NHH

Labour Economic Outcomes and the Returns to Language Acquisition

A Comparative Study of Refugees and Economic Immigrants in Germany

Beverly Faye Becker   **Lorenz Vignold-Majal**

Supervisor: Aline Bütikofer

Master Thesis, MSc, Economics

NORWEGIAN SCHOOL OF ECONOMICS

Labour Economic Outcomes and the Returns to Language Acquisition: A Comparative Study of Refugees and Economic Immigrants in Germany

Beverly Faye Becker   Lorenz Vignold-Majal

20th June 2017

Abstract

This paper analyses the differences in labour economic and language outcomes between refugees and economic immigrants in Germany. We use data from the German Socioeconomic Panel (GSOEP), a yearly household survey of the population of Germany. Analysing the period from 1994 to 2014, we study only immigrants working at the time of survey. We find that working refugees, on average, earn between 17.6 and 19.2 percent less than working economic immigrants in Germany. The reasons for this are that refugees work on average 10.9 percent less hours annually and earn 8.3 percent (though not a significant result) less hourly wages, than their counterparts.

We further set out to explore how important language proficiency is for the annual earnings of these two distinct groups of working immigrants. We find that for both groups, being classified as "Good" or "Very Good" in spoken German improves their earnings on average by around 11.9 percent. In addition, working refugees are not more likely to speak German well, or improve their language proficiency at a faster rate, compared to working economic immigrants. Our main models of choice use the Ordinary Least Squares (OLS) estimation method with robust standard errors.

Contents

List of Figures

1 Introduction

Germany has the second-most immigrants in the world (OECD, 2017). These immigrants come to Germany for a variety of reasons, from a large array of countries, and with many different motivations, qualities, and skills. As a result, immigrants should not be characterized as a homogeneous group. Instead, we categorize immigrants into two distinct groups; refugees and economic immigrants.

The UN Refugee Agency (UNHCR, 2017) defines refugees as individuals who fled from their home country due to the threat of persecution, and are in most cases unable to return home. By this definition, refugees would most likely settle permanently in their host country. In contrast, economic immigrants are defined as individuals who choose to work in Germany for the purpose of maximizing their economic outcomes, and are able to return to their home country at any desired time. In fact, many economic immigrants can be expected to eventually return home, whether it is to support family members, build a house, or retire (Cortes, 2004). As a result, these two groups differ with respect to their expected length of stay [1] in Germany.

As we will see from the literature review section of this paper, some academic literature suggests that due to their longer time horizons refugees have relatively greater incentives to invest in country-specific human capital than economic immigrants, such as learning the German language, to give themselves the best possible opportunities in Germany. By this reasoning, refugees should be in a position to have more favourable economic outcomes in Germany than economic immigrants (Cortes, 2004).

This paper seeks to answer the following research questions. First, we aim to determine whether there are significant differences in annual earnings between working refugees and working economic immigrants. We further investigate if these differences are due to the distinctions in hourly wages or annual working hours between the two groups.

Following that, we investigate the importance of language proficiency on immigrants' earnings. According to Dustmann and Van Soest (2002), language is likely the most important "single alterable factor" contributing to an immigrant's integration success. In addition, we explore whether one immigrant group has a greater likelihood than the other to speak German at a high proficiency level, and if there are differences in the rate of language improvement.

Analysing only migrants who are working at the time of survey, we find that refugees, on average, earn 17.6 to 19.2 percent less than economic immigrants in Germany. The direct reasons for this are that refugees, on average, work 10.9 percent fewer annual hours, and earn 8.3 percent (a non-significant result) less hourly wages, than economic immigrants.

---

[1]Or "time-horizon"

We find that most of these differences in annual income between economic immigrants and refugees are captured only by the male refugees.

Secondly, we find that being classified as "Good" or "Very Good" in spoken German improves the earnings of immigrants by 11.2 to 11.9 percent, with no distinction between the immigrant groups. Analysing only the male immigrant population, this effect is increased to between 23.9 and 24.5 percent. Finally, we find that refugees are neither more likely to speak German at a higher proficiency level, nor are they more likely to improve their language proficiency at a faster rate, than economic immigrants.

The rest of this paper is organized as follows. Chapter 2 provides a brief historical overview of immigration to Germany. In chapter 3, relevant findings from academic literature are presented. Our dataset is presented in chapter 4, while chapter 5 provides the methodology of our analysis. We present the empirical results in chapter 6, after which we proceed with a discussion of the results in chapter 7 and test the robustness of our models in chapter 8. Finally, we conclude in chapter 9.

2 Background

The following chapter summarises the history of migration into Germany. Additionally, it provides a brief overview of the historical development of the most important legal frameworks concerning migrants and refugees in particular. Hereby, we aim to not only describe current legislations, but also those of former German states since the GSOEP captures individuals who entered Germany as early as 1949. Certainly, early immigrants faced a different situation than those immigrating to Germany in more recent times.

2.1 The history of Germany as an immigration country

The modern history of migration into Germany was, in the beginning, mostly driven by the consequences of the First and Second World War. Before the First World War, a boom in the German industry and a lack of qualified workers led to the employment of 1.2 million foreigners, predominantly from Poland (Oltmer, 2005). During the Second World War, more than ten million foreign workers were forced to work for the German war industry, while after 1945, approximately 12 million expelled Germans from Eastern and Central Europe and 10 million foreign-displaced persons had to be resettled or repatriated to their home countries. By 1950, most of these people had returned home or emigrated overseas.

In 1961, during the construction of the Berlin wall that separated the former Democratic

Republic of Germany (GDR) and the Federal Republic of Germany (FDR), approximately 2.7 million people immigrated from East to West Germany. This led to a shortage of qualified workers in the GDR, which then recruited approximately 500 thousand "guest" workers from countries such as Vietnam, Poland, and Mozambique in order to compensate for the loss of labour (Butterwegge, 2005a). Millions of such guest workers were also recruited to absorb the loss of large parts of the work force during the war. In these years, the German government signed contracts with Mediterranean and North African countries, starting with Italy (1955), Spain and Greece (1960), Turkey (1961), Morocco (1963), Portugal (1965), Tunisia (1965) and finally, with Yugoslavia (1968). The programme, which was specifically designed to attract workers with skills and qualifications the German labour market was short of, expired in 1973. Due to this recruitment strategy, the share of foreigners living in Germany increased from 1.2 percent in 1960 to 4.9 percent in 1970 (Butterwegge, 2005b).

Many of these "guest" workers ended up settling in Germany, against the government's expectation that they would return to their countries of origin as soon as their work was no longer needed. Additionally, most workers were followed by their families after 1973. Today, people from the above-mentioned countries still represent the largest groups among the people with a migration background in Germany.

Another large group among the immigrants in Germany are Ethnic Germans, also known as resettlers. Resettlers are Germans who moved back to the FDR or GDR after having lived in former German territories after 1945. They are allowed to become citizens of Germany, and are entitled to different forms of integration assistance (BAMF - Bundesamt für Migration und Flüchtlinge, 2017). In total, approximately 3.3 million Ethnic Germans, mostly from Poland, the Soviet Union, and Romania have relocated back to Germany since 1950 (Hönekopp, 1997). After the fall of the Berlin Wall in 1989, the number of immigrants spiked again as people from the former Soviet Union migrated to Germany.

Figure 1 above gives a graphical overview of the immigration waves in Germany since 1953. From the figure, we observe three main peaks, where the first two correspond to the years Germany had an active recruitment strategy of "guest" workers (around the 1970s), and the year the Berlin wall fell (around the 1990s). More recently, we see a surge in immigrants after 2008, the year plagued by the financial crisis. This immigration wave consisted mostly of within-EU immigration, following the lifting of immigration restrictions in May 2011 from countries that entered the EU in 2004 (Beyer, 2016).

We replicate the same figure for our data sample, which is reported in Figure 2. We streamline economic immigrants in our study to include only individuals immigrating to Germany from abroad, who are neither Ethnic Germans nor of refugee status. The term "economic migrant" might be misleading because it subliminally indicates that the
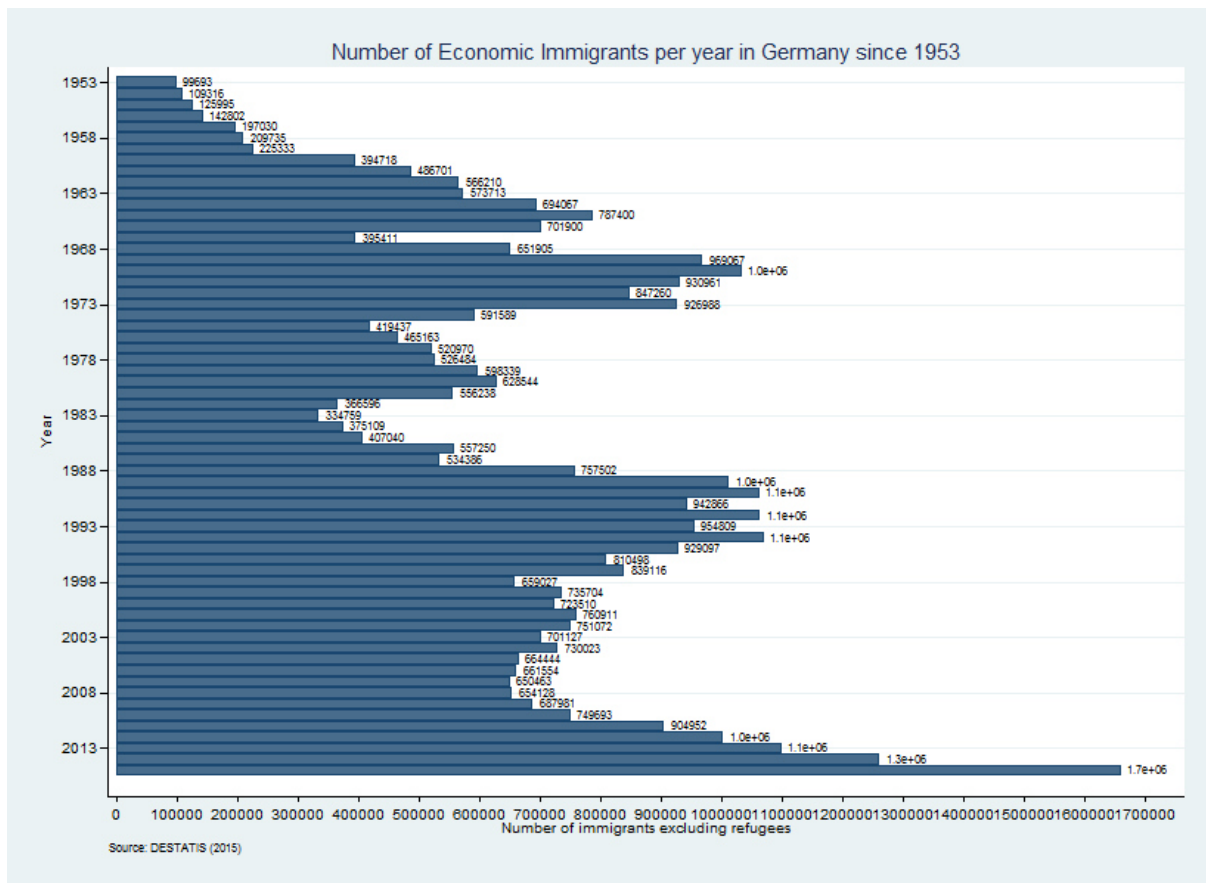
Figure 1: Migration Waves to Germany since 1953

only motivation for migration is to find work in Germany. However, immigrants can certainly have very different migration motivations. Hence, in this paper, we use the term "economic migrant" in a broader sense, namely for immigrants who work in Germany. Given the data we have, this is the closest approximation we have to an economic migrant.

We exclude Ethnic Germans for the following reasons. Firstly, Ethnic Germans might have an advantage with regards to mastering the German language, and do in fact have an advantage when it comes to obtaining a German citizenship, as mentioned above. In addition, they might have an already-existing network in Germany that could help them integrate back into the society more successfully than other types of immigrants. Lastly, they do not satisfy our definition of economic immigrants. Ethnic Germans returning to Germany are presumably more likely to stay long-term, and this in itself does not fulfil the relatively short implicit time horizon [2] we impose on our definition of an economic immigrant. By including them in our analysis, any earnings discrepancy found between economic immigrants and refugees might be upwards biased due to the sole presence of this immigrant group.

Figure 2 paints a less clear picture regarding the main economic immigration waves in our sample. However, we observe that there are generally two predominant waves, cor-

---

[2]Compared to refugees
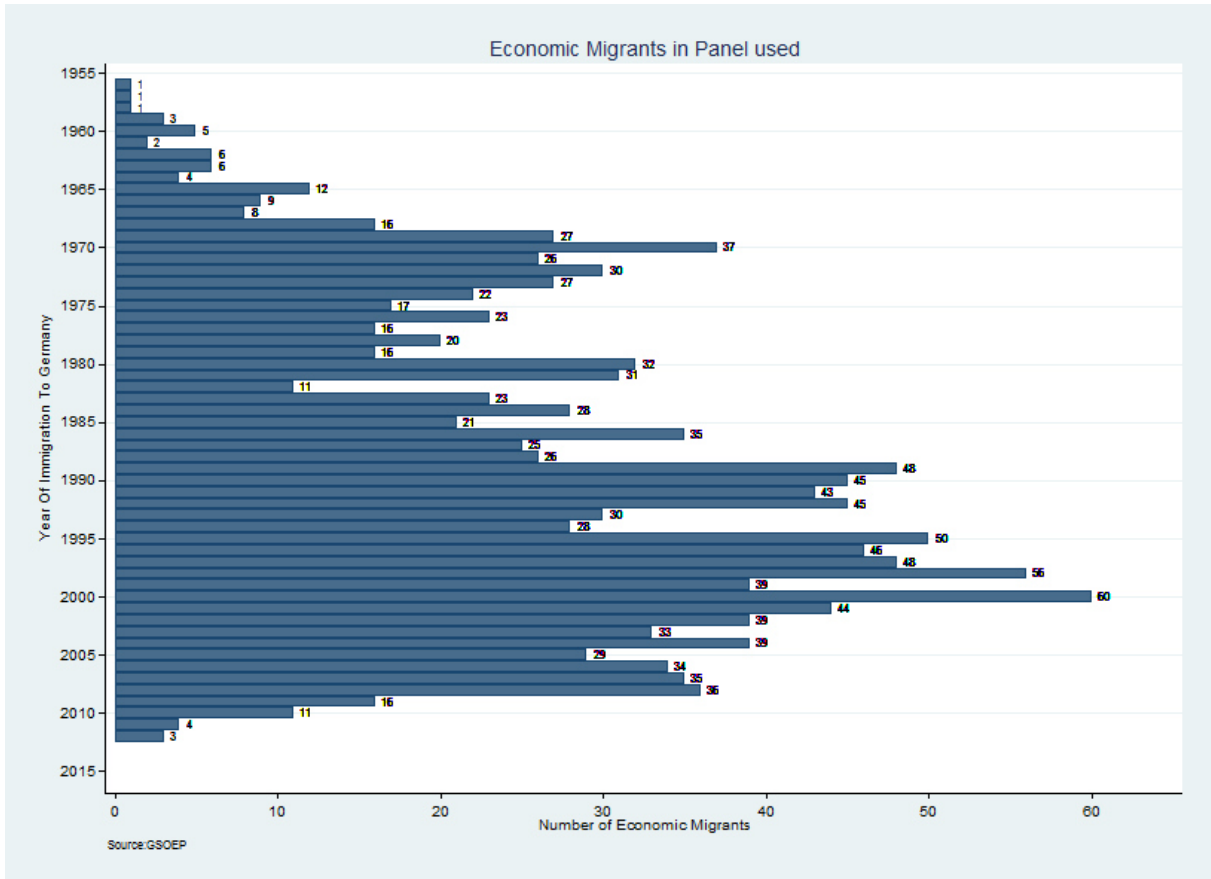
Figure 2: Immigration Distribution of Economic Immigrants to Germany in the sample used since 1953

responding to the first two waves highlighted from Figure 1. For completeness, Table 1 highlights the top 10 countries of origin the economic immigrants in our sample originate from.

2.2 Asylum in Germany

In Germany, the Asylum Act differentiates between four different forms of protection; Entitlement to asylum (Basic Law, 2017), refugee protection (Asylum Act, 2017a), subsidiary protection (Asylum Act, 2017b) , and ban on deportation (Residence Act, 2017). These various forms of protection refer to different rights, duties and entitlements, which are summarized in Table A.1 in the Appendix A.1. It is important to note that the GSOEP does not make this distinction between the different forms of protection.

The right to asylum was for the first time mentioned in Article 14 of the 1948 Universal Declaration of Human Rights from the United Nations (Assembly, UN General, 1948). Building up on that, the right to asylum was also incorporated in the German constitution in 1949. Article 16a (1) states that "politically persecuted persons have the right to asylum". The ulterior motive behind the article was twofold. On one hand, the German government wanted to distance itself from its national-socialistic heritage and show

Table 1: Top 10 Countries of Origin, Economic Migrants

Country Of Origin	Frequency	Percent	Cum.
1. Turkey	282	19.53	19.53
2. Poland	99	6.86	26.39
3. Italia	95	6.58	32.97
4. Austria	65	4.5	37.47
5. Ex-Yugoslavia	60	4.16	41.63
6. Russia	51	3.53	45.16
7. France	45	3.12	48.28
8. Greece	41	2.84	51.12
9. America	39	2.7	53.82
10. Great Britain	35	2.42	56.24

Source: GSOEP

acceptance for the Declaration of Human Rights. On the other hand, the German authorities thought that most of the incoming refugees would come from the Soviet zone of occupation, and thus wanted to create the necessary legal framework to receive these people (Gesemann and Roth, 2009).

Later on in 1951, the Convention Relating to the Status of Refugees (United Nations Conference of Plenipotentiaries on the Status of Refugees and Stateless Persons, 1951 Geneva, 1951) defined the minimum legal standards in the treatment of refugees as a reaction to the recent experiences from the national-socialism heritage, the Second World War, and it's refugee waves. In detail, the 1951 Refugee Convention defines who a refugee is, which rights refugees have after they have been offered asylum, and what the responsibilities of the countries granting asylum are. The convention from 1951 still forms the basis for international refugee rights today (Weinzierl, 2009).

After a period of low migration, the annual number of immigrants to Germany skyrocketed in the 1980s and 1990s, from some ten thousand to a couple of hundred thousand migrants. Reasons for this influx were the large refugee waves from the former Yugoslavia and the just-collapsed former Soviet Union, in addition to those who moved from East to West Germany. In 1992, the total number of asylum seekers in Germany was approximately 440 thousand, almost twice as many as in the year before (OECD, 2017).

This rapid increase in asylum seekers was, in many respects, challenging for the responsible municipalities, which struggled to provide accommodation and care. This led to resentment towards the asylum seekers by the local population, manifesting in the form of several racist attacks against them. To gain control over the situation, the government changed Article 16 by limiting the rights of refugees drastically (von Altenbockum, 2017). The new regulation prevented refugees who either entered through a *safe country*

[3] (Sicherer Drittstaat) or came from a *safe country of origin* [4] (Sichere Herkunftsländer), from applying for asylum [5] in Germany (Weinzierl, 2009). Consequently, the number of asylum seekers dropped significantly while the number of deportations increased. Figure 10 in Appendix A.2 provides a list of countries that are considered *safe countries of origin*, for the respective years.

In recent times, Europe has been characterized by the "refugee crisis"; a spike in the number of asylum seekers primarily from Syria, Afghanistan, Iran, Iraq, Albania, Eritrea, and Nigeria. Of all European countries, Germany accommodated the most number of refugees. Since 2008, more than five million asylum seekers entered the EU, with more than 1.7 million immigrating to Germany alone (Eurostat, 2017).

Apart from increasing engagement in border controls, the EU implemented various measures [6] to gain control over the situation and slow down the refugee influx. This led to a decrease in the number of refugees arriving in Germany. The result of Germany's migration history is not only an advanced and complex set of regulations and laws, but also an extensive system, which facilitates how refugees are accommodated and integrated.

Figure 3 gives a graphical overview of the refugee waves in Germany since 1953. Two sharp spikes of refugee migration into Germany stand out. The first occurred during the early 1990s, largely due to the influx of refugees from the former Yugoslavia and the just-collapsed former Soviet Union. The second spike, and also the largest influx of refugees encountered to date, occurred in 2016 for reasons pertaining to the migration wave most recently described as the "refugee crisis".

We replicate the same figure for our data, which we report in Figure 4. This table shows that the predominant wave in which refugees in our sample entered Germany corresponds to the first spike just explained from Figure 3. Therefore, the refugees in our sample are mainly originating from Eastern Europe. Again for completeness, Table 2 highlights the

---

[3]EU member states or neighbouring countries such as Poland, Austria and Czech Republic which had not been in the EU yet

[4]The law defines countries that have a democratic system and in which persecution is not to be feared, as safe countries of origin

[5]Exceptions exist for asylum seekers who can prove that they would face prosecution after returning to their home country

[6]The most important and prominent one is the EU-Turkey Statement (Council of the European Union, 2016) which is intended to end the irregular crossing of migrants from Turkey to Greece and break the business model of smugglers exploiting migrants and the situation faced by refugees. The statement includes that those who migrated irregularly to Greece will be sent back to Turkey, and in place, another Syrian will be resettled to the EU. Additionally, the EU Commission agreed on the resettlement of 160,000 refugees from Greece and Italy to other member states (European Commission, 2016a). However, only a few thousand refugees have been resettled yet due to disagreements about a compulsory quota system among the member states (European Commission, 2017) The EU also signed a contract with Libya to fight the reasons for flight in North Africa (European Commission, 2016b). The fourth measure was the implementation of hold-up camps in countries along the "Balkan-Route", the most used route for refugees from Greece towards central Europe.
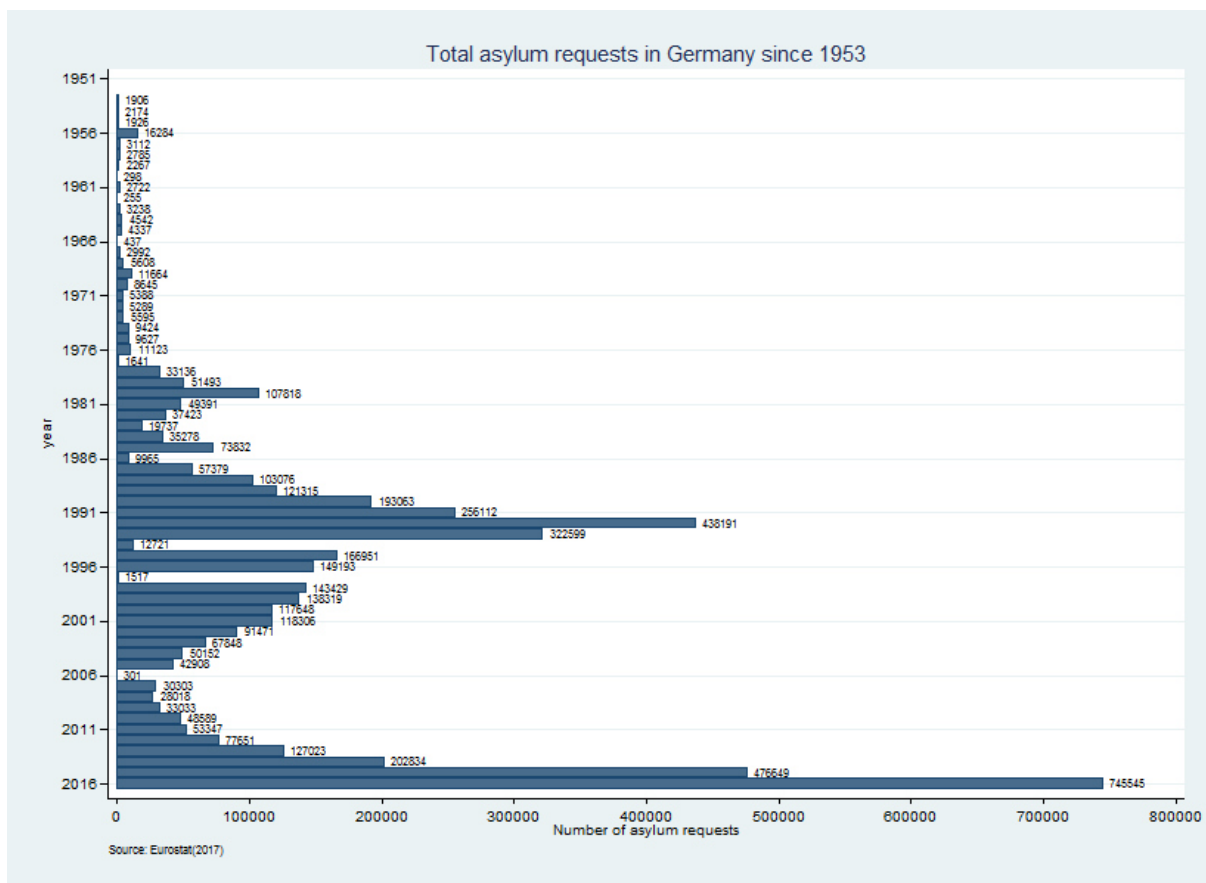
Figure 3: Immigration Distribution of Refugees to Germany since 1953

list of the top 10 countries of origin the refugees in our sample originate from.

3 Literature Review

This section provides an overview of past research related to our analysis. In particular, we focus on research pertaining to the earnings gap experienced by various immigrant groups, the returns to language on earnings, and the determinants of language acquisition among immigrants.

