STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic Year 2016-17

A Day Late and a Krona Short
Age at Immigration and Labour Market Integration

Weronika Brzuchalska (22798)
Charlie Ekberg Ehnlund (22131)

Abstract

The paper examines how age at immigration affects the labour market integration of refugees in Sweden. Four different outcomes are studied: educational attainment, employment, wage, and welfare benefits. The individual data from the register-based longitudinal database GEOSweden developed by Statistics Sweden is used in this study. In order to ensure that the studied relationship is causal, the results are compared within siblings by family fixed effects. The results show that the increase in the age at immigration has negative impact on the educational attainment, employment, wage and increases the received welfare benefits of an individual.

Keywords: Economic Integration, Family Fixed Effects, Refugees, Immigration, Labour Economics
JEL: J15, J60, F22

Supervisors: Kelly Ragan and Matz Dahlberg
Date submitted: January 2, 2017
Date examined: January 10, 2017
Examiner: Mark Sanctuary
Discussants: Adrian Rydén and Ingvar Ziemman

Acknowledgements

Prima Facie, the authors gratefully acknowledge grant support from Insitutet för Bostads och Urbanforskning in Uppsala. This thesis would never see the daylight without the great support of Professor Matz Dahlberg. The authors are also grateful for advice and for the continuous support of the supervisor Professor Kelly Ragan. Thank you, Tino Sanandaji for the long and full of excitement discussions. The authors are forever grateful for your unlimited enthusiasm and passion for the topic. The authors want to highlight the great contribution of Professor Anders Böhlmark. Without your great econometric skills and expertise in the area of immigrant integration, the empirical set up of this paper would be impaired dramatically.

Furthermore, great acknowledge is directed to the Faculty of the Economics Department at Stockholm School of Economics, especially to Anna Dreber Almenberg, Erik Linqvist, Federica Romei, Mark Sanctuary, Maria Perrotta, Rickard Sandberg and Örjan Sjöberg.

Last, but not least we want to thank our classmates for your mental support. Petter Svärd and Mara Solomon, the authors cannot express enough thanks for your existence. At the end, the authors want to thank you, Dear Reader. If you take time to read this line after the others, you at least read one page of this thesis. Thank you.

Contents

List of Figures

List of Tables

1 Introduction

"We live in the age of the refugee, the age of the exile"

Ariel Dorfman

This paper investigates the labour market integration of refugees in Sweden. Unambiguously, we study the relationship between age at immigration and educational attainment, employment, earnings and welfare benefits. It is important to study the effect of time of arrival to Sweden, due to the fact that it is a determinant of social and economic integration (Schaafsma and Sweetman, 2001). In order to implement feasible programs it is crucial to understand the differences between the cohorts arriving.

With raising conflicts and political crises, the number of migrants worldwide has increased drastically in recent decades. The Western world has experienced rapid changes in the demographic composition of migrants. Labour migration, motivated by economic aspects, has been replaced by immigration due to exogenous shocks (OECD, 2016). Although there is a large literature describing patterns of social integration and income performance of labour migrants, there are few studies that concern the economic integration of refugees. In this paper, the term "refugee" is defined in line with the definition of The UN Refugee Agency, as a person fleeing conflict or persecution.

The differing reasons behind immigration as well as characteristics of refugees and labour migrants are distinct. Refugees are forced to flee from the country of origin due to force majeure, while labour migrants have the possibility to plan and time the immigration. The broad generalisation of immigrants in terms of Northern/Western versus non-European that has been the standard in previous literature, is misleading in lots of aspects due to the wide heterogeneity among the immigrants from different countries (Katz and Österberg, 2013). The analysis of the performance of refugees on labour market is valuable in order to address the problems with novel and more efficient policies. We limit the sample to only siblings and conduct a sibling comparison study. Furthermore, the differences between men and women, two generations of immigrants and regions are also studied. By performing F-tests, existence of structural breaks is analysed. At the end, a sensitivity analysis is performed, where the consistency of the results for additional samples (labour refugees and parentless immigrants), as well as in later career life, is confirmed.

Our analysis builds on individual register data, which allows us to analyse the integration into the labour market using different variables: education, wage, employment and welfare benefits. Education is the intermediary variable that greatly influences the future results in the labour market and the quality of the acquired employment in adult life. Previous literature concludes that reduction in educational attainment leaves a permanent scar on earnings in adulthood (Schaafsma and Sweetman, 2001). Employment and income performance in early career mirrors to what extent the individual refugee has integrated into a society and the labour market. There is a general agreement that the gap in the income between immigrants and natives is mainly due to the employment gap - it is a huge challenge to enter the labour market at all. Nevertheless, different studies find that the the earnings are also differentiated between immigrants from different countries (Katz and Österberg, 2013). Therefore, the most important variable in the analysis is the wage. The earnings of the studied individuals can be interpreted as the indirect contribution of the refugees to public finance, due to the fact that we use taxable income

(Ruist, 2015). The main reason behind the decision to use an indirect measure instead of studying directly the taxes is the access to data. Secondly, the process of economic integration is more transparent and comparable between different groups of immigrants than the effects of taxation. Hence, it is easier to address the differences in economic integration by different policies. The last economic outcome studied, social welfare benefits, can be treated as an approximation of how much an individual costs the society, in the sense that it is the amount received from the state for the people that lack any other income sources.

A crucial fact that needs to be considered is the differentiation between the refugees arriving to Sweden. The refugee has a choice to make of whether to flee the country or not and, most importantly, when. Consequently, we analyse the impact of age at immigration using sibling pairs in order to isolate the effects of childhood environments on social integration among individuals with otherwise comparable background.
The aim is to answer the following research questions:

1. What is the impact of age at immigration to Sweden on:
   (i) the level of the educational attainment?
   (ii) the level of income?
   (iii) the number of unemployment days?
   (iv) the level of social welfare benefits received?

2. Are there any critical breakpoints in the age influencing economic integration of refugees?

In comparison with previous studies of labour migrants, this study finds stronger effects of age at immigration for refugees. There is negative impact of higher age at immigration for all the variables included in the analysis. We find significant differences between different generations and gender. Further, we find that the effect for certain years is much stronger than the others - the critical thresholds for age at immigration exist.

In all honesty, the fact that age at immigration matters for integration might seem trivial to the uninformed reader. However, previous research has shown that the exact effect is ambiguous. The study by Beckley (2015) has found evidence on the existence of "Swedish paradox" in the sense that immigrants arriving to Sweden at an older age are less likely to commit crime than immigrants arriving at a younger age. That contradicts the intuition that acculturation associated with more time in the receiving country leads to higher social integration that should have a negative effect on the probability of committing crime. In addition to that, some studies report evidence on probability of being unemployed increasing with years since immigration (Aguilar and Gustafsson, 1991). This finding suggests that we should see higher unemployment among the younger siblings in our analysis. Therefore, it is captivating and important to understand the exact mechanisms behind the age at immigration and the effect of time spent in the receiving country.

