Naturally, migrants anticipate their relocation and thus are often characterized by more applicable human capital than refugees. Considering these heterogeneous attributes, the pertinent question arises of whether this carries any economic implications, motivating the research question of this thesis.

In particular, we aim at studying the heterogeneous labor market outcomes for migrants and refugees in Germany. Given the previously mentioned large refugee inflows and the large share of working-age males, Germany offers an ideal setting to study the impact for refugees in a high-income country. To do so, we employ the household survey Socio Economic Panel (SOEP) for the waves between 2016 and 2019. With the increasing share of migrants and refugees in Germany, the core sample of the survey was expanded by a migrant and refugee sample in 2013 and 2016, respectively. Asking questions about socioeconomic background, employment situation and earnings, allows researchers to design representative studies examining heterogeneous labor market outcomes for both groups. For robustness we expand the time frame to the COVID-19 pandemic years 2020 and 2021, which does not alter our results notably.

Specifically, we apply a probit model to assess the employment probability between migrants and refugees. Furthermore, conditional on employment, we investigate potential wage gaps using an Ordinary Least Squares (OLS) analysis. We further explore gender-based disparities by analyzing the outcomes for each gender separately and examine differences across the 16 federal states.

Our findings reveal that being a refugee is associated with a 10 p.p. lower likelihood of employment, and 11 percent lower hourly wage, after controlling for socioeconomic and demographic characteristics such as length of stay, education, language proficiency, and age. Females exhibit a lower employment gap (-8 p.p.) than males (-11 p.p.), yet the difference between the two is statistically not different from zero. Regarding the wage gap, we observe that it is driven by males since the gap becomes insignificant for females after controlling for observables, occupational and sectoral sorting.

Moreover, we find heterogeneous outcomes for refugees across states with strong negative but also positive gaps. That is, the raw employment gap is negative and significant in 13 states. However, after controlling for the aforementioned observables, the gap remains significant in 10 states, with 8 states exhibiting negative gaps (-22 p.p. to -8 p.p.) and two states positive gaps (12 p.p. to 34 p.p.). Similarly, the raw wage gap analysis shows significant negative gaps in 14 states. Yet, after accounting for occupational sorting and including our explanatory variables, we find significant negative gaps (-44 percent to -10 percent) in 7 states and a significant positive gap in one state (30 percent). This implies, refugees are more likely to sort into low-paid occupations, which plays an important role in explaining the variation in wages between refugees and migrants.

To examine the relevance of the observable characteristics, we employ an Oaxaca decomposition. The results suggests that 63 percent of the negative employment gap is explained by differences in observables, especially by age. Yet, with respect to the wage gap, we can explain 45 percent with variations in observables, where full-time work experience and high education are the main variables driving the gap. In addition, we interpret these results as follows. To seek employment refugees are less likely to be constrained by unobserved factors, whereas in terms of earnings they experience a disadvantage due to unobserved factors or prejudice. Besides their importance and comprehension, our results cannot be interpreted as causal but serve as descriptive evidence.

Our paper contributes in various ways to the existing literature. First, we support the established theory that outcomes vary depending on the local context (S. O. Becker and Ferrara, 2019; Ruiz and Vargas-Silva, 2013) by exploiting variation across German federal states. To the best of our knowledge, this has not been done before. Second, while the literature studying the disparities between migrants and natives is broad, studies comparing migrants and refugees remain scarce. We aim to provide further evidence on the disparities between the two. Third, we provide descriptive evidence for a high-income host country while most of the refugee literature is focused on low-income countries.

The structure of this thesis is designed to provide a comprehensive understanding of the topic and is as follows. Section 2 provides an overview of the existing literature related to the topic and explains our contribution to it. Section 3 gives a brief overview of the evolution of the migrant and refugee population, information on the asylum process and the labor market access for refugees, in Germany. This is followed by the theoretical background motivating the empirical analysis in section 4. Explanation of the data set and the empirical methodology can be found in sections 5 and 6, respectively. The results and the robustness tests are presented in sections 7 and 8, respectively. Finally, section 9 reflects on the conclusions drawn from the analysis.

2 Literature Review

2.1 Migration Literature

The literature on migration and its impact on receiving economies is vast and covers different topics such as labor market outcomes for natives as well as immigrants (Borjas, 2003; Dustmann and Fabbri, 2003; Dustmann, Fabbri, and Preston, 2005; Dustmann, Hatton, and Preston, 2005), and the immigrant-native gap (Chiswick, 1978). Empirical research has identified the limited international transferability of human capital and less effective job matching among immigrants as potential explanations for the wage gap between immigrants and natives (Aldashev et al., 2012). For instance, using an Oaxaca-Blinder decomposition method and SOEP panel data Aldashev et al. (2012) find a significant wage gap between natives and immigrants in Germany between 1992 and 2009. Precisely, migrants earn on average 25.3 percent less, with 66 percent of that gap being explained by differences in characteristics and 34 percent account for unobserved influences. This result differs slightly from previous findings in Germany by Peters (2008). His findings reveal a significant discrimination effect for immigrants' earning 12 percent less on average in 2006 [1]. While most of the existing literature is focused on first-generation migrants, Algan et al. (2010) takes the discussion further by examining the native-immigrant gap for first- and second-generation immigrants in Germany, France, and the United Kingdom (UK) between 1993 and 2007. The findings indicate that across all three countries immigrants exhibit lower employment probabilities and earn lower wages, on average, compared to native-born individuals. However, notable differences emerge when comparing the two immigrant generations. In France, second-generation immigrants have a lower probability of employment compared to their first-generation counterparts. Whereas in the UK and Germany, there are no significant differences in employment outcomes between first- and second-generation immigrants. Regarding wage disparities, Germany and France show little variation between the two generations, whereas in the UK, second-generation immigrants earn, on average, higher wages than first-generation immigrants.

[1] For further literature with similar results for other countries, see Blackaby et al., 2002; Kee, 1995

2.2 Refugee Literature

The following section provides an overview of the most relevant literature on refugee economics. The existing literature on the effects of migration on host populations frequently does not differentiate between forced and voluntary migrants, despite the potential for significant differences between these groups (S. O. Becker and Ferrara, 2019). As highlighted by Cortes (2004) and Ruiz and Vargas-Silva (2018), this distinction is crucial for understanding the varying impacts of migration. Comparative international studies on the employment and wage impacts of refugee inflows on native populations show heterogeneous findings. Factors such as context, time frame, sample composition and substitutability between natives and displaced individuals help explain these varying results (S. O. Becker and Ferrara, 2019; Ruiz and Vargas-Silva, 2013).
Research on developing countries indicates that forced migration imposes significant adverse effects on those who are compelled to migrate (Ruiz and Vargas-Silva, 2013). For instance, in African countries the literature shows a negative effect on overall employment for natives but a positive effect for natives in the high-skilled sector and refugee unemployment (S. O. Becker and Ferrara, 2019). Heterogeneity across economic sectors is found in Italy. That is, the overall employment effect is zero but in the construction sector it is positive while it is negative in the hospitality sector (S. O. Becker and Ferrara, 2019).
In Turkey, natives residing in regions with a significant influx of Syrian refugees in 2012 experienced declines in both employment (1.8 p.p) and in the labor force participation rate (1.03 p.p). In other words, of those who lost their jobs following the increase in Syrian refugees, 43 percent remained unemployed while 57 percent exited the labor force (Tumen, 2016). Braun and Omar Mahmoud (2014) examined the expulsion of millions of Germans from Eastern Europe to West Germany to assess its impact on the labor market, after World War two (WWII). Given that they shared the same language and received education in German schools, the authors argue that forced migrants and natives were nearly perfect substitutes. Their findings indicate that a 10 p.p rise in a particular occupation and state was associated with a decline of 3 p.p in the native employment rate in that same occupation and state. A more recent view on the effect of a refugee inflow on the labor demand in Germany between 2010 and 2019 is presented by Auer and Götz (2023). Using an IV approach where the expected share of refugees for each state is used as an instrument, the authors find that an increase of 1000 refugees expands the number of jobs available by 300 in a district. This result is mainly driven by a rise in demand for female labor in service, social work and public administration.

Furthermore, the literature finds that refugees have worse labor market outcomes than other migrants. The literature identifies different potential reasons why refugees, on average, have a lower employment probability compared to other migrants. First, refugees in high-income countries often have a lower level of education and thus experience a reduction in transferability of human capital

into the host country's labor market (Brell et al., 2020).
Second, labor market entry restrictions make it difficult for refugees to seek employment. Hence they might experience a depreciation of human capital.
Third, refugees might suffer physical and psychological damages upon arrival to the host country, affecting their overall health status which in turn can lead to a lower employment probability compared to other migrants (Brell et al., 2020).

The first paper distinguishing between migrants and refugees was published by Cortes (2004). Using US census data from 1980 and 1990, she finds that upon arrival, refugees earned 6 percent less and worked 14 percent fewer hours than other migrants. However, over time, these differences reverse, with refugees achieving better labor market outcomes than other migrants. Lastly, the results suggest that, in the long run refugees accumulate higher levels of country-specific human capital investment compared to other migrants. Bedaso (2021) compares the employment probability gap among non-EU refugees, non-EU other migrants, and natives in Germany. Using the SOEP panel data from 1995 to 2019 she finds a substantial employment gap between natives and both migrant groups, with the gap being more pronounced for refugees. Additionally, female refugees exhibit poorer labor market outcomes than their male counterparts. When health status is included as an explanatory variable, the results do not support the hypothesis that it accounts for the disparity. Whereas, proficiency in German writing skills is associated with a 6 p.p. reduction in the employment probability gap. In contrast, Ruiz and Vargas-Silva (2018) find that heterogeneity in health status is a potential factor explaining the gaps between refugees, other migrants and natives in the UK. Using the UK labor force survey between 2010 and 2017 with information on the reason to migrate to the UK, the authors find that asylum seekers are 19 p.p. less likely to be employed than their native counterparts. The employment gap is also existent for other migrants, yet the magnitude is lower with the gaps converging over time. Similarly, OLS estimates suggest that refugees earn on average between 15 and 19 percent less than other migrants in Switzerland. The employment gap for refugees is lower ranging between 7 and 11 percent (Wong, 2024). Applying an Oaxaca-Decomposition, Wong (2024) finds that discrimination (unexplained factors) against refugees decrease over time. The potential sources of the aforementioned labor market gaps are studied by many scholars with a special focus on the impact of language proficiency. Most of the studies analyzing the language proficiency and refugees' labor market outcomes are descriptive rather than quantitative (Auer, 2018). Auer (2018) expands the literature by providing causal evidence on how refugees' language proficiency affects their employment probability in Switzerland. The distribution of refugees across Switzerland is random - similar to the process in Germany. Exploiting this random variation, Auer (2018) finds that refugees in regions where the local language closely resembles their own are 14 percent more likely to secure employment within the first two years.

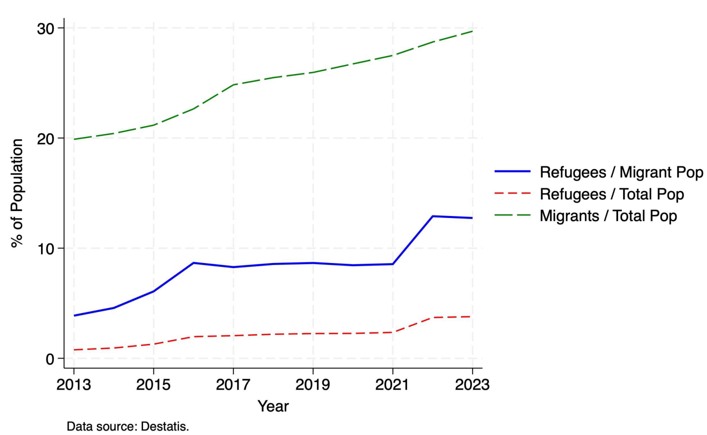
This thesis contributes to the existing literature by expanding the insights into the refugee migrant gap with respect to labor market outcomes in high-income countries. While our results shall not be interpreted as causal, we aim to provide descriptive evidence on the disparities by gender and geographic location. Specifically, we examine these differences at the state level, showing significant variation within Germany. To the best of our knowledge this has not been done before. Furthermore, we apply an Oaxaca decomposition to reveal how much the observable characteristics explain the gaps. Building upon findings from previous literature, we include language and length of stay in our set of observables.

3 Institutional Background

3.1 Migrant and Refugee Population in Germany

The German Federal Agency for Migration and Refugees (BAMF) classifies refugees, depending on their legal status, as (I) asylum seekers, (II) asylum applicants, or (III) persons entitled to protection. According to the BAMF definition, (I) are persons who intend to apply for asylum and are not yet registered as asylum applicants, (II) have applied for asylum but are still in process and (III) received asylum or who are allowed to stay in Germany due to a ban on deportation (BAMF, 2024b). For simplicity, in the remainder of this paper, refugees is a generic term and includes individuals who are asylum seekers, who have been granted protection status, or whose asylum application has been rejected but received a tolerated stay. Furthermore, the terms refugees and asylum seekers are used interchangeable.

Figure 2: Evolution of Migrant and Refugee Shares in Germany



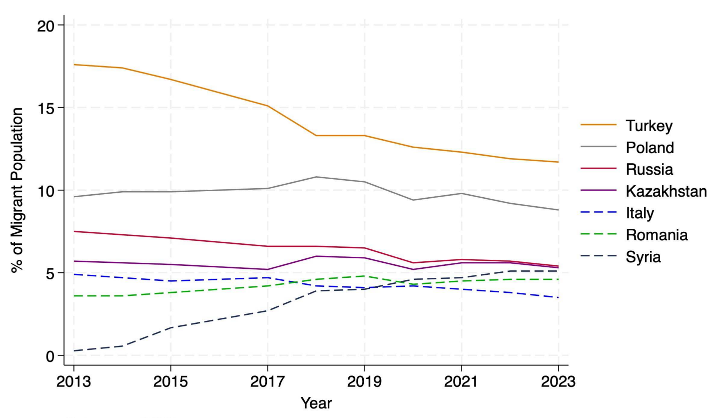
Notes: The figure displays the evolution of the migrant and refugee shares between 2013 and 2023 in Germany. The green line shows the migrant share to total population. The red line shows the refugee share to total population. The blue line shows the refugee share to total migrant population.

Between 2013 and 2023, the share of refugees in relation to total population has increased from 1 percent to 4 percent, representing an increase of 3 p.p., as illustrated by the red line of Figure 2. A more pronounced increase in the share of refugees is observed when looking at the total migrant population. That is, within ten years the refugee share with respect to total migrant population increased by 9 p.p., reaching an unprecedented level of 13 percent in 2023, as highlighted by the blue line of Figure 2. The sharpest increases across the years can be observed between 2015 and 2016 and 2021 and 2022. The former, was due to the large influx of refugees from Syria while the latter was driven by the Russian invasion of Ukraine.

Additionally, the total migrant population reached a new high, with every third person having a migration background, as highlighted by the green line of Figure 2. The total migrant population includes (I) foreigners (individuals without German citizenship), (II) first-generation migrants (individuals with German citizenship but who were born abroad) and (III) second-generation migrants (individuals with German citizenship who were born in Germany but have at least one parent born abroad) (BAMF, 2024a).

Individuals from European countries represented a substantial majority of 62 percent. This was followed by 24 percent of individuals from the Asian continent. Precisely, the largest groups within the migrant population originated from Turkey (12 percent), Poland (9 percent), Russia (5 percent), Kazakhstan (5 percent), and Syria (5 percent) (BAMF, 2024a). The latter one has been considered to be among the five most common countries of origin since 2020 as shown in figure 3.

Figure 3: Evolution Migrant Shares of the Top 5 Countries of Origin

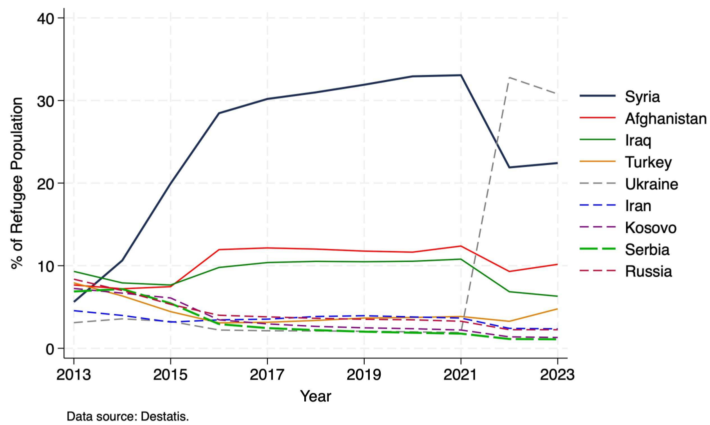


Notes: The figure displays the annual shares of migrants of selected countries of origin with respect to total migrant population in Germany between 2013 and 2023. Since the ranking of the countries change across the years we include more than 5 countries. The countries that are constantly within the top 5 are represented as solid lines, otherwise they are represented by dashed lines.