3.1 The earnings gap between immigrant groups

In recent years, more and more academic literature has focused on the earnings gap between economic immigrants and refugees (also known as the "refugee gap" [7]). Interestingly, there is no consensus regarding the refugee gap, as research observes contradicting results. Given the vast number of studies [8] available that analyse this earnings gap, we

---

[7] In general, this could also refer to the differences in earnings between refugees and natives

[8] For an extensive list and summary of the research conducted on the refugee gap see Table 1 of Ott (2013)
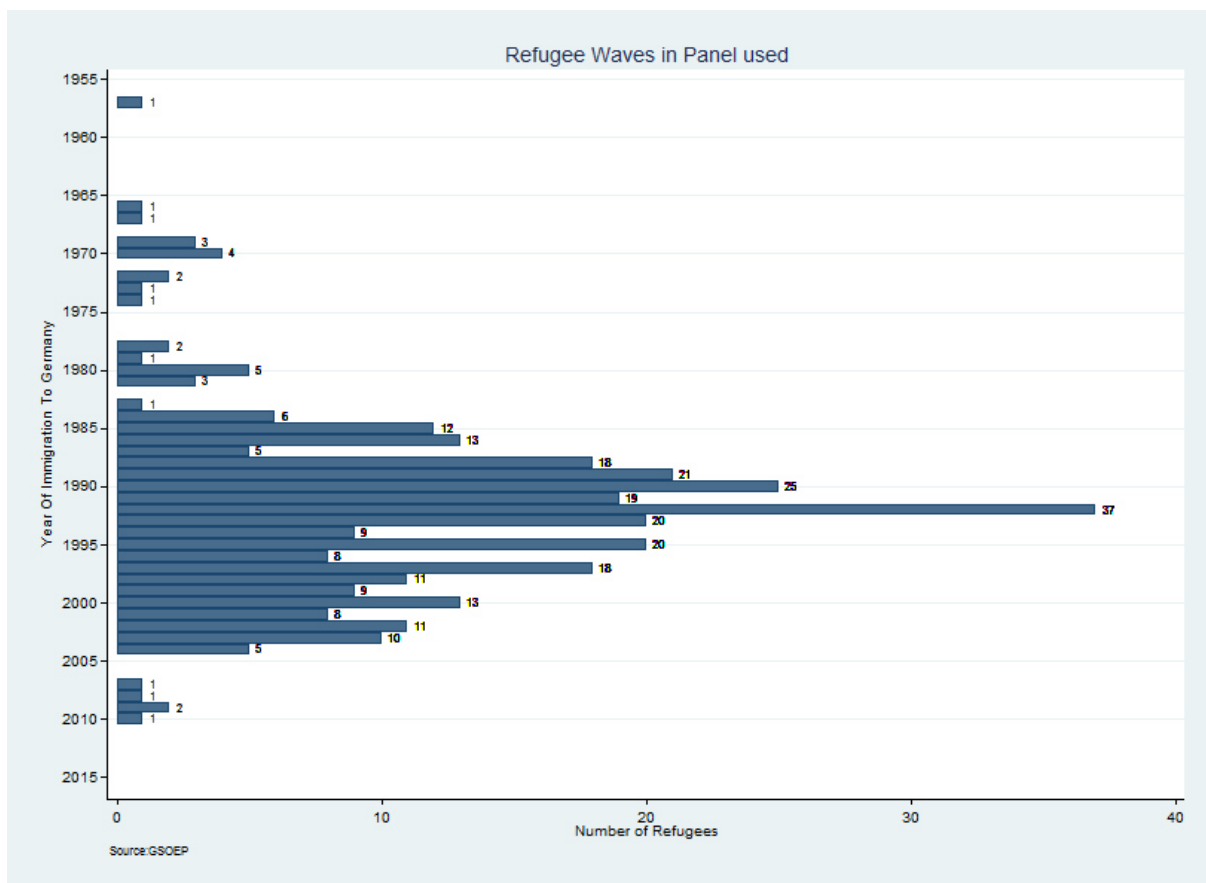
Figure 4: Distribution of Refugees in Sample used

focus our attention on two studies that illustrate such a conflict in findings. To the best of our knowledge, a study on the refugee gap has not been conducted for immigrants in Germany.

The first study is one by Cortes (2004), who simulates a panel from two census datasets on the years 1980 and 1990 to study the earnings gap between refugees and economic immigrants in the U.S.. She finds that, at the time of arrival, refugees have on average lower annual earnings, receive lower wages, and speak English at a lower proficiency level than economic immigrants. She observes, however, that ten years after arrival refugees report 20 percent higher annual earnings than economic migrants due to working 4 percent more hours. Furthermore, she finds that refugees increase their level of English proficiency by 11 percent more than economic migrants.

Duleep et al. (2001) analyse also census data from the U.S. in order to compare Indochinese (Vietnamese, Cambodian, and Laotian) refugees to economic migrants from Thailand, India, Korea, China, and the Philippines. They find that refugees have lower average earnings than the group of economic migrants. It is interesting to note that Cortes (2004) also paid mind to the Asian refugees in her sample, as it has been argued that Asian immigrants are more successful in the U.S. compared to other immigrant groups. She finds that the Asian contribution to the finding that refugees' earnings eventually

9

Table 2: Top 10 Countries of Origin, Refugees

Country Of Origin	Frequency	Percent	Cum.
1. Kosovo-Albania	35	10.23	10.23
2. Turkey	29	8.48	18.71
3. Ex-Yugoslavia	27	7.89	26.60
4. Bosnia-Herzegovina	22	6.43	33.03
5. Iran	18	5.26	38.29
6. Iraq	18	5.26	43.55
7. Albania	10	2.92	49.39
8. Portugal	9	2.63	52.02
9. Lebanon	9	2.63	54.65
10. Poland	8	2.34	56.99

Source: GSOEP

surpasses those of economic migrants is relatively small. These results show that the existence of the refugee gap is very country- and data-dependent.

There is also plenty of research existing on the gap between immigrants as a whole, and the native population. Again, there are conflicts in the findings. While many studies find a declining native-immigrant earnings gap, other studies observe the opposite. Namely, that the gap remains unclosed regardless of the duration the immigrants stay in the respective destination country.

A study on white native and foreign men [9] by Chiswick (1978) finds that although immigrants to the U.S. initially earn less than natives, their earnings increase more rapidly with additional years worked in the country. After 10 to 15 years, their earnings equalize and sometimes even exceed those of locals. Chiswick identifies the acquisition of knowledge of the destination language, customs and nature of the labour market, as well as post-school training as drivers of the growth in earnings. Anderson (2015) confirms these findings for the U.S. in a recent study. However, she identifies that the earnings gap does not close fully for some immigrant groups.

Focusing on Europe, Adsera and Chiswick (2007) use panel data from 15 European countries and observe that the earnings of immigrants and natives converge, and that the earnings gap closes fully after approximately 18 years.

Among others (Constant and Massey, 2005; Pischke, 1992; Dustmann, 1993; Schmidt, 1997; Bauer et al., 2005) Beyer (2016) explores the immigrant-native earnings gap in Germany. By using data from the GSOEP, he finds that recent immigrant workers earn on average 20 percent less than German natives. However, the gap declines over time and is smaller for immigrants with a good command of the German language, among other

---

[9]Chiswick (1982) observes a similar catching-up process by looking at male immigrants to the U.S. of all races and ethnicities from all countries of origin.

things. Furthermore, he identifies that immigrants face tough challenges in the German labour market, particularly due to skills downgrading and persistent wage gaps. This is further supported by Dustmann et al. (2016).

3.2 The returns on earnings

There exists abundant literature on the returns of immigrants' investments into country-specific human capital on earnings. In general, research shows that proficiency in the host-country's language has a positive effect on the immigrant's labour market success.

Using panel data for Germany, Dustmann (1994) finds that language proficiency is an important determinant of immigrants' earnings. In detail, he identifies that writing fluency for male immigrants is a more important determinant of earnings than speaking proficiency. However, speaking proficiency alone has a significant effect on earnings. For female workers, "speaking fluency alone is not sufficient for having an earnings advantage". Chiswick and Miller (1995) find further evidence that language proficiency is particularly important to explain the differences in earnings between natives and immigrants. They conclude that among immigrants, earnings are greater for those more proficient in the destination language. We will explore this study in greater detail in section 3.3 of the literature review.

Borjas (1995) uses age at arrival as an additional explanatory variable in the earnings regressions, and finds that an increase in the entry age from 20 to 30 years decreases earnings of migrants by 5 percent. He attributes this difference to the greater incentives younger migrants have in investing in country-specific human capital. Similarly, Bleakley and Chin (2004) use age at arrival as an instrument for language proficiency, as the critical-period hypothesis states that the younger [10] a migrant arrives in the host country, the more quickly he or she picks up a new language and hence, have a better language proficiency than those who arrive at an older age. They find that improving English-speaking ability by one unit increases wages by 33.35 percent.

Similar to Dustmann (1994), Dustmann and Van Soest (2002) find that the language proficiency of immigrants in Germany is a very important predictor of earnings, with a return of 14.12 percent on earnings based on their most optimistic model. Relying as well upon the GSOEP panel dataset for the years between 1984 and 1993, they highlight two main sources of bias faced in their analysis; measurement errors and unobserved heterogeneity. In order to tackle the latter issue to some extent, they include partner and household variables in their OLS models. In addition, they argue that by applying the Instrumental Variable (IV) estimation approach by instrumenting language proficiency

---

[10]Specifically migrants arriving under the age of 12

with father's education, the bias from both sources can be reduced. The validity of their instrument is highly debated, and therefore their results should be interpreted with caution. Regardless of their model specification, they find a positive relationship between language proficiency and earnings of immigrants. Further evidence for this relationship is provided by Dustmann and Fabbri (2003), who focus their analysis on the United Kingdom and find that in simple regressions, language proficiency is associated with higher employment probabilities and higher earnings.

3.3 The differences between language acquisitions

The determinants of language acquisition among immigrants are complex. We now explore several studies that identify factors that cause differences in investments into country-specific human capital among immigrant groups.

Looking deeper into the paper by Chiswick and Miller (1995), which we have already briefly introduced, they claim that the language proficiency of migrants can be expressed as a rising function of three input variables: Economic incentives, exposure to the language, and efficiency.

Economic incentives due to language fluency are in general considered to be important drivers of language acquisition. Incentives arise, for instance, from an increment in the market wage rate, a higher rate of employment, or a decrease in the cost of consumption and job search. These incentives differ with the expected duration in the destination country (Chiswick and Miller, 1995). For instance, a migrant who intends to stay for a longer duration has more incentives to invest in learning the local language than someone who does not intend to stay as long in the country.

Exposure refers to the extent an immigrant is exposed to the language of destination prior and after migration, as well as to the intensity of exposure. Exposure prior to migration will be larger if the language of destination is similar to the mother tongue of the migrant, have the same roots, or share the same linguistic structure. The exposure to the language of destination is strongly dependent on the area the migrant is living post-migration. In general, the more people there are who speak the language of destination in the migrant's neighbourhood, the more fluent the migrant would be. This is supported by Case and Katz (1991), who study the effect of living in an enclave or a neighbourhood on the behaviour of inner city youth. They find that the behaviour of the neighbourhood has an important influence on the individual. For instance, they find that individuals are more likely to be involved in crime if they are surrounded by a lot of crime. According to Chiswick and Miller (1998), Chiswick and Miller (1995) argument follows a similar line of thought.

The exposure to the destination language in one's own home is particularly determined by the spouse and children, though the effects are somewhat unclear. If the spouse has the same linguistic background as the migrant in question, the exposure to language of the destination country is decreased and fluency would consequently suffer. In contrast, the effect on language proficiency is more difficult to determine if the spouse speaks the local language, as the spouse can take on two different roles: that of a teacher or a translator. In the first case, the effect on language fluency would be positive, while in the latter case the expected effect would be negative.

Similarly, exposure is influenced by the presence of school-going children in the household. Since children are better equipped to learn a new language than adults and are more exposed to the local language through compulsory schooling, they can support their parents in two ways. Either the parents acquire fluency from their children, which would have a positive effect on their language proficiency, or, the children take on the role of interpreters and help their parents communicate in the language of destination. The latter would again limit the exposure of the parents to the local language and hence, have a negative effect on their level of fluency (Chiswick and Miller, 1995).

Efficiency refers to the ability to achieve fluency with a given amount of destination language exposure. As mentioned, this ability is particularly high among those who migrate at a very young age, but seems to diminish with an increase in age at arrival (Bleakley and Chin, 2004). Generally, a high level of education is also associated with the ability to learn a language more efficiently.

The factors mentioned above explain partly the differences we expect in the level of language acquisition between refugees and other migrants. The economic incentives for acquiring the language spoken in the country of destination are higher for refugees, as their possibilities to return to their country of origin are lower than those of economic migrants. However, the general economic immigrant is expected to be favourably selected since he can choose the country of destination based on country characteristics that enable him to integrate successfully. For instance, he could choose a country whose language is closer to his mother tongue and increase his exposure to that language prior to migration. [11] In contrast, we do not expect refugees to have invested into country of destination specific human capital prior to migration, as they leave their home countries unwillingly and sometimes very suddenly. Both effects combined suggest a lower *initial* endowment of country-specific human capital for refugees, relative to economic immigrants. On the other hand, we could expect refugees to investment more into country-specific human capital *post-migration*, and hence have a steeper improvement in the language proficiency over time.

---

[11]We will elaborate more on the favourable selection of migrants in the Discussion chapter.

As further support, Chiswick and Miller (2007) find that Mexican immigrants are less successful in terms of language acquisition than migrants of other origins in the U.S., due to their higher probability of returning home to Mexico. Dustmann (1999), who studies the determinants of speaking and writing German proficiency by way of an ordered probit analysis, shows that an immigrant's language fluency is significantly and negatively affected by a greater return propensity. Finally, we turn our attention again to Cortes (2004), who attributes the higher investment by refugees into the acquisition of language proficiency than economic immigrants to the longer time horizons refugees have in Germany. These findings reaffirm our hypothesis that, given their supposed longer time horizons in the host country than economic immigrants, refugees would likely invest more in Germany-specific human capital.

Our first research question builds largely upon Cortes (2004). However, in contrast to her, we use the GSOEP panel data, which has two advantages over Cortes' census data. Firstly, panel data allow us to potentially better control for individual specific effects that would have otherwise led to biased estimates. Secondly, our dataset allows us to confidently identify refugees from economic immigrants, as the immigrants declare their migration status upon entry into Germany. Cortes' dataset did not contain such information, which led her to define certain countries as only refugee- or economic immigrant-sending countries. The implication of her approach is that one country is not able to send both a refugee and an economic migrant to the U.S.. This might be a strong assumption to make in light of our dataset, whereby a refugee and an economic immigrant could originate from the same country. Our second research question follows closely the methodology by Dustmann and Van Soest (2002), where we attempt to replicate their main models.

4 Data

This study analyses panel data from the German Socioeconomic Panel (GSOEP) [12], which began in 1984 and is collected by the Deutsches Institut für Wirtschaftsforschung (DIW) in Berlin. The GSOEP provides information on members of households situated in Germany, consisting of both locals and foreigners. In total, the GSOEP entails information on more than 11 thousand households with over 30 thousand individuals. Members of the panel are re-interviewed annually, with the survey taking on a different theme each year. Some of these themes include household composition, occupational biographies, earnings, health, and satisfaction indicators.

Those interviewed individually include persons above the age of 16 years. Every household is represented by the head of the household, who answers household-level questions.

---

[12]SOEP (2017); Wagner et al. (2007)

Individuals exit the panel by moving abroad, or by death. On the other hand, individuals enter the panel by birth into a household captured in the panel, by reaching the age of 16 years, by the formation of new households, or by moving into Germany from abroad.

We follow 343 distinct refugees, and 1452 distinct economic immigrants in our analysis, which spans from year 1994,[13] up until 2014. We restrict our analysis to only individuals who report themselves as having a positive annual labour earning, positive labour hours worked, and a positive wage. Therefore, we follow the approach by Dustmann and Van Soest (2002) and focus only on the intensive margin of the sample. [14] We do this because we want to study the effect of key variables on earnings, and therefore exclude those with no income. As a consequence of this sample restriction, our sample contains only individuals who are between the ages of 17 and 64, at the time of the interview.

We investigate both working male and female immigrants. Most studies in this field consider only men, as the mechanisms governing women's language acquisition may not be as straightforward as those for men. As explained by Dustmann and Van Soest (2002), some extra considerations made by women on their language acquisition include fertility decisions. In addition, women are more likely to enter the panel as dependants, and so their decision to invest in human capital, and hence to work might be influenced by their husbands' decisions. Finally, a woman's non-random participation might be problematic to obtain unbiased estimates. However, we still proceed to consider women in our analysis, but bear in mind that we base the analysis on a very specific female sample, namely working female immigrants in Germany, and that we are not able to generalise these results to all female immigrants.

4.1 Comparability of refugees with economic immigrants

To give a better understanding of the comparability between the two immigrant groups, we provide the descriptive statistics and histograms for key variables, across four years. (See appendix A.3 and A.4, respectively). These particular years were chosen in order to provide a broad visual overview of our dataset. We are especially concerned with ensuring that the groups are comparable prior to their arrival to Germany, which would then allow us to make post-entry comparisons between the groups. Available variables for this comparison include own and father's education levels obtained prior to arriving in Germany, and age at arrival. We show the histograms for these three specific variables in Figures 5, 6 and 7 .

---

[13]The first refugee recording in our dataset appeared in 1994.

[14]We are aware that this could be a potential source of bias, as the decision to work may be systematically correlated with unobservables that affect the wage offer (Wooldridge, 2013). Hence our results cannot be generalised.

According to Cortes (2004), it is reasonable to expect refugees to be closer to having an even distribution of age at arrival due to the spontaneous nature of their migration. However, one could argue that younger, more able-bodied people would find it easier to escape their plight and migrate to Germany. As for economic immigrants, we expect them to arrive mostly at the working ages of between 18 and 35 years (Cortes, 2004). For the four years presented, we see that the age at arrival of economic migrants have a bell-curved distribution, while the mode is concentrated around the working age. As expected, the same result is generally seen for refugees.

For the education variable, we expect economic immigrants to have, on average, more years of education than refugees, as the group is favourably selected and, in the case of third country immigrants, the German migration system favours highly educated immigrants. We will explore these ideas in greater detail in the discussion section of the paper. From the histograms, we can see that there are generally a lot of overlaps between the two groups. Yet, we observe more economic immigrants in the right tail of the figure, while there are more refugees on the left tail. However, it is worth noticing that we find both groups present in the extremes.

The last variable we focus on is father's education. From the respective histograms, we observe a good degree of overlap between the two immigrant groups, for the years after 1995. The key takeaway from these histograms is that even though the two immigrant groups are not identical in every aspect, they display similarities to a sufficient degree.
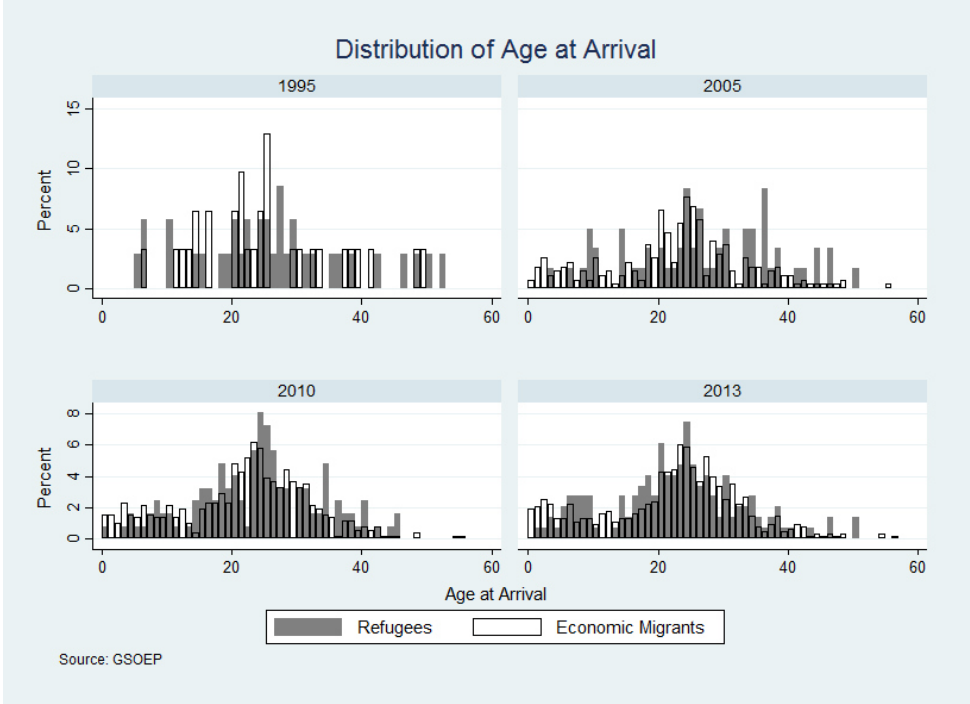
Figure 5: Histogram Age at arrival
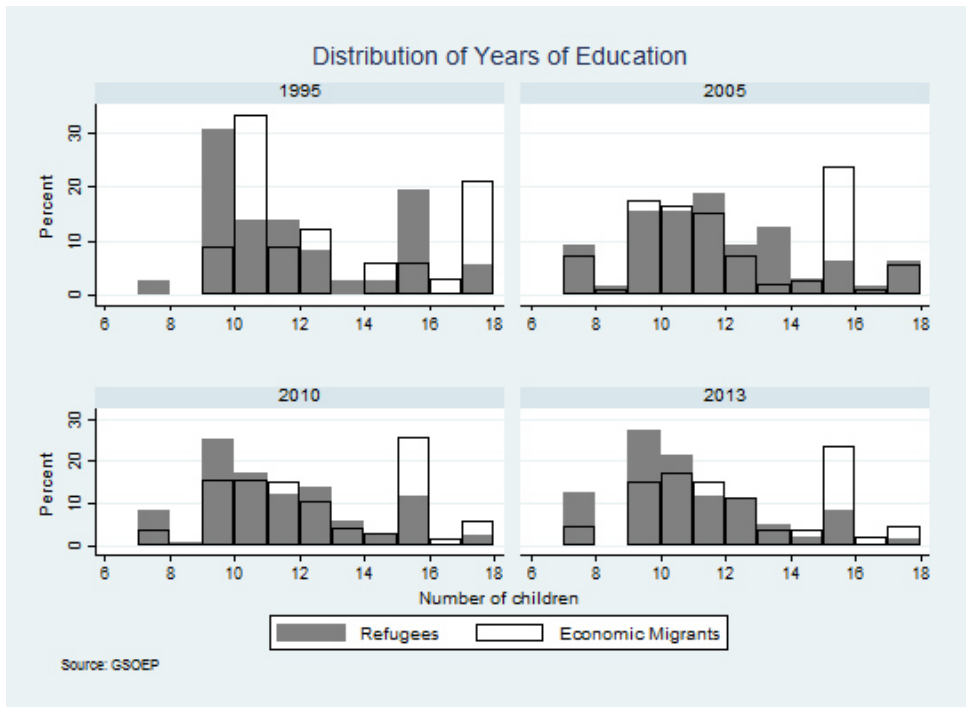
Figure 6: Histogram Years of Education
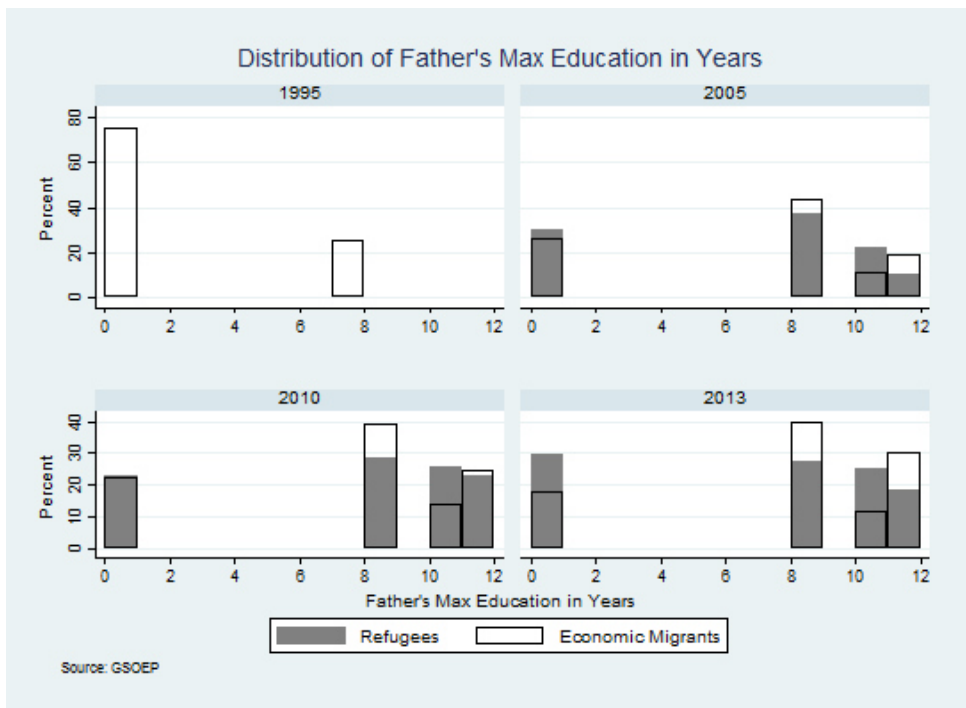
Figure 7: Histogram Years of Father's Education

5 Methodology

5.1 To determine the earnings gap between refugees and economic immigrants

In order to answer the first research question, we utilise the OLS estimation method by first regressing the natural logarithms of immigrants' annual earnings on the dummy for refugee status, with a standard set of controls [15] (Model 1 [16] in Tables 9, 10 and 11 in the Empirical results chapter). We build upon this model by adding controls for language proficiency, and subsequently partner and household characteristics [17] (Models 2 and 3 respectively, in Tables 9, 10 and 11). Ideally, we would have used the Fixed Effects method to obtain unbiased estimates and utilize the panel structure of our data. Unfortunately, the method does not permit us to identify the effect of time-invariant variables such as the effect of refugee status on annual earnings. Thus, this dummy variable will drop out of the regression following the Fixed Effects transformation. Additionally, limited within-group variation would lead to weak estimates.

Our model is thus specified as follows:

$$lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it} \tag{1}$$

Y_{it} is the dependent variable, which in our case is annual earnings. The vector of explanatory variables X_{it} includes the refugee dummy, potential labour market experience and its square, years of education, age and its square, marital status, years since migration, and a male-gender dummy. Every model also controls for the year of survey, the countries of origin, and the states in Germany where the immigrants reside. [18] α_i represents the unobserved heterogeneity, while v_{it} is the idiosyncratic error term.

To obtain fully robust standard errors and test statistics considering serial correlation and heteroskedasticity, we cluster the observations with respect to individuals based on their unique identifier (Wooldridge, 2013). Appendix A.5 further provides an overview and description of the variables we include in our regression.

As an expansion to this first research question, we also attempt to determine whether the differences in earnings are due to differences in annual hours worked or differences in hourly wages. We replicate all previous models, but interchange the dependent variables

---

[15] See Appendix A.5 for list of standard controls

[16] Model 1 corresponds to Column 1 in all the regression tables, respectively.

[17] A further explanation on the reasons for including partner variables is provided in the subsection *Identification*

[18] This is the most detailed variable that the GSOEP provides regarding the area the individuals are living in. Ideally, we would control for municipality and neighbourhood effects.

with the natural logs of immigrants' hourly wages and annual hours worked. (Tables 25 to 30 in Appendix A.6).