To the best of our knowledge, no previous study has studied the economic performance of the recent refugee waves and analysed the regional differences. Secondly, our paper studies different methods to deal with business cycle problems related to the income performance in a wide time span. Lastly, we make an attempt to distinguish different types of immigrants to compare assimilation time in the labour market. This paper contributes important results that can help form efficient integration policies in order to improve the performance of refugees in the labour market. Hence, this study aims to minimize the costs of political migration for the overall Swedish state.

There is no other study that attempts to analyse the differences between arriving to Sweden as a family versus arriving as an orphan.

The rest of the study is laid out as follows. After a brief background, Section 3 presents research in the area of immigration and income performance and Section 4 describes the limitations. Sections 5 and 6 present the data and descriptive statistics. Section 7 outlines the empirical strategy and Section 8 presents the results. Section 9 and Section 10 describe the sensitivity of the results and discuss the overall analysis. Finally, Section 11 concludes the main findings.

2 Background

In this section we present the past and ongoing conflicts in the world that we use as exogenous shocks in the empirical analysis. Secondly, we discuss the structure of the dependent variables analysed in the following sections. Next, we present the second language acquisition theory in order to gain an understanding of the mechanisms that can aggravate the integration. Further, we present the immigration policies that facilitate the process of integration.

In the recent 30 years the composition of immigrants coming to Sweden has changed. Previously, the main drivers of the decision to emigrate were the economic aspects and hope for greater opportunities in the labour market. Currently we observe a greater proportion of immigrants from non-OECD countries that are most likely to come due to political reasons. Immigrants from these countries earn less and are less likely to be employed upon arrival. This shift in the immigrant composition indicates that the gap between the non-native and native population will increase in the future. Non-OECD immigrants start with severe income disadvantages and the current studies find little evidence of convergence to the income of native population. In 2002 the employment rate among immigrants moving to Sweden from outside of Europe was 15.8 percent lower than among European immigrants and 23.3 percent lower than among natives. While immigrants of the European heritage had income almost identical to the average of natives, the disparity between the average income of natives and non-European immigrants equalled 3,200 Swedish kronor/person (Åslund and Skans, 2014). The differences in labour market performance between the two different types of immigrants; economic and political, are largest over the first five years in Sweden (Edin et al., 2000).

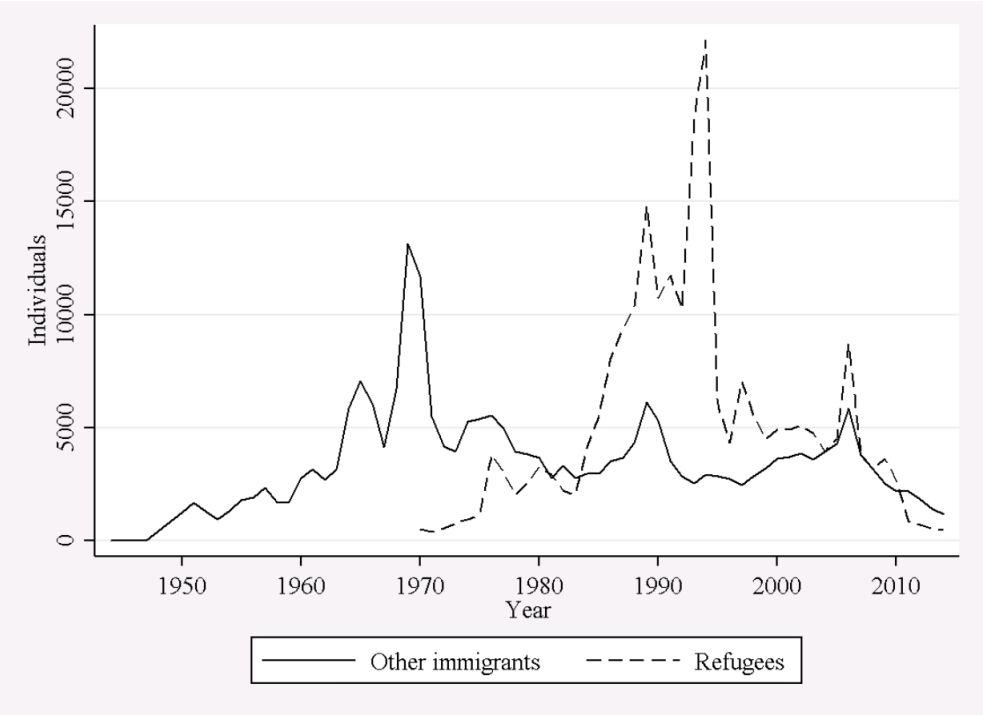
2.1 Exogenous shocks

Exogenous shocks have been defined in this paper as instances of political turbulence that have generated massive influxes of immigrants in the years 1950-2012. We see clear patterns of large immigration waves in the data, in the points in time, where important political conflicts took place historically.

During the 20th century, Europe has experienced several exogenous shocks due to the existing conflicts. To begin with, a large influx of immigrants had taken place after the Second World War when Sweden imposed laws that facilitated immigration for a number of countries to create a higher inflow of labour force. Later on, in 1969, Sweden restrained the labour migration law and imposed stricter regulations in order to reduce the number of labour migrants (Migrationsverket, 2016). After the restrictions have been implemented the labour migration has diminished, leaving more space for political/refugee migration. Nevertheless, these regulations did not have any significant impact on the immigration

flows from Northern and Baltic countries. Therefore, the inflows of immigrants from this region have been constant since the beginning of 20th century until now.

Figure 1: Trends in immigration to Sweden [1]



Note: The trends illustrate individuals below age 26 that turn 28 to 30 during the time span 1990 to 2014, immigration year for second-generation immigrants is the immigration year of the parent. Definition of other immigrants, see appendix I.

Since 1960, the number of first-generation immigrants living in Sweden has grown from 300,000 to 1,400,000 in 2011. The number of unparented children arriving to Sweden from countries such as Colombia and Sri Lanka, has risen dramatically during this period. Currently, more than 40 percent of the first-generation immigrant population living in Sweden originates from Non-Western countries outside of Europe (Migrationsverket, 2016).

After the economic turbulence in the 1970s, the main inflows of immigration consist of refugees. As previously mentioned, the labour market position of the immigrant population has deteriorated during the last thirty years in European countries. The military dictatorship and other conflicts in South America (1973-1990) have generated a grand influx of immigrants from the unstable political regimes. In the beginning of the 1970s, the turbulence in Africa have forced the minorities living in countries such as Uganda to flee, some of them moved to Sweden (Byman, 2005). The ongoing Arab Israeli conflict has been a source of large immigration waves since it started in 1948. The original refugees stemming from this conflict totaled 750,000 (Global Jewish Advocacy, 2016).