In contrast, the largest groups within the refugee population originated from, Ukraine (31 percent), Syria (22 percent), Afghanistan (10 percent), Iraq (6 percent), and Turkey (5 percent) (BAMF, 2024a). As shown in Figure 4, prior to 2021, Syria has been the most common country of origin for refugees in Germany since 2014. Since the invasion of Ukraine, this has changed and a significant share of the refugees were Ukrainian in 2023.

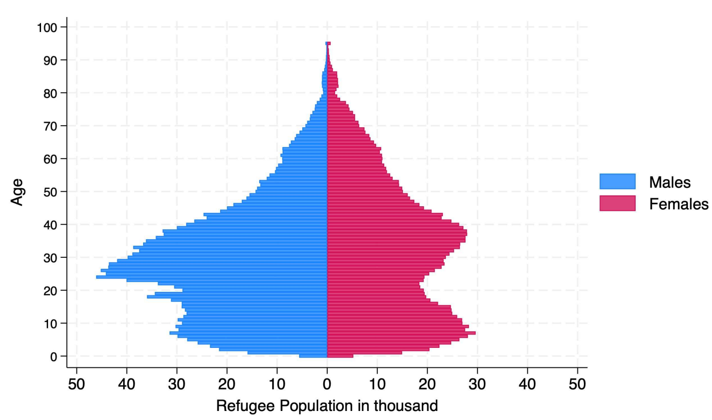
Most of the refugees in Germany are males, and between 20 and 30 years old as shown in Figure 5. The male population is particularly high in the working age group (20 to 50 years). Furthermore, the age distribution between males and females becomes more balanced in older age cohorts, between 50 and 80 years.

Figure 4: Evolution Refugee Share of the Top 5 Countries of Origin



Notes: The figure displays the annual shares of refugees of selected countries of origin with respect to total refugee population in Germany between 2013 and 2023. Since the ranking of the countries change across the years we include more than 5 countries. The countries that are constantly within the top 5 are represented as solid lines, otherwise they are represented by dashed lines.

Figure 5: Refugee Population by Gender and Age



Notes: The figure displays the gender and age distribution of refugees in Germany in 2023.

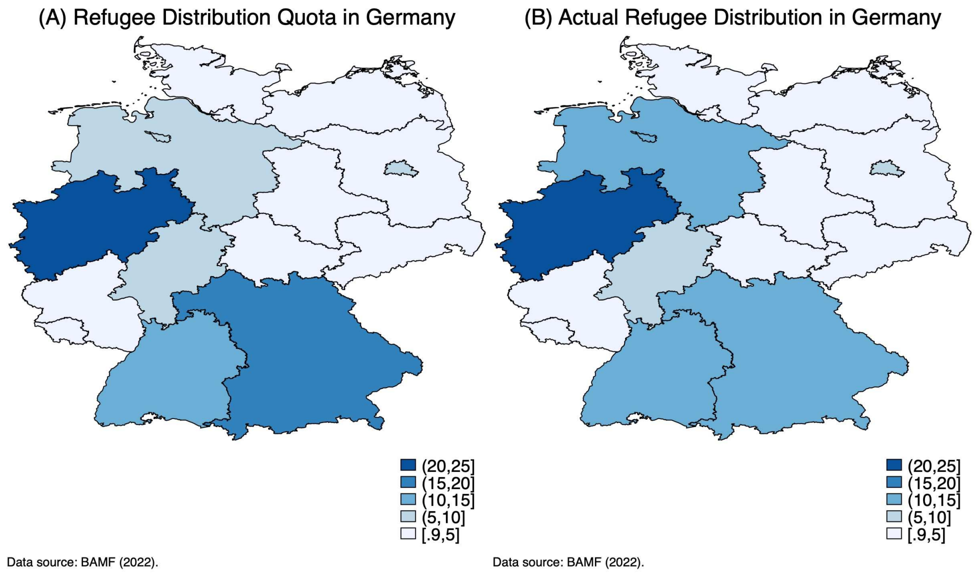
3.2 Asylum Process

At the moment of arrival individuals have to register either at the border authority or at government agencies. Everyone who applies for asylum will be registered in an official database. Since 2016, refugees receive a certificate of arrival as proof of their registration, allowing individuals to receive state benefits such as healthcare, accommodation and food. Individuals who are registering as asylum seeker upon arrival, are forwarded to the nearest initial reception center where the authorities then determine whether the refugee can be hosted by the state or needs to be forwarded to another state. This decision is based on a quota called ”Königsteiner Schlüssel” set by the German government since 1993. As a result, refugees are not free to choose their place of residence for the duration of the asylum procedure in Germany and are randomly, independently from individual characteristics, allocated. The purpose is to ensure a reasonable and fair allocation among the states and it is calculated as one third of the population and two thirds of tax revenue of each state. The distribution quota is annually recalculated. Yet, the shares have not differed significantly over the years. For instance, North Rhine-Westphalia, being the most populated and having the highest tax revenue, has been receiving the largest share of refugees, as shown in Figure 6. In reality, however, the expected shares might differ slightly from the actual shares each state receives. As highlighted in figure 6 Bavaria (Southeast) received less than supposedly, whereas Lower-Saxony (Northwest) received more. Key reasons for this are availability in housing, infrastructural capacities, or to maintain public safety (Auer and Götz, 2023).

In general, asylum seekers are required to live in a reception center for at least three months and a maximum amount of 18 months post arrival. For families with underage children a maximum of 6 months post arrival applies. As long as the asylum seeker is obliged to live in a reception center, the so-called *residence obligation* applies. This means that asylum seekers require permission to leave the district of the responsible immigration authority. Additionally, as long as refugees receive social welfare benefits, they are required to reside in the designated district. After the time limit has expired they should be allocated to shared accommodations or apartments.

When asylum is granted, individuals have residence permit valid for three years. Depending on the specific form of protection, this permit may be extended for an additional three years, or individuals may apply for a settlement permit after five years. To obtain a settlement permit, refugees must meet several requirements, such as securing an independent means of livelihood and demonstrating sufficient proficiency in the German language (BAMF, 2024b).

Figure 6: Refugee Allocation: Theoretical vs. Actual Distribution



Notes: Panel (A) of the figure shows the theoretical allocation of refugees across German states, based on a quota from 2022. Panel (B) of the figure shows the actual refugee distribution across German states in 2022. The shares are in relation to total refugee population. The values shown are percentage.

3.3 Labor market access

The labor market access for refugees differs according to their residence status. Asylum seekers who are not obligated to live in a reception center are granted access to the labor market after 3 months post arrival to Germany. Afterwards, refugees may request authorization from both the Immigration Office and the Federal Employment Agency, which evaluates applications according to a precisely defined set of criteria. If potential employment requires relocation within Germany, such a move is permitted given that the individual subsequently becomes independent of welfare benefits. The aforementioned time constraint is longer for asylum seekers who are obligated to live in a reception center, namely 9 months post arrival. The same condition applies to tolerated individuals, except if concrete measures to end the stay are imminent. Lastly, asylum seekers belonging to safe countries[2] of origin such as Albania, Bosnia and Herzegovina, Georgia, Ghana, Kosovo, Montenegro, North Macedonia, Republic of Moldova, Senegal and Serbia, face employment restrictions.
Unemployed asylum seekers receive social benefits that are paid in kind as well as in monetary payment. To cover essential needs such as food, housing, clothing, health care, and education they receive on average a monthly payment of 670 Euros per person, with the exact amount varying based on family status and age (Auer and Götz, 2023).

[2]Safe country means that the country's safety situation is considered stable enough to justify return.

4 Theoretical Background

This section outlines the theoretical framework that explains the heterogeneous labor market outcomes observed between refugees and migrants.

The first theoretical argument explaining variation between refugees and migrants in the labor market can be traced back to the theory developed by G. S. Becker (1957), namely the theory of discrimination. In his book, he argues that discrimination in the labor market occurs when individuals belonging to two different groups, in this case migrants and refugees, are treated differently due to certain characteristics, besides demonstrating the same observable attributes. In the case of migrants and refugees, the employment gap may arise because employers discriminate against refugees due to stereotypical thinking, statistical statements, or preference based discrimination. For instance, assume a simple world with only two workers. One is a refugee and the other is a migrant. We observe the same characteristics for both, such as education, age, and language proficiency. Yet, because of personal prejudice the employer is more inclined to hire the migrant.

The second argument in the literature is derived from the general human capital theory, that explores disparities in earnings. Accordingly, earnings are associated with a person's productivity which in turn depends on a person's skills (G. S. Becker, 1975). With respect to migration, authors argue that gaps in the labor market between refugees and migrants may be explained by a gap in human capital. That is, most asylum seekers originate from low-income countries whereas migrants from more economically advantaged backgrounds are more likely to relocate. This implies they potentially find it less challenging to integrate in the host country's job market. Lastly, individuals with higher levels of education may face lower migration costs and possess a greater capacity to adjust to new labor markets, further influencing their decision to migrate (Ruiz and Vargas-Silva, 2018). Additionally, migrants anticipate their relocation allowing them to develop host country specific human capital and specific labor market knowledge. Whereas refugees, relocate without any previous preparation for the host country labor market leading to a disadvantage. Nevertheless, this does not imply that refugees cannot catch up over time with migrants. In contrast, as their time living in the host country increases, they are able to invest in host country specific human capital and thus earn higher wage, converging with their native counterparts (Chiswick, 1978; Cortes, 2004).

The aforementioned theory, however, can only partially explain the potential gaps between refugees and other migrants. That is why, I also follow the rationale from Ruiz and Vargas-Silva, 2018 who argues that the return to migration, r, depends on three major elements. First, the difference in earnings between the host country (hc) and the country of origin (cor), expressed as $(W_{hc} - W_{cor})$. Second, the difference in safety levels between the host and home country, $(S_hc - S_{cor})$. Third the cost of migration (C). To put into context, if $W_{hc} > W_{cor}$ and $S_{hc} = S_{cor}$, an individual is expected to realize economic benefits and, thus, be more inclined to migrate. Although, even if an economic loss is anticipated, migration may still occur, if $S_{hc} > S_{cor}$, in anticipation of a positive r, as is often the case with refugees who face significant security concerns in their home countries (Ruiz and Vargas-Silva, 2018).

5 Data

This section of the paper presents the data employed for the analysis. The data set consists of the Socio-Economic-Panel (SOEP) from the German Institute for Economic Research Berlin (DIW). It is a longitudinal survey of private households, covering almost 160 thousand individuals, which is ongoing since 1984 in Germany. Respondents are interviewed annually. With the increasing share of migrants and refugees in Germany, alongside a growing scholarly interest in studying these populations, the SOEP was expanded to include migration and refugee subsamples in 2013 and 2016, respectively. Thus, we use the waves between 2016 and 2021 for our analysis. The survey provides information on socio-demographic information such as, country of origin, immigration year, educational level, and labor force status. Those who are working are also asked about their monthly gross income in the previous month and their average actual time worked per week.

5.1 Migration Sample

The migration sample was added to the SOEP core sample in 2013 and covers individuals who immigrated after 1995 or second-generation immigrants. This sample was extended in 2015 with mainly individuals who immigrated between 2010 and 2013 from the new EU member states in Eastern Europe. The most recent expansion of the sample occurred in 2020 with individuals who arrived to Germany between 2016 and 2018 from Poland, Bulgaria, and Romania. In the dataset, migrants are defined as individuals who were born outside of Germany. The survey further categorizes migrants into direct migration background (not born in Germany), and indirect migration background (at least one of the parents were not born in Germany).

5.2 Refugee Sample

The refugee subsamples were created in collaboration with the Institut für Arbeitsmarkt- und Berufsforschung (IAB), the DIW, and the Research Center on Migration, Integration, and Asylum of the Federal Office of Migration and Refugees (BAMF-ZF). The first sample was included in 2016 and covers refugees who applied for asylum between 2013 and 2016. To counteract a high dropout of respondents, a boost sample was included in 2020 (Steinhauer et al., 2022). The study distinguishes between refugees and other migrants by assessing whether individuals have direct, indirect or no evidence of a refugee experience, conditional on having a direct or indirect migration

background (Kroh et al., 2016).
The data set resembles the actual refugee and migration situation in Germany well. For example, in 2021 migrants represented approximately 23 percent of the total sample population, while according to official statistics 23 percent have a migration background. Similarly, refugees represent approximately 3 percent of the total sample population. In reality, refugees represented 2 percent of the total population. With respect to total migrant population, refugees represent 11 percent in the sample, while they represented 9 percent in the official statistics (BAMF, 2024a).

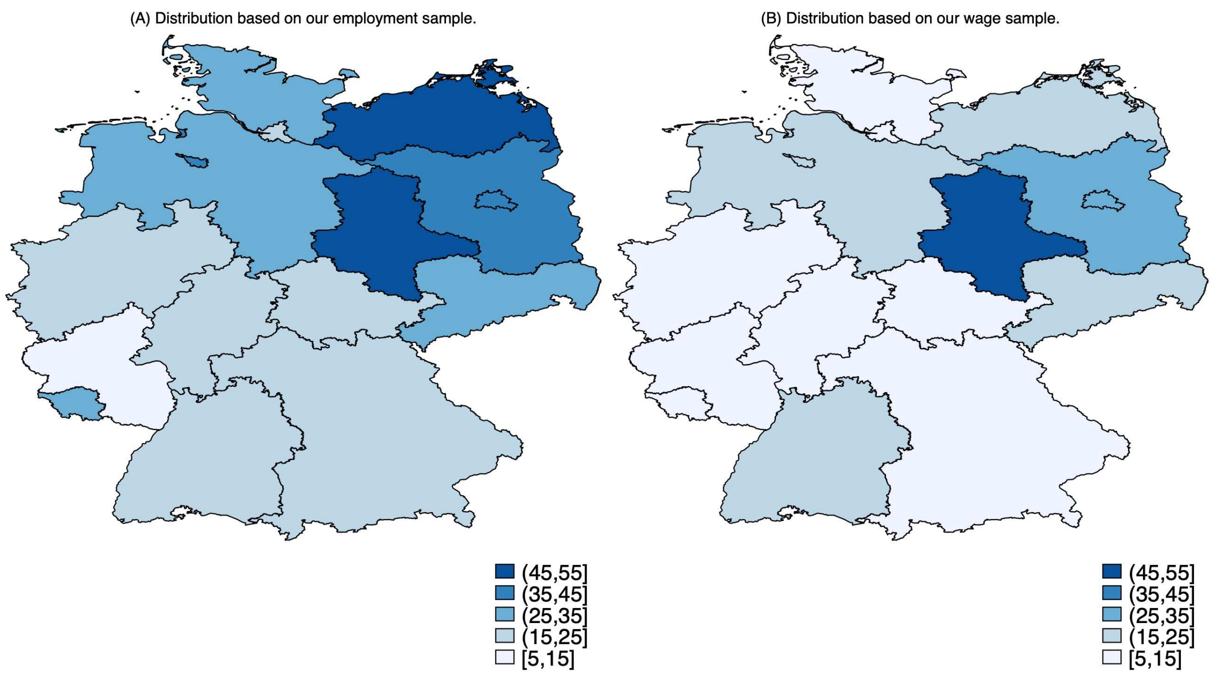
5.3 Sample and Variables construction

We restrict our sample to first-generation migrants and refugees[3] and the waves between 2016 and 2019[4]. For robustness tests, we expand the analysis by including 2020 and 2021. We focus on first-generation individuals as they face unique immigration barriers which might lead to different labor market outcomes, compared to second-generation individuals(Aldashev et al., 2012). Nevertheless, examining possible variation between the two generations may be interesting for future research.
Furthermore, we focus only on working-age individuals who are between 18 and 64 years old[5]. Our sample includes individuals who migrated between 1956 and 2019[6]. Finally, we exclude observations with missing values[7] for the variables of interest[8].
We are left with an unbalanced panel with 18,784 person-year observations for refugees and 11,980 person-year observations for migrants (total N = 30,764) for the employment analysis and 2,585 person-year observations for refugees and 8,439 person-year observations for migrants (total N= 11,024) for the wage analysis. As a result, we have two samples, (I) the employment sample which covers everyone after cleaning the data and (II) the wage sample covering all individuals from sample (I) conditional on being employed, working at least 1 hour per week and reporting gross monthly income.
After cleaning the data, refugees account for 20 percent of the sample. The distribution of refugees in relation to migrants across German states, is shown in Figure 7. Panel A shows the distribution of refugees based on our employment sample, whereas panel B shows the distribution of refugees based on our wage sample. The dark blue states indicate a relatively large share between 45 and 55 percent. The brighter states represent a lower share of refugees ranging between 5 and 15 percent. Panel A shows a higher share of refugees in East Germany, suggesting regional differences.