To get a better idea of the drivers of the differences in earnings between refugees and economic immigrants, we introduce refugee interaction terms to key variables from Models 1 to 3. The regression results are displayed in Models 4 to 6 in Tables 9, 10 and 11. Models 7 and 8 build upon Models 3 and 6 respectively, and present the Fixed Effects [19] regressions results. Thus, the new model specification includes interaction terms:

$$lnY_{it} = \beta_0 + \beta X_{it} + \delta[X_{it} * Refugee_i] + \alpha_i + v_{it} \quad (2)$$

5.2 Construction of the spoken German language dummy

Pertaining to our second research question, we focus on the German speaking proficiency of immigrants. In the GSOEP, there are two different variables available that specify an immigrant's level of spoken German proficiency. One of the two variables was used in the years 1984 to 1987 and every other year between 1989 and 2005. The other variable replaced the former in year 2007, and continued annually thereafter except for the years 2012 and 2014. Both measures report language proficiency along a five-point scale: 5 for "Very Poor", 4 for "Poor", 3 for "Intermediate", 2 for "Good", and 1 for "Very Good". We reverse the scale to make it more intuitive.

We proceed to merge the two language variables in order to form a single language variable for spoken German ability. To the best of our knowledge, the two original language variables are measured via the same self-reporting method. Our analysis would suffer if one measure significantly differs from the other. As an extreme example of this, imagine an individual who constantly reports his language proficiency as "Very Good". If there are systematic differences in the way the two language measures are reported, this individual might be classified as "Poor" in the following period. This unsystematic jump in language proficiency level would create noise in the regression results if it does not reflect the true fluctuations of this individual's language proficiency.

We conduct a chi-square test on the language variable for two subsamples, before and after the introduction of the new language variable, in order to ensure that the two measures are in fact not significantly different to each other. The chi-statistic test value was 0.44 (with 1239 degrees of freedom and a p-value of 0.5074). This implies that we could not reject the null hypothesis of equal coefficients across models. In order to ensure that both

---

[19]We will not discuss the results of the Fixed Effect model. The results are mainly to highlight the problems we face with this methodology, especially concerning the removal of key variables of interest.

samples are comparable, we only include individuals who were present in both subsamples. Thus, we conclude that the two measures are indeed similar.

The question on language proficiency was not asked annually. Therefore, years exist with missing information on German language abilities. Ideally, we would like to have language skills reported in every year of our dataset. In order to include as many observations as possible in our regressions, we assume that language ability does not change drastically over two consecutive years, and copy downwards the variable for the years with missing information, until the next year the original language variable reappears. For instance, an individual's language proficiency in the year 2000 (which is missing information on language proficiency) would therefore take the language proficiency level reported in the year 1999. This is a fairly acceptable assumption, as we will illustrate in the following tables.

Table 3: Cross tabulation of German speaking proficiency levels, Economic Migrants

t/t-1	Very Poor	Poor	Intermediate	Good	Very Good	Total
Very Poor	**1**	0	2	0	0	3
Poor	2	**33**	16	1	1	53
Intermediate	0	13	**81**	34	6	134
Good	0	0	15	**178**	56	249
Very Good	0	1	5	55	**294**	355
Total	3	47	119	268	357	794

Numbers refer to 2007-2011. Columns entries: previous year. Row entries: current year. 116 above the diagonal, 99 below diagonal

Table 4: Cross tabulation of German speaking proficiency levels, Refugees

t/t-1	Very Poor	Poor	Intermediate	Good	Very Good	Total
Very Poor	**0**	0	0	0	0	0
Poor	0	**3**	5	4	0	12
Intermediate	0	7	**33**	15	1	56
Good	0	2	14	**60**	7	83
Very Good	0	0	2	6	**29**	37
Total	0	12	54	85	37	188

Numbers refer to 2007-2011. Columns entries: previous year. Row entries: current year. 32 above the diagonal, 31 below diagonal

Tables 3 and 4 present the cross tabulation of both immigrant groups' speaking proficiency levels, for consecutive years 2007 to 2011.[20] Note that the cross tabulation includes only the years in which the language question was *originally* asked. It is therefore not affected by the downwards copying of the language variable from the year before. The column entries refer to the language proficiency in the previous year, t-1, while the row entries

---

[20]The smallest "t" year is thus 2008, and "t-1" is 2007

refer to language proficiency for the current year, t. The numbers in the table refer to the number of distinct observations that correspond to the immigrant's level of speaking ability in both time periods.

Both tables show that most observations are found along the diagonal. This suggests that the majority of immigrants have had the same language proficiency over two consecutive years. This pattern supports our assumption that it is acceptable to copy downwards the language proficiency level from the previous year, if the language variable were missing.

Entries found above the diagonal suggest an improvement in language skills from the previous year, while entries found below the diagonal imply that there has been a decline in language proficiency from the previous year. A priori, we would expect to find the majority of entries on, or above, the diagonal. On the other hand, we would expect to see few to none observations below the diagonal. However, we find a similar number of observations above and below the diagonal in both Tables 3 and 4. This implies that some individuals improved or regressed in their language proficiency.

Having brought this to attention, we accept the possibility that immigrants can deteriorate in their level of language proficiency over time. For instance, this could be due to the lack of practice over the years. It could even be explained by a simple difference in an immigrant's level of confidence towards his or her own language proficiency, which could undoubtedly vary from year to year. In extreme cases, the language proficiency gap could be larger than just by one level, which is represented by observations found *even further* from the diagonal, and not immediately surrounding the diagonal. These extreme observations are, however, few. Nevertheless, these tables might highlight a possible misclassification error that occurs over time, as long as the fluctuations in language proficiency are not reflecting the true variability in the language proficiency of the immigrant.

Besides the presence of measurement errors that are independent over time, our language variable is also faced with measurement errors that are time-persistent. For instance, there could be constant under- or over-reporting of one's language skills. These two types of measurement errors thus lead us to explore alternative methods to measure the impact of language on earnings. One way to minimize the implications of these potentially noisy fluctuations is to generate a dummy for good German ability, coded 1 if language proficiency is reported as "Good" or "Very Good", and 0 otherwise. The reference category would thus be working immigrants whose proficiency level is "Intermediate", "Poor", or "Very Poor". We do the same cross-tabulation analysis for this language dummy, and present the results in Tables 5 and 6.

As before, we find the majority of observations along the diagonal, with a relatively similar number of observations below and above it. One takeaway from these tables is that if the decrease and increase in language skills is not capturing the true language fluctuations

Table 5: Cross tabulation of German speaking proficiency levels, Refugees

t/t-1	Good German ability=0	Good German ability=1	Total
Good German ability=0	**148**	42	190
Good German ability=1	21	**583**	604
Total	169	625	794

Numbers refer to 2007-2011. Columns entries: previous year. Row
entries: current year.

Table 6: Cross tabulation of German speaking proficiency levels, Refugees

t/t-1	Good German ability=0	Good German ability=1	Total
Good German ability=0	**48**	20	68
Good German ability=1	18	**102**	120
Total	66	122	188

Numbers refer to 2007-2011. Columns entries: previous year. Row
entries: current year.

for each individual, then we are forced to avoid using Fixed Effects estimation method, which exploits this within-group variation. In addition, there are only a small number of observations representing fluctuations in language proficiency. This would create noisy estimates, and is further motivation for us to avoid the Fixed Effects approach.

In addition to the two measurement errors just discussed, the study on the returns to human capital investment on earnings faces the issue of endogeneity (Dustmann and Van Soest, 2002). Our case is no exception. One could argue for the presence of the unobserved variables, such as ability, that hides in the error term. These unobserved variables could concurrently affect the level of human capital investment and the level of earnings. By not controlling for them, our models would suffer from omitted variable bias, making our estimates upward biased. The next section will explain how we attempt to deal with these potential sources of biases.

5.3   Identification

To deal with the issue of unobserved heterogeneity, our first approach was to apply the Fixed Effects regression method. However, and as mentioned previously, the high degree of uncertainty regarding the variation of an immigrant's language proficiency leads us to explore other methods and move away from the Fixed Effects method (nevertheless, we report the Fixed Effects findings in the respective tables). Instead, we follow the approach by Dustmann and Van Soest (2002) that includes partner variables and household characteristics of the individual. The argument is such that if mating were assortative, then controlling for these variables would capture at least parts of the unobserved individual characteristics (Dustmann and Van Soest, 2002). Thus, we specifically include two

partner variables in our regression. Both are indicator variables that equal one when an immigrant has a partner who speaks German well and when the immigrant has a partner who is a German citizen, respectively. We are careful not to exclude individuals who have no partner, which would have reduced our sample size drastically. Thus, we work with partner dummy variables to include as big a sample as possible.[21] As per all previous models, we also enhance the model by including refugee interaction terms for interesting variables.

Another way we could potentially deal with endogeneity is to use the IV method, which was, as briefly introduced beforehand, done in the paper by Dustmann and Van Soest (2002). To qualify as an instrument, the variable must satisfy two assumptions. The first is that the instrument should be sufficiently correlated with the language variable. The second is that the instrument must be uncorrelated to any other determinants of an immigrant's annual earnings (also known as the exclusion restriction) (Wooldridge, 2013). We explore two potential instruments for language proficiency: father's education and age at arrival.

Father's education and the two assumptions: As argued for in the paper by Dustmann and Van Soest (2002), attitudes towards the valuation of language skills are likely developed at a young age within a family, and how much a child values languages is related to the parents' level of educational attainment. Children of highly educated parents are also more likely to be exposed to different languages and cultures at a younger age, and could therefore have already been exposed to the German language prior to moving to Germany. This satisfies the first assumption. As for the second assumption, Dustmann and Van Soest (2002) briefly introduced an argument that suggests that better-educated parents may have networks that could help the offspring at the start of their career. However, they counter this argument by arguing that this network link is most likely broken when the offspring migrates to another country. That said, past research has found a causal link between the father's education level and the son's education level (Ermisch and Pronzato, 2010). Therefore, we conclude that the exclusion restriction is not met.

Age of arrival and the two assumptions: Following closely the methodology by Bleakley and Chin (2004) as we have already introduced in the literature review, the critical-period hypothesis suggests that the younger the immigrants move to the host country, the better they grasp the host country's language. Hence, there is a well-supported correlation between language skills and age at arrival. However, for this instrument to satisfy the exclusion restriction, the decision to move has to be exogenous. This is to ensure that

---

[21]We introduce two dummy variables. The first indicates those individuals who have a partner with a good German proficiency, while the second indicates those with a German partner. By doing so, we implicitly create a baseline group that entails not only those whose partner is neither a German citizen nor fluent in German, but also single immigrants.

whatever unobserved factor affecting the decision to migrate is not also related to the income, or determinants of the income, of the individual. Bleakley and Chin (2004) restrict their sample to only immigrants who arrived at the age of 11 or younger. This way, they argue that the decision to migrate would have been made by the parents of these young immigrants, making it an exogenous decision. Note that if we were to do the same, we would have to exclude a large share of our sample. However, it is often impossible to determine whether the second assumption is satisfied as there are no tests we can conduct to prove it.

One key issue we face with using these instruments arises because age at arrival and father's education are time-constant variables. As a result, there are difficulties in establishing a strong correlation between the time-varying endogenous regressor language proficiency, and the time-constant instruments. To force the instruments to vary, we interact them with the *year of survey* [22] variable.

Table 7 provides the first-stage regression results [23] using both instruments just described respectively. Particularly, we look to the partial R-square value, which shows the correlation between the instrument and the language dummy (Shea, 1997), and the p-value, which tells us whether or not we can reject the null hypothesis of having weak instruments. As we can see, the partial R-square values in both cases are extremely small, suggesting that the first assumption for a good instrument is not satisfied for either of the two instruments. Secondly, the p-values reported in the tables prevent us from rejecting the null hypothesis that the instruments are weak at a 5 percent significance level. These results lead us to avoid using our chosen instruments. It is also generally advised to avoid the use of weak instruments, as they may lead to inconsistent (Chao and Swanson, 2005) and, in small samples, even biased estimates (Bound et al., 1995). For all these reasons mentioned above, our main models will rely solely upon OLS estimation methods.

6 Empirical results

Table 9 presents the results for regression equation (1) in columns 1 to 3. Model 1 includes only the standard controls, whereas we include the language dummy in Model 2, and variables for partner and household characteristics in Model 3. We rerun these models separately for each gender, as presented in Tables 10 and 11.

---

[22]The years in which the survey was taken

[23]We run this regression on equation (1)

Table 7: First stage regression summary statistics using age at arrival as an instrument for language proficiency. Observations are restricted only to those who immigrated at an age of 11 or younger.

Instrument	R-Squared	Adj. R-Squared	Part. R-Squared	Robust F(1,785)*	Prob > F
Father's Education	0.3764	0.3506	0.0039	2.73208	0.0988

Instrument	R-Squared	Adj. R-Squared	Part. R-squared	Robust F(1,244)**	Prob > F
Age at arrival	0.3053	0.2216	0.0003	0.117362	0.7322

*)F statistic adjusted for 786 clusters in individual identifier
**)F statistic adjusted for 245 clusters in individual identifier

6.1 Interpretation of the earnings gap

Regardless of the specification in Table 9, the coefficient on the refugee dummy is negative. While the coefficient in Model 1 is not significant, it becomes significant after adding a control for language proficiency, and continues to remain significant at a 5 percent level when variables for partner and household characteristics are added. Depending on the model, a refugee earns between 17.6 percent and 19.2 percent less than economic migrants. Similar results are observed for the male sample [24] in Table 10, although the coefficient is not statistically significant in Model 3. However, regardless of the specification shown in the female sample (Table 11), the coefficients are not significantly different from zero, though the signs are negative. These results are contrary to the findings by Cortes (2004), who finds that typical male and female refugees earn between 3 to 21 percent *more* than comparable male and female economic immigrants, ten years after arrival in the United States. Cortes' results are also supported by academic theory, as we have presented in the literature review section, which predicts a positive link between the expected duration in the country-of-destination and investments into country-of-destination-specific human capital, and thus implicitly also between expected duration and future earnings. The presented results from this study fail to provide evidence for this relationship. One potential reason Cortes (2004) finds a positive earnings gap in favour of U.S. refugees could be due to the presence of the welfare system in Germany. Refugees in Germany generally benefit greatly from state-provided welfare programmes, whereas the U.S. supports refugees to a much smaller extent (Alesina et al., 2001). Therefore, refugees in the U.S. have a much greater incentive to work hard in order to "make it" on their own. This could thus explain the difference in our results to those of Cortes'.

---

[24]A male refugee earns between 16.5 percent and 18 percent less than a male economic migrant.

Table 9: Y=Log Annual Earnings, Pooled Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Refugee	-0.145	-0.192*	-0.176*	0.458	0.311	0.459	0	0
	(0.0805)	(0.0863)	(0.0887)	(0.374)	(0.421)	(0.446)	(.)	(.)
Age	0.0555**	0.0520*	0.0672**	0.0571**	0.0523*	0.0681**	-0.0259	-0.0360
	(0.0204)	(0.0226)	(0.0229)	(0.0206)	(0.0227)	(0.0230)	(0.0449)	(0.0444)
Age^2	-0.000915***	-0.000860**	-0.00105***	-0.000917***	-0.000860**	-0.00106***	-0.000369	-0.000255
	(0.000251)	(0.000273)	(0.000279)	(0.000251)	(0.000273)	(0.000278)	(0.000413)	(0.000415)
Male	0.809***	0.808***	0.821***	0.807***	0.805***	0.819***	0	0
	(0.0552)	(0.0607)	(0.0603)	(0.0548)	(0.0603)	(0.0598)	(.)	(.)
Experience	0.0750***	0.0732***	0.0704***	0.0768***	0.0757***	0.0726***	0.175***	0.183***
	(0.00894)	(0.00950)	(0.00940)	(0.00949)	(0.00999)	(0.00995)	(0.0273)	(0.0288)
Experience^2	-0.000852***	-0.000782**	-0.000731**	-0.000886***	-0.000818***	-0.000760**	-0.00166***	-0.00185***
	(0.000228)	(0.000239)	(0.000239)	(0.000230)	(0.000241)	(0.000242)	(0.000364)	(0.000381)
YSM	0.0112***	0.00950**	0.0106***	0.0113***	0.0104**	0.0116***	0	0
	(0.00272)	(0.00317)	(0.00317)	(0.00284)	(0.00339)	(0.00339)	(.)	(.)
Married	-0.0220	-0.0148	0.0481	-0.00116	0.00720	0.0775	0.0323	0.0498
	(0.0466)	(0.0521)	(0.0603)	(0.0508)	(0.0572)	(0.0663)	(0.0755)	(0.0784)
Education	0.104***	0.0985***	0.0958***	0.107***	0.103***	0.102***	0	0
	(0.0101)	(0.0114)	(0.0114)	(0.0111)	(0.0124)	(0.0124)	(.)	(.)
Naturalized	0.0305	0.0460	0.0447	0.0133	0.0356	0.0440	-0.0437	0.0313
	(0.0558)	(0.0619)	(0.0631)	(0.0608)	(0.0656)	(0.0670)	(0.0739)	(0.0843)
Good German ability		0.112*	0.119*		0.0773	0.0856	0.00351	-0.0231
		(0.0524)	(0.0518)		(0.0587)	(0.0578)	(0.0477)	(0.0592)
Partner Good German ability			0.0410			0.0416	0.104*	0.0464
			(0.0521)			(0.0602)	(0.0525)	(0.0669)
Partner German Citizen			-0.0301			-0.0592	-0.0422	-0.00510
			(0.0524)			(0.0558)	(0.0743)	(0.0759)
Number of Persons in HH			-0.0457			-0.0456	-0.0257	-0.0234
			(0.0262)			(0.0298)	(0.0256)	(0.0284)
Number of Children in HH			-0.0208			-0.0171	-0.0201	-0.0621
			(0.0306)			(0.0357)	(0.0352)	(0.0407)
Ref*Age				-0.00462	0.000188	-0.000244		0.00836
				(0.00993)	(0.0108)	(0.0110)		(0.0609)
Ref*Male				-0.271	-0.250	-0.295		0
				(0.148)	(0.160)	(0.176)		(.)
Ref*Experience				-0.00214	-0.00468	-0.00361		-0.0205
				(0.00942)	(0.0106)	(0.0106)		(0.0640)
Ref*YSM				-0.00873	-0.0113	-0.0118		0
				(0.00865)	(0.00943)	(0.00937)		(.)
Ref*Married				-0.102	-0.107	-0.119		-0.0646
				(0.120)	(0.135)	(0.163)		(0.213)
Ref*Education				-0.0164	-0.0244	-0.0349		0
				(0.0233)	(0.0256)	(0.0269)		(.)
Ref*Naturalized				0.0965	0.0477	-0.00458		-0.218
				(0.148)	(0.166)	(0.173)		(0.181)
Ref*Good German ability					0.143	0.141		0.0980
					(0.127)	(0.127)		(0.103)
Ref*Partner German Citizen						0.129		-0.103
						(0.142)		(0.203)
Ref*Persons in HH						0.00362		-0.0124
						(0.0577)		(0.0575)
Ref*Children in HH						-0.0275		0.157*
						(0.0672)		(0.0769)
Ref*Partner Good German ability						-0.0358		0.148
						(0.124)		(0.111)
Observations	6744	5493	5493	6744	5493	5493	5493	5493
Adjusted R^2	0.396	0.401	0.404	0.397	0.402	0.405	0.076	0.080

Standard Errors in Parentheses are clustered on individual level, Included years: 1994 to 2014, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

(4) $lnY_{it} = \beta_0 + \beta X_{it} + \delta[X_{it} * Ref_i] + \alpha_i + v_{it}$

(5) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \delta[(X_{it} + Good\ German\ ability_{it}) * Ref_i] + \alpha_i + v_{it}$

(6) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \delta[(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) * Ref_i] + \alpha_i + v_{it}$

26

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Refugee	-0.165*	-0.180*	-0.140	0.515	0.496	0.596	0	0
	(0.0800)	(0.0818)	(0.0830)	(0.417)	(0.458)	(0.482)	(.)	(.)
Age	0.0598*	0.0572*	0.0597*	0.0761**	0.0723*	0.0764**	0.0333	0.0196
	(0.0264)	(0.0289)	(0.0287)	(0.0267)	(0.0293)	(0.0291)	(0.0745)	(0.0723)
Age^2	-0.000958**	-0.000937**	-0.000975**	-0.00102***	-0.000986**	-0.00105**	-0.000534	-0.000485
	(0.000299)	(0.000329)	(0.000328)	(0.000297)	(0.000329)	(0.000329)	(0.000828)	(0.000821)
Experience	0.0509***	0.0512***	0.0480**	0.0403**	0.0412**	0.0373*	0.112*	0.134**
	(0.0138)	(0.0149)	(0.0149)	(0.0147)	(0.0159)	(0.0157)	(0.0451)	(0.0432)
Experience^2	-0.000496	-0.000471	-0.000366	-0.000473	-0.000467	-0.000314	-0.000803	-0.000958
	(0.000297)	(0.000321)	(0.000321)	(0.000296)	(0.000323)	(0.000323)	(0.000790)	(0.000830)
YSM	0.0113***	0.0115**	0.0129***	0.0101**	0.0108**	0.0116**	0	0
	(0.00307)	(0.00353)	(0.00349)	(0.00328)	(0.00377)	(0.00375)	(.)	(.)
Married	0.249***	0.279***	0.277***	0.291***	0.322***	0.316***	0.103	0.0806
	(0.0574)	(0.0631)	(0.0703)	(0.0633)	(0.0696)	(0.0795)	(0.0995)	(0.0693)
Education	0.117***	0.126***	0.127***	0.110***	0.119***	0.121***	0	0
	(0.0126)	(0.0141)	(0.0145)	(0.0140)	(0.0158)	(0.0161)	(.)	(.)
Naturalized	0.0189	0.0269	0.0212	0.0579	0.0947	0.101	-0.116	-0.0122
	(0.0670)	(0.0759)	(0.0745)	(0.0690)	(0.0750)	(0.0761)	(0.0810)	(0.0628)
Good German ability		0.239***	0.245***		0.231**	0.242***	-0.0321	-0.0770
		(0.0586)	(0.0583)		(0.0707)	(0.0701)	(0.0572)	(0.0626)
Partner Good German ability			0.100			0.0799	0.108	-0.0258
			(0.0537)			(0.0621)	(0.0675)	(0.0559)
Partner German Citizen			-0.0463			-0.0647	0.0888	0.164*
			(0.0576)			(0.0605)	(0.0898)	(0.0693)
Number of Persons in HH			-0.0747*			-0.0768*	-0.00646	-0.00929
			(0.0295)			(0.0355)	(0.0322)	(0.0336)
Number of Children in HH			0.0750*			0.0929*	0.0567	0.0463
			(0.0347)			(0.0410)	(0.0380)	(0.0442)
Ref*Age				-0.0321*	-0.0300*	-0.0275		0.0418
				(0.0133)	(0.0143)	(0.0144)		(0.0756)
Ref*Experience				0.0316*	0.0302*	0.0265		-0.0646
				(0.0125)	(0.0137)	(0.0138)		(0.0763)
Ref*YSM				0.00573	0.00293	0.00559		0
				(0.00792)	(0.00827)	(0.00809)		(.)
Ref*Married				-0.183	-0.175	-0.140		0.176
				(0.142)	(0.153)	(0.165)		(0.333)
Ref*Education				0.0161	0.0168	0.00797		0
				(0.0327)	(0.0362)	(0.0386)		(.)
Ref*Naturalized				-0.151	-0.248	-0.297		-0.239
				(0.153)	(0.169)	(0.174)		(0.214)
Ref*Good German ability					0.00866	-0.0180		0.128
					(0.119)	(0.119)		(0.125)
Ref*Partner German Citizen						0.0471		-0.185
						(0.142)		(0.234)
Ref*Persons in HH						0.00322		0.00312
						(0.0643)		(0.0772)
Ref*Children in HH						-0.0664		0.0442
						(0.0763)		(0.0818)
Ref*Partner Good German ability						0.0311		0.318*
						(0.110)		(0.159)
Observations	3618	2910	2910	3618	2910	2910	2910	2910
Adjusted R^2	0.359	0.389	0.394	0.363	0.394	0.399	0.080	0.087

Standard Errors in Parentheses are clustered on individual level, Included years: 1994 to 2014, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

(4) $lnY_{it} = \beta_0 + \beta X_{it} + \delta[X_{it} * Ref_i] + \alpha_i + v_{it}$

(5) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \delta[(X_{it} + Good\ German\ ability_{it}) * Ref_i] + \alpha_i + v_{it}$

(6) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \delta[(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) * Ref_i] + \alpha_i + v_{it}$

Table 11: Y=Log Annual Earnings, Female Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Refugee	-0.0152	-0.0750	-0.105	0.416	0.510	1.068	0	0
	(0.145)	(0.152)	(0.153)	(0.786)	(0.887)	(0.870)	(.)	(.)
Age	0.0640*	0.0578	0.0973**	0.0681*	0.0649*	0.110***	-0.0197	-0.0328
	(0.0294)	(0.0329)	(0.0330)	(0.0288)	(0.0322)	(0.0320)	(0.0463)	(0.0485)
Age^2	-0.00100**	-0.000933*	-0.00143***	-0.00106**	-0.00103**	-0.00158***	-0.000494	-0.00105*
	(0.000355)	(0.000389)	(0.000394)	(0.000350)	(0.000384)	(0.000385)	(0.000446)	(0.000476)
Experience	0.0860***	0.0843***	0.0748***	0.0860***	0.0823***	0.0711***	0.147***	0.209***
	(0.0132)	(0.0136)	(0.0135)	(0.0134)	(0.0140)	(0.0141)	(0.0424)	(0.0406)
Experience^2	-0.000931**	-0.000840*	-0.000678*	-0.000910**	-0.000780*	-0.000578	-0.00161***	-0.00178***
	(0.000342)	(0.000347)	(0.000345)	(0.000345)	(0.000355)	(0.000356)	(0.000411)	(0.000433)
YSM	0.0125**	0.0118*	0.0130*	0.0154**	0.0148**	0.0161**	0	0
	(0.00462)	(0.00524)	(0.00519)	(0.00471)	(0.00547)	(0.00540)	(.)	(.)
Married	-0.238***	-0.217**	-0.161	-0.243***	-0.220**	-0.155	-0.166	0.0171
	(0.0697)	(0.0769)	(0.0920)	(0.0735)	(0.0821)	(0.0994)	(0.0851)	(0.107)
Education	0.0908***	0.0838***	0.0771***	0.0959***	0.0907***	0.0857***	0	0
	(0.0153)	(0.0165)	(0.0166)	(0.0162)	(0.0176)	(0.0178)	(.)	(.)
Naturalized	-0.0761	-0.0433	-0.0294	-0.140	-0.123	-0.108	-0.0388	0.0615
	(0.0915)	(0.0980)	(0.102)	(0.0954)	(0.103)	(0.106)	(0.105)	(0.170)
Good German ability		-0.0695	-0.0487		-0.0701	-0.0494		0.0353
		(0.0904)	(0.0881)		(0.0941)	(0.0914)		(0.0999)
Partner Good German ability			-0.0206			-0.0343		0.109
			(0.0871)			(0.0995)		(0.137)
Partner German Citizen			0.0328			0.00952		-0.171
			(0.0938)			(0.0981)		(0.120)
Number of Persons in HH			0.0120			0.0103		-0.0669
			(0.0394)			(0.0411)		(0.0471)
Number of Children in HH			-0.160***			-0.158**		-0.189**
			(0.0467)			(0.0523)		(0.0626)
Ref*Age				0.00349	0.00283	-0.00150		-0.0568
				(0.0182)	(0.0197)	(0.0190)		(0.0721)
Ref*Experience				-0.00105	0.00512	0.00692		0.0664
				(0.0210)	(0.0235)	(0.0225)		(0.0774)
Ref*YSM				-0.0376*	-0.0382*	-0.0426*		0
				(0.0167)	(0.0173)	(0.0176)		(.)
Ref*Married				0.0536	-0.0211	-0.0325		-0.271
				(0.215)	(0.239)	(0.265)		(0.198)
Ref*Education				-0.00872	-0.0198	-0.0429		0
				(0.0371)	(0.0393)	(0.0398)		(.)
Ref*Naturalized				0.520*	0.543*	0.557		-0.0238
				(0.249)	(0.271)	(0.291)		(0.226)
Ref*Good German ability					-0.0147	-0.00322		0.0919
					(0.254)	(0.241)		(0.178)
Ref*Partner German Citizen						0.218		-0.136
						(0.237)		(0.247)
Ref*Persons in HH						-0.0506		0.00893
						(0.0992)		(0.0822)
Ref*Children in HH						-0.0836		0.198
						(0.124)		(0.124)
Ref*Partner Good German ability						0.228		-0.0773
						(0.208)		(0.163)
Observations	3126	2583	2583	3126	2583	2583	3126	2583
Adjusted R^2	0.290	0.289	0.300	0.296	0.294	0.308	0.087	0.117

Standard Errors in Parentheses are clustered on individual level, Included years: 1994 to 2014, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

(4) $lnY_{it} = \beta_0 + \beta X_{it} + \delta[X_{it} * Ref_i] + \alpha_i + v_{it}$

(5) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \delta[(X_{it} + Good\ German\ ability_{it}) * Ref_i] + \alpha_i + v_{it}$

(6) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \delta[(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) * Ref_i] + \alpha_i + v_{it}$

From Table 9, we further observe that the coefficients on the variables for age, [25] male, labour market experience, [25] years since migration, and years of education have the expected positive effects and are significant. Additionally, we observe a large difference between the earnings of male and female immigrants, whereby a male immigrant earns on average around 80.8 to 82.1 percent more than a female immigrant.