The 1980s had been a decade of asylum seekers in Sweden. When a new system for refugee reception was inaugurated in 1975, the number of refugees from the Middle East (mainly from Iran, Iraq, Lebanon, Syria and Turkey) had risen. The end of communist regimes induced a large wave of immigration due to the fact that living conditions were initially impaired after the independence and people received the freedom of movement

[1] Own production. Data retrieved from GeoSweden on 2016-12-01.

(Migrationsverket, 2014).

The next decade, the 1990s, had been turbulent. On the positive side, it was a period when many wars came to an end (for example in Lebanon, Eritrea, Iran and Iraq). On the negative side, the Balkan War took place. For the first time since the Second World War, Europe experienced record high numbers of people migrating. More than 100,000 refugees from the Balkan area found their new home in Sweden. The vote to join European Union resulted in Schengen cooperation and the migration to Sweden from other European countries has been facilitated (Migrationsverket, 2014).

2.2 The economic outcomes

In order to fully understand the impact of age on the economic outcomes, it is necessary to understand the mechanisms behind them explained in the existing models and research.

Traditional human capital theories assume that there is positive return to education in terms of earnings, due to signalling and increase in human capital. Furthermore, educational attainment is strongly correlated with the family background (Acemoglu and Autor, 2011).

Since earnings are a return on cumulated net investments in education, the income also increases at a diminishing rate over the working life of an individual. It declines when the investments become negative, when the individual ages. The typical (logarithmic) wage profile during the entire working life is therefore concave (Mincer, 1975). There are also different trends between individuals with varying educational background. In the Ben-Porath model, during adolescence less able individuals tend to earn more, but more educated individuals have rapidly rising earning profiles that usually overtake those of the less educated (Lillard, 1977).

The use of current income as a proxy for long-run income has been vividly criticized due to the error-in-variables biases. The most famous examples are the two studies by Modiglani and Brumberg (1954) and Brady and Friedman (1947), where the properties of consumption functions estimated with current income and permanent income have been compared. The study by Haider and Solon (2006) has concluded that the use of short run proxies is well-grounded if they are measured between early thirties and forties.

Lastly, it has also been documented that there are substantial differences in variation of life-cycle wage growth among countries. Lagakos et al.(2016) has found that experience-wage profiles are on average twice as steep in rich countries as in poor countries.

These findings will be used in order to decide on the optimal time of observation of the effects and in order to analyse the results in depth.

2.3 Second language acquisition

In language learning the rule of thumb is: the younger the better. The Critical Period Hypothesis (CPH) derived from biology by Penfield and Roberts in 1959, states that there are structural breaks in the function between learners' ages and their susceptibility to second language. In other words, the process of second language acquisition becomes more difficult after a certain age. The empirical evidence for this hypothesis has been found for both pronunciation (for example: Asher and Garcia, 1969; and Oyama, 19761) and grammar learning (for example: Harley Hart, 1997; Johnson Newport,1989; Patkowski, 1980). DeKeyser (2000) claims that there is a significant negative correlation between age of acquisition and ultimate attainment of the second language. The author has found

evidence on CPA and fundamental difference hypothesis, which states that adults have largely lost their ability to learn a second language implicitly. The author emphasizes the importance of specially designed language teaching policies that focus on adult acquirers.

According to the theory developed by Scovel in 1988, there is a biologically constrained period for the accent acquisition, ending at around age 12. After this critical period, it is impossible for learners to acquire such good pronunciation in a non-native language that they can pass themselves off as native speakers.

It is important to be critical to the existing theory. Despite all the existing evidence, CPH and "the younger, the better" hypothesis evoke a lot of scepticism. Some recent studies have found no evidence of them and claim that they are formed due to the application of wrong empirical methods. The critics contend that the effect of age on language acquisition is very individual. Notwithstanding, it is a factum that adults encounter more difficulties in the second language acquisition (Scovel, 2000; Vanhove, 2013). These findings will be used in order to understand differences in acculturation.

2.4 Immigrant integration policies

Once a refugee has entered Sweden, his or her rights do not distinguish from the rights of natives, despite the right to vote in parliamentary elections (Gustafsson and Zheng, 2006). The main improvements that are planned to be made in Sweden to facilitate integration of refugees are: (i) increase cooperation of Swedish institutions and (ii) implement acts and actions against discrimination (Government Offices of Sweden, 2015).

The Swedish Public Employment Service is required to produce a personal introduction plan together with the incoming immigrant in order to help with the preparation for the future employment and provision of required language skills. The immigrants who actively participate in the introductory process receive monetary introductory benefits. Furthermore, there is a program of specially subsidized employment introduced in 2007, so called step-in-jobs. These types of jobs can be offered to unemployed newly arrived immigrants and are to be combined with courses in the Swedish language. The subsidy amounts to 75 percent of employer wage costs. There are also programs being implemented in order to facilitate entrepreneurship among people with foreign backgrounds, such as events in order to help finding financing for company growth (Government Offices of Sweden, 2015). In addition to the existing labour market policies, all the migrants who received a residence permit are entitled to receive lessons in Swedish for immigrants (Migrationsverket, 2015). Single adults are given a benefit of 2,159 Swedish kronor per month, with more available for those with families (The Local, 2016).

A side effect of generous Swedish transfer systems is the lack of incentives for the newcomers to find a job. There are several new proposals for integration policies on the floor that claim to be an improvement of the existing ones. Since more than 40% of all the people considered as unemployed by the Swedish Public Employment Service are born outside of Europe, efficient tools should be used to achieve a change. In June 2016, the Swedish Parliament decided to implement one of the strictest laws for asylum-seekers in the European Union. The new law concerns the residence permits and for now it is temporal. The refugees should find employment in order to receive the permanent residence. Furthermore, the new law limits the family reunification immigration. The goal is to give additional incentives to improve integration in the labour market (Dagens Nyheter, 2016). The right-wing political parties in Sweden claim that the economic support for the immigrants needs to be reformed, so that it will become profitable to

take the step from the welfare benefit to employment. In a recent program proposal, the Swedish political party "Liberalerna" has proposed to implement a limit for maximum welfare benefit. Furthermore, the possibility to prolong the compulsory school attendance to the age 20 for immigrants is currently debated (Liberalerna, 2016).

The current political debate and the labour market situation demonstrate that the stimulation to work for immigrants in Sweden needs to increase. In order to introduce relevant policies, it is crucial to study the mechanisms behind the integration of refugees.

3 Previous research

In this section, we present the highlights from the topic on integration of immigrants in the labour market.