[3]Dropped 1 million observations, reaching a total of 253.763 observations.
[4]Dropped 170 thousand observations.
[5]Dropped 27 thousand observations
[6]Dropped 6 thousand observations
[7]Dropped 20 thousand observations
[8]Except for the variables write and speak. These questions are asked every second year. We impute the missing values by either using the previous or subsequent value to avoid a significant loss in observations.

Additionally, there is a higher concentration across Germany for the employment sample than for the wage sample, except for one eastern state Sachsen-Anhalt. This implies a large share of the refugees in our employment sample are not employed, leading to lower shares in the wage sample. Although, a high concentration of refugees remains in the eastern part of Germany for panel B.
Moreover, we include individuals from 135 countries. The top 5 countries of origin for each group is represented in table 1. In line with the official statistics between 2016 and 2019 shown in section 3.1, the most common countries of origin for refugees are Syria, Kosovo, Iraq, Afghanistan and Iran. For migrants those are Poland, Kazakhstan, Turkey, Russia and Romania.

Figure 7: Distribution of Refugees in Germany (2016 - 2019)



Notes: The figures shows the distribution of refugees. Panel A shows the distribution based on our employment sample. Panel B shows the distribution based on our wage sample. The values represent percentage share relative to total sample population. Observations are weighted.

The main variables used for this analysis are presented in table 2. The employment variable is based on the self-reported current employment status and differentiates between working and non-working.

Table 1: Top 5 Countries of Origin for Refugees and Migrants

Country of Origin	% Migrants	% Refugees	% of Sample
Panel A: Top 5 countries of origin for refugees			
Syria	9.50	90.50	22.28
Iraq	8.51	91.49	10.17
Kosovo	44.53	55.47	8.62
Afghanistan	4.21	95.79	7.94
Iran	22.79	77.21	6.83
Panel B: Top 5 countries of origin for migrants			
Poland	98.41	1.59	13.55
Kazakhstan	100	0	11.61
Turkey	87.68	12.32	9.79
Russia	96.72	3.28	7.85
Romania	99.27	0.73	5.44

Notes: The values displayed represent percentage shares.

The second outcome variable, log hourly real wages, is the logarithm of gross hourly wages in 2020 prices. This variable is obtained by dividing the self-reported gross monthly wage in the month prior to the interview by the reported average working hours per week at the time of the survey that are extrapolated to monthly hours (with factor 4.33). Over time payments are included in the gross monthly earnings. To account for outliers, we trimmed the 1st and 99th percentiles of the wage distribution (Aldashev et al., 2012).

The key explanatory variable *Refugee* is a binary variable indicating whether the individuals is a first-generation refugee or first-generation migrant without refugee background. For the identification of refugees and other migrants, the survey provides information on whether the migrant (not born in Germany) has a direct, indirect or no refugee background. That is, first generation refugees, second generation refugees and other migrants, respectively. In the data, the migration background variable is time-invariant and an individual with migration background always remains a migrant even though the person might officially be a German resident.

The language variable *Language* is a categorical variable that takes the value 1 if an individual speaks and writes German very well, 2 if an individual either speaks or writes very well, or 0 if an individual neither speaks nor writes very well. This variable reflects the respondents' answers about their self-assessed language proficiency based on a five-point scale, where (1) ”Very well”, (2) ”Well”, (3) ”Okay”, (4) ”Badly”, (5) ”Not at all”. The questions regarding the language proficiency are only asked every second year and thus we impute missing values by taking either the previous or subsequent value.

The length of stay in Germany is given by the *Length of Stay* variable and is calculated as the difference between the survey year and the immigration year, thus is a continuous variable.

The *Age* variable is self-calculated and is the difference between the survey year and the birth year.

The highest achieved educational level of individuals is measured as International Standard Classification of Education (ISCED11) categories, which we separate in three different groups, namely (I) ”High Education” (ISCED11 levels 5 to 8), (II) Middle Education (ISCED11 levels 3 and 4), and (III) Low Education (ISCED11 levels 1 and 2).

For the occupation and sector variables, respondents are asked about their current situation and the answers are then transformed into International Standard Classification of Occupation (ISCO-08) and Nomenclature statistique des activités économiques dans la Communauté européenne (NACE2) codes, respectively. In the data we have 486 and 82 distinct occupations and sectors. To avoid multicollinearity and complexity for the empirical models we aggregate those distinct observations into broader categories, based on the ISCO-08 and NACE2 codes. Further information and descriptive statistics with respect to occupation and sector can be found in Appendix A.1. Similarly, there are 135 distinct countries of origin in the data which we aggregate into 6 geographic regions. Using disaggregated information, i.e. the country of origin, for the empirical analysis constraints our employment estimation. In some cases, all observed refugees from a particular country are either employed or unemployed. Aggregating countries into broader geographic regions mitigates this issue, ensuring a sufficient degree of variation for statistical estimation. For the classification of the countries into regions, we follow the World Bank definition of geographic regions (Worldbank, 2017). Detailed information on the classification can be found in Appendix A.2.

Table 2: Variables Description

Variable	Description
Outcome Variables	
Employed	Dummy = 1 if individual is working (conditional on labor force status)
Log hourly real wage	Log of real hourly wage (measured in 2020 prices)
Explanatory Variables	
Refugee	Dummy = 1 if direct refugee background,
	0 if direct migrant background without refugee background
Language	Categorical variable, 1 = very well speak & write,
	2 = either very well speak or write, 0 = neither speak nor write
Length of stay	Year since migration in Germany
Male	Dummy = 1 if male, 0 otherwise
Age	Age of individual
Married	Dummy = 1 if married, 0 otherwise
Nchild	Number of children in household
Loweduc	Dummy = 1 if primary or lower secondary, 0 otherwise
Middleduc	Dummy = 1 if upper secondary or post-secondary non-tertiary, 0 otherwise
Higheduc	Dummy = 1 if short-cycle tertiary or higher, 0 otherwise
FT work experience	total length of full-time employment measured in years with months as decimals
PT work erpxerience	total length of part-time employment measured in years with months as decimals.
Weekly working hours	average weekly working hours including over time

5.4 Descriptive Statistics

The descriptive statistics are presented in table 3, where panel A highlights the descriptives for the employment sample and panel B for the wage sample.
The refugee subsample exhibits notable differences from the migrant sample. For instance, refugees are on average 30 p.p. less likely to be employed than other migrants. Among those who are employed, we observe that refugees earn, on average, 25 percent less than their migrant counterparts. Looking at the differences by gender, we observe that the employment gap is similar for both gender. Yet, the wage gap is more pronounced for males earning approximately 34 percent less than migrants, on average. Analyzing the differences in the covariates of the descriptive statistics already reveals potential sources for the aforementioned heterogeneity in the dependent variables. For instance, panel A in table 3 shows a large gap in educational attainment with almost every second refugee having attained less than high school. In contrast, other migrants are, on average, 13 p.p more likely to have completed high school or to hold a university degree. This implies, refugees in our sample have a lower educational level than other migrants.
We also find differences in the language proficiency. Migrants have a better language proficiency than refugees. 62 percent of the migrants can speak and write German very well, as opposed to 46 percent of refugees. Likewise, the gender samples show the same pattern with migrants having a better command of German than refugees.
In addition, length of residence in Germany varies, with migrants in the sample having lived in the country for approximately 6 years longer than asylum seekers, on average.
The demographic composition of the two groups also diverges with the majority of the refugee sample being male, whereas the migrant sample is consists mainly of females.
For the remaining explanatory variables such as age, immigration age, marital status, number of children, and geographical location the outcomes are very similar for refugees and other migrants. Interestingly, we observe a 13 p.p higher share of refugees speaking and writing German very well for the wage sample in panel B of table 3. Simultaneously, the share of individuals without good command of the language decreased. This highlights the importance between language proficiency and employment. Particularly, for female refugees the share increases by 18 p.p., whereas for male refugees the share increases by 11 p.p., indicating that language might be more important for female refugees.
Considering the previously mentioned differences in education, we can also observe heterogeneity for the wage sample. Yet, the differences for the middle educated individuals become insignificant. That is, the distribution of middle educated individuals is similar among refugees and migrants in the wage sample.

Table 3: Summary Statistics with Differences in Means

	Mean (SD)						Diff. mean (SE)			Observations
	Pooled Sample		Females		Males		Ref.- Mig.	Females	Males	
	Ref.	Mig.	Ref.	Mig.	Ref.	Mig.				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Outcome variables										
Employed	0.43	0.74	0.33	0.67	0.49	0.81	-0.30***	-0.34***	-0.33***	30,764
							(0.01)	(0.02)	(0.02)	
Log real hourly wage	2.35	2.64	2.33	2.51	2.36	2.77	-0.29***	-0.18***	-0.41***	11,024
	(0.58)	(0.55)	(0.53)	(0.53)	(0.60)	(0.55)	(0.03)	(0.05)	(0.04)	
Explanatory variables										
Panel A: Employment Sample										
Language categories										
Very well speak & write	0.46	0.62	0.45	0.63	0.46	0.61				
Either very well speak or write	0.14	0.13	0.14	0.12	0.15	0.15				
None	0.40	0.24	0.42	0.25	0.39	0.24				
Length of stay	13.30	19.30	14.35	19.28	12.72	19.33	-6.01***	-4.93***	-6.60***	
	(11.00)	(10.15)	(10.81)	(9.84)	(11.06)	(10.52)	(0.35)	(0.49)	(0.49)	
Male	0.65	0.45					0.20***			
							(0.01)			
Age	37.92	42.26	39.65	42.24	36.99	42.28	-4.33***	-2.59***	-5.29***	
	(12.75)	(11.33)	(12.12)	(11.29)	(12.98)	(11.38)	(0.39)	(0.59)	(0.53)	
Married	0.55	0.66	0.68	0.67	0.49	0.64	-0.10***	0.01	-0.16***	
							(0.01)	(0.02)	(0.02)	
Number of children	1.01	0.73	1.37	0.75	0.81	0.71	0.28***	0.62***	0.11**	
	(1.30)	(1.05)	(1.35)	(1.06)	(1.23)	(1.05)	(0.03)	(0.05)	(0.04)	
Urban	0.74	0.79	0.76	0.79	0.73	0.79	-0.05***	-0.03	-0.06**	
							(0.01)	(0.02)	(0.02)	
High education	0.17	0.29	0.18	0.30	0.16	0.28	-0.13***	-0.13***	-0.12***	
							(0.01)	(0.02)	(0.02)	
Middle education	0.35	0.48	0.33	0.44	0.36	0.53	-0.13***	-0.11***	-0.17***	
							(0.02)	(0.02)	(0.02)	
Low education	0.48	0.23	0.50	0.25	0.48	0.19	0.26***	0.24***	0.28***	
							(0.01)	(0.02)	(0.02)	
Panel B: Wage sample										
Language categories										
Very well speak & write	0.59	0.65	0.63	0.67	0.57	0.64				
Either very well speak or write	0.17	0.13	0.17	0.11	0.16	0.16				
None	0.25	0.22	0.20	0.22	0.27	0.21				
Length of stay	18.16	19.75	21.41	20.16	16.87	19.34	-1.59***	1.26*	-2.47***	
	(9.19)	(9.98)	(7.42)	(9.47)	(9.50)	(10.46)	(0.42)	(0.58)	(0.56)	
Male	0.72	0.50					0.22***			
							(0.02)			
Age	39.62	42.50	42.46	42.74	38.49	42.24	-2.88***	-0.29	-3.76***	
	(11.11)	(10.82)	(9.45)	(10.91)	(11.51)	(10.72)	(0.55)	(0.88)	(0.69)	
Married	0.66	0.67	0.75	0.66	0.63	0.68	-0.00	0.09*	-0.04	
							(0.02)	(0.04)	(0.03)	
Number of children	0.91	0.66	0.89	0.58	0.92	0.75	0.25***	0.32***	0.17**	
	(1.13)	(0.95)	(1.00)	(0.87)	(1.18)	(1.02)	(0.05)	(0.08)	(0.06)	
Urban	0.73	0.79	0.77	0.79	0.72	0.79	-0.05***	-0.03	-0.07*	
							(0.01)	(0.05)	(0.03)	
High education	0.17	0.32	0.22	0.34	0.15	0.30	-0.15***	-0.12**	-0.15***	
							(0.02)	(0.05)	(0.02)	
Middle education	0.49	0.50	0.51	0.46	0.49	0.54	-0.01	0.05	-0.05	
							(0.03)	(0.05)	(0.03)	
Low education	0.34	0.18	0.27	0.20	0.36	0.16	0.16***	0.07	0.20***	
							(0.02)	(0.04)	(0.03)	
FT work experience	10.89	13.61	5.48	9.81	13.04	17.42	-2.72***	-4.33***	-4.37***	
	(10.42)	(11.20)	(7.16)	(9.48)	(10.72)	(11.50)	(0.55)	(0.60)	(0.69)	
PT work experience	3.19	3.68	6.56	6.11	1.84	1.25	-0.49	0.46	0.59***	
	(5.12)	(5.62)	(6.55)	(6.59)	(3.65)	(2.80)	(0.29)	(0.71)	(0.18)	
Weekly working hours	34.74	36.50	28.15	30.97	37.35	42.05	-1.77**	-2.82*	-4.70***	
	(12.58)	(13.06)	(12.86)	(13.50)	(11.46)	(9.85)	(0.58)	(1.10)	(0.62)	

For columns (1)-(6) standard deviation in parentheses. For columns (7)-(9) standard errors in parentheses. Female employment sample N = 14,627. Male employment sample N = 16,137. Female wage sample N = 4,997. Male wage sample N = 6,027. Observations are weighted. Dummy variables are interpreted as percentage.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Additionally, refugees and migrants in the wage sample are, on average, more similar with respect to their length of stay in Germany as the differences between means decreased.
Regarding the differences in full-time work experience, we observe, on average, refugees have 2.7 years less experience while this gap is smaller and statistically not significant for part-time work experience with 0.49 years less.
Lastly, refugees work less hours per week than migrants. This effect is even stronger for males with male refuees working, on average, 4.7 hours less. For the remaining explanatory variables we observe similar outcomes between the groups and panel A and B.
The descriptive statistics with respect to economic sector and occupations in Appendix A.1, further suggests heterogeneity for the distribution of refugees and migrants in economic sectors. The most common sectors for refugees in our sample are manufacturing (29 percent) and wholesale and retail trade (22 percent). 28 percent of female refugees, however, are working in the public service sector. Similarly, occupational distribution reveals that refugees are disproportionately employed in low-skilled jobs, particularly in elementary occupations, and are underrepresented in high-skilled professions.