The effects of marriage and naturalization on earnings are less clear. The coefficient on the marriage status in Models 1 and 2 in Table 9 is negative and not significantly different from zero at a 5 percent significance level. However, it becomes positive after controlling for partner and household effects in Model 3. A clearer picture of the effect of marriage is provided by the regression results in Table 10 and Table 11. From the male sample (Table 10), we observe positive and highly significant coefficients on marriage for all three model specifications. In contrast, the same coefficients in the women sample (Table 11) are negative and significant at a one percent level for the first two specifications. Although, the effect becomes non-significant after controlling for partner and household variables, it remains negative. These findings hint at a rather classical role distribution between married men and women, whereby the man is the main breadwinner of the household while the woman spends more hours on things other than work, such as parenting and housework.

Academic theory predicts a positive effect of naturalization on earnings (Steinhardt, 2012) as the German citizenship provides, for instance, unlimited access to the labour market for immigrants from outside the EU, and access to certain jobs in the public sector such as judical and civil positions, which are not accessible to foreigners. However, we observe no statistically significant effect on the returns to naturalization on earnings, regardless of the model specification and the sample.

Looking at Models 4 to 6 in Table 9, we observe no significant coefficients in any of the interaction terms, suggesting that there are no significant differences between refugees and economic immigrants at a 5 percent significance level for the variables we control for. The regression results for the male sample, as shown in Table 10, show a similar picture as in Table 9, as most of the coefficients on the interaction terms are not significant. However, we observe in Models 4 and 5 in Table 10 that the earnings of male refugees increase at a slower rate per additional year of age, compared to male economic migrants. Though, this effect becomes non-significant in Model 6.

Looking at the coefficient on the interaction term between the refugee indicator and years of labour market experience, we observe that male refugees benefit to a larger extent from additional years of experience, than economic immigrants. While the earnings of male economic migrants increase by between 3.73 percent to 4.12 percent (Models 4 to 6 in

---

[25]At a decreasing rate

Table 10) with each additional year of labour market experience, Models 4 and 5 in the same table show that the earnings of male refugees increase by an additional 3.02 percent to 3.16 percent. Again, this result is not significant in Model 6. We do not identify this same effect between economic migrants and refugees for the pooled or the female samples.

Turning our attention to the female sample, as shown in Table 11, we observe a significant and positive effect of naturalization on the earnings of female refugees in Models 4 and 5. More specifically, a female refugee earns 52 percent to 54.3 percent more than a female economic immigrant who is also naturalized. The overall effect of naturalization on the earnings of female refugees is 38 percent and 42 percent in Models 4 and 5 respectively. However, after controlling for partner and household characteristics, the coefficient becomes non-significant. We do not observe a similar relationship between naturalization and earnings for the male and the pooled samples.

Furthermore, we observe that the earnings of female refugees grow at a slower rate than those of female economic migrants, with each additional year of residence in Germany. Note that this interaction variable can be interpreted as the difference in the earnings growth rate between the two immigrant groups. While the earnings of female economic migrants increase by between 1.48 percent and 1.61 percent each year in Germany, the earnings of female refugees grow at a rate that is 3.76 percent to 4.26 percent slower than those of their female counterparts. In fact, the growth rates of earnings for female refugees is negative. Regardless of the specification of the model, the coefficients on this variable are significantly different from zero at a 5 percent significance level. This is a striking result, as refugees across all three samples either do not increase their earnings at a faster rate than economic immigrants, or even have decreasing earnings over time. This implies that the earnings gap between refugees and economic immigrants in Germany would never close, and sometimes, it could even widen with years.

Since annual earnings is the product of hourly wage and annual hours worked, the growth in annual earnings can be decomposed into the sum of the growth in annual hours, and the growth in hourly wages. [26] To identify whether the differences in earnings between refugees and economic migrants result from differences in the hourly wages or annual working hours, we rerun Models 1 to 3 from Tables 9, 10 and 11 with new dependent variables Log Annual Working Hours and Log Hourly Wages. The full results are displayed in Tables 28 and 25 in Appendix A.6, respectively. However, we present specifically the coefficients on the refugee identifier in Tables 12 and 13 below.

Although the refugee coefficient is not significant in Models 1 and 3 in Table 12, we observe that refugees work 10.9 percent fewer hours than economic migrants in Model 2. We do not observe a significant difference between the annual working hours of refugees

---

[26]This results from the linearisation of the earnings formula by the log-transformation.

and economic migrants in the male- and female-only samples.

Table 12: Coefficients on Refugee Identifier from Log Annual Hours Worked Regressions

	(1)	(2)	(3)
Pooled Sample	-0.0816	-0.109*	-0.0964
	(0.0506)	(0.0537)	(0.0551)
Male Sample	-0.0771	-0.0949	-0.0697
	(0.0471)	(0.0493)	(0.0480)
Female Sample	-0.0288	-0.0605	-0.0869
	(0.0920)	(0.0944)	(0.0962)

Standard Errors in Parentheses are clustered on individual level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

Table 13: Coefficients on Refugee Identifier from Log Hourly Wage Regressions

	(1)	(2)	(3)
Pooled Sample	-0.0637	-0.0830	-0.0801
	(0.0459)	(0.0474)	(0.0479)
Male Sample	-0.0883	-0.0853	-0.0709
	(0.0530)	(0.0525)	(0.0552)
Female Sample	-0.0136	-0.0146	-0.0184
	(0.0747)	(0.0791)	(0.0785)

Standard Errors in Parentheses are clustered on individual level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

Looking at Table 13, we observe no statistically significant differences between the hourly wages of refugees and economic migrants across all three samples. It is still worth noting that all coefficients are negative. These findings hint that the differences in earnings between refugees and economic migrants result directly from the fact that refugees work fewer annual hours than economic migrants.

In her study, Cortes (2004) observes that the increase in earnings of refugees over the ten-year period was particularly driven by an increase in the number of hours worked and, to a lesser extent, by an increase in the hourly wage. In order to observe similar changes in the growth rate of the components of annual earnings, we look at the coefficient of the interaction term between the refugee identifier and the variable indicating the years since migration, which gives us the differences in the growth rates between refugees

and economic migrants per additional year of residence in Germany. The coefficients on the interaction terms for the different samples and dependent variables are presented in Tables 14 and 15. We observe from Table 14 that the differences between the growth rates of wages of refugees compared to economic immigrants are negative and significant only in the female sample. This suggests that female refugees grow their hourly wages at a slower rate than female economic immigrants. On the other hand, the growth rates of hours worked shown in Table 15 is not significantly different between the two immigrant groups across all three samples. Overall, in contrast to Cortes (2004), we find no significant differences in the growth rates of wages [27] and hours worked between refugees and economic immigrants. To wrap up the findings of our first research question, we therefore find that refugees generally earn less than economic immigrants, and this is mainly due to fewer annual hours worked.

Table 14: Coefficients on Ref*YSM Interaction Term from Log Hourly Wage Regressions

	(1)	(2)	(3)
Pooled Sample	-0.00567	-0.00675	-0.00616
	(0.00464)	(0.00512)	(0.00502)
Male Sample	0.00295	0.00142	0.00215
	(0.00542)	(0.00584)	(0.00586)
Female Sample	-0.0233**	-0.0228**	-0.0221**
	(0.00743)	(0.00841)	(0.00844)

Standard Errors in Parentheses are clustered on individual level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

Table 15: Coefficients on Ref*YSM Interaction Term from Log Annual Hours Worked Regressions

	(1)	(2)	(3)
Pooled Sample	-0.00306	-0.00455	-0.00562
	(0.00549)	(0.00582)	(0.00589)
Male Sample	0.00278	0.00151	0.00344
	(0.00450)	(0.00449)	(0.00436)
Female Sample	-0.0143	-0.0154	-0.0205
	(0.0118)	(0.0117)	(0.0119)

Standard Errors in Parentheses are clustered on individual level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

---

[27]Except for the negative and significant result found in the female sample

6.2 Interpretation of the language coefficients

The relationship between language proficiency and earnings of immigrants has been frequently discussed in Labour Economics, as already described in the literature review section. By looking at Models 2 and 3 of Tables 9 and 10, we find a positive and significant relationship of language proficiency and earnings. In detail, the general effect of having a good proficiency in spoken German lies between 11.2 percent and 11.9 percent. The inclusion of controls for the partner's language proficiency and citizenship as well as for household characteristics should reduce the downward bias on the language coefficient resulting from unobserved heterogeneity Dustmann and Fabbri (2000), as we also observe. For men, the effect of language skills on earnings is even larger, as presented in Table 10. A working male immigrant with a good speaking proficiency earns between 23.9 percent and 24.5 percent more than a male immigrant with a lower speaking proficiency. These findings are consistent with the results of past research by Cortes (2004), Dustmann and Van Soest (2002) and Chiswick and Miller (1995). Surprisingly, being fluent in the German language does not appear to benefit female immigrants, regardless of whether they are a refugee or an economic immigrant. Looking at the interaction term between the refugee dummy and the language dummy, as reported in Models 5 and 6 across Tables 9 to 11, we find no significant differences to the returns of language between the two immigrant groups.

Following the argument that refugees have a greater incentive to invest in country-specific human capital, we run a logistic regression to further investigate the likelihood for a refugee to speak German well compared to an economic immigrant. We further explore the rate of improvement in German proficiency between the two immigrant groups. We choose a logistic regression model over a linear probability model in order to restrict the predicted probabilities to lie in the interval [0,1]. If we were to proceed with a linear probability model, our predicted probabilities could lie outside of that interval, which does not satisfy the definition of a probability. Section A.7 of the Appendix provides reasons to support the inclusion of key variables in the logistic regression, as well as our expectation of the outcomes. This approach follows somewhat the model used by Chiswick and Miller (1995), which focuses on the determinants of language proficiency. We first obtain the logit model (not reported), where we only focus on interpreting the sign of the coefficient, and not the magnitude as different probability models use different scales for the coefficients. In order to meaningfully interpret the magnitude, we further obtain the marginal effects when all other variables are equal to their means (Marginal Effects at the Mean or MEM)(Scott Long, 1997; Carmeron and Triveda, 2010; Long and Freese, 2006), which is displayed in Table 16.

We observe that refugees are not more likely to speak German well than economic immig-

Table 16: Logit Model, Y=Good German ability

	Good German ability=1
Refugee	0.0340
	(0.0253)
YSM	0.00267*
	(0.00110)
Ref*Years of Residence	-0.000762
	(0.00122)
Male	-0.0341*
	(0.0145)
Married	-0.0612***
	(0.0145)
Naturalized	0.00922
	(0.00958)
Partner German Citizen	0.0136
	(0.00964)
N	5520

Marginal effects; Standard Errors in Parentheses are clustered on individual level

For discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

rants. In addition, a refugee is also not more likely to speak German well than economic immigrants, with every additional year spent in Germany. We are, however, unable to conclude whether a refugee in fact does not invest more in learning the German language. Rather, we can say that there are no differences in improvement rates between the two groups. Note that we are also unable to determine the improvement rates at the point in time when the immigrant first entered Germany. At the time of our analysis, many immigrants would already have resided in Germany for many years. Hence, our approach only captures the rate of improvement at the time of our analysis.

It is worth mentioning a few other interesting findings from Table 16. Being male or married, an immigrant is less likely to speak German well, than being female or unmarried. Also, an immigrant who naturalized is not more likely to speak German well, than an immigrant who has not naturalized. This is surprising, given the criteria for one to become a German citizen. [28] It is also interesting that an immigrant whose partner is a German citizen has no significant effect on the likelihood for that immigrant to speak German well. All other variables in Table 16 have the expected signs, as we have described in Appendix A.7.

To further support our finding of the lack of differences in German proficiency improvement rates between immigrant groups, we calculate the average number of years since migration it takes an immigrant to proceed from not speaking German "well" to speaking

---

[28]Since 2005, immigrants have to obtain a proficiency level of B1 in order to naturalize in Germany

German "well". These results, and the respective sample sizes, are displayed in Table 17. We find that, on average, refugees take 16.08 years, while economic immigrants take just about the same amount of time of 16.64 years, to progress to a high level of speaking proficiency. A note of caution regarding this result is that this sample size is not representative of the entire sample we analyse, and this finding is not insightful for those who already belong to the language proficiency categories "Good" or "Very Good", and become even more proficient in German to a level that is not captured by the five-point scale. In addition, we are also not able to comment on immigrants who progressed from "Poor" to "Intermediate", for instance.

Table 17: Average number of years to reach good language proficiency

Status	Years taken to improve	Sample Size
Refugee	16.08	36
Economic Immigrant	16.64	59

Summarizing, we find that there are no significant differences between the groups in their returns to language on annual earnings. In addition, refugees are neither more likely to speak German well than economic immigrants, nor are they more likely to improve their language proficiency over time.

7 Further discussion of results

The following section discusses the main findings presented above. In detail, we first aim to present possible explanations for the differences in labour market outcomes in favour of economic immigrants. Thereafter, we will explore potential reasons behind the observed similarities in the returns to language on earnings, as well as language improvement rates, for both immigrant groups.

7.1 Labour economic success of economic immigrants over refugees in Germany

We identified four possible explanatory factors for the differences in earnings between refugees and economic migrants: Favourable selection of economic immigrants in Germany, cultural disposition resulting from different countries of origin, as well as network, and enclave, effects. A discussion of these factors is presented in the following.

7.1.1 Favourable selection of economic immigrants in Germany

In this section we will present a more in-depth exploration of the favourable selection of immigrants, which was briefly introduced in the literature review section of the paper. Depending on the foreigner's citizenship, access to the German labour market is restricted differently. From 1999, citizens from all EU or EEA member states and Switzerland have unrestricted labour market access and do not need a visa to reside indefinitely in Germany. In contrast, immigrants from a third country (countries besides those just mentioned) who intend to work in Germany need to have a "residence title for the purpose of employment" [29] in order to begin an occupation. In general, an immigrant who wants to work in Germany needs to have a passport and sufficient funds to support his stay. Are these prerequisites met, the requirements are further differentiated between applicants who have already secured a job offer and those who are intending to search for employment while in Germany. Additionally, the future level of salary, the educational qualification, and the industry the applicant is working in determine which residence title the respective person can apply for. Applicants with high qualifications and a job offer with a reasonably high salary can apply for the EU Blue Card, which entitles them to work in Germany for a maximum of four years. As of 2016, the required yearly salary by an immigrant seeking the EU Blue Card stands at EUR 49,600 (BAMF - Bundesamt für Migration und Flüchtlinge, 2016). If certain prerequisites are met, the right to residence can be extended to an open-ended permission. Job seekers from a third country with a low level of education are only allowed to stay for a maximum of six months to find work.

However, and as mentioned at the very start of the background chapter, this recent legislation cannot be applied to the entry situation of all economic immigrants in our sample, as many of them arrived in Germany in much earlier times. From Figure 2, we noted that the most prominent economic waves affecting our sample were in the 1970s and 1990s. The 1970s provide a good example of a very different immigration situation than what immigrants face in Germany today. As mentioned, the 1970s was a period where Germany had an active recruitment strategy of "guest" workers. As a result, Germany provided visas generously to many low-qualified workers. Hence, it should be noted that the strict requirements immigrants in Germany face today is not applicable to all the economic immigrants we analyse.

Considering the presented recent regulations regarding the immigration of third country citizens, it is reasonable to conclude that the EU system certainly favours well-educated and skilled applicants. This is supported by findings from a survey (Heß, 2009) among third country work immigrants, conducted in 2012, which shows that the immigrants were

---

[29]There are four kinds of residence titles, Residence permit, EU Blue Card, Settlement permit, Permanent EU residence permit, which differ with respect to the requirements which have to be met by the applicant.

overall highly qualified; 86.8 percent held a university degree as their highest educational degree, 7.5 percent were vocationally trained, 4.4 percent held a high school degree, and only 1.3 percent of all survey participants had no degree. To see if our sample displays a similar education pattern as seen from this survey, we reconstruct the histogram on "years of education" from Figure 6, but instead limit the economic immigrants to only third country citizens. The new histogram is presented in Figure 8 below. We see that our sample differs quite drastically from that in the aforementioned survey, suggesting that the third country citizens we analyse are generally not as highly educated as suggested by the survey results above.
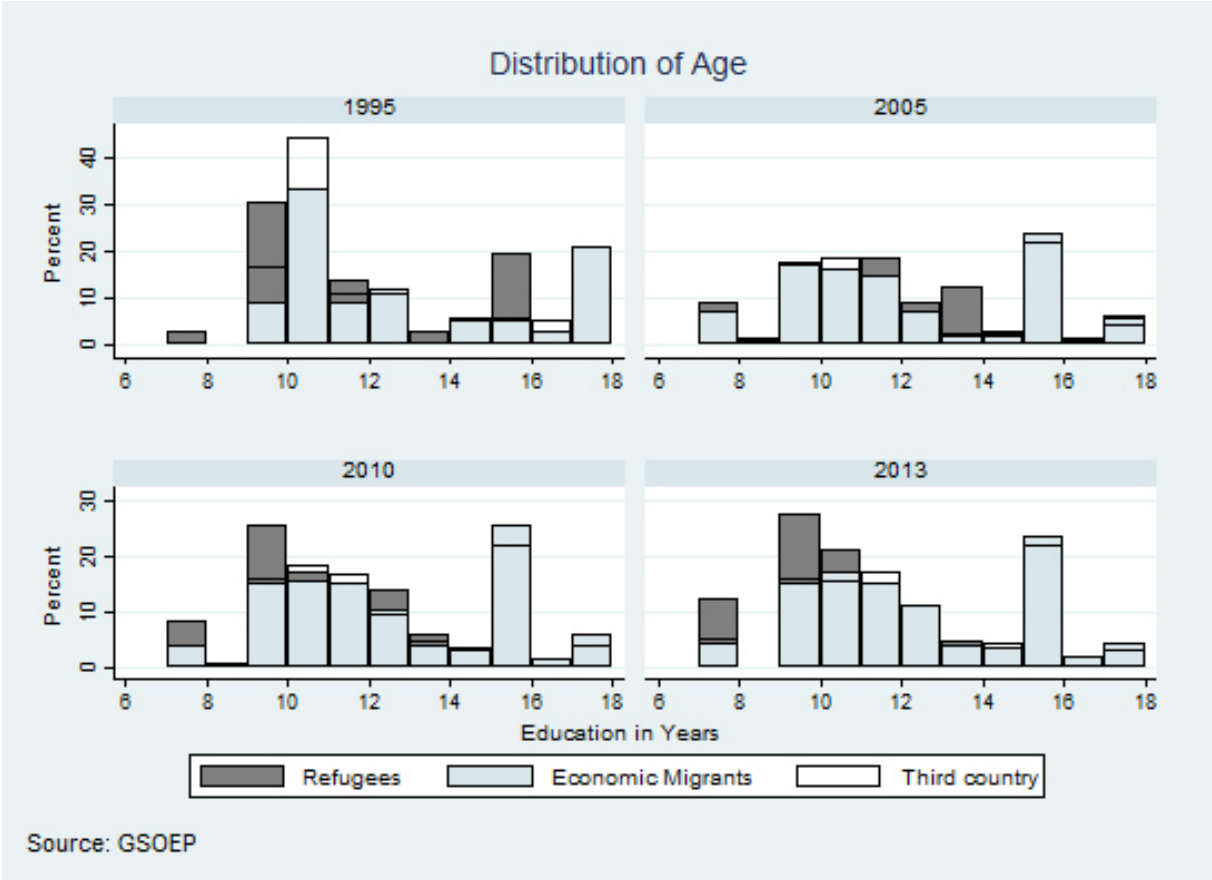
Figure 8: Histogram Education of Immigrant Groups

In addition, we analyse the types of jobs held by both immigrant groups in our sample to provide additional evidence for the favourable selection of economic migrants. Specifically, the GSOEP provides information on whether an immigrant holds a white-collar job. White-collar jobs typically require higher qualifications and expertise in the field. On the other hand, blue-collar jobs require manual labour and to some extent, less educational requirements than white-collar jobs. Thus, it is not unexpected that the salaries of white-collar workers are higher in general than blue-collar workers, so as to compensate for the time spent in gaining knowledge (Rutkowski, 1996). The results are displayed in Figure 9. This figure shows the share of white-collar workers in our sample among the total number of people for the respective years and immigrant groups. It is apparent that for

the majority of survey years included in our analysis, the share of those who are employed in white-collar jobs is larger for economic migrants than for refugees, particularly due to economic immigrants from EU countries. As the number of economic immigrants are greater than the number of refugees in our sample, Figure 9 also suggests that economic immigrants make up a larger share of all the people holding white-collar jobs in our sample.

The results from this analysis hint that the differences in annual earnings between refugees and economic migrants could potentially be driven by the higher share of economic immigrants in higher income jobs. One would also expect that white-collar jobs provide, in general, more stable working hours and higher hourly wages than blue-collar jobs. This further supports our additional findings that refugees work fewer hours, and earn a smaller hourly wage [30] on average, than economic migrants.
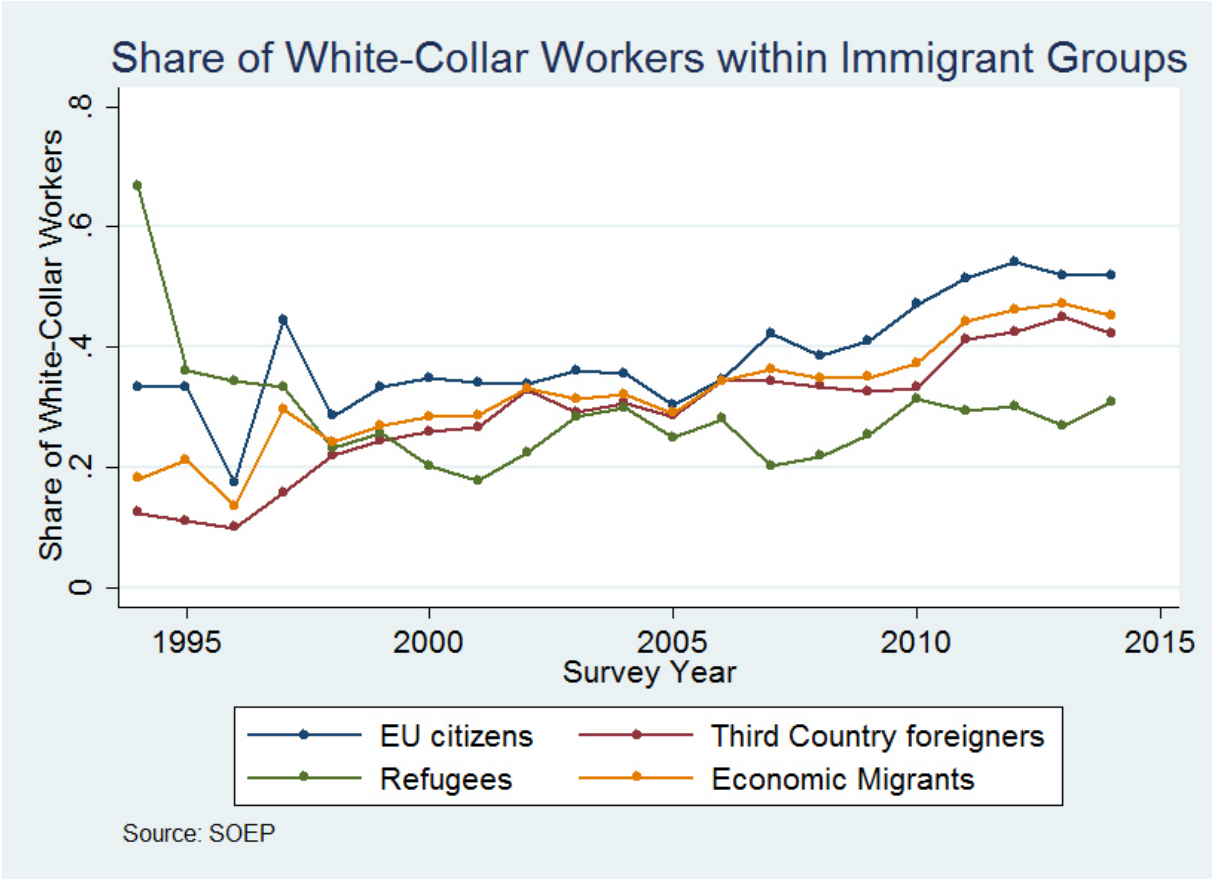
Figure 9: Share of White-Collar Workers Among Different Immigrant Groups

Another argument for the favourable selection hypothesis is based on the assumption of rational decision making, and the fact that economic immigrants enter Germany voluntarily. More precisely, in contrast to a refugee, an economic immigrant has the possibility to inform himself in advance about the requirements that need to be met in order to reside and work in Germany. Following that line of thought, a rational decision maker

---

[30]Though this was not a significant result

would only migrate to Germany if he expects to successfully maximise his economic outcome. This argument explains partly why the share of highly-educated immigrants is particularly high for the group of third country immigrants.