Previous research has focused on the reason behind the poor labour market situation of immigrants. There is no evidence that immigrants are less active in job-searching activities and wage discrimination has not been proved (Bantekas, 1992; Le Grand and Szulkin, 1999; Vilhelmson, 2002; Le Grand et al., 2004). An important observation is that immigrant earnings are generally positively related to the years since immigration. The majority of the previous research concludes that time in the host country matters for income performance (Rashid, 2004). However, some studies report evidence on probability of being unemployed increasing by years since immigration (Aguilar and Gustafsson, 1991). Borjas (1985, 1995) emphasizes that the skill characteristics of immigrants are strongly related to the country of origin. For example, country of origin dummies explain 30% of the variation in average education levels among immigrants in the 2000 United States Census. The studies of the earnings of immigrants encounter a lot of challenges, due to the fact that newer cohorts can differ from the older ones and the earnings are sensitive to business cycle effects (Gustafsson and Zheng, 2006). However, Borjas (1985, 1995) finds out that the cohort groups of immigrants are not very different from each other, reflecting the relative stability of immigrant inflow composition since the late 1970s.

Böhlmark (2009) is the first to apply the family fixed effects in the context of Swedish immigration. The author applies the family fixed method in order to analyse the effects of age at immigration on education among immigrants. The analysis exploits within-family variation in a large set of register data on immigrant siblings (and native children) graduating from compulsory school (usually at age 16) between the years 1988 and 2003. The reasoning behind this choice of the model is to remove the unobserved heterogeneity and to handle the potential for reversed causality and omitted variables. The main source of the bias is the fact that the timing of immigration may be affected by different patterns, for example the age of children in the family. Additionally, the author finds out that the birth order within family matters, since the children who are born first have better prospects for high educational attainment. Therefore, Böhlmark (2009) includes a control variable *Firstborn* in the estimation. The main finding of the paper is that there is a strong association between age at immigration and the level of Sweden-specific acquired skills and knowledge of new subjects at the time of graduation from primary school. Böhlmark (2008) observes that there is a threshold at the age 9 that can be called "the critical age at immigration". Above the age 9, the negative impact of coming late to Sweden is significant in magnitude and gets larger the shorter time before graduating.

Despite the existence of the threshold, the individual seems to catch up at the end of the educational path. However, the effect of immigration at sensitive ages seems to impair the results in the labour market for the individual, in terms of greater risk of being unemployed and having a lower wage. According to the author, a potential explanation for this phenomenon could be the fact that immigration above certain ages leads to permanent limitations in Swedish-language proficiency. This can lead to penalization of the lack of language skills in the form of discrimination. In turn, this indicates that the policies that have been implemented in order to close the educational gap between natives and foreign born have been successful, but there is room for improvement of the labour market interventions (Böhlmark, 2009).

Åslund et al. (2015) take the study by Böhlmark (2009) into the next level and study social integration of immigrants in Sweden for the years 1991-2005. Social integration is studied in terms of exposure to natives in the workplace, residential and marriage. The economic outcomes (education, employment and logarithm of wage) are mainly used in order to explain the effect of age on social segregation. The authors use mother identification number in order to define family fixed effects and cluster the standard errors at cohort and regional level. The analysis is limited to children arriving to Sweden in the ages 0-15 and the second-generation immigrants. The sample studied consists in 60% of the Nordic migrants, and within this group Finland constitutes 85%. Yugoslavia, Turkey, Germany and "Southern Europe" are other groups in the studied sample. The strategy used in the paper in order to cope with macroeconomic influences on the outcomes is to use the same region age-zero migrants to identify calendar effects. The results are focused on presenting the social segregation at workplace, marriage and residential integration as effects of age at immigration. The authors find that both education and labour market outcomes fall with age at immigration. The effect is stronger for the more geographically distant countries. They conclude that time is more important if the initial cultural distance is greater. There is no significant effect for the wages of immigrants. The main conclusion of the paper regarding the economic outcomes is that processes related to skills and economic outcomes are unlikely explanation of the negative effect of age at immigration on integration into the marriage market and workplace as well as residential segregation. In this research, there is no specific evidence of a critical age for the effect of age at immigration on social segregation, as it is often found in studies studying language acquisition. Additionally, the results indicate that parental assimilation is important for the outcomes of second-generation immigrants (Åslund et al., 2015).

In addition to that, another observation of the studies conducted in Sweden, is that there are substantial differences between immigrant men and women. On an aggregate level, the performance deficit of male immigrants relative to male natives exceeds the performance deficit of female refugees relative to native female employees. In other words, foreign-born women perform better among the female employees than foreign-born men in comparison with other male employees (Lundborg, 2013). This can be explained by the fact that immigrant female workers always face the disadvantage in earning of being female, in addition to being foreign. The disadvantage is shared with the reference group, the native women (Le Grand and Szulkin, 2002). It is plausible that in a country like Sweden, where the lower part of the wage distribution is more compressed compared to many other countries, there is less room for discrimination at the lower end of the scale and therefore the differences in performance within respective gender group are lower for women (Katz and Österberg, 2013).

The studies that analyse the relationship between the age at immigration and eco-

nomic outcomes outside of Sweden, are different due to the limitations in existing data. The typical empirical strategy applied in these studies is to project the difference between the immigrant's actual earnings and the predicted earnings that the worker would earn if born in the host country. Later on, the authors analyse if the gap between the natives and immigrant earnings is a function of age at immigration. The results do not differ to a larger extent with the results observed in Sweden. Educational attainment and earnings vary systematically across age at immigration with those arriving around age 15 to 18 receiving fewer years of education and consequently lower earnings (Schaafsma and Sweetman, 2001). Other studies, like the one by Börjas (1985, 1995) focus on identifying assimilation profiles, and therefore apply cohort fixed effects, using the variation within the different cohorts immigrating.

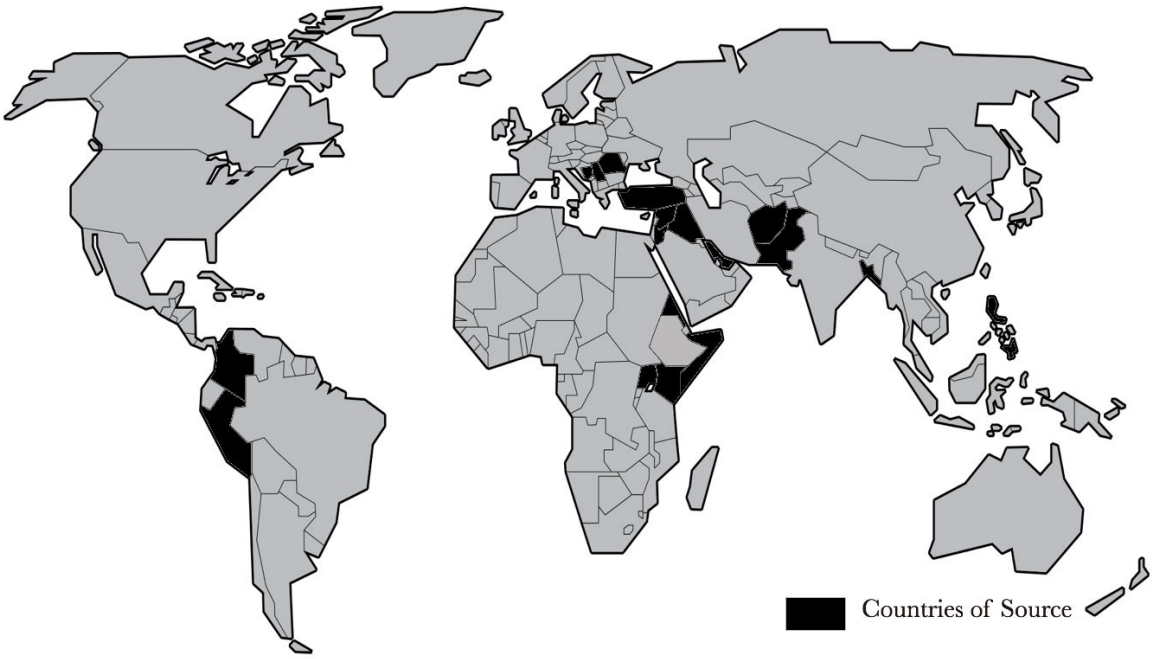
The overview of the existing research on the topic clearly indicates that there is room for more careful analysis of the sample of recent immigrants from more distant countries and their integration in the labour market.