6 Methodology

6.1 Employment Probability Analysis

For the empirical analysis of the employment probability gap between refugees and other migrants, we apply a probit model. Generally, for a binary outcome it is suggested to estimate a non-linear model, confining the probability to values between 0 and 1, conditional on a set of observables. Moreover, it assumes that the unobserved propensity to success, i.e. $P = 1$, follows the cumulative normal distribution function, Φ (Horowitz and Savin, 2001). In our case we motivate the probit model as a latent variable model, where E^*_{it} is the unobserved binary variable expressing the employment propensity for individual i in year t, with auxiliary random variables as follows,

$$
\begin{aligned}
E^*_{itrs} &= \beta_1 R_{itrs} + \beta_2 X'_{itrs} + \tau_t + \theta_s + \eta_r + \epsilon_{itrs} \\
\text{where } E_{itrs} &= \begin{cases} 1, & \text{if } E^*_{itrs} > 0 \\ 0, & \text{otherwise} \end{cases}
\end{aligned} \tag{1}
$$

where E_{itrs} represents the probability of employment (employed = 1, not employed = 0) for individual i, in year t, from region of origin r, living in state s. β_1 is the coefficient of interest that captures the observed difference in employment probability between refugees and other migrants. R_{itrs} is a binary variable for whether an individual is a refugee or a migrant. X'_{itrs} is the transpose of a vector of observables including language proficiency, length of stay, gender, education level, age and its square, marital status, number of children in the household, and whether the individual lives in an urban or rural area. Survey year fixed effects, (τ_t), account for time-specific factors across the survey years that could influence the employment probability across different years. To account for regional labor market heterogeneity, state fixed effects (θ_s) are applied. Region of origin fixed effects (η_r) capture time-invariant macroeconomic conditions that potentially affect the employment probability based on an individual's origin. Finally, ϵ_{itrs} is the error term. We cluster the standard errors at the region of origin to allow for correlation across regions.

Then, we estimate the following probit model using maximum likelihood,

$$Pr(E_{itrs} = 1|R_i, X'_{itrs}) = \Phi(\beta_1 R_{itrs} + \beta_2 X'_{itrs} + \tau_t + \theta_s + \eta_r) \tag{2}$$

To interpret the results obtained from this model, we compute average marginal effects (AME), providing an interpretable measure of how being a refugee influences the probability of employment, on average. We estimate the model for the pooled sample and separately by gender. For the state analysis, we use the pooled sample and estimate the effect for each state separately.

6.2 Wage Analysis

In the second step of the analysis, we estimate the wage gap between refugees and migrants. To do so, we follow a standard OLS wage equation as follows,

$$LogRealHourlyWage_{itrso} = \alpha + \beta_1 R_{itrso} + \beta_2 X'_{itrso} + \tau_t + \theta_s + \eta_r + \lambda_o + \epsilon_{itrso} \tag{3}$$

where the dependent variable is the log real hourly wage for individual i in year t, coming from region of origin r, living in state s, and working in sector or occupation o. Real hourly wages are used to make wages across years comparable and adjust for inflation. The log of those wages is used to account for proportional wage differences and to adjust for highly skewed data. The obtained estimates can then be interpreted as approximate percentage changes in wages instead of absolute differences.

The coefficient of interest is β_1, capturing the effect of being a refugee, R_{itrso}, on the log hourly real wage. A negative coefficient would indicate a wage penalty for refugees. That is, refugees earn less than migrants per hour, conditional on observables. Conversely, a positive sign would imply a wage premium for refugees. X'_{itrso} is the transpose of a vector of observables as already mentioned for equation 2, yet we extend the covariates for the wage gap with full-time work experience, part-time work experience and their squares, and average weekly working hours. The observables included account for differences in individual characteristics, human capital endowments, and labor market experience. Furthermore, we include sector or occupation fixed effects, λ_o, separately. The former is included to account for the effect on wage differences due to industry differences. To capture specific job-related wage differences, we include occupation fixed effects. That is, refugees might sort into low-paid occupations while migrants might be more present in high-paid occupations. Thus, omitting occupation fixed effects could lead to an overestimation of the wage gap between refugees and migrants. Finally, ϵ_{itrso} is the error term. We cluster the standard errors at the region of origin to allow for correlation across regions. We estimate the model for the pooled sample and separately by gender. For the state analysis we use the pooled sample and estimate each state separately.

6.3 Oaxaca Decomposition

In the final step of our empirical strategy we decompose the mean employment and wage gap between refugees and migrants into explained and unexplained factors using the Oaxaca decomposition method. The explained factors are given by the differences in the distribution of the aforementioned observables between the two groups. In contrast, the unexplained part captures the share of the gap that is not attributable to the average difference in observables. This is often interpreted as discrimination or other factors that cannot be observed (Oaxaca, 1973). Decomposing the differences in outcomes is done by estimating separate regressions for each group (refugees and migrants) as follows,

$$Y_R - Y_M = \hat{\beta}_R(X_R - X_M) + X_M(\hat{\beta_R} - \hat{\beta_M}) \tag{4}$$

where Y_R is the respective outcome for refugees and Y_M for migrants. $\hat{\beta_R}$ and $\hat{\beta_M}$ are the conditional expectations of the variables of interest for refugees and migrants, respectively. The first term of equation 4 captures the explained part (differences in observables), while the second term represents the unexplained part (differences in coefficients).

The decomposition is based on a counterfactual outcome structure, which depends on the choice of reference group. Consistent with the previous analysis, we use refugees as the reference group.

6.4 Limitations

This section discusses the limitations of our empirical strategies that potentially may provide avenues for further research.

One of the primary limitations is that our results are rather descriptive than causal due to endogeneity concerns. That is, individuals' probability of employment and wages might be correlated with unobserved factors that we do not control for in our models (Bleakley and Chin, 2004; Chiswick and

Miller, 1995). For instance, motivation, social networks, or employer preferences could influence both outcomes studied in this thesis, leading to omitted variable bias.
Another issue in our analysis is selection bias in survey participation. For instance, if refugees who do not find employment leave the host country to seek employment in other countries, our results might underestimate the employment gap. In other words, if refugees with more applicable human capital stay in the host country, the estimated employment gap would be too small to capture the true underlying effect. Regarding the wage gap, our results might be overestimated if migrants their earnings do not match with their pre-migration expectations, and hence leave the country. This would imply that we mainly include migrants with well paid job, skewing the wage gap estimates. Furthermore, we are not able to identify the reasons why migrants chose to relocate to Germany. The results may potentially differ if most of the migrants included in our analysis came for family reunification instead of work reasons, as family migrants are often faced with worse labor market outcomes than work-oriented migrants (Ruiz and Vargas-Silva, 2018). In particular, this may potentially underestimate the true employment probability gap between refugees and migrants.

7 Results

7.1 Employment Probability Analysis

7.1.1 Germany (Pooled and by Gender)

Table 4 presents the probit results for the employment probability gap between refugees and other migrants. Overall, we find a significant negative gap, regardless of the sample and specification used. This implies that being a refugee is associated with a lower probability of being employed, in Germany. Yet, the magnitude of the estimated gap varies across specifications. Specifically, the unadjusted employment gap is the largest, ranging from 27 p.p. in the pooled sample to 32 p.p. in the female sample. This is not surprising, as with this specification we simply compare the average employment of refugees with the average employment of migrants, not adjusting for differences in age, education, or language proficiency. After accounting for these differences, the gap reduces significantly as shown in columns (2), (5), and (8). This indicates that differences in observable characteristics help explaining parts of the gap. In addition, the smallest gap is found after controlling for observables and region of origin, state and year fixed effects. In other words, controlling for socioeconomic, origin and state-level factors, help to reduce the employment difference between refugees and migrants. Interestingly, the employment probability gap for males is the largest with male refugees having a 11 p.p. lower probability of being employed, compared to their migrant counterparts, ceteris paribus. For females, this gap is lower with refugee women being

8 p.p. less likely to be employed, keeping everything else constant. Yet, we have to be more careful with the interpretation of the difference since it is not statistically significant. The results suggests, that the employment gap is reduced by accounting for differences in observables between refugees and migrants. Furthermore, the pooled gap is mainly driven by a disadvantage for refugee males.

Table 4: Employment Probability Gap - Probit Results

	Dependent variable: Employed								
	Panel A: Pooled sample			*Panel B: Females*			*Panel C: Males*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Refugee	-0.27***	-0.19***	-0.10***	-0.32***	-0.17**	-0.08**	-0.28***	-0.18***	-0.11***
	(0.0801)	(0.0607)	(0.0386)	(0.1097)	(0.0684)	(0.0346)	(0.0131)	(0.0506)	(0.0376)
Controls:									
Observables	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Region FE	No	No	Yes	No	No	Yes	No	No	Yes
Year FE	No	No	Yes	No	No	Yes	No	No	Yes
State FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	30,764	30,764	30,764	14,627	14,627	14,627	16,137	16,137	16,137

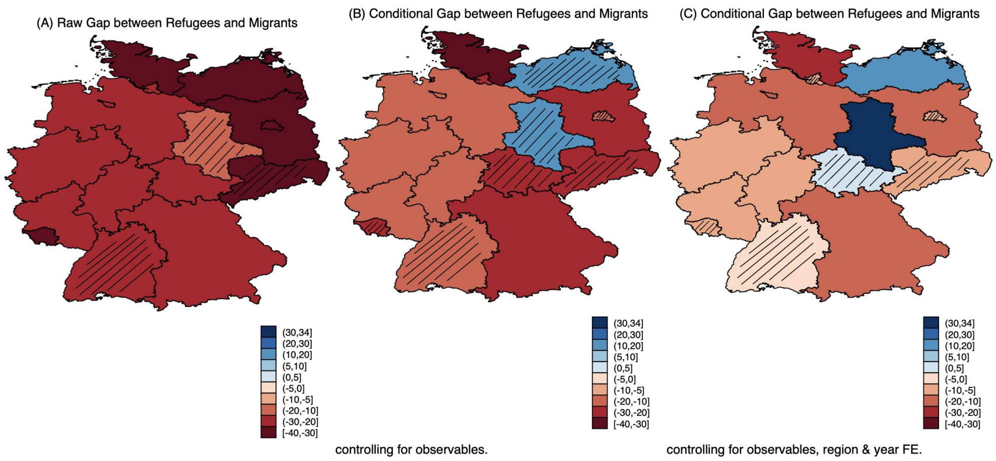
Standard errors (clustered by region of origin) in parentheses. The values displayed represent the average marginal effects. Observations are weighted.
* p< 0.1, ** p < 0.05, *** p < 0.01

7.1.2 State Analysis (Pooled Sample)

Next, we also observe heterogeneity in the results for refugees in different states in Germany, as presented in Figure 8. The corresponding result tables can be found in Appendix 7.1.2. We can observe a significant negative raw employment gap for all states. After controlling for observables this gap becomes less pronounced for most of the states and even positive, but not statistically significant, for others. In line with the findings from table 4, the gap decreases notably after including all explanatory variables and fixed effects. Interestingly, in two East states, the negative gap is reversed and significant, indicating that refugees have a higher likelihood of 12 p.p. and 34 p.p of being employed than migrants, respectively. The findings clearly show that the effect of being a refugee on the employment probability varies across states, implying that regional labor market conditions, and state-level factors play a crucial role in shaping the employment outcome for refugees in Germany.

Economic conditions may play an important role in shaping the employment probability for individuals. Hence, we assess the relationship between our results from panel C of Figure 8, where we control for observables and region of origin and survey year fixed effects, with GDP per capita growth and unemployment rate per state. Figure 10 in Appendix A.3 highlights a moderate positive relationship. That is, an increase in GDP per capita per state is associated with an improvement for refugees' labor market outcomes. Yet, the low R^2 imply that other factors are more likely to determine the variation in employment.

Figure 8: Employment Gap in Germany (2016 - 2019)



Notes: This figure shows the probit estimates of the employment gap between refugees and migrants across German states. The values displayed represent the average marginal effect for the refugee coefficient, expressed in percentage points. The shaded states do not show significant coefficients. The observations are weighted.

7.2 Wage Gap Analysis

7.2.1 Germany (Pooled and by Gender

Table 5 displays the results for the wage gap between refugees and migrants in Germany. In accordance with the findings for the employment gap, the difference in log real hourly wages between refugees and migrants is negative. Meaning, on average, being a refugee is associated with a lower hourly real wage compared to migrants. Remarkably, while this gap is significant across different specifications for the pooled and male sample, the significance diminishes for females when controlling for sector or occupation fixed effects. This implies, that female migrants and refugees within the same sector or occupation earn similar wages, after controlling for observables. On the contrary, panel (C) shows that the negative wage gap is most pronounced for males. The raw gap suggests that male refugees earn, on average, 34 percent less per hour than male migrants. This wage difference decreases by including observables and fixed effects, as shown in column (3). The smallest gap between the two groups shows column (5) controlling for occupational differences. That is, a male refugee working in the same occupation as their migrant counterpart earns, on average, 14 percent less per hour, ceteris paribus. In addition, by including occupation fixed effects, we can achieve the highest R^2 among the different specifications, suggesting that occupational differences play a significant role in explaining the wage gap between refugees and migrants in Germany. Nevertheless, the R^2 across the specifications remain below 0.5, meaning a substantial part of the variation remains unexplained setting the stage for further research. Lastly, the wage gaps observed in the pooled sample are driven by differences between males.

Table 5: Wage OLS Regression Results

	Dependent variable: Log hourly real wage				
	(1)	(2)	(3)	(4)	(5)
Panel A: Pooled sample					
Refugee	-0.29***	-0.19***	-0.17***	-0.17***	-0.12***
	(0.0716)	(0.0234)	(0.0126)	(0.0160)	(0.0111)
R-squared	0.03	0.29	0.30	0.34	0.38
Observations	11024.00	11024.00	11024.00	11024.00	11024.00
Panel B: Females					
Refugee	-0.18*	-0.10*	-0.07*	-0.06	-0.03
	(0.0778)	(0.0471)	(0.0281)	(0.0312)	(0.0283)
R-squared	0.01	0.24	0.27	0.31	0.34
Observations	4997.00	4997.00	4997.00	4997.00	4997.00
Panel C: Males					
Refugee	-0.41***	-0.23***	-0.21***	-0.21***	-0.15***
	(0.0699)	(0.0178)	(0.0095)	(0.0129)	(0.0097)
R-squared	0.07	0.30	0.32	0.37	0.42
Observations	6027.00	6027.00	6027.00	6027.00	6027.00
Controls:					
Observables	No	Yes	Yes	Yes	Yes
Region FE	No	No	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
State FE	No	No	Yes	Yes	Yes
Sector FE	No	No	No	Yes	No
Occupation FE	No	No	No	No	Yes

Standard errors (clustered by region of origin) in parentheses. Observations are weighted.
* p< 0.1, ** p< 0.05, *** p< 0.01

7.2.2 State Analysis (Pooled Sample)

Figure 9 presents the OLS estimates of the wage gap between refugees and migrants by states between 2016 and 2019. The values displayed show the percentage difference between the two groups. The result tables can be found in Appendix A.4.

In line with the findings on the employment probability gap, our results for the wage gap between refugees and migrants suggest substantial regional differences, as highlighted by Figure 9.

The raw gap, as shown in panel A, is statistically significant and negative in 14 out of 16 states. Particularly, in Berlin and Saarland the differences are not different from zero. Considering the magnitudes of the coefficients, the smallest (largest) gap is found in Rhineland-Palatinate (Mecklenburg-West Pomerania), where refugees earn on average 14 percent (64 percent) less per hour than migrants.

Next, controlling for observable characteristics such as language proficiency, length of stay, and education reduces the gap across most states and even becomes positive in West Germany. There, being a refugee is associated with an hourly wage premium of 13 percent, ceteris paribus. Again, the largest negative gap is found in East Germany, where refugees experience a wage penalty of 41 percent, everything else equal. Furthermore, in the southern states Baden-Württemberg (-12 percent) and Bavaria (-38 percent) the hourly wage gap remains negative and statistically different from zero.