Following the same line of argument, a rational immigrant's success is more likely realized when he or she comes from a country that has a similar culture and language to Germany. [31] As an example, one would expect an immigrant from Austria or Switzerland to find it easier to integrate in Germany due to their strong cultural and linguistic ties, than an immigrant from, say, Turkey. If there is a disproportionately higher number of economic immigrants originating from countries similar to Germany, this could potentially explain the higher average annual earnings enjoyed by this immigrant group, compared to refugees. Indeed, we see from Table 1 that Austria and Italy [32] are one of the top economic immigrant-sending countries. Future research could instead control for countries similar to Germany, rather than for every country of origin as we did. A note should be made that an implication of this favourable selection of immigrants in Germany is that our results cannot be generalised to other countries. However, the sample we analyse is nevertheless representative of the economic immigrants in Germany.

7.1.2 Cultural disposition

Another explanation for the higher earnings of economic migrants could be due to differences in the cultural disposition to work and earnings among immigrants. Turning to Table 2, we see that the top ten refugee-sending countries are largely made up of countries less-developed than the top ten immigrant-sending countries as seen in Table 1. To reaffirm this thought, we look to the Human Development Index (HDI) (Jahan et al., 2016) of these countries. The HDI provides a weighted statistic of key components with regards to life expectancy, education, and per capita income indicators. A country ranks highly in the HDI when the life expectancy, education levels, and GDP per capita of the country are high. Table 18 gives an overview of the HDI of the immigrant-sending countries. Indeed, we observe that a majority of the economic immigrant-sending countries have a "Very High" HDI, while there are fewer of such observations in the list of refugee-sending countries.

One argument for the lower average earnings of refugees could be due to this very distinction. A refugee who is used to a lower standard of living would possibly not be as driven to aim for higher wages than an economic migrant from a more developed country, who is used to higher relative wages, and would thus seek at least the same level of wages as in the country of origin before migration is considered. This thought is further supported

---

[31] This line of thought is further supported by Adsera and Chiswick (2007)

[32] People from Northern Italy speak German to some extent

Table 18: Overview of HDI scores of immigrant's and refugee's countries of origin

Eco. Immigrants- sending Countries	HDI	Refugee- sending Countries	HDI
Turkey	0.767**	Kosovo	0.739**
Poland	0.855***	Turkey	0.767**
Italy	0.887***	Former Yugoslavia	
Austria	0.893***	Bosnia- Herzegovina	0.750**
Former Yugoslavia		Iran	0.774***
Russia	0.804***	Iraq	0.649*
France	0.897***	Albania	0.764**
Greece	0.866***	Portugal	0.843***
United States	0.920***	Lebanon	0.763**
Great Britain	0.909***	Poland	0.855***

Source: Human Development Index (Jahan et al., 2016)

*** Very high HDI ** High HDI * Medium HDI

by the possibility that refugees might have other goals, such as to ensure personal safety, rather than to maximise their economic outcomes. Yet, Aiyar et al. (2016) share that current refugees have a strong preference for countries that exhibit high employment rates, such as Germany. However, given that we have controlled for the countries of origin in our sample, we control for some aspects of these cultural differences to some extent.

7.1.3 The Network Effect

The "network effect" could also explain the greater success enjoyed by economic immigrants over refugees. Introduced by Massey and España (1987) and Massey et al. (1990), the network approach is a social factor for the probability of migration. The reasoning is that having an existing network in the host country lowers the risks and costs associated with migration and integration. Reduced costs and risks lead to the higher net return of migrating, and thus to a higher migration probability. It is reasonable to assume that economic migrants are more likely to have an already established network in the country of destination prior to arrival, than refugees. This is because refugees would, per definition, not have considered to migrate in the absence of the event that forced them to leave their home country. In that respect, one advantage an economic immigrant has over a refugee is the ability to make plans and prepare for their move. Having a strong network prior to arrival in Germany could help the immigrant find well-paying and stable jobs, possibly explaining the superiority in wages and hours worked, and hence in annual earnings. On the other hand, the spontaneous migration by refugees could prevent them from making any concrete future plans and benefiting from networks they may have in Germany. Further studies could examine the role networks play in the income differentials between the two immigrant groups.

7.1.4 Enclaves

On a similar note, it would also be interesting to study the effect of enclaves on earnings for these two distinct immigrant groups. Refugees are free to choose where to reside in Germany after successfully applying for asylum. Past research by Borjas (2000), and Edin et al. (2003) find, among other things, that living in an enclave has an effect on immigrants' wages. Whether this effect is of positive or negative nature is dependent on the enclave's quality. In general, members of high-income ethnic groups gain more from living in an enclave than members of low-income ethnic groups. Hence, the earnings gap between economic migrants and refugees could potentially be explained by a negative enclave effect if refugees were among the low-income ethnic groups who live in enclaves.

7.2 Similarities in language acquisition and returns to language

Our analysis identifies the same returns to language for both immigrant groups. In addition, we identify no differences in the probabilities for refugees to achieve a good German proficiency and to improve in the language, compared to their counterparts. In the following we present explanations for these findings.

7.2.1 Similarities in the returns to language on earnings

We find the expected result of a positive and significant return to language on annual earnings in the pooled and male samples. This is in line with a vast number of supporting research, as being proficient in spoken German provides more opportunities for work and networking, among many other things (Chiswick and Miller, 1995). We would not have expected, and did not observe, a difference in the return to language between the two immigrant groups.

7.2.2 Conflicting forces: More able economic immigrants and more hard-working refugees?

Reiterating the argument presented in section 7.1, economic immigrants make rational decisions by also taking into account information on cultural and linguistic distances to their country of origin and mother tongue. At the same time, they could also invest in country of destination-specific human capital, such as attending language courses, prior to migration. This makes it reasonable to expect the average language proficiency of economic migrants to be of a relatively high standard. Whereas, refugees would invest more in country-specific human capital only *after* arrival in Germany. Some academic

literature finds that the language learning curve for refugees is steeper than for economic migrants and explains this with larger investments into human capital (Cortes, 2004). These theoretical reasonings might hint that economic migrants have a lead over refugees regarding German proficiency before, and shortly after, migration. However, we expect refugees to eventually close the language gap due to potentially larger investments into language skills. These reasonings might explain why the probability for a refugee to speak German well is not significantly different from the probability for an economic immigrant to speak German well. However, we did not find that refugees were more likely than economic immigrants to improve their language proficiency over time. To confirm this thought, further research with more detailed data on investments into learning could identify specifically whether there are in fact differences in hours invested into learning the German language both prior- and post-migration, between immigrant groups.

7.2.3 No difference in time horizons between the two immigrant groups

At the very heart of our analysis we relied on a key distinction between a refugee and an economic immigrant; the difference in time horizons in Germany. This distinction was the basis for our hypothesis that refugees would have a greater incentive to invest in country-specific human capital than economic migrants, who have shorter time horizons. Yet, we find no difference in the probability for a refugee to speak German well over time, compared to an economic immigrant. This led us to analyse the intended length of stay for each immigrant group. Our data provides us with a variable indicating one's intention to stay permanently in Germany. We identify that a smaller share of refugees (23.32 percent) intended to stay permanently in Germany relative to the share of economic immigrants (33.61 percent). This finding potentially contradicts our assumption that economic immigrants have a shorter time horizon in Germany than refugees, and provides reasoning that economic migrants also have large incentives to invest in country-specific human capital. Note that one's intention to stay permanently in Germany might not accurately reflect the actual length of stay in Germany. However, this is the closest we can get in gaining an idea about the expected length of stay in Germany for each immigrant group.

7.2.4 Compulsory language classes for both immigrant groups

Yet another reason that could explain the similarities found in the likelihood to speak German well and the language improvement rates is the availability of integrational courses for migrants in Germany that apply to both immigrant groups.

In 2005, Germany introduced new measures to facilitate the integration process of mi-

grants. One of the core initiatives was the implementation of integration courses for all migrants who receive a residence permission after 2004. Course participation is compulsory for migrants who lack a sufficient proficiency of the German language. Migrants from an EU member state, or Ethnic Germans, are excused from this obligation. At that time, asylum applicants were also permitted to participate in language courses (Residence Act, 2016). Originally, the language course was supposed to be 660 hours long, but after an evaluation in late 2006 that identified that most participants needed longer courses to reach the required proficiency, the courses were extended. Today, depending on the specialization, the language course is between 430 and 900 hours long.

In order to accelerate the application process, asylum applicants from countries with good prospects to remain gained the right to participate in the courses free-of-charge even though their case was still pending. In certain cases, asylum seekers and refugees with residence permits can also be forced to participate in the course (Residence Act, 2016).

The law differentiates here between immigrants who received a residence title before and after 2005, Ethnic German resettlers, and EU citizens. However, whether someone is entitled to attend the course is dependent on multiple criteria, which are summarized in Table 32 in Appendix A.8. As indicated in Table 32, participation in the integrational courses is not free-of-charge for economic migrants, unless they are Ethnic Germans. For every lesson, the participant must pay a fee of 1.95 EUR. Hence, the total individual cost lie between 838,50 EUR and 1950 EUR. To pass the language course, a level of B1 must be achieved in the final test. Between 2009 and 2015, approximately half of the test takers passed the test. Over the same period, the fraction of course participants who passed the test followed a positive trend (BAMF-Bundesamt für Migration und Flüchtlinge, 2016).

Overall, the provision of language courses could partly explain why we observe no significant differences in the likelihood of refugees and economic migrants to achieve a good level of language proficiency. This is because Germany provides both immigrant groups, to some extent, equal language training opportunities. It is thus reasonable to assume that successful course participants end up with a similarly high level of German proficiency.

8 Robustness tests

This section presents three robustness tests in order to determine whether our findings survive variations, and extra considerations, made to the model and sample.

8.1 Are those who leave the panel positively selected?

Our sample is unbalanced partly due to migration out of Germany. Panel attrition could likely bias the estimation results if the remaining sample is not randomly, but instead positively, selected. A positive selection of immigrants could occur if those immigrants who have the lowest earnings in Germany have also the highest probability to leave Germany. This would bias our results because we would expect migrants who remain in Germany to invest more in country-specific human capital and thus earn more, speak German more fluently, or both (Dustmann, 1999). Our data provides information on why individuals leave the sample, which allows us to determine whether we indeed suffer from selection bias due to attrition. To do so, we create a dummy variable that equals to 1 if the respondent has moved abroad in later years, and 0 if the respondent has not moved abroad by 2014. Next, we estimate a logit model for the year 2002, with the indicator variable as the dependent variable. Again, we prefer a logistic model over a linear probability model as the predicted probabilities are between the range of zero and one, which is necessary when we want to study the likelihood of an event occurring. We further include five independent variables; the natural logarithm of annual earnings, the language proficiency dummy, years since migration, age, and country of origin. The marginal effects of the logit regression are presented in Table 19 below. Ideally, we would have conducted this regression on the first year of our analysis, 1994. However, our data does not provide sufficient observations to conduct the regression for years earlier than 2002. Even for 2002, the number of available observations is considered very small. This method of testing whether migration out of Germany is influenced by positive selection was also conducted in the paper by Dustmann and Van Soest (2002).

The coefficients on annual earnings and language proficiency turned out to be nonsignificant. This implies that earnings and language proficiency is not affecting the probability of moving abroad, and therefore, we can conclude that attrition by migrating out of Germany does not bias our results. However, this analysis does not take into account immigrants who left the sample before the year 2002. Secondly, this approach does not allow us to determine whether selective migration out of Germany took place before the panel began. If it did in fact take place, then our sample would still suffer from some degree of non-random labour force participation.

8.2 Effect of excluding EU economic immigrants on earnings gap

In the following section, we aim to test whether the earnings gap between refugees and economic migrants is particularly influenced by the presence of EU citizens. This could

Table 19: Logit Model, Y=Moved abroad

	Moved abroad=1
Log Earnings	0.0301
	(0.304)
Good German ability	-0.435
	(0.207)
Age	-0.00298
	(0.0112)
YSM	0.000225
	(0.00478)
N	134

Marginal effects; Standard Errors in Parentheses are clustered on individual level

For discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

be the case if EU citizens and third country migrants are very different from each other and hence are not a homogeneous group. Heterogeneity could, for instance, arise from differences in migration restrictions. In our specific case, EU citizens have, in contrast to third country immigrants, unrestricted movement into Germany and are allowed to stay for an unlimited period of time. Thus, EU citizens could migrate to Germany for a variety of reasons and work could be of secondary importance to them. Recall that, as per our definition, an economic immigrant migrates to Germany to work. On the other hand, we could also expect EU citizens to demand as high a salary as they have enjoyed in their home country. This is especially likely for immigrants who come from other highly developed EU countries. This idea is supported by the earlier finding in the discussion section that many EU immigrants hold white-collar jobs. In short, EU economic immigrants might end up behaving very much like native Germans. To see if there would be a difference in the result on the earnings gap between refugees and economic migrants, we rerun Model 3 from Table 9, 10 and 11, while excluding EU citizens from the sample. The key results for the pooled, male, and female sample are displayed in Table 20 above. We observe that there is still a significant difference between the earnings of refugees and third country immigrants, in favour of the latter group. However, removing all the EU citizens from our sample removes 429 distinct individuals, effectively reducing the pool of economic immigrants by one third. Therefore, we are unable to compare the coefficients between the two regressions as they are based on different subsamples. However, we can comment that the magnitude of the refugee gap increases very slightly, as seen from the pooled sample. This hints that the inclusion of EU citizens in the group of economic migrants has only little influence on the earnings gap.

Table 20: Coefficients on Refugee Identifier, Y=Log Annual Earnings, Excluding EU citizens

	Pooled	Male	Female
Refugee	-0.190*	-0.140	-0.0714
	(0.0928)	(0.0876)	(0.158)
N	4285	2238	2047
adj. R^2	0.404	0.382	0.341

Standard Errors in Parentheses are clustered on individual level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

8.3   Is controlling for spoken German proficiency sufficient?

This final test is based on the concern that the coefficient on spoken language ability might be upward biased as it also captures other aspects of language proficiency, such as writing ability. To investigate this possibility, we rerun our regressions on Models 3 and 6 of Tables 9, 10 and 11 while controlling for written German proficiency. We do so in the same way as we did for spoken German proficiency. In other words, we create a dummy variable for good German writing skills, equals to 1 if the immigrant reports his proficiency level as "Good" or "Very Good", and 0 otherwise. In addition, we include also partner's writing proficiency to the list of controls. Furthermore, we add refugee interaction terms for all the writing variables. The results for the pooled, male and female samples are reported in Table 33 in Appendix A.9.

For the pooled sample in Table 33, we find that there are no longer statistically significant and positive returns to high levels of proficiency in spoken and written German language for both immigrant groups. However, the earnings gap between economic migrants and refugees sample remain at around 17.3 percent. Interestingly, specification 1 in the male-only sample shows a positive and significant return of 19.4 percent on annual earnings when an immigrant is fluent in spoken German. In addition, specification 2 shows that written and spoken German proficiency leads to a 17.2 and 17.3 percent increase in earnings respectively, at the 5 percent significance level. There are no reported differences between immigrant groups in the pooled and the male sample. Finally, we observe that female working refugees benefit more from written German than their economic migrant counterparts. Female refugees increase their earnings on average by 60.9 percent more than female economic immigrants, and this result is significant at the 5 percent significance level.

9   Conclusion

In recent times, there has been an increased interest in the differences in economic outcomes between refugees and economic immigrants. This is especially so, considering the refugee crisis that is at the forefront of global issues today. Germany has played, and will continue to play, an important role in the integration of refugees into Europe.

Our results show that immigrants to Germany are in fact not a homogenous group. Particularly, there exists an earnings gap between refugees and economic immigrants. This gap is especially prominent among the male immigrants, while it is statistically non-significant in the female sample. It is important to reiterate that our findings are very country- and sample-specific, and can only be applied to the case of working immigrants in Germany or countries with similar welfare systems and immigration practices to Germany. We explain this gap by establishing that refugees differ from economic immigrants in some crucial aspects that make their labour economic outcomes inferior to those of economic migrants. Looking deeper, we identify that refugees work on average fewer hours than economic migrants, which partly explains the earnings differences between the two groups.

The presence of such an earnings gap is almost always expected, though sometimes it is only observed at the beginning of the immigrant's residence in a new country. The review of academic literature hints that natives are generally always going to have somewhat superior earnings to immigrants (at least in the beginning). Our results present further reason to expect that refugees will be the least successful group among immigrant workers in Germany. This suggests that Germany is generally a tough environment for immigrants, and especially for refugees, to integrate into and catch up to the earnings level of natives. Despite the presence of important integrational courses, there are some aspects of an immigrant, such as their cultural disposition to their old way of life, that the German government can never alter. These innate differences between refugees and economic migrants might prevent the gap from ever closing completely.

Furthermore, we find that language is an important and significant contributor to the earnings of both immigrant groups. Hence, emphasis should be placed on ensuring that the immigrants achieve a high standard of German proficiency. In Germany today, there exists German language courses with strict participation and fulfilment requirements tailored to most immigrants. This effort should therefore be continued, and further enforced to help the immigrant master the language.

Academic literature assumes that investments into country-specific human capital and language are largely dependent on the expected duration the immigrant stays in the country of destination (Chiswick and Miller, 1995; Cortes, 2004). This study finds that there are only limited differences in the expected time horizons between refugees and economic

migrants, as many economic migrants report the desire to stay in Germany permanently. This partly explains why we do not observe differences in language improvement rates, or the likelihood to speak German well, between the two immigrant groups.

In our effort to determine the returns of language on earnings, we meet with one key problem that usually plagues studies like ours. Namely, the presence of the endogenous variable *ability* that affects both the earnings and the language proficiency of immigrants simultaneously. We attempted to overcome this problem by exploring IV and Fixed Effects estimation methods. However, we were unsuccessful in obtaining strong instruments, which led us to avoid the IV method. Due to limited within-variation of key variables of interest, and the fact that time-invariant variables do not survive the differencing process, we were also unable to benefit from the Fixed Effects estimation method. Instead, we utilize OLS estimation methods. To obtain correct standard errors, we cluster with respect to the individual identifier and allow observations to be independent across individuals, but not within an individual. Certainly, we still expect some degree of endogeneity bias in our results. Given this, the true returns to language might be smaller than what we reported.

Future research could further examine the effect of enclaves in Germany on the earnings of refugees and economic immigrants. As per our discussion, we noted that the presence of enclaves could potentially be an explanation for the earnings differentials. As we lack information on enclaves at this point in time, we are unable to study this effect. Future research could also change the approach we use in controlling for the countries of origin. In our study, we control for every country of origin present in our sample. However, it would be interesting to control specifically for countries that are culturally and linguistically similar to Germany.

References

Adsera, A. and Chiswick, B. R. (2007), 'Are there gender and country of origin differences in immigrant labor market outcomes across european destinations?', *Journal of Population Economics* **20**(3), 495.

Aiyar, S., Barkbu, B., Batini, N., Berger, H., Detragiache, E., Dizioli, A., Ebeke, C., Lin, H. H., Kaltani, L., Sosa, S. et al. (2016), The refugee surge in europe; economic challenges, Technical report, International Monetary Fund.

Alesina, A., Glaeser, E. and Sacerdote, B. (2001), Why doesn't the us have a european-style welfare system?, Technical report, National Bureau of Economic Research.

Anderson, K. H. (2015), 'Can immigrants ever earn as much as native workers?', *IZA World of Labor* .

Assembly, UN General (1948), 'Universal declaration of human rights', *UN General Assembly* .

Asylum Act (2017*a*), 'Asylum act in the version promulgated on 2 september 2008 (federal law gazette i, p. 1798), last amended by article 2 of the act of 11 march 2016 (federal law gazette i, p. 394)', (Section 3 Recognition of refugee status).

Asylum Act (2017*b*), 'Asylum act in the version promulgated on 2 september 2008 (federal law gazette i, p. 1798), last amended by article 2 of the act of 11 march 2016 (federal law gazette i, p. 394)', (Section 4 Subsidiary protection).

BAMF - Bundesamt für Migration und Flüchtlinge (2016).
`http://www.bamf.de/EN/Willkommen/Aufenthalt/WichtigeInformationen/wichtigeinformationen-node.html`

BAMF - Bundesamt für Migration und Flüchtlinge (2017), 'Ethnic german resettlers'.
`http://www.bamf.de/EN/Migration/Spaetaussiedler/spaetaussiedler-node.html`

BAMF-Bundesamt für Migration und Flüchtlinge (2016), 'Schlüsselzahlen integrationskurse 2015'.

Basic Law (2017), 'Basic law for the federal republic of germany in the revised version published in the federal law gazette part iii, classification number 100-1, as last amended by article 1 of the act of 23 december 2014 (federal law gazette i p. 2438)', (Article 16a Right of asylum).

Bauer, T., Dietz, B., Zimmermann, K. F. and Zwintz, E. (2005), 'German migration: Development, assimilation, and labour market effects', *European Migration: What do we know* pp. 197–261.

Beyer, R. C. (2016), *The Labor Market Performance of Immigrants in Germany*, International Monetary Fund.

Bleakley, H. and Chin, A. (2004), 'Language skills and earnings: Evidence from childhood immigrants', *Review of Economics and statistics* **86**(2), 481–496.

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

Borjas, G. J. (2000), 'Ethnic enclaves and assimilation', *Swedish Economic policy review* **7**(2), 89–122.

Bound, J., Jaeger, D. A. and Baker, R. M. (1995), 'Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak', *Journal of the American statistical association* **90**(430), 443–450.

Butterwegge, C. (2005*a*), 'Migration in Ost- und Westdeutschland von 55 bis 04 ', *Bundesbehörde für politische Bildung* .

Butterwegge, C. (2005*b*), 'Von der Gastarbeiter-Anwerbung zum Zuwanderungsgesetz', *From the recruitment of gastarbeiter to the Immigration Law* .

Carmeron, C. and Triveda, P. (2010), 'Microeconomics using stata revised edition', *StataCorp LP College Station, Texas* .

Case, A. C. and Katz, L. F. (1991), 'The company you keep: The effects of family and neighborhood on disadvantaged youths'.

Chao, J. C. and Swanson, N. R. (2005), 'Consistent estimation with a large number of weak instruments', *Econometrica* **73**(5), 1673–1692.

Chiswick, B. R. (1978), 'The effect of americanization on the earnings of foreign-born men', *Journal of political Economy* **86**(5), 897–921.

Chiswick, B. R. (1982), 'The economic progress of immigrants: some apparently universal patterns.'.

Chiswick, B. R. and Miller, P. W. (1995), 'The endogeneity between language and earnings: International analyses', *Journal of labor economics* **13**(2), 246–288.

Chiswick, B. R. and Miller, P. W. (1998), 'English language fluency among immigrants in the united states', *Research in Labor Economics* **17**(9), 151–200.

Chiswick, B. R. and Miller, P. W. (2007), Modeling immigrants' language skills, *in* 'Immigration', Emerald Group Publishing Limited, pp. 75–128.

Constant, A. and Massey, D. S. (2005), 'Labor market segmentation and the earnings of german guestworkers', *Population Research and Policy Review* **24**(5), 489–512.

Cortes, K. E. (2004), 'Are refugees different from economic immigrants? some empirical evidence on the heterogeneity of immigrant groups in the united states', *Review of Economics and Statistics* **86**(2), 465–480.

Council of the European Union (2016), 'Eu-turkey statement'.
`http://www.consilium.europa.eu/press-releases-pdf/2016/3/40802210113_en.pdf`

DESTATIS (2014), 'Data base'.
`https://www.destatis.de/EN/FactsFigures/SocietyState/Population/Migration/Tables_/%lrbev07.html`

Duleep, H., Regets, M. and Sanders, S. (2001), 'A new look at human capital investment: A study of asian immigrants and their family ties'.

Dustmann, C. (1993), 'Earnings adjustment of temporary migrants', *Journal of Population Economics* **6**(2), 153–168.

Dustmann, C. (1994), 'Speaking fluency, writing fluency and earnings of migrants', *Journal of Population economics* **7**(2), 133–156.

Dustmann, C. (1999), 'Temporary migration, human capital, and language fluency of migrants', *The Scandinavian Journal of Economics* **101**(2), 297–314.

Dustmann, C. and Fabbri, F. (2000), 'Language proficiency and labour market performance of immigrants in the uk'.

Dustmann, C. and Fabbri, F. (2003), 'Language proficiency and labour market performance of immigrants in the uk', *The Economic Journal* **113**(489), 695–717.

Dustmann, C., Schönberg, U. and Stuhler, J. (2016), 'The impact of immigration: Why do studies reach such different results?', *The Journal of Economic Perspectives* **30**(4), 31–56.

Dustmann, C. and Van Soest, A. (2002), 'Language and the earnings of immigrants', *ILR Review* **55**(3), 473–492.

Edin, P.-A., Fredriksson, P. and Åslund, O. (2003), 'Ethnic enclaves and the economic success of immigrants—evidence from a natural experiment', *The quarterly journal of economics* **118**(1), 329–357.

Ermisch, J. and Pronzato, C. (2010), 'Causal effects of parents' education on children's education'.