4 Limitation of scope

We limit the analysis of the sample of immigrants to the refugees migrating to Sweden during the time period of 1950-2012. The effects on our dependent variables are studied during the following observation years: from 1990 to 2014. In this paper we use the following definitions: first-generation immigrants are the people who were born in one country and relocated to Sweden at a young age. Second-generation immigrants are born in Sweden, but with parents born outside of Sweden. Age-zero immigrants belong to the reference group and are the children to the immigrants who immigrated to Sweden just after or before their birth. Orphans are defined as parentless children (ages 0-5), arriving to Sweden solely with their siblings. The welfare benefits is the social assistance received from the Government; the benefits are given to people who have no other income (Socialstyrelsen, 2016). Even though the overall population of Swedish refugees is included in the analysis, we refer to it as the refugee sample, since we limit it to the sibling sample. Time in Sweden and age at immigration can be used reversibly. The conflicts below are used in the main analysis of the refugee sample.

Conflict	Period	Location
Cold War	year 1947-1991	America Asia Europe
Arab Israeli Conflict	year 1948-present	Africa Asia
Vietnam War	year 1955-1975	Asia
Eritrean War	year 1962-1993	Africa
Expulsion of Asians	year 1972	Africa
Military Dictatorship	year 1973-1990	South America
Soviet War in Afghanistan	year 1979-1989	Asia
Balkan War	year 1991-1999	Europe

Figure 2: Countries included in the main analysis[2]



The reason behind the exclusion of some of the countries that immigrated after the revolutions of 1989, such as Poland, Germany and Russia, is that there has been a constant influx of the labour migrants from these areas during the period studied, and no clear shocks (drastic increases in the number of incoming migrants) are visible in the data.

Considering the structure of the economic outcomes, we have decided to observe the effects on the dependent variables when the individuals are 28-30 years old. Taking an average of the observation in a time interval, reduce the noise of missing values or point-wise fluctuations in individual economic outcome. We want to analyse the effects of age at immigration during the early working life, since we suspect that at this point the effect of age at immigration is very transparent. However, in order to check the robustness of the estimates in the long-term, the effect when the individuals are 34-36 years old is also analysed. The optimal would be to analyse life-time earnings. Nonetheless, this is impossible due to limitations in the access to data and time scope of this study.

4.1 Hypotheses

Our initial speculation is that there are negative consequences due to arrival to Sweden at higher immigration ages. According to the existing literature, the time spent in Sweden increases the outcomes of refugees in the labour market and school results. Therefore, arriving to Sweden at younger ages should be more efficient for the population of refugees. The following hypotheses are being tested:

Hypothesis 1: The educational attainment within the age interval 28-30 decreases due to

[2]Adapted from *Simple world map*, by Wikimedia, December 1, 2016. Retrieved from: `https://commons.wikimedia.org/wiki/File:Simple_world_map.svg`.

increase in the age at immigration.
Hypothesis 2: The logarithm of earnings within the age interval 28-30 decreases due to increase in the age at immigration.
Hypothesis 3: The probability of being employed within the age interval 28-30 decreases due to increase in the age at immigration.
Hypothesis 4: The amount of received welfare benefits within the age interval 28-30 increases due to increase in the age at immigration.
Hypothesis 5: There is a critical age at which the effects of age at immigration on the economic outcomes are amplified.

5 Data

In order to answer the research questions, an internationally unique database is used: GEOSweden Data. This database contains yearly (longitudinal) geocoded data on all residents in Sweden, including demographic and socioeconomic characteristics with precise coordinates and neighbourhood area codes for homes and workplaces. The main original source is individual data from the register-based longitudinal database LISA LOUISE (Longitudinal Integration Database for Health Insurance and Labour Market Studies) collected and developed by Statistics Sweden. The sample analysed contains only the individuals who migrated to Sweden during the studied period and are still in the country when they turn age 30. Therefore, the studied samples do not include the people who emigrated from Sweden during the time frame of analysis. For the individuals included in the dataset, we can identify their demographic characteristics, labour market characteristics, education and social benefits. States of origin include region of birth, and year of birth for each individual and their parents. Additionally, the intergenerational data also contains information on year of immigration, gender and siblings. Furthermore, for each year we have information on obtained earnings, social welfare benefits and registered days of unemployment.

All the samples analysed in this paper consist of immigrants born between 1960 and 1984. The groups included in the analysed samples are generated either through first-generation migrants before age 28 or second-generation, whose parents immigrated 10 years before they were born. We measure age at immigration in the [-10,26] interval and include years of migration span from 1950 to 2012. The study of robustness of the results is conducted within the same time interval, which reduces the number of observations significantly. The focus lies within three different groups of migrants: second-generation and first-generation including (i) children and (ii) adults. The individual's region of origin in the analysis means implicitly place of birth. In order to identify sibling pairs of the second-generation immigrants, we use the identification number of the parents that are born outside of Sweden. The first-generation of immigrants' sample is identified through the household identification number. The orphan sample is generated through first-generation immigrant children (ages 0-5) without parents using the household number. This identification method gives us access to all sibling pairs (families with two or more children). In our main analysis we measure outcomes as average values of observations at the ages of 28-30; using several years to decrease the number of missing values for wages, social benefits and unemployment days. In the sensitivity analysis we analyse the effects within the age interval 34-36, in order to draw conclusions about long-term effects and trends from the main analysis.