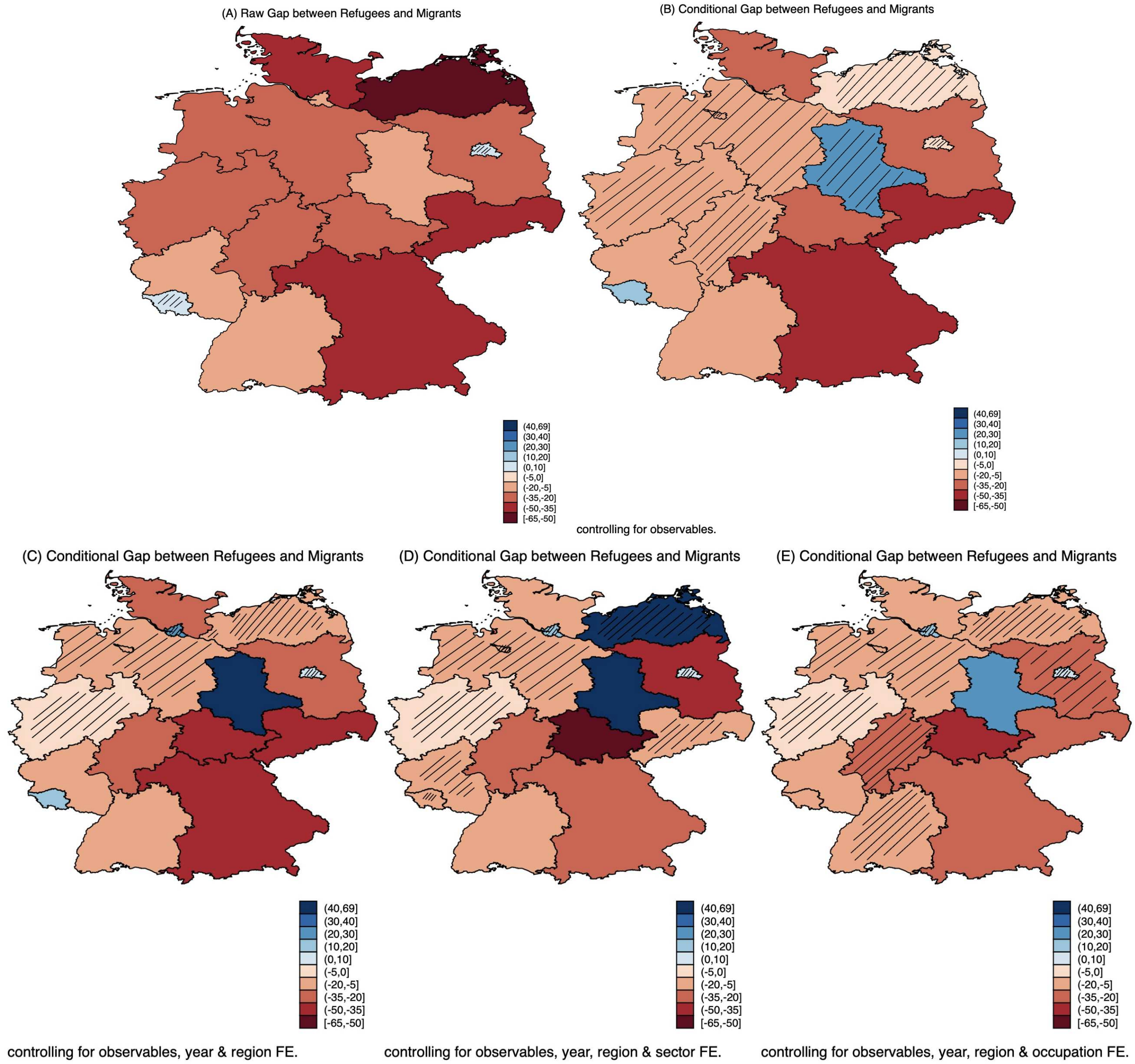
Accounting for time trends and differences due to region of origin, recovers the significance for some states, as illustrated by panel C in Figure 9. Besides recovering significance, Saxony-Anhalt also shows the largest positive wage premium for refugees, which is significant at the conventional levels. That is, refugees earn, on average, 46 percent more than their migrant counterparts, ceteris paribus. This gap diverges even further in that same state when accounting for differences between sectors, suggesting refugees earn on average 68 percent more per hour than migrants. Including occupation fixed effects, reduces the gap to 29 percent, ceteris paribus. Thus, the wage gap in this state remains ambiguous. However, the strong positive gap for refugees and the ambiguity may leave room for further research to explore the mechanisms behind this.

Furthermore, we observe the largest negative gap again in East and South Germany, with an hourly wage penalty of 47 percent in Thuringia, and 38 percent in Bavaria, everything else equal. Hence, recalling the previous results from panel B, the gap in Bavaria remains similar in panel C. Similarly, the findings in West and Southwest, in Baden-Württemberg (-10 percent), Saarland (13 percent), and Rhineland-Palatinate (-14 percent) exhibit little variation compared to the results in panel B.

Panel D of Figure 9 displays the results for the specification that controls for sector fixed effects. With this, the wage gap remains significant in less than half of the states. In line with the findings of panel D, in East Germany we observe the largest disparities between refugees and migrants in both terms, positive and negative. While Brandenburg and Thuringia demonstrate a negative gap of 42 percent and 51 percent, respectively, Saxony-Anhalt reveal a positive gap of 68 percent. Again, in accordance with the previous panels, we observe a relatively negative constant gap in the Southwest state Baden-Württemberg of 11 percent. These results suggest that in most of the states, the gap is largely explained by differences in the sectors each group works in. Yet, substantial regional variation persists, especially in East Germany where the largest positive and negative gaps are observed.

The wage gaps become statistically insignificant for every second state, after controlling for occupational differences, as presented in panel D of Figure 9. Therefore, there is evidence that occupational sorting is an important determinant in explaining the observed wage gap between refugees and migrants. In contrast to the aforementioned findings, the gap in Baden-Württemberg is no longer different from zero, indicating that after controlling for occupational sorting, refugees and migrants receive similar wages. To put this into perspective with the findings from panel (D), this suggests that in Baden-Württemberg occupational barriers preventing refugees from entering higher paying jobs may play an important role. Regarding the East states Thuringia and Saxony-Anhalt, the gap is reduced in both states, suggesting that additional factors determine the gap beyond occupational and sectoral sorting.

Figure 9: Wage Gap in Germany (2016 - 2019)



Notes: The figure shows the OLS estimates of the wage gap between refugees and migrants across German states. The values displayed represent the percentage difference between the two groups. Shaded states indicate coefficients that are not statistically significant. Observations are weighted.

Three final key points to highlight are, first, Berlin exhibits no significant wage gap across all specifications, mirroring the results of the employment gap analysis, where only the raw gap suggests a significant result.

Second, the wage gap in Bavaria persists negative and statistically significant across all specifications. The wage penalty in this state varies between 42 (panel A) and 32 (panel E) percent, ceteris paribus. This resembles the evidence from the employment gap analysis suggesting significant barriers for refugees to be employed. Even if employment is achieved, refugees experience a wage penalty. Third, although, being in the opposite geographical location, Schleswig-Holstein exhibits a similar pattern as Bavaria in terms of significance. For both, the wage and employment gap analysis, the results remain statistically different from zero and negative across all specifications. The former analysis, shows a wage penalty for refugees between 43 (panel A) and 11 (panel D) percent. As is the case with Bavaria, these findings suggest barriers for refugees to employment and then later on further exacerbate their disadvantage by experiencing a wage penalty.

Finally, to assess how key economic conditions correlate with the estimated wage disparities across states we present Figure 11 in Appendix A.3. In line with the findings for the employment gap, we find a positive moderate relationship between GDP per capita growth, mean unemployment rate per state and the estimated wage gaps. Meaning, as GDP per capita increases, the wage gap becomes less negative. While we observe that the gaps are smaller in wealthier states, there is variation around the fitted line indicating that besides economic conditions other factors are driving the gap.

7.3 Oaxaca Decomposition

Table 6: Results Oaxaca Decomposition

	Panel A: Employed			*Panel B: Log hourly real wage*		
	Pooled sample	Females	Males	Pooled sample	Females	Males
	(1)	(2)	(3)	(4)	(5)	(6)
Differential						
Refugees	0.43***	0.33**	0.49***	2.35***	2.33***	2.36***
	(0.1071)	(0.1317)	(0.1004)	(0.0756)	(0.0868)	(0.0736)
Migrants	0.74***	0.67***	0.81***	2.64***	2.51***	2.77***
	(0.0159)	(0.0232)	(0.0111)	(0.0110)	(0.0169)	(0.0169)
Difference	-0.30***	-0.34***	-0.33***	-0.29***	-0.18**	-0.41***
	(0.0924)	(0.1109)	(0.0919)	(0.0716)	(0.0778)	(0.0699)
Decomposition						
Explained	-0.19**	-0.23**	-0.21***	-0.13	-0.04	-0.21**
	(0.0794)	(0.1049)	(0.0683)	(0.0842)	(0.0665)	(0.0964)
Unexplained	-0.11***	-0.11***	-0.12***	-0.16***	-0.13***	-0.20***
	(0.0244)	(0.0336)	(0.0310)	(0.0208)	(0.0244)	(0.0539)
% Explained	63%	68%	64%	45%	22%	51%
Controls:						
Observables	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Occupation FE	No	No	No	Yes	Yes	Yes
N Refugees	18,784	7,623	11,161	2,585	543	2,042
N Migrants	11,890	7,004	4,976	8,439	4,454	3,985
N Total	30,764	14,627	16,137	11,024	4,997	6,027

Standard errors (clustered by region of origin) in parentheses. Observations are weighted. Refugees are used as reference.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6 displays the results from the Oaxaca decomposition method. Panel A of the table presents the decomposition results with respect to the employment probability gap. To decompose the employment probability gap we use an OLS estimation, as opposed to the main analysis. To show that OLS results are similar to our probit results, we show the OLS estimates for the main employment gap analysis in Appendix A.5. The decomposition of panel A shows that for the pooled sample, 63 percent of the employment gap is explained by socioeconomic and demographic factors, as shown in column (1). Similar results are observed for each gender separately. This suggests, that the employment gap between refugees and migrants can substantially be explained by differences in observables. In Appendix A.6 we show which observables help explain the gap. Interestingly, the length of stay and age contribute most to explain the employment gap.

In panel B of table 6, the results show a slightly different picture for the wage gap. Particularly, we observe that we are able to explain only 45 percent of the wage gap between refugees and migrants, yet statistically insignificant. For males, however, we are able to explain 51 percent of the wage gap through differences in observables. Although half of the wage penalty for male refugees can be explained by observables, the unobserved part suggests structural barriers for refugees to earn as much as their migrant counterparts. In other words, if male refugees have the same observable characteristics as male migrants, they would earn 19 percent higher hourly wages. Appendix A.6 shows that full-time experience and high education contribute the most to explaining the wage gap. For example, if refugee males had the same level of full-time experience as their migrant counterparts, the wage gap would be decreased by 16 percent.

Our results highlight that refugees experience a disadvantage in the labor market with respect to employment probability and wages. However, the former can be largely explained by heterogeneity in socioeconomic and demographic characteristics between migrants and refugees.

8 Robustness

8.1 Fixed Effects Interaction (2016-2019)

In the first step of our robustness analysis, we keep the main time period between 2016 and 2019 and include an interaction of state and year fixed effects for both, the employment probability and wage gap analysis. This is done to control for state-specific trends over time and analyze whether the main results change. The corresponding results are shown in Appendix A.7. As shown in table 23 and 24, the estimated coefficients remain unchanged when incorporating state-year interaction fixed effects, compared the results in table 4 and 5, respectively.

8.2 Employment Gap Analysis (2016 - 2021)

8.2.1 Germany (Pooled and by Gender)

In the second step, we investigate the robustness of our results, by expanding the time frame to 2021. Due to the COVID pandemic in 2020 and 2021, including those years may potentially alter the results.

Table 7 displays the probit results for the employment gap across the pooled, female and male sample. Overall, the gaps decreased slightly compared to our main results. For females the specification in column (7) shows no statistically significant gap.

Table 7: Employment probit results, pooled, female, male, 2016 - 2021

	Dependent variable: Employed								
	Panel A: Pooled sample			*Panel B: Females*			*Panel C: Males*		
	(1)	(2)	(3)	(5)	(6)	(7)	(9)	(10)	(11)
Refugee	-0.25***	-0.16***	-0.09***	-0.30***	-0.14**	-0.06	-0.26***	-0.17***	-0.11***
	(0.0652)	(0.0495)	(0.0296)	(0.1090)	(0.0714)	(0.0366)	(0.0122)	(0.0308)	(0.0240)
Controls:									
Observables	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Region FE	No	No	Yes	No	No	Yes	No	No	Yes
Year FE	No	No	Yes	No	No	Yes	No	No	Yes
State FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	42,243	42,243	42,243	19,925	19,925	19,925	22,318	22,318	22,318

Standard errors (clustered by region of origin) in parentheses. The values displayed show the average marginal effects. Observations are weighted
* p< 0.1, ** p< 0.05, *** p< 0.01

8.2.2 State Analysis

Similarly, the raw employment gap across states does not differ remarkably. We show the corresponding results in Figure 12 in Appendix A.7. Overall, for 11 states the magnitude of the coefficients decreased slightly while for 3 states we observed an increase. In Sachsen, however, refugees now face a signifcantly lower chance of employment by 31 p.p., opposing the result in our main analysis where we did not find a significant gap.

In line with the findings from panel B in Figure 8, we observe a decline in the difference between refugees and migrants, after controlling for observable characteristics. Yet, the differences are significant in more states now, namely in 13 as opposed to 9.

Although, incoporating region and year fixed effects makes the gap in Berlin significant, we lose significance in Hessen and Niedersachsen.

Thus, we can conclude that the pandemic years did not worsen the employment gap for refugees, instead it decreased slightly further confirming our previous results.

8.3 Wage Gap Analysis (2016 - 2021)

8.3.1 Germany (Pooled and by Gender)

Table 8: Wage Regression Results, 2016 - 2021

	Dependent variable: Log hourly real wage				
	(1)	(2)	(3)	(4)	(5)
Panel A: Pooled sample					
Refugee	-0.28***	-0.16***	-0.13***	-0.13***	-0.08***
	(0.0629)	(0.0197)	(0.0081)	(0.0093)	(0.0122)
R-squared	0.03	0.27	0.30	0.34	0.38
Observations	15,981	15,981	15,981	15,981	15,981
Panel B: Females					
Refugee	-0.17**	-0.07***	-0.05***	-0.03**	-0.01
	(0.0475)	(0.0115)	(0.0097)	(0.0124)	(0.0121)
R-squared	0.01	0.24	0.28	0.33	0.36
Observations	6,914	6,914	6,914	6,914	6,914
Panel C: Males					
Refugee	-0.40***	-0.20***	-0.17***	-0.18***	-0.11***
	(0.0662)	(0.0254)	(0.0088)	(0.0147)	(0.0130)
R-squared	0.07	0.28	0.31	0.36	0.41
Observations	9,067	9,067	9,067	9,067	9,067
Controls:					
Observables	No	Yes	Yes	Yes	Yes
Region FE	No	No	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
State FE	No	No	Yes	Yes	Yes
Sector FE	No	No	No	Yes	No
Occupation FE	No	No	No	No	Yes

Standard errors in parentheses. Observations are weighted.

* p< 0.1, ** p< 0.05, *** p< 0.01

The robustness analysis for the wage gap reveals similar patterns as with the robust employment analysis. Overall, the gap for the pooled, female and male samples is reduced after incorporating 2020 and 2021 to the analysis. What is more, while we did not find a significant gap for female asylum seekers after accounting for sectoral sorting in table 5, we observe a significant negative gap in table 8, with female asylum seekers earning 3 percent less than their migrant counterparts between 2016 and 2021, ceteris paribus.

8.3.2 State Analysis

Additionally, the gaps for the different specifications across states decreased slightly by expanding the time frame, as shown by Figure 13 in Appendix A.7. Morever, the significance for Schleswig-Holstein diminishes for the conditional specifications, contradicting our main results where it remained different from zero throughout all specifications. Nevertheless, the robustness tests do not drastically change our findings for the wage gap.

9 Conclusion

This thesis explores the labor market differences between refugees and migrants in Germany, providing descriptive evidence on heterogeneity across gender and states. The analysis is conducted using the SOEP survey data between 2016 and 2019. As robustness test we expand the time frame to 2021, which did not alter our results. We estimate the employment probability with a probit model, and the wage gap with an OLS regression. Estimating the models separately for each gender reveals a stronger effect for males. Yet, the gaps decrease after controlling for socioeconomic and demographic characteristics. For instance, male refugees face a 11 p.p. lower probability of employment and earn 14 percent less per hour than their migrant counterparts. What is more, sectoral and occupational sorting plays an important role for female refugees, since the gap becomes insignificant after accounting for these and observables.