European Commission (2016*a*), 'Die EU und die Flüchtlingskrise'.
`https://ec.europa.eu/germany/sites/germany/files/na0416628den_002.pdf`

European Commission (2016*b*), 'Eu - libya relations, factsheet'.

European Commission (2017), 'Press release: Relocation and resettlement: Member states need to build on encouraging'.

Eurostat (2017), 'Asylum and first time asylum applicants - annual aggregated data (rounded)'.
`http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&plugin=1&`
`language=en&pcode=tps00191`

Gesemann, F. and Roth, R. (2009), *Lokale Integrationspolitik in der Einwanderungs- gesellschaft: Migration und Integration als Herausforderung von Kommunen*, Springer-Verlag.

Heß, B. (2009), 'Zuwanderung von hochqualifizierten aus drittstaaten nach deutschland: Ergebnisse einer schriftlichen befragung'.

Hönekopp, E. (1997), *Labour migration to Germany from Central and Eastern Europe: old and new trends*, number 23, Institut für Arbeitsmarkt-und Berufsforschung der Bundesanstalt für Arbeit.

Jahan, S. et al. (2016), 'Human development report 2016', *United Nations Development Programme* .

Long, J. S. and Freese, J. (2006), *Regression models for categorical dependent variables using Stata*, Stata press.

Massey, D. S., Alarcón, R., Durand, J. and González, H. (1990), *Return to Aztlan: The social process of international migration from Western Mexico*, Vol. 1, Univ of California Press.

Massey, D. S. and España, F. G. (1987), 'The social process of international migration', *Science* **237**, 733–739.

OECD (2017), 'Finding their way, labour market integration of refugees in germany'.

Oltmer, J. (2005), 'Deutsche Migrationsgeschichte seit 1871'.
`http://www.bpb.de/themen/Q0DB0G2C0%2CDeutsche_Migrationsgeschichte_`
`seit_1871.html`

Ott, E. (2013), 'The labour market integration of resettled refugees'.

Pischke, J.-S. (1992), Assimilation and the earnings of guestworkers in germany, Technical report, ZEW Discussion Papers.

Residence Act (2016), 'Residence act (aufenthaltsgesetz), section 44 subs. 4 sentence 2 nos. 1 to 3 gesetz über den aufenthalt, die erwerbstätigkeit und die integration von

ausländern im bundesgebiet § 44 berechtigung zur teilnahme an einem integrationskurs'.
https://www.gesetze-im-internet.de/aufenthg_2004/__44.html

Residence Act (2017), 'Residence act in the version promulgated on 25 february 2008 (federal law gazette i p. 162), last amended by article 1 of the act of 11 march 2016 (federal law gazette i p. 394)', (Section 60 Prohibition of deportation).

Rutkowski, J. (1996), 'High skills pay off: the changing wage structure during economic transition in poland', *Economics of Transition* **4**(1), 89–112.

Schmidt, C. M. (1997), 'Immigrant performance in germany: Labor earnings of ethnic german migrants and foreign guest-workers', *The Quarterly Review of Economics and Finance* **37**, 379–397.

Scott Long, J. (1997), 'Regression models for categorical and limited dependent variables', *Advanced quantitative techniques in the social sciences* **7**.

Shea, J. (1997), 'Instrument relevance in multivariate linear models: A simple measure', *Review of Economics and Statistics* **79**(2), 348–352.

SOEP (2017), 'Socio-economic panel (soep), data for years 1984-2019, version 36, soep, 2020, doi:10.5684/soep.v36.'.

Steinhardt, M. F. (2012), 'Does citizenship matter? the economic impact of naturalizations in germany', *Labour Economics* **19**(6), 813–823.

UNHCR (2017), 'What is a refugee'.
http://www.unrefugees.org/what-is-a-refugee/

United Nations Conference of Plenipotentiaries on the Status of Refugees and Stateless Persons, 1951 Geneva (1951), *Final Act and Convention Relating to the Status of Refugees*, New York: United Nations.

von Altenbockum, J. (2017), 'Zwanzig Jahre Asylkompromiss: Die Brutstätte des Rechtsextremismus'.
http://www.faz.net/aktuell/politik/harte-bretter/
zwanzig-jahre-asylkompromiss-die-brutstaette-des-rechtsextremismus-12195833.
html

Wagner, G. G., Frick, J. R. and Schupp, J. (2007), 'The german socio-economic panel study (soep)-evolution, scope and enhancements'.

Weinzierl, R. (2009), 'Der Asylkompromiss 1993 auf dem Prüfstand', *Gutachten zur Vereinbarkeit der deutschen Regelungen über sichere EU-Staaten und sichere Drittstaaten mit der Europäischen Menschenrechtskonvention, dem EU-Recht und dem Deutschen Grundgesetz* .

Wooldridge, J. M. (2013), *Introductory econometrics: A modern approach*, Nelson Education.

A Appendix

A.1 Overview protection measures

	Legal Basis	Target group	Entitlement
Refugee protection	Section 3 subs. 1 of the Asylum Act (AsylG)	Persons who, because of a well-founded fear of being persecuted by state or non-state players for race, nationality, political opinion, fundamental religious conviction or belonging to a particular social group, are outside their country of origin and nationality, or as stateless individuals are outside of their country of habitual residence.	• Three year residence permit • Settlement permit possible after five years if reconditions are met (ability to secure own living, sufficient language proficiency) • Unrestricted labor market access • Entitled to privileged family reunification
Entitlement to Asylum	Article 16a para.1 of Constitution	Persons who cannot return to their country of origin because they would be subject to serious human rights violation and who are deemed to have been persecuted on political grounds because of their race nationality, political opinion, funamental relegious conviction or belonging to a particular social group.	• Three year residence permit • Settlement permit possible after five years if reconditions are met (ability to secure own living, sufficient language proficiency) • Unrestricted labor market access • Entitled to privileged family reunification
Subsidary protection	Section 4 subs. 1 of the Asylum Act (AsylG)	"Persons, who put forward substantial grounds for the presumption that they are at risk of serious harm, [33] from both governmental or non-governmental players, in their country of origin and that they cannot take up the protection of their country of origin or do not wish to take it up because of that threat."	• One year residence permit • if extended: two more years in each case • Settlement permit possible after five years if reconditions are met (ability to secure own living, sufficient language proficiency) • Unrestricted labor market access • **Not** entitled to privileged family reunification
National ban on deportation	Section 60 subs. 5 of the Residence Act (AufenthG)	"A person who is seeking protection may not be returned if return to the destination country constitutes a breach of the European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR), or a considerable concrete danger to life, limb or liberty exists in that country. A considerable concrete danger can be considered to exist for health reasons if a return would cause life-threatening or serious diseases to become much worse." country constitutes a breach of the European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR), or a considerable concrete danger to life, limb or liberty exists in that country.	• At least one year residence permit (repeated extension possible) • Settlement permit possible after five years if reconditions are met (ability to secure own living, sufficient language proficiency) • Employment possible, permission required • **Not** entitled to privileged family reunification

	1993-1995	1995-2000	2000-2007	2007-2014	2014-2015	Since 2015
Albania						▓
Bosnia and Herzegovina					▓	▓
Bulgaria*	▓	▓	▓			
Czech Republic*	▓	▓	▓			
Hungary	▓	▓	▓			
Gambia	▓					
Ghana	▓	▓	▓	▓	▓	▓
Kosovo						▓
Macedonia, Former Yugoslavia					▓	▓
Montenegro						▓
Poland*	▓	▓	▓			
Romania*	▓	▓	▓			
Senegal	▓	▓	▓	▓	▓	▓
Serbia					▓	▓
Slovakia*	▓	▓	▓			

* Country became EU member state and hence automatically a safe country of origin

Figure 10: Safe Countries of Origin defined by Section 29a Safe country of origin (2) Annex II Asylum Act

A.3 Descriptive statistics

Table 21: Descriptive statistics, pooled sample

| | 2005 | | | | | | 2010 | | | | | | 2013 | | | | | |
| | Refugee | | | Economic Migrants | | | Refugee | | | Economic Migrants | | | Refugee | | | Economic Migrants | | |
Independent Variable	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.
Age	42.38	10.91	64	41.85	10.59	282	39.98	9.33	124	40.83	9.15	519	41.24	9.60	148	42.17	9.11	641
Years of Residence in Germany	14.60	6.06	60	19.62	11.54	276	16.52	7.02	124	19.21	12.11	516	18.93	6.01	147	20.69	12.07	628
Age at immigration	26.48	11.19	60	22.27	10.64	276	23.46	9.72	124	21.56	10.24	516	22.27	10.19	147	21.65	10.76	628
Max Years of Education	11.41	2.76	64	11.84	2.93	278	11.00	2.50	122	12.23	2.76	510	10.49	2.35	146	12.06	2.70	629
Married	0.78	0.42	64	0.78	0.41	282	0.83	0.38	124	0.77	0.42	519	0.81	0.39	148	0.74	0.44	641
Log Annual Earnings	9.34	1.13	64	9.72	1.21	282	9.46	0.96	124	9.72	1.18	519	9.39	1.12	148	9.75	1.09	641
LogHHlaborincome	10.12	0.75	64	10.50	0.97	282	9.95	0.91	124	10.45	1.00	519	9.95	1.04	148	10.50	0.93	641
Annual Work Hours of Individual	1605.58	844.86	64	1803.85	866.77	282	1658.94	882.68	124	1759.07	919.31	519	1622.55	850.20	148	1738.97	858.59	641
No. Children in Household	0.91	1.02	64	0.96	1.13	282	1.95	1.55	124	1.49	1.16	519	2.10	1.34	148	1.49	1.17	641
Household Size	3.55	1.32	64	3.17	1.46	282	4.19	1.60	124	3.59	1.25	519	4.53	1.71	148	3.57	1.39	641
Naturalized	0.31	0.47	64	0.26	0.44	282	0.51	0.50	55	0.36	0.48	255	0.45	0.50	148	0.31	0.46	641
Labor Market Experience	15.17	10.81	64	16.96	10.43	282	14.65	8.71	124	15.90	10.18	519	15.49	9.24	148	16.26	10.10	641
Good German ability	0.54	0.50	59	0.71	0.46	264	0.73	0.45	55	0.78	0.42	241	0.72	0.45	146	0.76	0.42	598
Spoken German Ability	3.63	1.00	59	4.00	0.96	264	3.95	0.76	55	4.15	0.916	241	3.92	0.82	146	4.16	0.86	698
Oral Ability: Language Country of Origin							1.36	0.62	53	1.28	0.57	221	1.31	0.55	144	1.31	0.65	580
Father's Education in Years	6.45	4.45	40	6.88	4.38	209	7.60	4.44	35	7.46	4.32	183	6.86	4.69	44	8.02	4.11	246
Partner Good German ability	0.31	0.47	64	0.37	0.48	282	0.23	0.43	124	0.16	0.37	519	0.41	0.49	148	0.34	0.47	641
Partner Years of Residence	16.32	9.12	38	20.98	10.44	118	14.81	8.24	86	19.13	11.57	215	17.41	7.09	94	20.36	11.64	249
Partner Age	42.71	11.87	56	51.01	127.00	241	37.96	10.59	112	40.22	9.92	455	40.24	9.71	128	53.16	148.42	530
Partner Good German ability with missing	0.47	0.50	43	0.74	0.44	139	0.71	0.46	41	0.78	0.42	108	0.58	0.50	103	0.76	0.43	290
Partner Max Years of Education	10.80	2.39	40	12.10	3.12	175	10.73	2.32	94	12.65	3.09	367	10.51	2.41	116	12.47	3.07	429
Partner German Citizen	0.37	0.49	54	0.59	0.49	230	0.27	0.45	108	0.34	0.48	435	0.42	0.50	124	0.70	0.46	491

Table 22: Descriptive statistics male sample

| | 2005 | | | | | | 2010 | | | | | | 2013 | | | | | |
| | Refugee | | | Economic Migrants | | | Refugee | | | Economic Migrants | | | Refugee | | | Economic Migrants | | |
Independent Variable	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.
Age	42.40	11.75	40	41.23	10.56	148	40.52	9.42	88	41.32	9.12	284	41.76	9.12	97	42.93	9.09	309
Years of Residence in Germany	12.84	4.42	37	19.73	11.75	143	16.75	7.07	88	19.49	12.63	282	18.94	5.72	96	21.10	13.04	305
Age at immigration	27.97	11.39	37	21.59	10.83	143	23.77	9.53	88	21.77	10.39	282	22.77	9.41	96	21.93	11.03	305
Max Years of Education	11.39	2.52	40	11.80	2.82	146	10.76	2.27	86	11.96	2.81	279	10.06	2.09	97	11.88	2.75	304
Married	0.83	0.38	40	0.80	0.40	148	0.89	0.32	88	0.82	0.38	284	0.88	0.33	97	0.80	0.40	309
Log Annual Earnings	9.68	0.95	40	10.17	1.03	148	9.64	0.85	88	10.19	0.92	284	9.64	0.97	97	10.26	0.81	309
LogHHlaborincome	10.19	0.66	40	10.52	0.98	148	9.89	0.89	88	10.48	0.91	284	9.92	0.96	97	10.58	0.76	309
Annual Work Hours of Individual	1983.80	686.67	40	2122.29	751.06	148	1821.35	834.75	88	2174.79	790.98	284	1831.13	738.77	97	2167.21	722.13	309
No. Children in Household	1.03	1.12	40	1.01	1.08	148	2.16	1.54	88	1.65	1.17	284	2.38	1.40	97	1.61	1.21	309
Household Size	3.63	1.29	40	3.24	1.36	148	4.39	1.59	88	3.78	1.25	284	4.84	1.72	97	3.73	1.40	309
Naturalized	0.28	0.45	40	0.25	0.43	148	0.54	0.51	35	0.36	0.48	123	0.43	0.50	97	0.31	0.46	309
Labor Market Experience	17.05	11.08	40	19.07	10.13	148	16.50	8.47	88	18.35	10.10	284	17.62	8.18	97	19.58	9.84	309
Good German ability	0.46	0.51	37	0.71	0.45	136	0.71	0.46	35	0.76	0.43	117	0.70	0.46	96	0.73	0.45	284
Spoken German Ability	3.49	0. 93	37	3.95	0.99	136	3.94	0.72	35	4.08	0.94	117	3.85	0.81	96	4.04	0.89	284
Oral Ability: Language Country of Origin							1.30	0.64	33	1.30	0.56	104	1.26	0.51	95	1.32	0.62	275
Father's Education in Years	5.46	4.51	26	6.76	4.24	116	6.61	5.17	23	6.66	4.60	97	5.08	4.94	26	7.66	4.34	112
Partner Good German ability	0.38	0.49	40	0.45	0.50	148	0.24	0.43	88	0.18	0.39	284	0.43	0.50	97	0.43	0.50	309
Partner Years of Residence	16.08	9.03	25	20.00	9.84	68	13.98	8.02	63	18.10	11.23	138	17.20	7.85	66	18.72	10.67	145
Partner Age	40.22	11.23	36	54.78	172.79	130	35.53	9.71	83	37.63	8.69	265	37.92	8.51	90	53.69	167.42	278
Partner Good German ability with missing	0.50	0.51	30	0.77	0.42	86	0.70	0.47	30	0.85	0.36	61	0.56	0.50	75	0.75	0.44	179
Partner Max Years of Education	10.37	2.58	23	11.82	2.91	91	10.48	2.31	68	12.47	3.07	214	10.13	2.15	80	12.44	3.05	232
Partner German Citizen	0.34	0.48	35	0.52	0.50	126	0.24	0.43	80	0.27	0.44	256	0.41	0.49	90	0.68	0.47	265

Table 23: Descriptive statistics female sample

| | 2005 | | | | | | 2010 | | | | | | 2013 | | | | | |
| | Refugee | | | Economic Migrants | | | Refugee | | | Economic Migrants | | | Refugee | | | Economic Migrants | | |
Independent Variable	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.	Mean	Std. Dev.	No. Obs.
Age	42.33	9.59	24	42.54	10.62	134	38.64	9.11	36	40.23	9.17	235	40.24	10.47	51	41.46	9.09	332
Years of Residence in Germany	17.43	7.27	23	19.50	11.34	133	15.94	6.97	36	18.86	11	234	18.92	6.57	51	20.30	11.07	323
Age at immigration	24.09	10.66	23	23.00	10.42	133	22.69	10.26	36	21.31	10.07	234	21.31	11.55	51	21.38	10.51	323
Max Years of Education	11.46	3.17	24	11.89	3.05	132	11.60	2.92	36	12.55	2.68	231	11.36	2.61	49	12.24	2.65	325
Married	0.71	0.46	24	0.77	0.42	134	0.69	0.47	36	0.70	0.46	235	0.69	0.47	51	0.68	0.47	332
Log Annual Earnings	8.77	1.21	24	9.22	1.20	134	9.01	1.07	36	9.14	1.20	235	8.92	1.25	51	9.28	1.10	332
LogHHlaborincome	9.99	0.89	24	10.48	0.96	134	10.08	0.94	36	10.43	1.10	235	9.99	1.18	51	10.43	1.07	332
Annual Work Hours of Individual	975.21	701.58	24	1452.14	852.28	134	1261.94	881.79	36	1256.66	806.81	235	1225.84	912.32	51	1340.40	780.76	332
No. Children in Household	0.71	0.81	24	0.91	1.17	134	1.44	1.46	36	1.31	1.11	235	1.57	1.04	51	1.37	1.12	332
Household Size	3.42	1.38	24	3.10	1.57	134	3.69	1.53	36	3.36	1.22	235	3.94	1.54	51	3.42	1.36	332
Naturalized	0.38	0.49	24	0.27	0.44	134	0.45	0.51	20	0.36	0.48	132	0.47	0.50	51	0.31	0.46	332
Labor Market Experience	12.04	9.78	24	14.62	10.28	134	10.14	7.65	36	12.95	9.49	235	11.44	9.83	51	13.17	9.33	332
Good German ability	0.68	0.48	22	0.70	0.46	128	0.75	0.44	20	0.79	0.41	124	0.76	0.43	50	0.80	0.40	314
Spoken German Ability	3.86	1.08	22	4.06	0.92	128	3.95	0.83	20	4.22	0.89	124	4.06	0.84	50	4.27	0.82	314
Oral Ability: Language Country of Origin							1.45	0.60	20	1.26	0.59	117	1.39	0.61	49	1.31	0.68	305
Father's Education in Years	8.29	3.83	14	7.03	4.56	93	9.50	1.24	12	8.37	3.80	86	9.44	2.81	18	8.33	3.89	134
Partner Good German ability	0.21	0.41	24	0.28	0.45	134	0.22	0.42	36	0.14	0.34	235	0.35	0.48	51	0.26	0.44	332
Partner Years of Residence	16.77	9.64	13	22.32	11.17	50.00	17.09	8.59	23	20.97	12.03	77	17.93	4.94	28	22.63	12.58	104
Partner Age	47.20	11.94	20	46.59	11.94	111	44.93	10.03	29	43.84	10.41	190	45.74	10.28	38	52.58	124.48	252
Partner Good German ability with missing	0.38	0.51	13	0.70	0.46	53	0.73	0.47	11	0.68	0.47	47	0.64	0.49	28	0.77	0.43	111
Partner Max Years of Education	11.38	2.02	17	12.40	3.32	84	11.40	2.23	26	12.90	3.13	153	11.38	2.75	36	12.51	3.10	197
Partner German Citizen	0.42	0.51	19	0.66	0.47	104	0.36	0.49	28	0.45	0.50	179	0.44	0.50	34	0.73	0.44	226