The performance of immigrants in the labour market is approximated with the following economic outcomes; (i) wage (logarithm), (ii) obtained social welfare and (iii) registered unemployment days. We also study the mediator effect of educational attainment (iv). The outcome variables are defined and created in the following way: the highest educational attainment is yearly based using the following 7 categories; preschool (corresponding to 0 years of education), primary school (equal to 9 years of education), secondary school (equivalent to 12 years of education), higher education without degree (13-15 years of education), higher education with degree (15 years of education), master's degree (17 years of education). The index representing the educational attainment is translated to the years of education for interpretational reasons. Regarding the direct economic outcomes, each individual in the dataset is registered with tax income filled by the employer on a yearly basis. The annual earnings dataset contains information from the first to the last month each year. We use this information to generate the (i) logarithm of the mean of yearly wages by each individual over the span of three years. If the mean annual wage is below 10,000 Swedish kronor during this period, the individual is excluded from the analysis of the income. In the analysis there is no lower limit for the included observations for unemployment and welfare benefits. The reason for measuring the dependent variable of choice as the logarithm of wage and not simply the wage itself, is due to the fact that we solely want to include the earnings of the people who are part of the labour force, and therefore the part of the sample that has never worked is excluded (all the zero observations are excluded). It is also in order to make the income distribution more normally distributed. The same approach is used when creating the average social welfare benefits and unemployment days based on the same time span over three years' time. The unemployment days are generated through the number of days the individuals are registered at the Swedish Public Employment Service.

6 Descriptive statistics

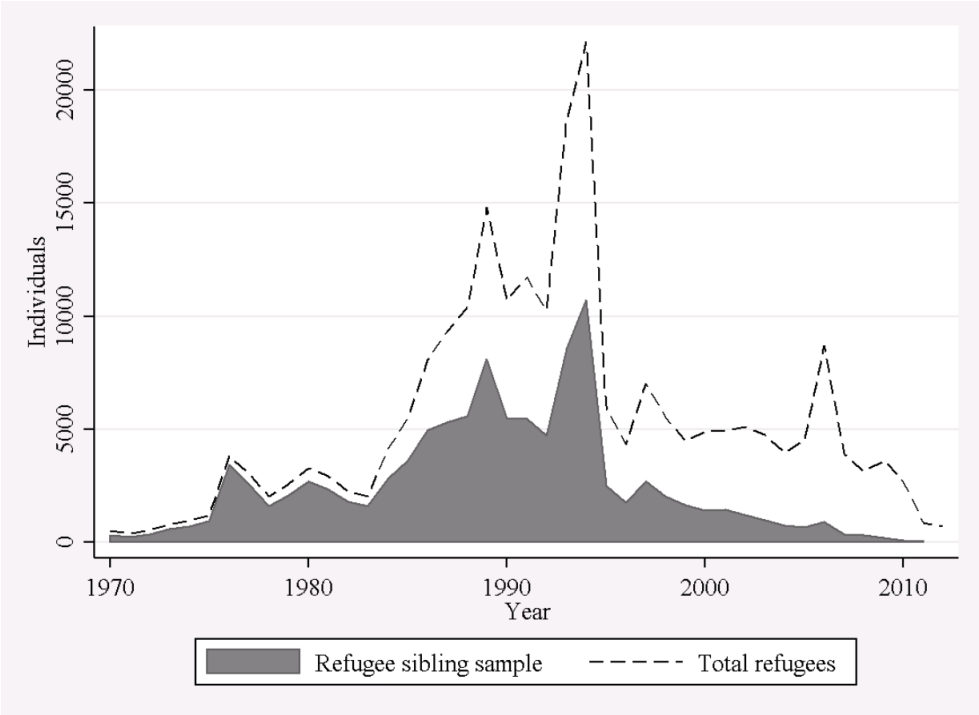
This section includes a description of the population sample that is analysed. We include sample statistics on total immigration on a country specific level. Furthermore, the statistical picture of the total population, measured during the same time period as explained in the section above, is presented.

The requirements that we put on the countries included in the analysis are two-fold. Firstly, all the countries included has experienced an exogenous shock that has generated a substantial influx of immigrants. Secondly, we decided to exclude geographically close countries in the analysis, for example Germany, even if an exogenous has occurred. The main reason behind the elimination is that the composition of these groups and root causes of immigration have been widely heterogeneous. The focus is to capture immigration from exogenous shocks in more geographically and culturally distant countries incoming to Sweden exclusively due to the shocks. This is the reason behind the exclusion of immigrants from Finland, Germany, and Poland - countries that have a history of labour migration prior to the restrained immigration policy.

The figure below presents the total inflow of immigrants and the individuals arriving in sibling pairs aged between 0 and 25 years. In addition, the sample also includes second-generation immigrants, children of parents who arrived to Sweden less than 10 years prior to their birth. We find that the majority of individuals in our sample came to Sweden after the restrained labour migration policies, which is in line with the purpose to not

focus on labour migrants and solely include refugees. The visual immigration inflow per country of origin can be found in the appendix H.

Figure 3: Refugee immigration to Sweden



Note: Second-generation immigrations use the immigration year of the parent.

Further, we analyse the sibling sample and its characteristics within the age interval 28-30 years. As seen in the table below, the average age in the population of Swedish refugees is 13.5 years. The sample is equally divided between females and males. The average refugee in the sample has approximately 13 years of education in total, which is equivalent to the level of higher education in the Swedish school system. The control variable used in the regressions, parental education, has the average value of 10 years, equivalent to secondary school level. However, the educational level of refugees should be interpreted with precaution; older immigrants may have completed their education in the country of origin and the reported value at immigration to Sweden may not transfer to the exact years, the same applies to the education of parents. We therefore take this into consideration in the analysis and exclude older immigrants, aged 16-26, when we analyse the attained education.

The average number of unemployment days for a refugee in the sample is 45 days (including all the days in a year, not only working days). Regarding the mean wage of refugees, we see that on average the refugee who is part of the labour force earns monthly 11,417 Swedish kronor and the native earns monthly 13,908 Swedish kronor. The sum of all the welfare benefits received annually by an immigrant in Sweden is 4,917 Swedish kronor. The source countries, generating refugee inflows, have the GDP level of 7.4 percent of Swedish GDP compared to 82 percent in the sample consisting of labour migration countries. Furthermore, English proficiency is below the level of the host country, Sweden.

Table 1: Descriptive statistics - refugee sample

VARIABLE	Obs.	Mean	Std. Dev	Min	Max
Birth year	111,721	1976	7.291	1960	1984
Immigration year	111,721	1989	7.414	1970	2010
Age at immigration	111,721	13.50	8.959	-10	26
Female==1	111,721	0.50	0.499	0	1
Attained education	86,567	12.64	3.339	0	17
Unemployment days	108,536	44.92	70.08	0	365
Wage(log)	86,203	6.827	1.257	2.303	10.38
Welfare benefts	111,721	49.17	121.05	0	2350
Year firstborn	42,523	1971	7.000	1960	1984
Number of siblings	111,721	2.80	1.157	2	6
Parent education	65,414	9.79	2.85	0	17
GDP / Sweden (perc.)	99,269	0.074	0.04	0.01	0.24
Corruption measure	111,721	37.95	21.60	8	87
English proficiency	78,564	50.47	9.441	38.02	70.94
Gender Equality	9,226	2.442	3.039	0	9.39

Note: The sibling sample of refugees used in the main analysis. The time span of observation lies between 1990 to 2014, the earliest possible immigration is 1950. Immigration year refer to the mother immigration year if the individual is second-generation. Outcomes measured in age 28 to 30. The average country characteristics are divided by number of country in the refugee sample, source of variable, see appendix A.