Furthermore, our state-level analysis shows heterogeneity across German states. That is, refugees are 34 p.p. more likely to be employed in one state, while simultaneously 22 p.p. less likely to be employed in another one. Similarly, the wage gap analysis shows a maximum wage premium of approximately 70 percent for refugees in one state, while a maximum wage penalty of approximately 51 percent is found in another state. How much each of the observable characteristics, in fact, explain those gaps is examined by employing an Oaxaca decomposition method. With respect to the employment gap, we are able to explain 63 percent of that gap with age being the main driver. For the wage gap, 45 percent can be attributed to differences in observables, with full-time work experience and high education being the main variables driving the gap. This can be interpreted as evidence for labor market barriers for refugees, either through discrimination or other unobserved factors.

Nevertheless, our results cannot be interpreted as causal but serve as descriptive evidence. Furthermore, potential selection bias in survey participation and differences in migration motives pose considerable limitations to this thesis. Despite these constraints, our analysis sheds light on several opportunities for further research. First, more detailed data on migrants' motives to relocate may counteract the aforementioned limitations. Second, analyzing the mechanisms underlying the large disparities across states could provide more targeted policies to improve labor market integration for refugees in Germany. Second, looking at differences between first- and second-generation refugees and migrants may potentially present an interesting framework to study, analyzing potential generational mobility issues.

Our results, albeit descriptive, offer valuable insights into potential policy measures to improve the labor market integration of refugees in Germany. That is, easing restrictions on labor market access may facilitate employment opportunities for refugees. Furthermore, by law, refugees are geographically constrained possibly forcing them to integrate into less favorable labor markets. Hence driving the employment gap between refugees and other migrants. Moreover, the substantial unexplained gap between male refugees and migrants underscores the urge for anti-discrimination policies. Finally, supporting refugees in accessing higher-paid employment opportunities may help to reduce the wage gap for males.

References

Aldashev, A., Gernandt, J., & Thomsen, S. L. (2012). The immigrant-native wage gap in germany. Jahrbücher für Nationalökonomie und Statistik, 232(5), 490–517. https://doi.org/10.1515/jbnst-2012-0502

Algan, Y., Dustmann, C., Glitz, A., & Manning, A. (2010). The economic situation of first and second-generation immigrants in france, germany, and the united kingdom. The Economic Journal, 120(542), F4–F30. https://doi.org/10.1111/j.1468-0297.2009.02338.x

Auer, D. (2018). Language roulette – the effect of random placement on refugees' labour market integration. Journal of Ethnic and Migration Studies, 3(44), 341–362. https://doi.org/10.1080/1369183X.2017.1304208

Auer, D., & Götz, L. (2023). Refugee migration, labor demand and local employment. Socio-Economic Review, 22(2). https://doi.org/10.1093/ser/mwad040

BAMF. (2024a). Das bundesamt in zahlen 2023 (tech. rep.) (Accessed: 2025-03-03). Bundesamt für Migration und Flüchtlinge (BAMF). https://www.bamf.de/SharedDocs/Anlagen/DE/Forschung/Migrationsberichte/migrationsbericht-2023.pdf?__blob=publicationFile&v=22

BAMF. (2024b). Das deutsche asylverfahren – verfahrensablauf, zuständigkeiten, statistiken (tech. rep.) (Accessed: 2025-02-11). Bundesamt für Migration und Flüchtlinge (BAMF). https://www.bamf.de/SharedDocs/Anlagen/DE/AsylFluechtlingsschutz/Asylverfahren/das-deutsche-asylverfahren.pdf?__blob=publicationFile&v=34

Becker, G. S. (1957). The economics of discrimination. University of Chicago press.

Becker, G. S. (1975). Human capital: A theoretical and empirical analysis, with special reference to education, 2nd NBER. https://www.nber.org/books-and-chapters/human-capital-theoretical-and-empirical-analysis-special-reference-education-second-edition

Becker, S. O., & Ferrara, A. (2019). Consequences of forced migration: A survey of recent findings. Labour Economics, 59, 1–16. https://doi.org/10.1016/j.labeco.2019.02.007.

Bedaso, F. (2021). The labor market integration of refugees and other migrants in germany (tech. rep.). GLO Discussion Paper.

Blackaby, D. H., Leslie, D. G., Murphy, P. D., & O'Leary, N. C. (2002). White/ethnic minority earnings and employment differentials in britain: Evidence from the lfs. Oxford Economic Papers, 54(2), 270–297. http://www.jstor.org/stable/3488780

Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. The Review of Economics and Statistics, 86(2), 481–496. http://www.jstor.org/stable/3211642

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. The American Economic Review, 77(4), 531–553. http://www.jstor.org/stable/1814529

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). https://doi.org/10.1162/003355303322552810

Braun, S., & Omar Mahmoud, T. (2014). The employment effects of immigration: Evidence from the mass arrival of german expellees in postwar germany. The Journal of Economic History, 74(1), 69–108. https://doi.org/10.1017/S0022050714000035

Brell, C., Dustmann, C., & Preston, I. (2020). The labor market integration of refugee migrants in high-income countries. Journal of Economic Perspectives, 34(1), 94–121. https://doi.org/10.1257/jep.34.1.94

Chiswick, B. R. (1978). The effect of americanization on the earnings of foreign-born men. Journal of Political Economy, 5(86), 897–921. http://www.jstor.org/stable/1828415

Chiswick, B. R., & Miller, P. W. (1995). The endogeneity between language and earnings: International analyses. Journal of Labor Economics, 13(2), 246–288. http://www.jstor.org/stable/2535104

Cortes, K. E. (2004). Are refugees different from economic immigrants? some empirical evidence on the heterogeneity of immigrant groups in the united states. The Review of Economics and Statistics, 86(2), 465–480. https://doi.org/10.1162/003465304323031058

Dustmann, C., & Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the uk. The Economic Journal, 113(489), 695–717. http://www.jstor.org/stable/3590195

Dustmann, C., Fabbri, F., & Preston, I. (2005). The impact of immigration on the british labour market. The Economic Journal, 115(507), F324–F341. https://doi.org/https://doi.org/10.1111/j.1468-0297.2005.01038.x

Dustmann, C., Hatton, T., & Preston, I. (2005). The labour market effects of immigration. The Economic Journal, 115(507), F297–F299. http://www.jstor.org/stable/3590382

Eurostat. (2025). Asylum applicants by citizenship and age group [[Online; accessed 26-February-2025]]. https://doi.org/https://doi.org/10.2908/MIGR_ASYAPPCTZA

Horowitz, J. L., & Savin, N. E. (2001). Binary response models: Logits, probits and semiparametrics. Journal of Economic Perspectives, 15(4), 43–56. https://doi.org/10.1257/jep.15.4.43

Kee, P. (1995). Native–immigrant wage differentials in the netherlands: Discrimination? Oxford Economic Papers, 47(2), 302–317. http://www.jstor.org/stable/2663561

Kroh, M., Brücker, H., Kühne, S., Liebau, E., Schupp, J., Siegert, M., & Trübswetter, P. (2016). Das studiendesign der iab-bamf-soep-befragung von geflüchteten. [Last accessed 12 October 2024]. SOEP Survey Papers 365: Series C., (4). https://www.diw.de/documents/publikationen/73/diw_01.c.571019.de/diw_ssp0365.pdf

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. International Economic Review, 14(3), 693–709. http://www.jstor.org/stable/2525981

Peters, H. (2008). Development of wage inequality for natives and immigrants in germany-evidence from quantile regression and decomposition. SOEP Paper, (113). https://doi.org/10.2139/ssrn.1151194

Ruiz, I., & Vargas-Silva, C. (2013). The economics of forced migration. The Journal of Development Studies, 49(6), 772–784. https://doi.org/10.1080/00220388.2013.777707

Ruiz, I., & Vargas-Silva, C. (2018). Differences in labour market outcomes between natives, refugees and other migrants in the uk. Journal of Economic Geography, 18(4), 855–885. https://doi.org/10.1093/jeg/lby027

Steinhauer, H. W., Siegers, R., Siegert, M., Jacobsen, J., & Zinn, S. (2022). Sampling, nonresponse, and weighting of the 2020 refreshment sample (m6) of the iab-bamf-soep refugee panel [last accessed 5 December 2024]. SOEP Survey Paper, Series C(1104). https://www.diw.de/documents/publikationen/73/diw_01.c.841080.de/diw_ssp1104.pdf

Tumen, S. (2016). The economic impact of syrian refugees on host countries: Quasi-experimental evidence from turkey. American Economic Review, 106(5), 456–60. https://doi.org/10.1257/aer.p20161065

UNHCR. (1951). Convention and protocol relating to the status of refugees (tech. rep.). United Nations High Commissioner for Refugees. https://www.unhcr.org/media/convention-and-protocol-relating-status-refugees.

UNHCR. (2014). Global trends: War's human cost. (tech. rep.). United Nations High Commissioner for Refugees. https://www.unhcr.org/media/unhcr-global-trends-2013.

UNHCR. (2024). Global trends: Forced displacement in 2023. (tech. rep.). United Nations High Commissioner for Refugees. https://www.unhcr.org/global-trends-report-2023.

Wong, L. (2024). The labour market differences between the asylum population and other migrants. Migration Studies, 12(4). https://doi.org/10.1093/migration/mnae036

Worldbank. (2017). The world by region [online, accessed 05-March-2025]. https://datatopics.worldbank.org/sdgatlas/archive/2017/the-world-by-region.html

Appendix

A.1 Economic Sectors and Occupations

Table 9: Summary Statistics for Economic Sectors and Occupations

	Mean					
	Pooled Sample		Female Sample		Male Sample	
	Ref.	Mig.	Ref.	Mig.	Ref.	Mig.
Sectors						
Primary	0.01	0.01	0.00	0.00	0.01	0.01
Manufacturing	0.29	0.23	0.20	0.13	0.32	0.33
Utilities	0.03	0.01	0.00	0.00	0.04	0.01
Construction	0.08	0.04	0.01	0.01	0.10	0.08
Trade	0.22	0.18	0.25	0.16	0.20	0.20
Hospitality	0.11	0.07	0.08	0.07	0.12	0.07
Information	0.01	0.04	0.00	0.02	0.01	0.05
Services	0.10	0.14	0.12	0.18	0.09	0.10
Public Services	0.14	0.24	0.28	0.36	0.09	0.11
Culture	0.01	0.02	0.02	0.02	0.01	0.02
Miscellaneous	0.02	0.03	0.04	0.05	0.01	0.01
Occupations						
Manager	0.02	0.04	0.03	0.02	0.02	0.05
Professionals	0.06	0.16	0.07	0.15	0.05	0.18
Technicians	0.08	0.19	0.16	0.23	0.05	0.15
Clerical Support	0.05	0.06	0.05	0.07	0.05	0.06
Services & Sales	0.20	0.18	0.30	0.24	0.16	0.11
Skilled Workers	0.00	0.00	0.00	0.00	0.00	0.01
Craft & Trade	0.15	0.11	0.04	0.03	0.19	0.20
Plant Operator	0.18	0.10	0.04	0.04	0.23	0.16
Elementary Occupations	0.26	0.16	0.32	0.22	0.24	0.10
Armed Forces	0.00	0.00	0.00	0.00	0.00	0.00
Observations	2,585	8,439	543	4,454	2,042	3,985

A.2 Geographical Regions

Region/ Country of Origin	Migrants	Refugees	Total
[1] Europe & Central Asia			
[8] Albania	59,273	100,012	159,285
[31] Azerbaijan	27,530	44,357	71,887
[40] Austria	650,817	151	650,968
[56] Belgium	24,357		24,357
[70] Bosnia and Herzegovina	429,015	295,134	724,149
[100] Bulgaria	395,920	8,546	404,466
[112] Belarus	72,258	18,839	91,097
[191] Croatia	356,176	174,026	530,202
[203] Czech Republic	130,406	6,959	137,365
[208] Denmark	46,257		46,257
[233] Estonia	17,674		17,674
[246] Finland	20,859		20,859
[250] France	380,708	64	380,772
[268] Georgia	41,214	13,106	54,320
[300] Greece	341,804	45	341,849
[348] Hungary	466,474		466,474
[372] Ireland	29,307		29,307
[380] Italy	1,018,616	644	1,019,260
[398] Kazakhstan	2,875,062	43	2,875,105
[417] Kyrgyzstan	137,254	7,458	144,712
[428] Latvia	64,902	195	65,097
[440] Lithuania	101,896		101,896
[442] Luxembourg	10,776		10,776
[498] Moldova	49,115	2,501	51,616
[499] Montenegro	12,090	6,488	18,578
[528] Netherlands	510,071		510,071
[578] Norway	21,958		21,958
[616] Poland	3,353,921	54,032	3,407,953
[620] Portugal	143,097	14,764	157,861
[642] Romania	1,347,578	9,934	1,357,512
[643] Russia	1,944,714	65,888	2,010,602
[688] Serbia	200,992	136,526	337,518
[703] Slovakia	107,444		107,444
[705] Slovenia	102,147	1,760	103,907
[724] Spain	190,430		190,430
[752] Sweden	42,415		42,415
[756] Switzerland	177,795	5,088	182,883

Region/ Country of Origin	Migrants	Refugees	Total
[762] Tajikistan	69,938	5,792	75,730
[792] Turkey	2,424,316	340,606	2,764,922
[795] Turkmenistan	79,043	110	79,153
[804] Ukraine	486,714	75,127	561,841
[807] North Macedonia	226,486	59,172	285,658
[826] United Kingdom	193,518	367	193,885
[860] Uzbekistan	74,408	16,329	90,737
[900] Kosovo	444,828	554,202	999,030
Total	19,901,573	2,018,265	21,919,838
[3] Latin America & Caribbean			
[32] Argentina	41,203		41,203
[68] Bolivia	7,373		7,373
[76] Brazil	233,031		233,031
[152] Chile	93,207		93,207
[170] Colombia	139,668	1,040	140,708
[188] Costa Rica	90	5,742	5,832
[192] Cuba	121,951		121,951
[214] Dominican Republic	1,812		1,812
[218] Ecuador	6,726		6,726
[222] El Salvador	15,480		15,480
[340] Honduras	1,322		1,322
[388] Jamaica	39,397		39,397
[484] Mexico	122,460		122,460
[558] Nicaragua	4,474		4,474
[604] Peru	80,795		80,795
[740] Suriname	131,125		131,125
[858] Uruguay	15,182		15,182
[862] Venezuela	16,507	526	17,033
Total	1,071,803	7,308	1,079,111

Region/ Country of Origin	Migrants	Refugees	Total
[4] Middle East & North Africa			
[12] Algeria	2,402	10,975	13,377
[275] Palestine	148,456	14,253	162,709
[364] Iran	129,535	438,731	568,266
[368] Iraq	60,815	653,554	714,369
[376] Israel	6,926		6,926
[400] Jordan	26,429	6,975	33,404
[414] Kuwait	6,573	5,781	12,354
[422] Lebanon	116,111	160,959	277,070
[434] Libya	2,808	15,362	18,170
[504] Morocco	304,056	36,411	340,467
[634] Qatar		699	699
[682] Saudi Arabia	354	4,761	5,115
[760] Syria	150,362	1,431,776	1,582,138
[784] United Arab Emirates		4,903	4,903
[788] Tunesia	145,144	2,464	147,608
[818] Egypt	42,681	12,367	55,048
[887] Yemen		5,026	5,026
Total	1,142,652	2,804,997	3,947,649
[5] North America			
[124] Canada	12,813		12,813
[840] United States of America	398,281		398,281
Total	411,094		411,094
[6] South Asia			
[4] Afghanistan	22,410	510,031	532,441
[50] Bangladesh	4,476	77,723	82,199
[51] Armenia	24,703	68,882	93,585
[144] Sri Lanka	125,635	84,329	209,964
[356] India	284,405	12,425	296,830
[524] Nepal	56,329	465	56,794
[586] Pakistan	262,944	174,655	437,599
Total	780,902	928,510	1,709,412

Region/ Country of Origin	Migrants	Refugees	Total
[7] Sub-Saharan Africa			
[24] Angola	30,501		30,501
[72] Botswana	26,666		26,666
[120] Cameroon	48,282	30,668	78,950
[148] Chad		6,212	6,212
[180] Democratic Republic of the Congo	24,901	8,360	33,261
[231] Ethiopia	4,951	72,394	77,345
[232] Eritrea		157,838	157,838
[270] Gambia		41,007	41,007
[288] Ghana	91,873	14,340	106,213
[324] Guinea	53,159	18,586	71,745
[384] Côte d'Ivoire	5,500	3,305	8,805
[404] Kenya	30,641	2,000	32,641
[426] Lesotho	66,976		66,976
[466] Mali		7,446	7,446
[508] Mozambique	868		868
[516] Namibia	1,103		1,103
[562] Niger	4,532	4,688	9,220
[566] Nigeria	112,825	140,067	252,892
[646] Rwanda		335	335
[686] Senegal		2,457	2,457
[694] Sierra Leone		885	885
[706] Somalia		66,509	66,509
[710] South Africa	54,877		54,877
[716] Zimbabwe		1,638	1,638
[729] Sudan	8,760	12,640	21,400
[768] Togo	13,608	31,697	45,305
[800] Uganda		1,720	1,720
[854] Burkina Faso		5,498	5,498
[894] Zambia	4,341		4,341
Total	584,364	630,290	1,214,654