A.4   Histograms
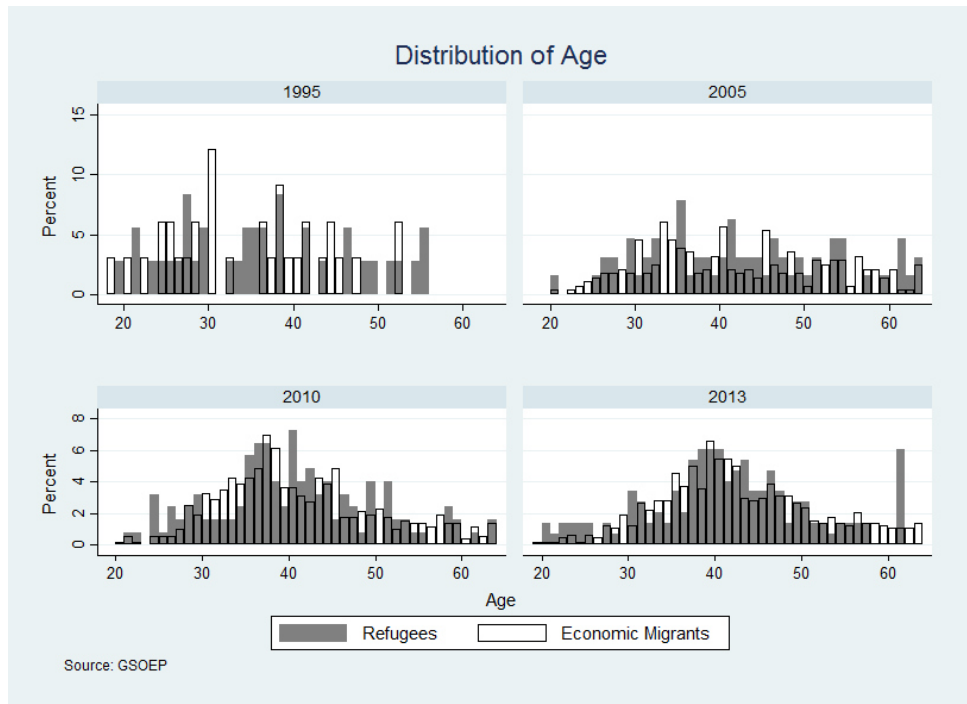
Figure 11: Histogram Age
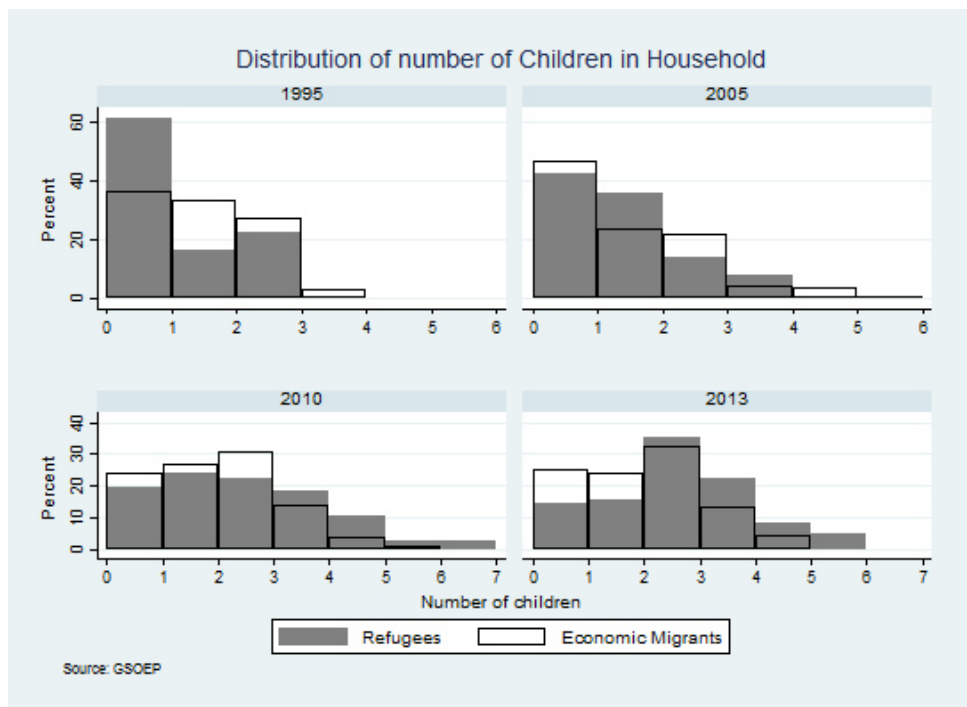
Figure 12: Histogram Number of Children in HH
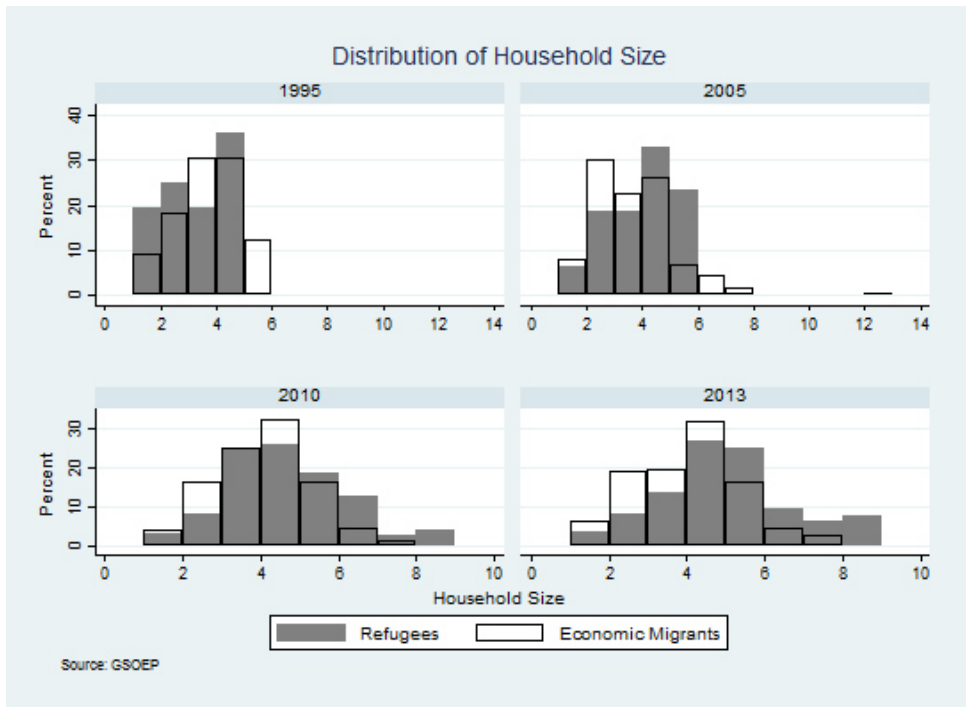
Figure 13: Histogram Household Size
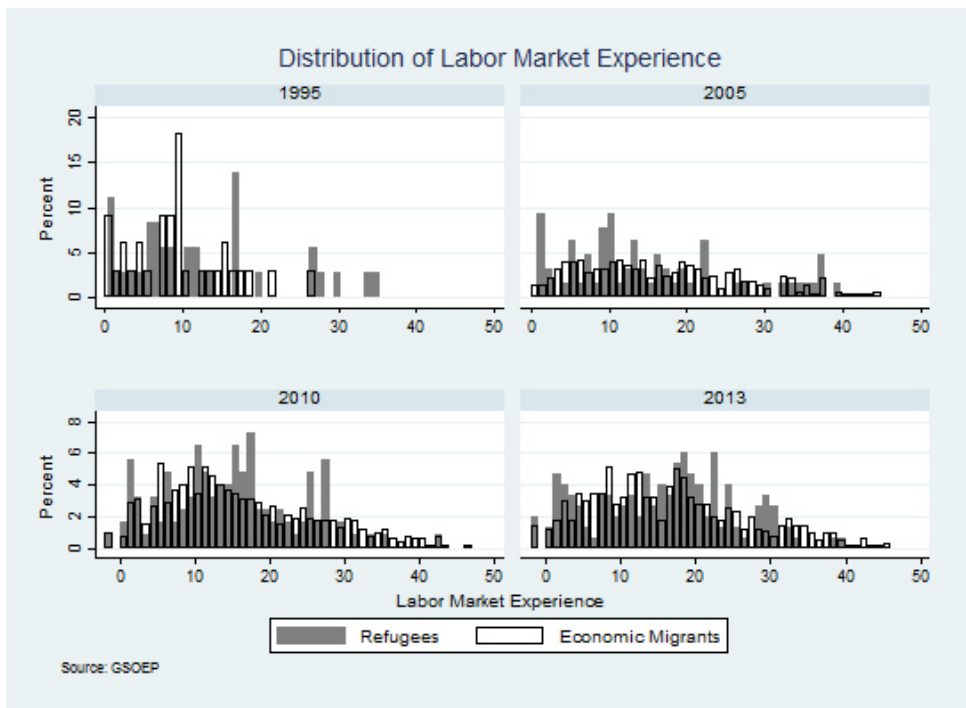
Figure 14: Histogram Years of Labor Market Experience
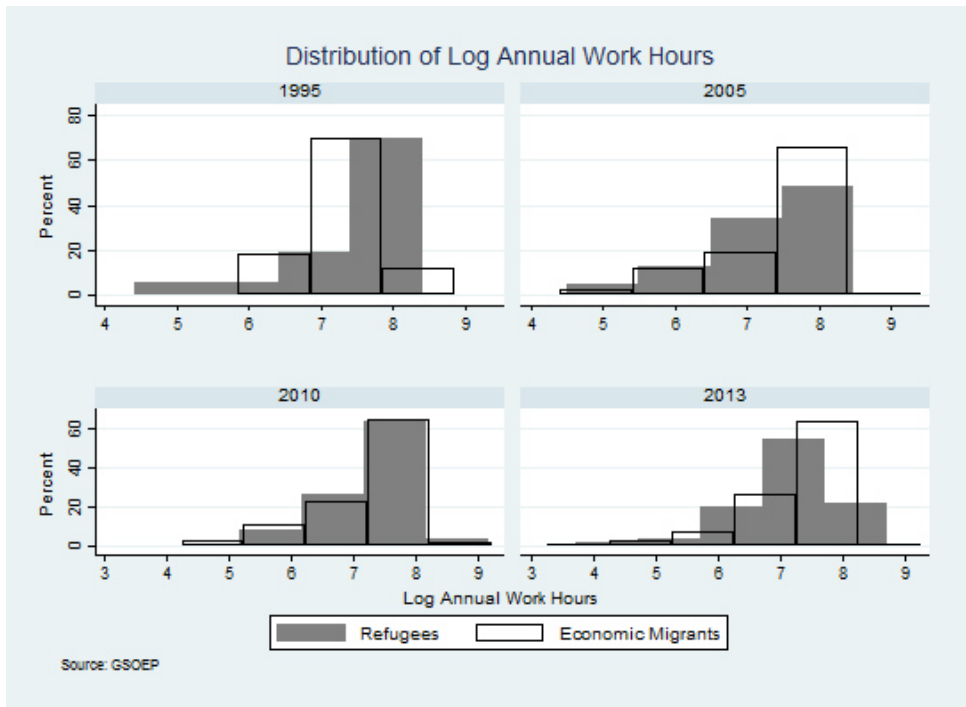
Figure 15: Histogram Log Annual Hours
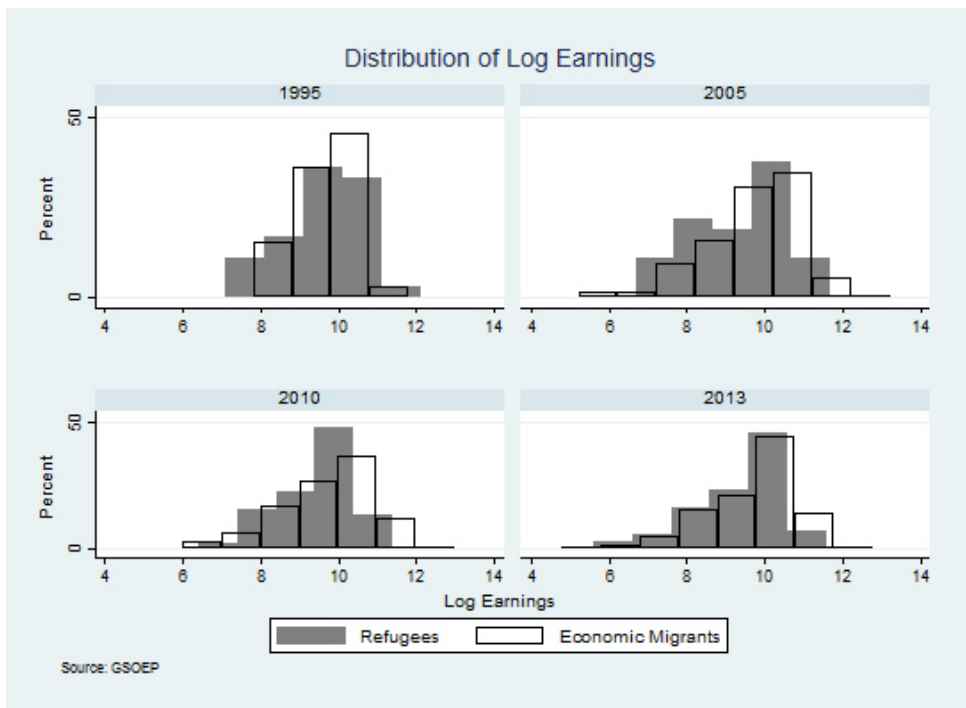
Figure 16: Histogram Log Annual Earnings
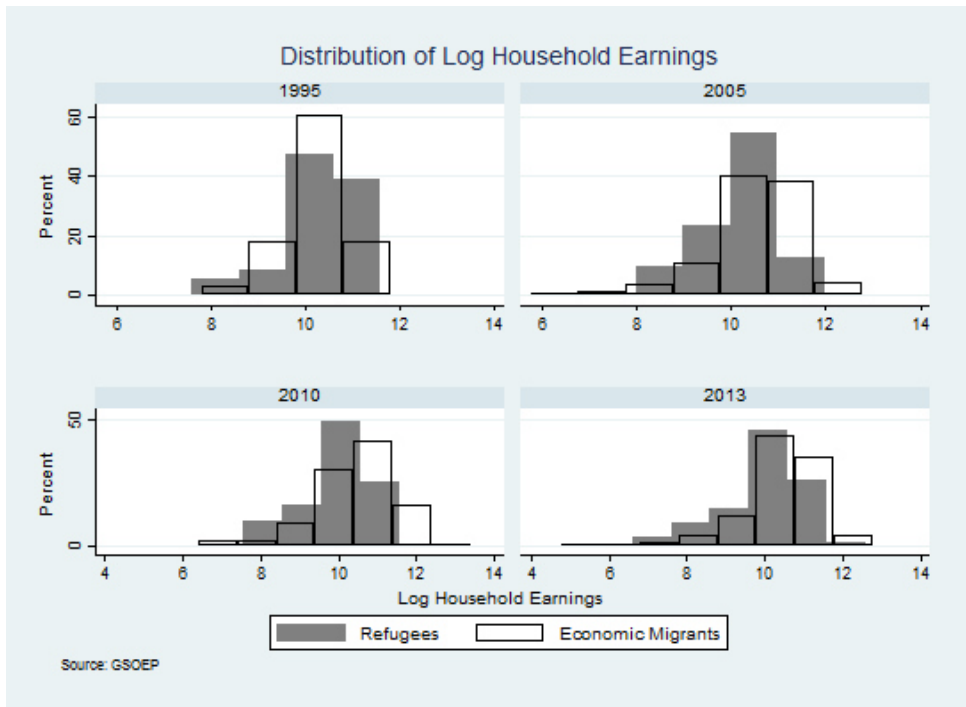
Figure 17: Histogram Log Annual Household Income
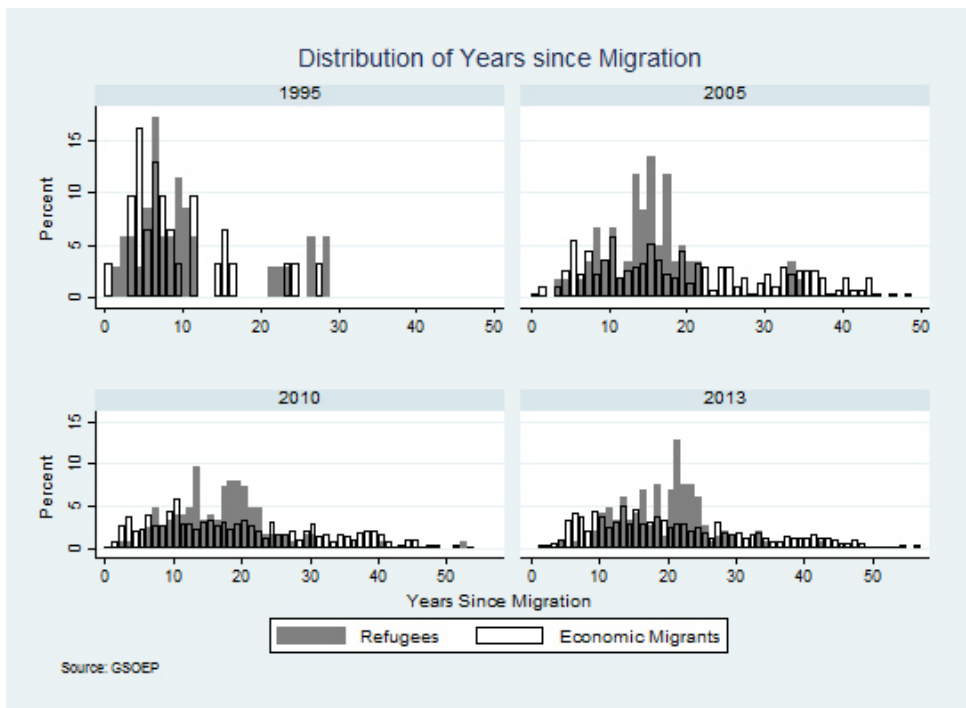
Figure 18: Histogram Years since Arrival
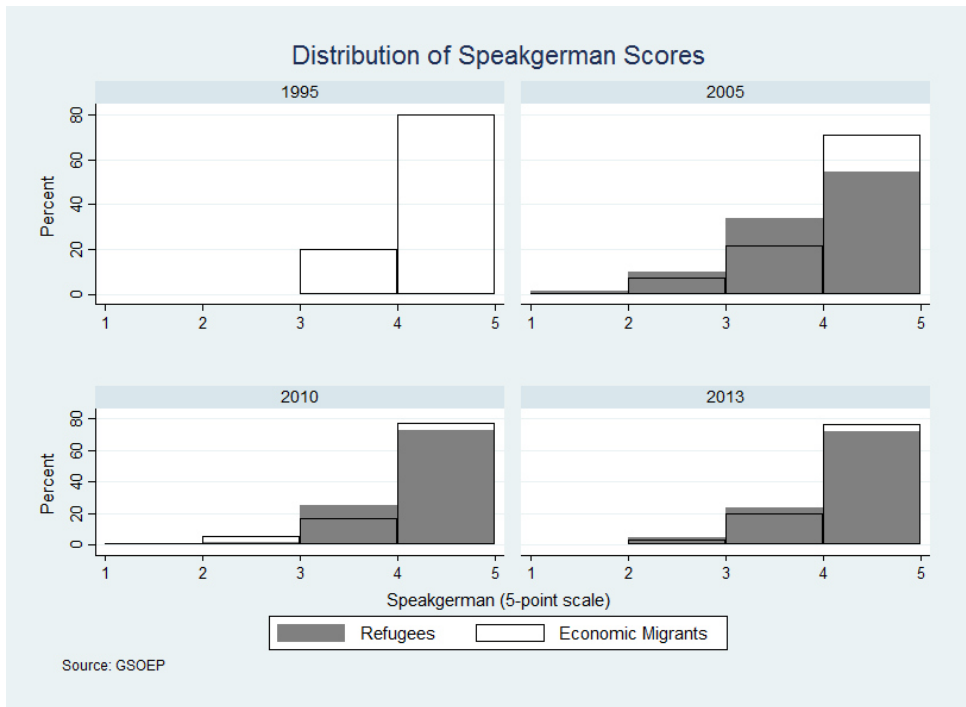
Figure 19: Histogram Spoken German Ability scores

A.5   List of variables

Table 24: List of variables in regressions

Variable	Description	Definition
Dependent Variables		
Log Annual Earnings	Natural Logarithm of Annual Earnings	log(i11110)
Log Hourly Wage	Natural Logarithm of Hourly Wage	log(i11110/e11110)
Log Annual Working Hours	Natural Logarithm of Annual Hours Worked	log(e11101)
Standard Controls		
Age	Age at survey year	SOEP variable
Age^2	Control for non-linear Age effects	Squared version of Age
Experience	Years of potential labour market experience	Sum of years in full-time and part-time employment
Experience^2	Control for non-linear Experience effects	Squared version of Experience
Education	Completed years of schooling	SOEP variable
Male	1 if male, 0 if female	
Married	1 if married, 0 if otherwise	SOEP variable
YSM	Years since migration	Survey Year-Immigration Year
Naturalized	1 if naturalized, 0 if otherwise	SOEP variable
Survey year	Dummy for every year i the sample	SOEP variable
Country of origin	Dummy for every country of origin	SOEP variable
Region of Residence	Dummy for every federal state of residence	SOEP variable
Additional Control Variables		
Partner Good German ability	1 if immigrant's partner speaks "Good" or "Very Good" German, 0 if otherwise	

Partner German Citizen	1 if immigrant's partner is a German citizen, 0 if otherwise	
Number of Persons in HH		SOEP variable
Number of Children in HH		SOEP variable
Instrumental Variables		
Father's Education*Survey Year	Years of Father's Education interacted with the year variable	
Age at Arrival*Survey Year	Age at Arrival interacted with the Year variable	
Variables of interest		
Good German ability	1 if immigrant speaks "Good" or "Very Good" German, 0 if otherwise	Combination of Own Opinion of Spoken German + Oral ability: German (copied down for years with no observations)
Good German writing	1 if immigrant writes "Good" or "Very Good" German, 0 if otherwise	Combination of Own Opinion of Written German + Written ability: German (copied down for years with no observations)
Refugee	1 if Immigrant is a Refugee, 0 if otherwise	SOEP variable

Interaction terms with Refugee get the prefix "Ref"*

A.6 Regression results

Table 25: Y=Log Annual Working Hours, pooled sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Refugee	-0.0816	-0.109*	-0.0964	0.0482	0.0388	0.172	0	0
	(0.0506)	(0.0537)	(0.0551)	(0.237)	(0.263)	(0.274)	(.)	(.)
Age	0.0196	0.0149	0.0311*	0.0172	0.0120	0.0284	-0.0513	-0.0553
	(0.0140)	(0.0154)	(0.0155)	(0.0142)	(0.0154)	(0.0155)	(0.0396)	(0.0396)
Age^2	-0.000436**	-0.000391*	-0.000599**	-0.000409*	-0.000362*	-0.000574**	-0.000189	-0.000130
	(0.000167)	(0.000182)	(0.000186)	(0.000167)	(0.000181)	(0.000184)	(0.000334)	(0.000335)
Male	0.497***	0.490***	0.500***	0.496***	0.488***	0.498***	0	0
	(0.0337)	(0.0367)	(0.0361)	(0.0334)	(0.0362)	(0.0356)	(.)	(.)
Experience	0.0600***	0.0632***	0.0609***	0.0624***	0.0661***	0.0636***	0.152***	0.153***
	(0.00604)	(0.00640)	(0.00629)	(0.00630)	(0.00657)	(0.00653)	(0.0246)	(0.0274)
$Experience^2$	-0.000773***	-0.000814***	-0.000773***	-0.000806***	-0.000852***	-0.000807***	-0.00152***	-0.00162***
	(0.000146)	(0.000156)	(0.000155)	(0.000147)	(0.000157)	(0.000156)	(0.000311)	(0.000321)
YSM	0.00177	0.00139	0.00254	0.00186	0.00173	0.00305	0	0
	(0.00162)	(0.00192)	(0.00193)	(0.00169)	(0.00202)	(0.00205)	(.)	(.)
Married	-0.129***	-0.133***	-0.0900*	-0.137***	-0.135***	-0.0950*	-0.0467	-0.0281
	(0.0300)	(0.0331)	(0.0386)	(0.0323)	(0.0354)	(0.0419)	(0.0577)	(0.0630)
Education	0.0295***	0.0299***	0.0268***	0.0315***	0.0330***	0.0307***	0	0
	(0.00624)	(0.00703)	(0.00698)	(0.00673)	(0.00760)	(0.00751)	(.)	(.)
Naturalized	0.0132	0.0183	0.0147	0.00187	0.00561	0.00621	-0.0553	-0.00193
	(0.0314)	(0.0349)	(0.0362)	(0.0344)	(0.0376)	(0.0387)	(0.0530)	(0.0583)
Good German ability		-0.0107	-0.00373		-0.0228	-0.0150	-0.000361	-0.0175
		(0.0328)	(0.0322)		(0.0376)	(0.0368)	(0.0335)	(0.0422)
Partner Good German ability			0.0505			0.0588	0.0675	0.0835
			(0.0308)			(0.0358)	(0.0461)	(0.0645)
Partner German Citizen			-0.000887			-0.00938	-0.0259	-0.00849
			(0.0326)			(0.0339)	(0.0552)	(0.0595)
Number of Persons in HH			-0.0257			-0.0254	-0.0106	-0.0141
			(0.0170)			(0.0192)	(0.0213)	(0.0226)
Number of Children in HH			-0.0443*			-0.0401	-0.0716*	-0.0922**
			(0.0197)			(0.0230)	(0.0281)	(0.0340)
Ref*Age				0.000369	0.00279	0.00196		0.00644
				(0.00688)	(0.00735)	(0.00749)		(0.0412)
Ref*Male				-0.00420	0.0296	0.0290		
				(0.0928)	(0.0986)	(0.104)		
Ref*Experience				-0.00562	-0.00776	-0.00631		-0.00390
				(0.00643)	(0.00700)	(0.00705)		(0.0434)
Ref*YSM				-0.00306	-0.00455	-0.00562		0
				(0.00549)	(0.00582)	(0.00589)		(.)
Ref*Married				0.0651	0.0251	0.0419		-0.121
				(0.0877)	(0.0968)	(0.110)		(0.147)
Ref*Education				-0.00675	-0.0127	-0.0199		0
				(0.0151)	(0.0168)	(0.0173)		(.)
Ref*Naturalized				0.0719	0.0687	0.0431		-0.204
				(0.0846)	(0.0950)	(0.104)		(0.137)
Ref*Good German ability					0.0594	0.0606		0.0805
					(0.0762)	(0.0750)		(0.0705)
Ref*Partner German Citizen						0.0474		-0.0243
						(0.0947)		(0.153)
Ref*Persons in HH						0.00200		0.00924
						(0.0415)		(0.0526)
Ref*Children in HH						-0.0220		0.0782
						(0.0466)		(0.0596)
Ref*Partner Good German ability						-0.0497		-0.0593
						(0.0761)		(0.0880)
Observations	6744	5493	5493	6744	5493	5493	5493	5493
Adjusted R^2	0.291	0.301	0.308	0.292	0.301	0.309	0.059	0.060

Standard Errors in Parentheses are clustered on individual level, Included years: 1994 to 2014, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

(4) $lnY_{it} = \beta_0 + \beta X_{it} + \delta[X_{it} * Ref_i] + \alpha_i + v_{it}$

(5) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \delta[(X_{it} + Good\ German\ ability_{it}) * Ref_i] + \alpha_i + v_{it}$

(6) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \delta[(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) * Ref_i] + \alpha_i + v_{it}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Refugee	-0.0771	-0.0949	-0.0694	0.193	0.286	0.511	0	0
	(0.0471)	(0.0493)	(0.0480)	(0.242)	(0.259)	(0.277)	(.)	(.)
Age	0.0537***	0.0484**	0.0513**	0.0587***	0.0528**	0.0570***	0.0859	0.0884
	(0.0155)	(0.0166)	(0.0169)	(0.0158)	(0.0167)	(0.0169)	(0.0521)	(0.0511)
Age^2	-0.000796***	-0.000743***	-0.000783***	-0.000800***	-0.000748***	-0.000810***	-0.00131*	-0.00134*
	(0.000174)	(0.000189)	(0.000195)	(0.000172)	(0.000188)	(0.000193)	(0.000555)	(0.000540)
Experience	0.0281***	0.0310***	0.0305***	0.0246**	0.0283***	0.0273**	0.0478	0.0471
	(0.00771)	(0.00817)	(0.00818)	(0.00810)	(0.00839)	(0.00843)	(0.0341)	(0.0359)
Experience^2	-0.000259	-0.000324	-0.000301	-0.000262	-0.000331	-0.000270	0.0000546	0.0000750
	(0.000166)	(0.000177)	(0.000178)	(0.000164)	(0.000176)	(0.000178)	(0.000546)	(0.000545)
YSM	0.00295	0.00316	0.00415*	0.00235	0.00255	0.00308	0	0
	(0.00155)	(0.00179)	(0.00182)	(0.00164)	(0.00186)	(0.00194)	(.)	(.)
Married	0.0595	0.0754*	0.0657	0.0612	0.0875*	0.0561	0.0303	-0.00211
	(0.0325)	(0.0350)	(0.0395)	(0.0353)	(0.0372)	(0.0433)	(0.0596)	(0.0478)
Education	0.0271***	0.0332***	0.0318***	0.0248***	0.0324***	0.0332***	0	0
	(0.00682)	(0.00733)	(0.00748)	(0.00750)	(0.00797)	(0.00797)	(.)	(.)
Naturalized	0.00929	0.0176	0.00165	0.0247	0.0455	0.0353	-0.122*	-0.0499
	(0.0323)	(0.0368)	(0.0383)	(0.0349)	(0.0383)	(0.0395)	(0.0501)	(0.0386)
Good German ability		0.0713*	0.0742*		0.0797*	0.0870*	0.000364	-0.0411
		(0.0298)	(0.0297)		(0.0352)	(0.0351)	(0.0378)	(0.0412)
Partner Good German ability			0.0681*			0.0637*	0.0346	0.0192
			(0.0277)			(0.0314)	(0.0491)	(0.0546)
Partner German Citizen			0.0300			0.0223	0.0814	0.138**
			(0.0303)			(0.0309)	(0.0525)	(0.0448)
Number of Persons in HH			-0.0287			-0.0232	-0.00563	-0.0182
			(0.0179)			(0.0209)	(0.0267)	(0.0268)
Number of Children in HH			0.0153			0.0302	-0.00739	0.0146
			(0.0196)			(0.0234)	(0.0287)	(0.0349)
Ref*Age				-0.0140	-0.0110	-0.00982		0.00711
				(0.00833)	(0.00858)	(0.00865)		(0.0474)
Ref*Experience				0.0121	0.00920	0.00598		-0.00532
				(0.00729)	(0.00778)	(0.00778)		(0.0498)
Ref*YSM				0.00278	0.00151	0.00344		0
				(0.00450)	(0.00449)	(0.00436)		(.)
Ref*Married				0.0132	-0.0377	0.0487		0.0844
				(0.0991)	(0.102)	(0.104)		(0.210)
Ref*Education				0.00523	-0.00314	-0.0155		0
				(0.0189)	(0.0202)	(0.0218)		(.)
Ref*Naturalized				-0.0603	-0.103	-0.134		-0.195
				(0.0799)	(0.0909)	(0.101)		(0.141)
Ref*Good German ability					-0.0353	-0.0509		0.114
					(0.0682)	(0.0662)		(0.0795)
Ref*Partner German Citizen						0.0153		-0.144
						(0.0801)		(0.146)
Ref*Persons in HH						-0.0170		0.0215
						(0.0448)		(0.0646)
Ref*Children in HH						-0.0635		-0.0411
						(0.0494)		(0.0594)
Ref*Partner Good German ability						0.00187		0.00792
						(0.0634)		(0.104)
Observations	3618	2910	2910	3618	2910	2910	2910	2910
Adjusted R^2	0.165	0.187	0.191	0.166	0.188	0.196	0.066	0.068

Standard Errors in Parentheses are clustered on individual level, Included years: 1994 to 2014, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

(4) $lnY_{it} = \beta_0 + \beta X_{it} + \delta[X_{it} * Ref_i] + \alpha_i + v_{it}$

(5) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \delta[(X_{it} + Good\ German\ ability_{it}) * Ref_i] + \alpha_i + v_{it}$

(6) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \delta[(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) * Ref_i] + \alpha_i + v_{it}$