Table 2 presents the distribution of the refugees in the sample by the countries of birth. The distribution is skewed with half of the population originating from West Asia.

Table 3 presents the dependent variables used in our sample. The summary illustrates the situation of refugees compared to natives in Sweden. In addition, we also present the values for the European immigrants, classified in the previous research as "labour migrants" [3].

We find that the second-generation immigrants from the refugee sample have similar characteristics to the natives. The annual earnings of second-generation immigrants are even higher than for the natives. Regarding the first-generation arriving from refugee countries; we see that outcomes reflect a poor labour market situation compared to natives and second-generation immigrants. The difference in average unemployment days and received welfare benefits raises substantially. This indicates that later arrival to Sweden plays an important role. Further, when inspecting the first-generation immigrants aged 19 to 26, we see a substantial gap from the previous age group (see Table 3). This large gap indicates that older refugees in later teenage, are very sensitive to the time of arrival to Sweden.

[3]Countries included in the European sample: Denmark, Finland, Norway, Poland, Germany, Greece, United Kingdom and Italy.

Table 2: Distribution - refugee sample

REGION	Sibling immigrants	Distribution
Europe	17,954	0.16
West Asia	59,170	0.52
Central Asia	10,177	0.09
Africa	10,411	0.09
South America	14,009	0.13
Total	111,721	1.00

Note: The distribution by country can be found in the appendix C.

Table 3: Mean - dependent variables

	Natives	Other Immigrants	Refugee sibling sample		
OUTCOME VARIABLE	Aged [0;26]	Aged [-10;26]	Aged [-10;-1]	Aged [0;18]	Aged [19;26]
Attained Education	13.96	13.39	13.18	12.98	11.44
Unemployment days	20.83	27.74	23.05	34.66	66.63
Wage (log)	7.193	7.051	7.320	7.088	6.171
Welfare benefits	8.123	16.67	10.97	28.76	88.86
Number of observations	2,437,772	80,342	7,582	63,974	40,166

Note: The table with values for "natives" includes values for individuals with Swedish parents. Other immigrants and refugee sibling sample are individuals with siblings in the analysis.

Looking at the country specific characteristics of the sample, see appendix C, the refugees with the highest educational attainment in the sample come from Iran, where the norm is to have higher than high school education (almost 14 years of education). Another country with high educational attainment among immigrants is Bolivia and Peru. An interesting fact is that the young refugees (average age is 3) that often arrive parentless from Sri Lanka or Colombia have high levels of educational attainment in comparison with other immigrants. The immigrants from Sri Lanka also have the lowest unemployment among refugees and highest earnings, as well as lowest amount of welfare benefits received in adulthood. Therefore, we decide to analyse the phenomenon of high performance of young, parentless immigrants arriving from Sri Lanka, Colombia or Chile in the sensitivity analysis.

6.1 Correlation between dependent variables

In order to interpret our results, it is important to understand how the dependent variables interact. If the dependent variables are perfectly correlated there should not be any interest in evaluating them separately.

The results for all individuals and the refugee samples go in the same direction and are of similar magnitude. Considering first the logarithm of wage, there is a positive correlation between attained education. In other words, lower educational attainment is correlated with lower earnings. Moreover, receiving more welfare benefits and unemployment days is negatively associated with the logarithm of wage. Higher education has

a negative impact on unemployment. Looking at welfare benefits, there is a negative association with the level of education. As expected, more unemployment days implies higher amount of welfare benefits received. The correlation between unemployment and welfare benefits is not perfect, because the variable identifying unemployment has been defined ambiguously. The definition of unemployment leaves room for both receiving the welfare benefits during a part of a year and being unemployed, as well as earning income during another part of the same year. Interestingly, we find relatively weak correlation between all dependent variables, except unemployment days and wage.

Table 4: Correlation - dependent variables

VARIABLE	Attained Education	Unempl. days	Wage (log)	Welfare benefits
All subgroups				
Attained education	1.0000			
Unemployment days	-0.1149	1.0000		
Wage (log)	0.1344	-0.3622	1.0000	
Welfare benefits	-0.1202	0.2787	-0.2667	1.0000
Refugee sample				
Attained education	1.0000			
Unemployment days	-0.1460	1.0000		
Wage (log)	0.2047	-0.3873	1.0000	
Welfare benefits	-0.1467	0.2777	-0.3270	1.0000

Note: The correlation between the dependent economic outcomes used in the main analysis. The upper panel represent all individuals in Sweden, the lower panel present the refugee sibling sample. Observation time 1990 to 2014. For detailed explanation of the economic outcomes, see the data section.

7 The empirical strategy

In the following section the main strategy and empirical framework is set up, that is used to estimate the impact of age at immigration on outcomes within two different age intervals: (i) 28 to 30 years and (ii) 34 to 36 years. To begin with, the basic model is outlined to estimate the effect on four economic outcomes. Secondly, the strategy to eliminate the bias in estimations is presented. In addition to that, the test performed to find the structural breaks in the data that serve as a guideline for the division of the sample, is presented.

The age at immigration for the first-generation immigrants has been calculated in the following way:

$Age\ at\ Immigration\ =\ Year\ of\ arrival\ -\ Birth\ year\ [0, 26]$
Similar approach has been used for the second-generation immigrants:
$Age\ at\ Immigration\ =\ Year\ of\ arrival\ of\ the\ parent\ -\ Birth\ year$ [-10,0]

The estimations capturing the effect of age at immigration will be performed in two different ways:
1) using dummy variables. [4]
2) using spline function. In this function the effect of age at immigration is assumed to be linear. However, the slope is allowed to vary when the age equals zero and eighteen. [5]

The two different ways to use the explanatory variable of interest have both advantages and disadvantages. The estimation with dummies allows for a very precise analysis of the effect for every age at immigration, but the linear specification gives a broader picture of the effect of every extra year of the delay in the point of time of immigration. The bias due to measurement error is also excluded when using the linear specification, since the error is symmetric for the differences in arrival for both of siblings.

Previous research shows that there is a significant impact of birth order among the siblings that can influence the labour market outcomes. There is evidence on that being *Firstborn* among the siblings is beneficial for the educational performance (Böhlmark, 2009). Therefore, the model is augmented by inclusion of an indicator for being a first-born child and by specification by a gender dummy. Furthermore, there might be macroeconomic influences that could impact the analysed effect. Another important extension that is made to the existing methods is the control for the history of the business cycles every individual face, we include a variable capturing the business cycle when the individual gets out into the labour market at age 18. All the regression estimations performed include the clustered heteroscedasticity-consistent standard errors.