Region/ Country of Origin	Migrants	Refugees	Total
[8] East Asia & Pacific			
[104] Myanmar	4,875	4,874	9,749
[116] Cambodia	8,447		8,447
[156] China	115,209	715	115,924
[158] Taiwan	27,693		27,693
[344] Hongkong	2,795		2,795
[360] Indonesia	15,401		15,401
[392] Japan	27,291		27,291
[410] South Korea	37,163		37,163
[418] Laos	16,196		16,196
[458] Malaysia	31,920		31,920
[496] Mongolia		1,725	1,725
[554] New Zealand	52,422		52,422
[608] Philippines	142,327		142,327
[702] Singapore	31,747		31,747
[704] Vietnam	105,773	29,824	135,597
[764] Thailand	241,565		241,565
[882] Samoa	6,481		6,481
Total	867,305	37,138	904,443
Total	24,759,693	6,426,508	31,186,201

A.3 Probit Results by State

Table 10: Probit Results by State 2016 - 2019

	Independent variable: Employed															
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Refugee	-0.21	-0.27***	-0.30*	-0.32***	-0.40***	-0.28***	-0.29***	-0.36***	-0.28***	-0.26***	-0.21**	-0.37**	-0.32	-0.11	-0.39***	-0.25*
	(0.1541)	(0.0781)	(0.1323)	(0.0221)	(0.0678)	(0.0637)	(0.0345)	(0.0377)	(0.0715)	(0.0676)	(0.0764)	(0.1301)	(0.1718)	(0.1802)	(0.0355)	(0.1128)
Controls:																
Observables	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Region FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Year FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Observations	3870.00	4530.00	1208.00	922.00	438.00	723.00	2773.00	268.00	3292.00	7454.00	1620.00	478.00	768.00	544.00	1251.00	625.00

Standard errors (clustered by region) in parentheses. The values displayed represent the average marginal effects. Observations are weighted. Each column represents one state.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Probit Results by State, 2016 - 2019

	Independent variable: Employed															
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Refugee	-0.19	-0.21***	-0.13	-0.21***	-0.24***	-0.23**	-0.18***	0.13	-0.19***	-0.14**	-0.17***	-0.20	-0.23	0.18	-0.35***	-0.22
	(0.1359)	(0.0212)	(0.0690)	(0.0514)	(0.0314)	(0.0819)	(0.0311)	(0.1030)	(0.0528)	(0.0536)	(0.0368)	(0.1336)	(0.1406)	(0.1089)	(0.0397)	(0.1289)
Controls:																
Observables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Year FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Observations	3870.00	4530.00	1208.00	922.00	438.00	723.00	2773.00	268.00	3292.00	7454.00	1620.00	478.00	768.00	544.00	1251.00	625.00

Standard errors (clustered by region) in parentheses. The values displayed represent the average marginal effects. Observations are weighted. Each column represents one state.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

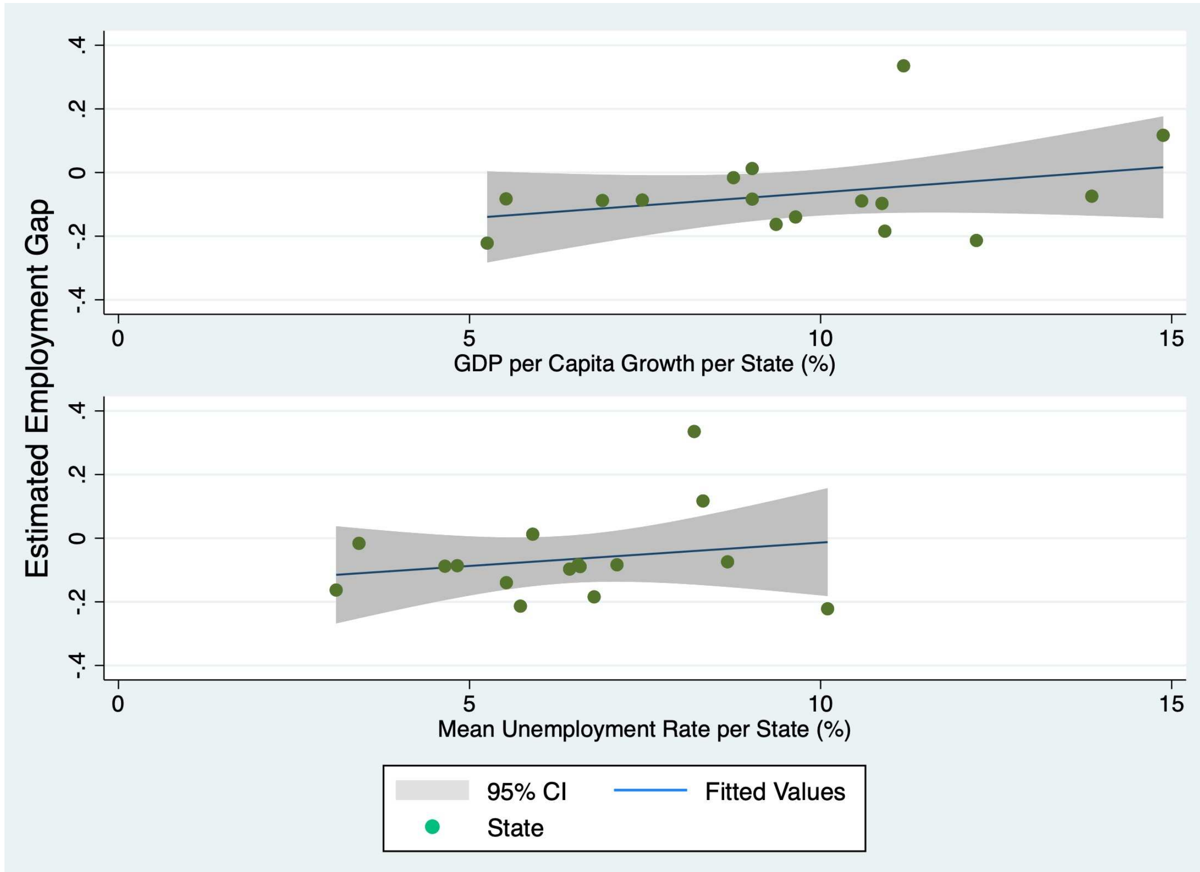
Table 12: Probit Results by State, 2016 - 2019

	Independent variable: Employed															
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Refugee	-0.02	-0.16***	-0.07	-0.18*	-0.22***	-0.09	-0.09***	0.12*	-0.14***	-0.08**	-0.09***	-0.08	-0.10	0.34***	-0.21**	0.01
	(0.1028)	(0.0430)	(0.0582)	(0.1099)	(0.0342)	(0.0720)	(0.0238)	(0.0617)	(0.0457)	(0.0401)	(0.0286)	(0.1733)	(0.2025)	(0.0777)	(0.0831)	(0.1662)
Controls:																
Observables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3863.00	4530.00	1200.00	922.00	435.00	723.00	2773.00	268.00	3288.00	7454.00	1620.00	475.00	763.00	540.00	1249.00	607.00

Standard errors (clustered by region) in parentheses. The values displayed represent the average marginal effects. Observations are weighted. Each column represents one state.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 10: Relationship between Key Economic Indicators and Employment Gap



Notes: Upper plot: R^2 = 0.10 Correlation Coefficient = 0.32. Bottom plot: R^2 = 0.04 Correlation Coefficient = 0.20.

Table 13: OLS Results by State, 2016 - 2019

	Independent variable: Log hourly real wage															
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Refugee	-0.19*	-0.54***	0.03	-0.36**	-0.35**	-0.20*	-0.36***	-1.01***	-0.28*	-0.23**	-0.15*	0.01	-0.62***	-0.19**	-0.56***	-0.38**
	(0.0865)	(0.0700)	(0.1675)	(0.1078)	(0.0810)	(0.0954)	(0.0835)	(0.1484)	(0.1349)	(0.0674)	(0.0639)	(0.1978)	(0.1094)	(0.0612)	(0.1395)	(0.1119)
Controls:																
Observables	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Region FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Year FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Sector FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Occupation FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
R-squared	0.02	0.06	0.00	0.08	0.07	0.02	0.04	0.51	0.03	0.02	0.01	0.00	0.11	0.03	0.14	0.08
Observations	1561.00	1996.00	377.00	196.00	124.00	247.00	903.00	51.00	1097.00	2754.00	779.00	124.00	154.00	102.00	335.00	224.00

Standard errors (clustered by region) in parentheses. Observations are weighted. Each column represents one state.

* p < 0.1, ** p < 0.05, *** p < 0.01

A.4 Wage Results by State

Table 14: OLS Results by State, 2016 - 2019

	Independent variable: Log hourly real wage															
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Refugee	-0.13*	-0.48***	-0.01	-0.32***	-0.06	-0.07*	-0.14	-0.02	-0.15	-0.07	-0.19**	0.13**	-0.53*	0.26	-0.32*	-0.34*
	(0.0608)	(0.1141)	(0.0268)	(0.0695)	(0.1287)	(0.0305)	(0.1283)	(0.3216)	(0.0988)	(0.0402)	(0.0554)	(0.0380)	(0.2395)	(0.2249)	(0.1361)	(0.1670)
Controls:																
Observables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Year FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Sector FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Occupation FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
R-squared	0.36	0.38	0.22	0.58	0.53	0.45	0.35	0.84	0.35	0.31	0.28	0.32	0.68	0.55	0.49	0.27
Observations	1561.00	1996.00	377.00	196.00	124.00	247.00	903.00	51.00	1097.00	2754.00	779.00	124.00	154.00	102.00	335.00	224.00

Standard errors (clustered by region) in parentheses. Observations are weighted. Each column represents one state.

* p < 0.1, ** p < 0.05, *** p < 0.01

Table 15: OLS Results by State, 2016 - 2019

	Independent variable: Log hourly real wage															
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Refugee	-0.11*	-0.50***	0.04	-0.36*	-0.20*	0.18	-0.24**	-0.13	-0.11	-0.04	-0.15**	0.13**	-0.45**	0.38**	-0.23***	-0.63**
	(0.0502)	(0.1067)	(0.0215)	(0.1653)	(0.0902)	(0.0907)	(0.0756)	(0.4026)	(0.1094)	(0.0393)	(0.0491)	(0.0434)	(0.1832)	(0.0936)	(0.0434)	(0.2218)
Controls:																
Observables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Occupation FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
R-squared	0.37	0.41	0.25	0.64	0.68	0.56	0.38	0.85	0.39	0.33	0.34	0.37	0.71	0.62	0.56	0.46
Observations	1561.00	1996.00	377.00	196.00	124.00	247.00	903.00	51.00	1097.00	2754.00	779.00	124.00	154.00	102.00	335.00	224.00

Standard errors (clustered by region) in parentheses. Observations are weighted. Each column represents one state.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 16: OLS Results by State, 2016 - 2019

	Independent variable: Log hourly real wage															
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Refugee	-0.12**	-0.43**	0.06	-0.55**	-0.23	0.12	-0.23***	0.44	-0.16	-0.02	-0.06	-0.11	-0.15	0.52***	-0.12*	-0.71***
	(0.0347)	(0.1176)	(0.0335)	(0.1567)	(0.1697)	(0.0611)	(0.0536)	(0.9260)	(0.1069)	(0.0426)	(0.0640)	(0.2196)	(0.2854)	(0.0914)	(0.0601)	(0.0952)
Controls:																
Observables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
R-squared	0.43	0.48	0.31	0.69	0.74	0.62	0.43	0.91	0.50	0.38	0.46	0.66	0.79	0.67	0.65	0.55
Observations	1561.00	1996.00	377.00	196.00	124.00	247.00	903.00	51.00	1097.00	2754.00	779.00	124.00	154.00	102.00	335.00	224.00

Standard errors (clustered by region) in parentheses. Observations are weighted. Each column represents one state.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

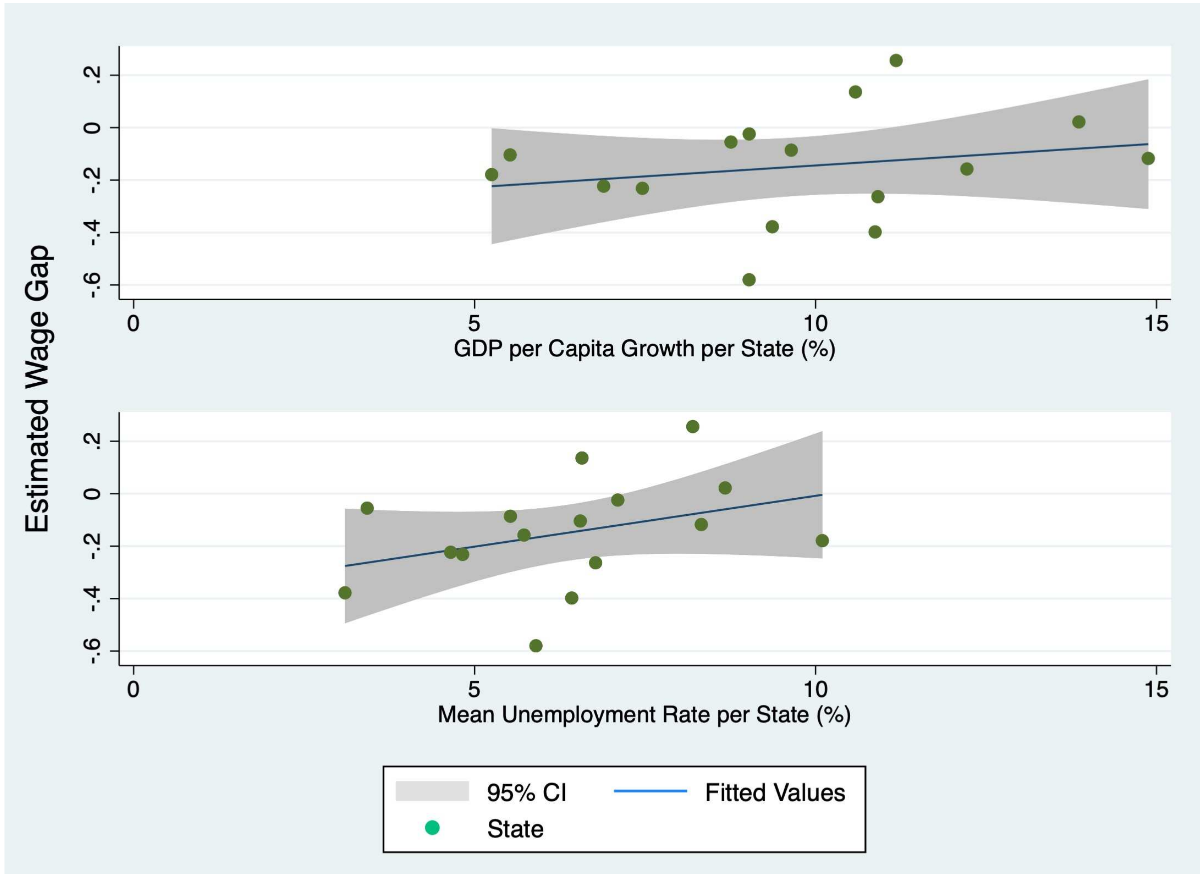
Table 17: OLS Results by State, 2016 - 2019

	Independent variable: Log hourly real wage															
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Refugee	-0.05	-0.38***	0.02	-0.26	-0.18*	0.14	-0.23	-0.12	-0.09	-0.02	-0.22**	-0.10**	-0.40***	0.26**	-0.16**	-0.58**
	(0.0390)	(0.1006)	(0.0874)	(0.1841)	(0.0828)	(0.1096)	(0.1428)	(0.6170)	(0.1083)	(0.0470)	(0.0640)	(0.0364)	(0.0992)	(0.0889)	(0.0502)	(0.1786)
Controls:																
Observables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Occupation FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.49	0.49	0.39	0.66	0.75	0.60	0.46	0.92	0.42	0.41	0.50	0.54	0.81	0.68	0.65	0.51
Observations	1561.00	1996.00	377.00	196.00	124.00	247.00	903.00	51.00	1097.00	2754.00	779.00	124.00	154.00	102.00	335.00	224.00

Standard errors (clustered by region) in parentheses. Observations are weighted. Each column represents one state.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 11: Relationship between Key Economic Indicators and Wage Gap



Notes: Upper plot: $R^2 = 0.05$ Correlation Coefficient = 0.22. Bottom plot: $R^2 = 0.13$ Correlation Coefficient = 0.36.