Table 27: Y=Log Annual Working Hours, female sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Refugee	-0.0288	-0.0605	-0.0869	-0.371	-0.400	-0.0375	0	0
	(0.0920)	(0.0944)	(0.0962)	(0.523)	(0.582)	(0.597)	(.)	(.)
Age	0.00430	0.000130	0.0393	0.00110	-0.00134	0.0405	-0.0197	-0.112**
	(0.0205)	(0.0224)	(0.0227)	(0.0204)	(0.0224)	(0.0227)	(0.0463)	(0.0428)
Age^2	-0.000250	-0.000220	-0.000716**	-0.000239	-0.000231	-0.000756**	-0.000494	-0.000287
	(0.000237)	(0.000258)	(0.000265)	(0.000236)	(0.000257)	(0.000263)	(0.000446)	(0.000416)
Experience	0.0707***	0.0728***	0.0635***	0.0730***	0.0749***	0.0646***	0.147***	0.211***
	(0.00939)	(0.00982)	(0.00958)	(0.00944)	(0.00996)	(0.00980)	(0.0424)	(0.0379)
$Experience^2$	-0.000868***	-0.000861***	-0.000705**	-0.000886***	-0.000872***	-0.000693**	-0.00161***	-0.00196***
	(0.000221)	(0.000231)	(0.000225)	(0.000222)	(0.000235)	(0.000230)	(0.000411)	(0.000419)
YSM	0.000107	0.000736	0.00209	0.00186	0.00267	0.00425	0	0
	(0.00309)	(0.00359)	(0.00360)	(0.00312)	(0.00367)	(0.00369)	(.)	(.)
Married	-0.272***	-0.268***	-0.199**	-0.281***	-0.270***	-0.194**	-0.166	-0.0550
	(0.0486)	(0.0522)	(0.0649)	(0.0516)	(0.0557)	(0.0691)	(0.0851)	(0.0883)
Education	0.0317**	0.0318**	0.0268*	0.0339**	0.0342**	0.0302**	0	0
	(0.00987)	(0.0108)	(0.0107)	(0.0103)	(0.0116)	(0.0115)	(.)	(.)
Naturalized	-0.0264	-0.0244	-0.000514	-0.0542	-0.0631	-0.0411	-0.0388	0.0454
	(0.0553)	(0.0607)	(0.0628)	(0.0566)	(0.0623)	(0.0652)	(0.105)	(0.116)
Good German ability		-0.117	-0.0922		-0.113	-0.0921		-0.00306
		(0.0635)	(0.0615)		(0.0673)	(0.0647)		(0.0716)
Partner Good German ability			0.0373			0.0339		0.122
			(0.0561)			(0.0636)		(0.128)
Partner German Citizen			-0.0299			-0.0400		-0.134
			(0.0620)			(0.0639)		(0.100)
Number of Persons in HH			0.00306			-0.00335		-0.0417
			(0.0282)			(0.0299)		(0.0379)
Number of Children in HH			-0.152***			-0.142***		-0.195***
			(0.0357)			(0.0404)		(0.0504)
Ref*Age				0.0141	0.0151	0.0120		0.0197
				(0.0120)	(0.0128)	(0.0124)		(0.0603)
Ref*Experience				-0.0126	-0.0130	-0.0109		-0.0158
				(0.0144)	(0.0158)	(0.0154)		(0.0665)
Ref*YSM				-0.0143	-0.0154	-0.0205		0
				(0.0118)	(0.0117)	(0.0119)		(.)
Ref*Married				0.0738	0.00539	-0.00748		-0.353*
				(0.166)	(0.182)	(0.194)		(0.164)
Ref*Education				0.00412	0.00747	-0.00717		0
				(0.0245)	(0.0280)	(0.0290)		(.)
Ref*Naturalized				0.245	0.303	0.317		-0.0508
				(0.169)	(0.179)	(0.186)		(0.230)
Ref*Good German ability					-0.0374	-0.0114		0.0998
					(0.169)	(0.161)		(0.132)
Ref*Partner German Citizen						0.113		-0.0280
						(0.167)		(0.215)
Ref*Persons in HH						-0.0179		0.0453
						(0.0771)		(0.0816)
Ref*Children in HH						-0.0628		0.111
						(0.0967)		(0.111)
Ref*Partner Good German ability						0.108		-0.140
						(0.148)		(0.159)
Observations	3126	2583	2583	3126	2583	2583	3126	2583
Adjusted R^2	0.219	0.233	0.255	0.221	0.236	0.259	0.087	0.094

Standard Errors in Parentheses are clustered on individual level, Included years: 1994 to 2014, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

(4) $lnY_{it} = \beta_0 + \beta X_{it} + \delta[X_{it} * Ref_i] + \alpha_i + v_{it}$

(5) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \delta[(X_{it} + Good\ German\ ability_{it}) * Ref_i] + \alpha_i + v_{it}$

(6) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \delta[(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) * Ref_i] + \alpha_i + v_{it}$

Table 28: Y=Log Hourly Wage, pooled sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Refugee	-0.0637	-0.0830	-0.0801	0.409	0.272	0.287	0	0
	(0.0459)	(0.0474)	(0.0479)	(0.220)	(0.241)	(0.264)	(.)	(.)
Age	0.0359**	0.0371**	0.0361**	0.0399***	0.0403**	0.0397**	0.0254	0.0193
	(0.0118)	(0.0124)	(0.0129)	(0.0119)	(0.0125)	(0.0130)	(0.0227)	(0.0228)
Age^2	-0.000479***	-0.000469**	-0.000455**	-0.000508***	-0.000498***	-0.000490**	-0.000180	-0.000125
	(0.000143)	(0.000149)	(0.000156)	(0.000144)	(0.000150)	(0.000157)	(0.000246)	(0.000257)
Male	0.312***	0.318***	0.321***	0.311***	0.317***	0.321***	0	0
	(0.0338)	(0.0370)	(0.0375)	(0.0340)	(0.0371)	(0.0377)	(.)	(.)
Experience	0.0150**	0.0101	0.00955	0.0144*	0.00958	0.00900	0.0235	0.0298
	(0.00536)	(0.00560)	(0.00558)	(0.00575)	(0.00600)	(0.00597)	(0.0162)	(0.0170)
Experience^2	-0.0000793	0.0000326	0.0000422	-0.0000794	0.0000335	0.0000470	-0.000142	-0.000236
	(0.000130)	(0.000134)	(0.000134)	(0.000133)	(0.000137)	(0.000136)	(0.000254)	(0.000271)
YSM	0.00939***	0.00811***	0.00807***	0.00945***	0.00865***	0.00851***	0	0
	(0.00170)	(0.00195)	(0.00193)	(0.00180)	(0.00209)	(0.00208)	(.)	(.)
Married	0.107***	0.119***	0.138***	0.136***	0.142***	0.172***	0.0790	0.0779
	(0.0299)	(0.0324)	(0.0354)	(0.0332)	(0.0364)	(0.0398)	(0.0413)	(0.0431)
Education	0.0745***	0.0686***	0.0690***	0.0756***	0.0700***	0.0710***	0	0
	(0.00632)	(0.00679)	(0.00692)	(0.00703)	(0.00759)	(0.00765)	(.)	(.)
Naturalized	0.0173	0.0278	0.0300	0.0114	0.0300	0.0378	0.0116	0.0332
	(0.0382)	(0.0415)	(0.0417)	(0.0424)	(0.0454)	(0.0461)	(0.0468)	(0.0522)
Good German ability		0.123***	0.123***		0.100**	0.101**	0.00387	-0.00562
		(0.0316)	(0.0315)		(0.0360)	(0.0357)	(0.0363)	(0.0408)
Partner Good German ability			-0.00950			-0.0173	0.0360	-0.0371
			(0.0327)			(0.0384)	(0.0443)	(0.0522)
Partner German Citizen			-0.0292			-0.0499	-0.0164	0.00339
			(0.0318)			(0.0356)	(0.0420)	(0.0433)
Number of Persons in HH			-0.0200			-0.0202	-0.0151	-0.00930
			(0.0158)			(0.0186)	(0.0157)	(0.0164)
Number of Children in HH			0.0234			0.0231	0.0515*	0.0300
			(0.0194)			(0.0224)	(0.0204)	(0.0224)
Ref*Age				-0.00499	-0.00260	-0.00220		0.00192
				(0.00522)	(0.00561)	(0.00566)		(0.0302)
Ref*Male				-0.267**	-0.280**	-0.324**	0	0
				(0.0815)	(0.0864)	(0.0988)	(.)	(.)
Ref*Experience				0.00348	0.00308	0.00270		-0.0166
				(0.00533)	(0.00566)	(0.00562)		(0.0322)
Ref*YSM				-0.00567	-0.00675	-0.00616		0
				(0.00464)	(0.00512)	(0.00502)		(.)
Ref*Married				-0.167**	-0.132	-0.161		0.0561
				(0.0633)	(0.0701)	(0.0862)		(0.114)
Ref*Education				-0.00966	-0.0117	-0.0150		0
				(0.0132)	(0.0141)	(0.0152)		(.)
Ref*Naturalized				0.0246	-0.0210	-0.0476		-0.0145
				(0.0889)	(0.0972)	(0.0967)		(0.114)
Ref*Good German ability					0.0838	0.0801		0.0174
					(0.0768)	(0.0779)		(0.0806)
Ref*Partner German Citizen						0.0819		-0.0785
						(0.0725)		(0.111)
Ref*Persons in HH						0.00162		-0.0217
						(0.0312)		(0.0416)
Ref*Children in HH						-0.00546		0.0786
						(0.0397)		(0.0480)
Ref*Partner Good German ability						0.0139		0.208*
						(0.0733)		(0.0935)
Observations	6744	5493	5493	6744	5493	5493	5493	5493
Adjusted R^2	0.288	0.299	0.299	0.290	0.300	0.300	0.036	0.039

Standard Errors in Parentheses are clustered on individual level, Included years: 1994 to 2014, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

(4) $lnY_{it} = \beta_0 + \beta X_{it} + \delta[X_{it} * Ref_i] + \alpha_i + v_{it}$

(5) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \delta[(X_{it} + Good\ German\ ability_{it}) * Ref_i] + \alpha_i + v_{it}$

(6) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \delta[(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) * Ref_i] + \alpha_i + v_{it}$

Table 29: Y=Log Hourly Wage, male sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Refugee	-0.0883	-0.0853	-0.0709	0.322	0.210	0.0851	0	0
	(0.0530)	(0.0525)	(0.0552)	(0.280)	(0.303)	(0.318)	(.)	(.)
Age	0.00606	0.00880	0.00837	0.0174	0.0195	0.0194	-0.0526	-0.0688
	(0.0171)	(0.0185)	(0.0187)	(0.0175)	(0.0189)	(0.0192)	(0.0393)	(0.0393)
Age^2	-0.000162	-0.000194	-0.000192	-0.000218	-0.000238	-0.000244	0.000773	0.000858
	(0.000200)	(0.000217)	(0.000222)	(0.000201)	(0.000220)	(0.000226)	(0.000434)	(0.000466)
Experience	0.0227*	0.0202*	0.0175	0.0157	0.0129	0.0100	0.0637*	0.0870**
	(0.00897)	(0.00944)	(0.00936)	(0.00987)	(0.0104)	(0.0103)	(0.0290)	(0.0283)
$Experience^2$	-0.000237	-0.000147	-0.0000653	-0.000211	-0.000135	-0.0000444	-0.000857	-0.00103
	(0.000197)	(0.000208)	(0.000207)	(0.000200)	(0.000212)	(0.000213)	(0.000477)	(0.000540)
YSM	0.00834***	0.00833**	0.00872***	0.00775***	0.00823**	0.00850**	0	0
	(0.00217)	(0.00257)	(0.00254)	(0.00232)	(0.00279)	(0.00276)	(.)	(.)
Married	0.189***	0.203***	0.211***	0.230***	0.234***	0.260***	0.0730	0.0827
	(0.0418)	(0.0451)	(0.0501)	(0.0471)	(0.0513)	(0.0579)	(0.0655)	(0.0637)
Education	0.0898***	0.0927***	0.0950***	0.0849***	0.0863***	0.0882***	0	0
	(0.00893)	(0.0101)	(0.0103)	(0.0101)	(0.0116)	(0.0119)	(.)	(.)
Naturalized	0.00962	0.00933	0.0196	0.0332	0.0492	0.0657	0.00543	0.0377
	(0.0479)	(0.0533)	(0.0521)	(0.0515)	(0.0563)	(0.0576)	(0.0570)	(0.0572)
Good German ability		0.168***	0.171***		0.151**	0.155**	-0.0324	-0.0359
		(0.0417)	(0.0412)		(0.0517)	(0.0506)	(0.0460)	(0.0490)
Partner Good German ability			0.0324			0.0161	0.0736	-0.0450
			(0.0378)			(0.0447)	(0.0616)	(0.0607)
Partner German Citizen			-0.0763			-0.0870*	0.00739	0.0261
			(0.0396)			(0.0438)	(0.0547)	(0.0476)
Number of Persons in HH			-0.0461*			-0.0537*	-0.000830	0.00893
			(0.0195)			(0.0242)	(0.0194)	(0.0204)
Number of Children in HH			0.0598*			0.0627*	0.0641*	0.0316
			(0.0250)			(0.0302)	(0.0261)	(0.0292)
Ref*Age				-0.0180*	-0.0191*	-0.0177		0.0347
				(0.00878)	(0.00946)	(0.00939)		(0.0395)
Ref*Experience				0.0195*	0.0210*	0.0205*		-0.0592
				(0.00868)	(0.00911)	(0.00913)		(0.0405)
Ref*YSM				0.00295	0.00142	0.00215		0
				(0.00542)	(0.00584)	(0.00586)		(.)
Ref*Married				-0.196*	-0.137	-0.189		0.0921
				(0.0846)	(0.0922)	(0.104)		(0.165)
Ref*Education				0.0109	0.0199	0.0234		0
				(0.0218)	(0.0232)	(0.0242)		(.)
Ref*Naturalized				-0.0906	-0.145	-0.163		-0.0443
				(0.101)	(0.108)	(0.105)		(0.142)
Ref*Good German ability					0.0440	0.0329		0.0136
					(0.0824)	(0.0821)		(0.0946)
Ref*Partner German Citizen						0.0319		-0.0410
						(0.0858)		(0.141)
Ref*Persons in HH						0.0202		-0.0184
						(0.0367)		(0.0459)
Ref*Children in HH						-0.00292		0.0853
						(0.0457)		(0.0556)
Ref*Partner Good German ability						0.0292		0.310*
						(0.0752)		(0.135)
Observations	3618	2910	2910	3618	2910	2910	2910	2910
Adjusted R^2	0.345	0.373	0.378	0.349	0.378	0.382	0.048	0.055

Standard Errors in Parentheses are clustered on individual level, Included years: 1994 to 2014, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

(4) $lnY_{it} = \beta_0 + \beta X_{it} + \delta[X_{it} * Ref_i] + \alpha_i + v_{it}$

(5) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \delta[(X_{it} + Good\ German\ ability_{it}) * Ref_i] + \alpha_i + v_{it}$

(6) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \delta[(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) * Ref_i] + \alpha_i + v_{it}$

Table 30: Y=Log Hourly Wage, female sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Refugee	0.0136	-0.0146	-0.0184	0.788*	0.910*	1.106**	0	0
	(0.0747)	(0.0791)	(0.0785)	(0.393)	(0.435)	(0.418)	(.)	(.)
Age	0.0597***	0.0577**	0.0581**	0.0670***	0.0662***	0.0690***	-0.0197	0.0790**
	(0.0169)	(0.0179)	(0.0189)	(0.0164)	(0.0174)	(0.0183)	(0.0463)	(0.0268)
Age^2	-0.000755***	-0.000713***	-0.000716**	-0.000825***	-0.000794***	-0.000825***	-0.000494	-0.000768*
	(0.000205)	(0.000215)	(0.000227)	(0.000201)	(0.000211)	(0.000222)	(0.000446)	(0.000318)
Experience	0.0153*	0.0115	0.0113	0.0130	0.00740	0.00644	0.147***	-0.00178
	(0.00756)	(0.00784)	(0.00807)	(0.00791)	(0.00820)	(0.00849)	(0.0424)	(0.0226)
Experience^2	-0.0000638	0.0000214	0.0000265	-0.0000239	0.0000915	0.000115	-0.00161***	0.000175
	(0.000193)	(0.000195)	(0.000199)	(0.000196)	(0.000199)	(0.000205)	(0.000411)	(0.000304)
YSM	0.0124***	0.0110***	0.0109***	0.0135***	0.0122***	0.0118***	0	0
	(0.00265)	(0.00299)	(0.00293)	(0.00275)	(0.00314)	(0.00308)	(.)	(.)
Married	0.0343	0.0511	0.0377	0.0384	0.0494	0.0396	-0.166	0.0721
	(0.0423)	(0.0467)	(0.0482)	(0.0450)	(0.0501)	(0.0528)	(0.0851)	(0.0557)
Education	0.0591***	0.0520***	0.0504***	0.0620***	0.0565***	0.0554***	0	0
	(0.00908)	(0.00940)	(0.00946)	(0.00989)	(0.0102)	(0.0102)	(.)	(.)
Naturalized	-0.0497	-0.0188	-0.0289	-0.0861	-0.0598	-0.0667	-0.0388	0.0160
	(0.0597)	(0.0637)	(0.0649)	(0.0644)	(0.0689)	(0.0699)	(0.105)	(0.0883)
Good German ability		0.0477	0.0434		0.0427	0.0427		0.0383
		(0.0492)	(0.0488)		(0.0517)	(0.0514)		(0.0641)
Partner Good German ability			-0.0579			-0.0683		-0.0129
			(0.0535)			(0.0618)		(0.0889)
Partner German Citizen			0.0628			0.0495		-0.0376
			(0.0532)			(0.0580)		(0.0685)
Number of Persons in HH			0.00893			0.0136		-0.0253
			(0.0230)			(0.0241)		(0.0253)
Number of Children in HH			-0.00867			-0.0161		0.00624
			(0.0298)			(0.0312)		(0.0344)
Ref*Age				-0.0106	-0.0123	-0.0135		-0.0765*
				(0.00907)	(0.00982)	(0.00984)		(0.0345)
Ref*Experience				0.0116	0.0181	0.0178		0.0822*
				(0.0105)	(0.0114)	(0.0113)		(0.0398)
Ref*YSM				-0.0233**	-0.0228**	-0.0221**		0
				(0.00743)	(0.00841)	(0.00844)		(.)
Ref*Married				-0.0202	-0.0265	-0.0251		0.0819
				(0.103)	(0.112)	(0.126)		(0.140)
Ref*Education				-0.0128	-0.0273	-0.0358		0
				(0.0184)	(0.0190)	(0.0199)		(.)
Ref*Naturalized				0.274	0.240	0.239		0.0270
				(0.140)	(0.154)	(0.162)		(0.164)
Ref*Good German ability					0.0227	0.00822		-0.00794
					(0.137)	(0.138)		(0.146)
Ref*Partner German Citizen						0.105		-0.108
						(0.121)		(0.156)
Ref*Persons in HH						-0.0327		-0.0363
						(0.0596)		(0.0743)
Ref*Children in HH						-0.0208		0.0865
						(0.0769)		(0.0743)
Ref*Partner Good German ability						0.120		0.0622
						(0.116)		(0.122)
Observations	3126	2583	2583	3126	2583	2583	3126	2583
Adjusted R^2	0.208	0.202	0.204	0.214	0.208	0.209	0.087	0.038

Standard Errors in Parentheses are clustered on individual level, Included years: 1994 to 2014, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(1) $lnY_{it} = \beta_0 + \beta X_{it} + \alpha_i + v_{it}$

(2) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \alpha_i + v_{it}$

(3) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \alpha_i + v_{it}$

(4) $lnY_{it} = \beta_0 + \beta X_{it} + \delta[X_{it} * Ref_i] + \alpha_i + v_{it}$

(5) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it}) + \delta[(X_{it} + Good\ German\ ability_{it}) * Ref_i] + \alpha_i + v_{it}$

(6) $lnY_{it} = \beta_0 + \beta(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) + \delta[(X_{it} + Good\ German\ ability_{it} + Partner_{it} + HH_{it}) * Ref_i] + \alpha_i + v_{it}$

71

A.7 List of variables in logit

Table 31: Variables of interest and control variables in logit analysis and expected effects

Variables of interest		
Refugee		
Ref*YSM		

Control Variables	Expected positive effect	Expected negative effect
Male	Family Investment model predicts that households should make investments into man's human capital as his earnings stream will not be interrupted by fertility	
Age at arrival		Ability to acquire language proficiency "effortless" declines with age
Experience	Experience in the labour market is associated with larger exposure to language	
YSM	The probability that a person acquires language proficiency is positively correlated with the duration of exposure	
Married	If married to a native of destination who takes on the "teacher role"	If married to a native of origin or to a native of destination, who takes on the "translator role"
Education	Higher educated are more efficient in learning new concepts and terminology.	
Naturalized	Minimum level of proficiency required for naturalization	
Partner Good German ability	If partner takes on "teacher role"	If partner takes on "translater role"
Partner German Citizen	If partner takes on "teacher role"	If partner takes on "translater role"

A.8 Integration courses in Germany

Table 32: Access to integration courses per immigrant title

Group	**Can** participate in integration course	**Must** participate in integration course	**Not legally entitled** to participate in integration course
Received residence title before the 1.1.2005	• If places are still available	• If receiving unemployment benefit II and the office which pays unemployment benefit II requires person to attend • If particular integration needs are present and the local immigration office requires person to attend.	

Received residence title after the 1.1.2005	• If in Germany:
	  – as an employee
	  – for the purposes of subsequent immigration of dependent family member
	  – as a long-term resident in another EU member state
	• If staying permanently in Germany and have for the first time received a settlement permit in accordance with section 23 subs. 2 (Residence Act)
		• If person cannot make herself understood in German at a simple, adequate level
		• If receiving unemployment benefit II and the office which pays unemployment benefit II requires person to attend
			• If person is a child, young person or young adult attending school in Germany,
			• If there is little recognized need for integration,
			• If person speaks German to an adequate level (possible to attend only orientation course).

Ethnic German Resettler	• When admitted into Germany after 1.1.2005. – (Spouse and children are also entitled to participate) • When admitted into Germany before 1 January 2005 and have not yet attended a language course organized by the Federal Agency for Employment (SGB III course)		
EU citizens, German Citizens	• If level of German is not adequate, • If present particular integration needs • If there places available		

A.9 Robustness test

Table 33: Controlling Writing Ability, Y=Log Annual Earnings, pooled, male and female sample

	Pooled		Male		Female	
	(3)	(6)	(3)	(6)	(3)	(6)
Refugee	-0.173*	0.858	-0.135	0.504	-0.104	2.377**
	(0.0875)	(0.506)	(0.0822)	(0.499)	(0.151)	(0.897)
Age	0.0683**	0.0300	0.0602*	0.0280	0.0991**	0.0834*
	(0.0229)	(0.0264)	(0.0288)	(0.0336)	(0.0329)	(0.0363)
Age^2	-0.00106***	-0.000587	-0.000975**	-0.000529	-0.00144***	-0.00122**
	(0.000279)	(0.000312)	(0.000329)	(0.000383)	(0.000393)	(0.000422)
Male	0.825***	0.991***				
	(0.0602)	(0.0721)				
Experience	0.0699***	0.0660***	0.0485**	0.0338*	0.0732***	0.0659***
	(0.00943)	(0.0109)	(0.0149)	(0.0172)	(0.0135)	(0.0154)
$Experience^2$	-0.000714**	-0.000734**	-0.000366	-0.000196	-0.000642	-0.000634
	(0.000240)	(0.000261)	(0.000322)	(0.000362)	(0.000345)	(0.000389)
YSM	0.00963**	0.0113**	0.0119***	0.00923*	0.0122*	0.0136*
	(0.00321)	(0.00382)	(0.00350)	(0.00414)	(0.00530)	(0.00652)
Married	0.0483	0.0622	0.274***	0.198	-0.157	-0.0999
	(0.0602)	(0.0930)	(0.0707)	(0.112)	(0.0918)	(0.139)
Education	0.0931***	0.102***	0.124***	0.115***	0.0742***	0.0966***
	(0.0115)	(0.0133)	(0.0147)	(0.0164)	(0.0168)	(0.0212)
Naturalized	0.0404	0.115	0.0173	0.171*	-0.0335	-0.00672
	(0.0629)	(0.0701)	(0.0745)	(0.0779)	(0.101)	(0.113)
Good German ability	0.0635	0.0457	0.194**	0.173*	-0.0992	-0.0188
	(0.0568)	(0.0700)	(0.0635)	(0.0818)	(0.0960)	(0.110)
Good German writing	0.101	0.109	0.0902	0.172*	0.0947	0.00213
	(0.0517)	(0.0627)	(0.0574)	(0.0733)	(0.0848)	(0.104)
Partner Good German ability	0.0469	0.0740	0.113	0.0790	-0.0285	-0.0527
	(0.0739)	(0.0922)	(0.0831)	(0.0935)	(0.0874)	(0.142)
Partner Good German writing	-0.00991	-0.0940	-0.0142	-0.0000903		-0.151
	(0.0756)	(0.0936)	(0.0897)	(0.0982)		(0.147)
Partner German Citizen	-0.0277	-0.155*	-0.0444	-0.139*	0.0353	-0.196
	(0.0522)	(0.0623)	(0.0579)	(0.0629)	(0.0933)	(0.123)
Number of Persons in HH	-0.0436	-0.0443	-0.0741*	-0.0371	0.0156	-0.00829
	(0.0263)	(0.0371)	(0.0296)	(0.0398)	(0.0397)	(0.0513)
Number of Children in HH	-0.0211	0.00628	0.0763*	0.0825	-0.165***	-0.128*
	(0.0307)	(0.0419)	(0.0346)	(0.0449)	(0.0469)	(0.0627)
Ref*Age		-0.0108		-0.0384*		-0.0102
		(0.0115)		(0.0154)		(0.0183)
Ref*Male		-0.298				
		(0.176)				
Ref*Experience		-0.00102		0.0289		0.00672
		(0.0110)		(0.0154)		(0.0217)
Ref*YSM		-0.0138		0.00899		-0.0514**
		(0.0105)		(0.00819)		(0.0191)
Ref*Married		-0.241		-0.198		-0.0806
		(0.189)		(0.175)		(0.322)
Ref*Education		-0.0377		0.0369		-0.101
		(0.0312)		(0.0381)		(0.0551)
Ref*Naturalized		-0.0229		-0.268		0.443
		(0.186)		(0.177)		(0.281)
Ref*Partner German Citizen		0.152		0.0696		0.299
		(0.151)		(0.139)		(0.267)
Ref*Persons in HH		0.0739		0.0784		-0.126
		(0.0727)		(0.0676)		(0.140)
Ref*Children in HH		-0.108		-0.125		-0.0328

		(0.0721)		(0.0714)		(0.144)
Ref* Good German writing		0.0619		-0.237		0.609*
		(0.160)		(0.131)		(0.266)
Ref*Good German ability		0.119		0.114		-0.339
		(0.147)		(0.137)		(0.272)
Ref*Partner Good German ability		-0.156		0.0256		-0.0584
		(0.160)		(0.166)		(0.226)
Ref* Partner Good German writing		0.106		-0.0829		0.427*
		(0.160)		(0.178)		(0.215)
Observations	5491	4502	2908	2516	2583	1986
Adjusted R^2	0.405	0.429	0.395	0.400	0.300	0.305

Standard Errors in Parentheses are clustered on individual level, Included years: 1994 to 2014

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$