A.5 OLS Employment Gap

A.5.1 Germany

Table 18: OLS Results Employment Probability Gap

	Dependent variable: Employed								
	Panel A: Pooled Sample			*Panel B: Females*			*Panel C: Males*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Refugee	-0.30**	-0.21**	-0.11**	-0.34**	-0.18**	-0.08**	-0.33***	-0.22**	-0.14*
	(0.0924)	(0.0697)	(0.0462)	(0.1109)	(0.0676)	(0.0292)	(0.0189)	(0.0644)	(0.0577)
Controls:									
Observables	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Region FE	No	No	Yes	No	No	Yes	No	No	Yes
Year FE	No	No	Yes	No	No	Yes	No	No	Yes
State FE	No	No	Yes	No	No	Yes	No	No	Yes
Rsquared	0.07	0.18	0.21	0.06	0.20	0.22	0.11	0.19	0.23
Observations	30,764	30,764	30,764	14,627	14,627	14,627	16,137	16,137	16,137

Standard errors (clustered at region of origin) in parentheses. Observations are weighted. The results are for the time period 2016 - 2019.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.5.2 State Analysis

Table 19: OLS Results Employment Probability Gap - Raw Gap

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
refugee	-0.24	-0.33**	-0.33*	-0.35***	-0.46**	-0.29***	-0.32***	-0.40***	-0.31**	-0.28***	-0.25**	-0.41**	-0.34	-0.11	-0.43***	-0.26*
	(0.1869)	(0.0991)	(0.1426)	(0.0274)	(0.1002)	(0.0718)	(0.0358)	(0.0521)	(0.0850)	(0.0753)	(0.0976)	(0.1568)	(0.2144)	(0.1840)	(0.0484)	(0.1257)
Controls:																
Observables	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Region FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Year FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Rsquared	0.05	0.07	0.06	0.12	0.21	0.06	0.07	0.16	0.08	0.05	0.04	0.14	0.11	0.01	0.16	0.04
Observations	3,870	4,530	1,208	922	438	723	2,773	268	3,292	7,454	1,620	478	768	544	1,251	625

Standard errors (clustered at region of origin) in parentheses. Observations are weighted. Each column represents one state in Germany. The results are for the time period 2016 - 2019.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 20: OLS Results Employment Probability Gap - Conditional Gap

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Refugee	-0.21	-0.26***	-0.13	-0.27***	-0.32***	-0.25**	-0.21***	0.06	-0.21**	-0.16**	-0.19***	-0.22	-0.26	0.19	-0.38***	-0.23
	(0.1670)	(0.0305)	(0.0751)	(0.0581)	(0.0167)	(0.0809)	(0.0349)	(0.0861)	(0.0618)	(0.0588)	(0.0462)	(0.1581)	(0.1707)	(0.1209)	(0.0352)	(0.1524)
Controls:																
Observables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Year FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Rsquared	0.16	0.17	0.19	0.36	0.53	0.34	0.21	0.58	0.22	0.19	0.21	0.37	0.17	0.40	0.35	0.32
Observations	3,870	4,530	1,208	922	438	723	2,773	268	3,292	7,454	1,620	478	768	544	1,251	625

Standard errors (clustered at region of origin) in parentheses. Observations are weighted. Each column represents one state in Germany. The results are for the time period 2016 - 2019.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 21: OLS Results Employment Probability Gap - Conditional Gap with FE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Refugee	-0.02	-0.19***	-0.08	-0.22	-0.27***	-0.14*	-0.10***	0.03	-0.16**	-0.10*	-0.09**	-0.10	-0.12	0.37**	-0.24**	0.10
	(0.1120)	(0.0461)	(0.0728)	(0.1112)	(0.0267)	(0.0606)	(0.0226)	(0.0659)	(0.0510)	(0.0454)	(0.0298)	(0.2104)	(0.2223)	(0.1248)	(0.0967)	(0.2222)
Controls:																
Observables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rsquared	0.22	0.21	0.24	0.38	0.55	0.40	0.28	0.59	0.24	0.20	0.26	0.41	0.28	0.47	0.41	0.49
Observations	3,870	4,530	1,208	922	438	723	2,773	268	3,292	7,454	1,620	478	768	544	1,251	625

Standard errors (clustered at region of origin) in parentheses. Observations are weighted. Each column represents one state in Germany. The results are for the time period 2016 - 2019.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.6 Oaxaca

Table 22: Oaxaca decomposition

	Panel A: Employed			*Panel B: Log hourly real wage*		
	Pooled sample	Females	Males	Pooled sample	Females	Males
	(1)	(2)	(3)	(4)	(5)	(6)
Explained						
Lstay	-0.05*	-0.04	-0.05**	-0.02	0.05	-0.00
	(0.0247)	(0.0337)	(0.0216)	(0.0323)	(0.0571)	(0.0109)
Language	-0.01***	-0.01**	-0.01**	0.00	-0.00	-0.00
	(0.0042)	(0.0049)	(0.0052)	(0.0004)	(0.0048)	(0.0011)
High Educ	-0.01***	-0.02***	-0.01**	-0.01**	-0.01	-0.02***
	(0.0049)	(0.0057)	(0.0053)	(0.0064)	(0.0081)	(0.0055)
Middle Educ	-0.02***	-0.02**	-0.02***	0.00	-0.00	-0.00
	(0.0051)	(0.0072)	(0.0049)	(0.0009)	(0.0044)	(0.0015)
Male	0.04***			0.01		
	(0.0074)			(0.0254)		
Age	-0.15**	-0.11	-0.15**	-0.13	-0.03	-0.08
	(0.0778)	(0.0942)	(0.0579)	(0.1333)	(0.1162)	(0.2004)
Age squared	0.16**	0.11	0.17***	0.16	0.07	0.13
	(0.0744)	(0.0903)	(0.0560)	(0.1357)	(0.1235)	(0.1794)
Married	-0.01	0.00	-0.02	-0.00	-0.00	-0.01
	(0.0104)	(0.0023)	(0.0199)	(0.0216)	(0.0129)	(0.0379)
Nchild	-0.01	-0.04	-0.00	-0.01***	-0.02	-0.01
	(0.0073)	(0.0283)	(0.0035)	(0.0036)	(0.0140)	(0.0052)
Urban	0.00	0.00	0.00	-0.00	-0.00	-0.00
	(0.0015)	(0.0025)	(0.0022)	(0.0023)	(0.0019)	(0.0028)
FT experience				-0.06**	0.08**	-0.17*
				(0.0236)	(0.0345)	(0.0905)
FT experience squared				0.01	-0.11***	0.07
				(0.0162)	(0.0302)	(0.0521)
Pt experience				0.01	-0.02	-0.01
				(0.0140)	(0.0697)	(0.0137)
PT experience squared				-0.01	0.01	0.01
				(0.0115)	(0.0408)	(0.0112)
Weekly working hours				0.01***	0.02	0.01
				(0.0019)	(0.0104)	(0.0189)
Controls:						
Observables	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Occupation FE	No	No	No	Yes	Yes	Yes
N Refugees	18,784	7,623	11,161	2,585	543	2,042
N Migrants	11,890	7,004	4,976	8,439	4,454	3,985
N Total	30,764	14,627	16,137	11,024	4,997	6,027

Standard errors (clustered at region) in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.7 Robustness Results

A.7.1 State & Year Fixed Effects Interaction

Table 23: Employment Probability Gap - Probit Results (2016-2019)

	Dependent variable: Employed		
	Panel A: Pooled Sample	*Panel B: Female Sample*	*Panel C: Male Sample*
	(1)	(2)	(3)
Refugee	-0.10***	-0.08**	-0.11***
	(0.0370)	(0.0321)	(0.0370)
Controls:			
Observables	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Year x State FE	Yes	Yes	Yes
Observations	30,764	14,627	16,137

Standard errors (clustered by region of origin) in parentheses. The values displayed show the average marginal effects. Observations are weighted. * p< 0.1, ** p< 0.05, *** p< 0.01

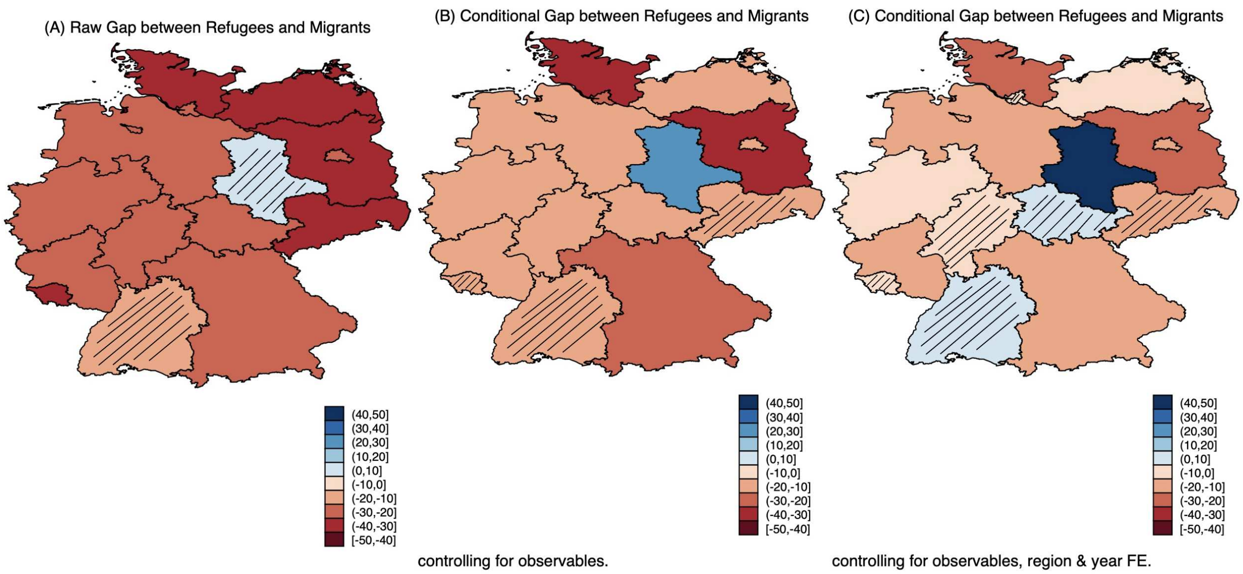
Table 24: Wage Regression Results (2016-2019)

	Dependent Variable: Log hourly real wage								
	Panel A: Pooled Sample			*Panel B: Female Sample*			*Panel C: Male Sample*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
refugee	-0.16***	-0.16***	-0.11***	-0.07*	-0.06	-0.03	-0.20***	-0.21***	-0.14***
	(0.0120)	(0.0148)	(0.0104)	(0.0292)	(0.0320)	(0.0291)	(0.0106)	(0.0111)	(0.0100)
Controls:									
Observables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year x State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.31	0.35	0.39	0.28	0.32	0.35	0.33	0.38	0.43
Observations	11,024	11,024	11,024	4,997	4,997	4,997	6,027	6,027	6,027

Standard errors (clustered by region of origin) in parentheses. Observations are weighted. * p< 0.1, ** p< 0.05, *** p< 0.01

A.7.2 Employment Gap by State

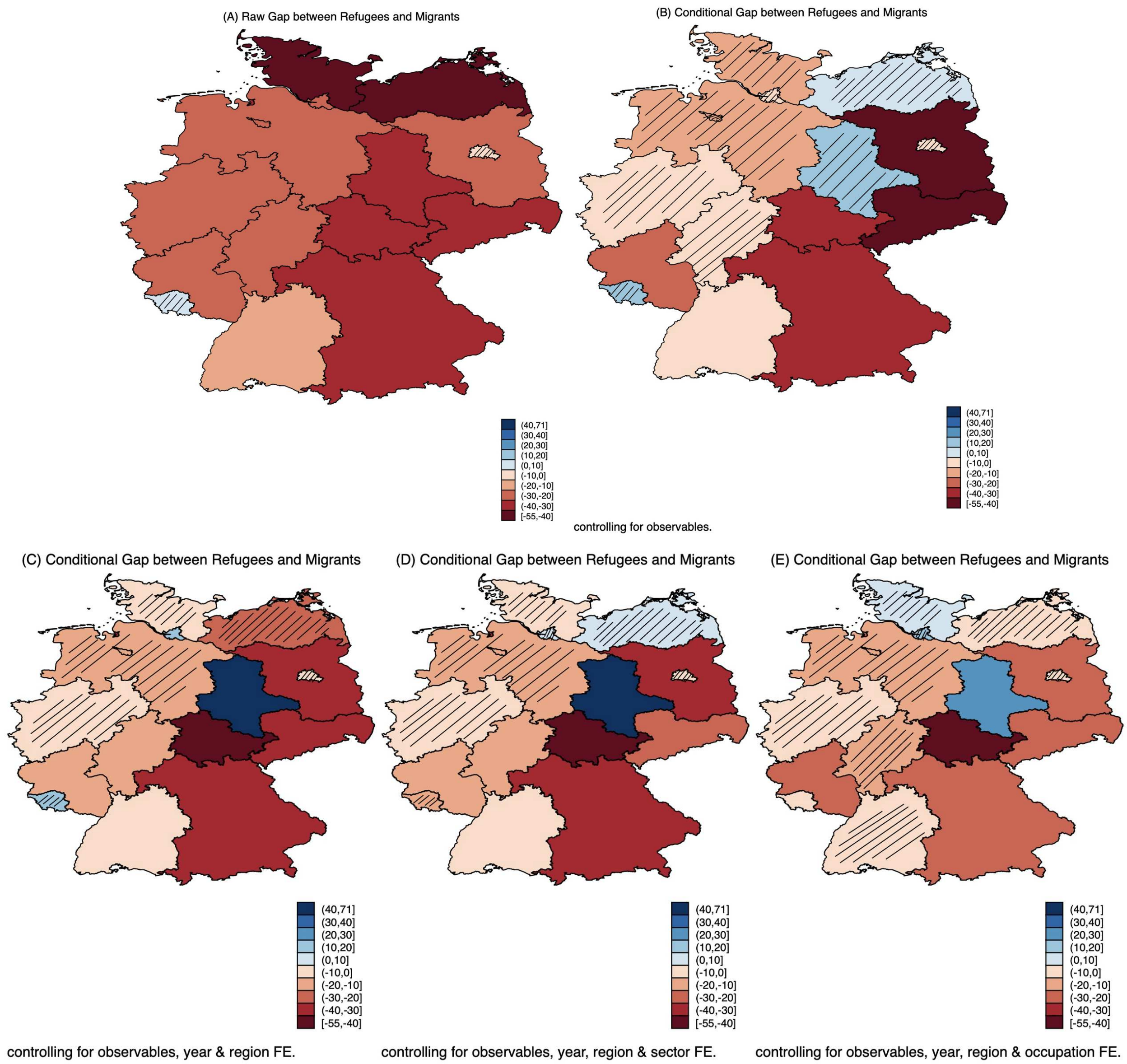
Figure 12: Employment Gap in Germany (2016 - 2021)



Notes: The figure shows the results for the probit estimation of the employment gap between refugees and migrants across German states. The values displayed represent the average marginal effect for the refugee coefficient, expressed in percentage points. Shaded states indicate coefficients that are not statistically significant at the conventional levels. Observations are weighted.

A.7.3 Wage Gap by State

Figure 13: Wage Gap Analysis (2016 - 2021)



Notes: The figure shows the OLS estimates of the wage gap between refugees and migrants across German states. The values displayed represent the percentage difference between the two groups. Shaded states indicate coefficients that are not statistically significant. Observations are weighted.