In other words, the following OLS regression equation is estimated:

$$Y_{ij} = a + \beta_1 Age_i^{Im} + \delta_t BCB_t + \beta_2 Firstborn_{ij} + \beta_3 Education_{Parent_{ij}} + \beta_4 BCB_{t|Age^{Im}=18} + \beta_5 Female_i + \epsilon_{ij}$$

where Y_{ij} is the economic outcome of interest (educational attainment, employment, wage, social welfare benefits) for the individual i in the household$j$, Age_i^{Im} is the dummy variable for each possible age at immigration, except the reference group. In addition, we also include BCB_t, Business Cycle Barometer [6] and $Firstborn_{ij}$, an indicator for being first-born child, as well as $BCB_{t|Age^{Im}=18}$ that is the Business Cycle when the individual is 18-year-old. $Education_{Parent_{ij}}$ is the highest level of educational attainment of the parent and ϵ_{ij} is the error term.

Nonetheless, in the OLS regression there is a problem of selection bias that occurs due to the fact that the parents have the possibility to choose when it is most beneficial for their children to immigrate. Some part of the bias is eliminated since we use exogenous shocks, different conflicts that force the individuals to immigrate. The use of stochastic events leaves no space to plan and time the age at arrival. However, in order to fully eliminate the bias, the analysis is performed by sibling comparison using family fixed effects outlined in the section below. There are significant differences between the families that decide to arrive with their new-born children and teenage children, as well as the adult refugees. Therefore, the within-family variation is isolated.

[4] In the binary variable estimation, the age span -10 to 18 is analysed in order to capture the effect of childhood integration to reduce the bias that emerges due to the wide time span between the siblings.

[5] Spline is applied on a wider age span (the age interval of -10 to 26, in order to be able to find the critical thresholds.

[6] Retrieved from National Institute of Economic Research, 2016-12-02.

The differences in the fixed age are compared between siblings arriving to Sweden at different ages. We define family fixed effects using household identification number. The sibling approach can be motivated by the presence of different sources of bias, such as neighbourhood effects. Quality of school and other external factors influence the effect of age at immigration on future income, which is eliminated while using family fixed effects. The main assumption of the model is that the siblings would perform equally well within the different dependent variables in the absence of immigration. The theory that birth order matters threatens this assumption, therefore we still include the control variable *Firstborn*. The family fixed-effect approach will be used simultaneously on each sample for comparison reasons (appendix include tests without fixed-effect estimates). The age-zero immigrants are used as the reference group.

Thus the main regression equation, for the age at immigration as a binary variable will be:

$$Y_{ij} = a + \beta_1 Age_i^{Im} + \delta_t BCB_t + \beta_2 Firstborn_{ij} + \beta_4 BCB_{t|Age^{Im}=18+} \beta_5 Female_i + \gamma_j f_j + \epsilon_{ij}$$

where f_j are the family characteristics influencing the effect of age at immigration on economic outcomes.

Next, we have categorized the immigration age into three buckets: (a) -10 to 0, (b) 0 to 18, and (c) 18 to 26. In the spline regression, the same equation as the main regression equation is estimated, but instead of $\beta_1 Age_i^{Im}$ we use:

$$I_a \beta_1^a \ Age_i^{Ima} + I_b \beta_1^b \ Age_i^{Imb} + I_c \beta_1^c \ Age_i^{Imc}$$

where the superscript denotes which group the person is categorized into, and I_a is the indicator function taking on the value 1 if person i is in group a and zero otherwise. Therefore, the interpretation of the coefficient will be the following:

$$\frac{dy}{dAge_i^{Im}} = \begin{cases} a_1, if Age_i^{Im} < 0; \\ a_2, if \ 0 \leq Age_i^{Im} < 18 \\ a_3, otherwise \end{cases}$$

In the analysis, there is a problem of the macroeconomic influences and calendar effects on the relationship between the age at arrival to Sweden and income. The strategy to apply year fixed effects is problematic. Since the outcomes are observed at a common age, the variation in age at immigration and time of observation is identical for siblings. This means that when including the family fixed effects, we have perfect correlation between age at immigration and time of observations. A frequently used strategy to deal with this issue, is to assume that observation time effects are the same for immigrants and natives. This approach has been questioned when studying social integration and economic outcomes by Barth et al (2004, 2006). The time effects are likely to differ between natives and immigrants, which has been confirmed by Gustafsson and Zheng (2006). To handle the effects of outcome years we instead follow the strategy used by Åslund et al. (2015). We define our assumption that differences in unobserved heterogeneity among age-zero immigrants are not systematically related to the age structure of the cohorts. Under this assumption we can derive consistent estimates through the transformation of the dependent variable. Furthermore, this assumption, can be seen as plausible, since the large waves of immigrants are outspread during the analysed period. Due to the symmetric and relatively constant over time influx of immigrants at age zero,

this group can be seen a good reference. This assumption is also backed by the strategies used in previous research (Åslund et al., 2015). As already mentioned, the comparison group consists of the immigrants arriving to Sweden at age zero, excluding those from geographically close countries. We do not separate the comparison based on the type of immigration in order to obtain sufficient number of observations for the reference group. As seen in the appendix D the trends for the refugees and the entire population of immigrants are overlapping. The benefits of this strategy in terms of lower standard errors outweigh the costs, since the results are consistent for both methods, independently of the composition of the reference group.

First, we calculate the average outcome among same-aged immigrants that arrived at age zero in the specific year of observation. Subsequently, the data is transformed by deducting this average from the individual outcome, as follows:

$$\bar{y}_t|_{Age}^{Im}{}_{=0} = \delta_t + \bar{f}_t|_{Age}^{Im}{}_{=0} \; \tilde{y}_{ijt} = \beta Age_i^{Im} + f_j + u_{ijt}$$
$$\text{Where } \bar{y}_{ijt} = y_{ijt} - \bar{y}_t|_{Age}^{Im}{}_{=0}$$

where δ_t= the time specific effect of the observation year on the outcome of interest.

Further, the same method is applied in order to adjust for the institutional changes in education. There is evidence of inflation of the grades in Sweden (Skolverket, 2016), which means that the estimation will be biased downward, since later arrival means also that it is easier to complete the school. Therefore, we adjust for that by controlling for the trends in educational attainment for age-zero immigrants.

The bias due to the effect of macroeconomic influences on the amount of welfare benefits received is eliminated by the control variable *Business Cycle Barometer*.

Next, in the sensitivity analysis the same empirical strategy is applied with minor adjustments. In the analysis of the orphan sample, year and regional fixed effects methods have been applied. The following regression equations has been estimated:

$$Y_i = a + \beta_1 Orphan + \beta_2 BCB_t + \gamma \mathrm{r} + \epsilon_i$$

where *Orphan* is the dummy variable for coming to Sweden with parents or alone and r are the region characteristics influencing the effect of age at immigration on economic outcomes.

$$Y_i = a + \beta_1 Orphan + \gamma \mathrm{Year}_t + \epsilon_i$$

where $Year_t$ are the time characteristics influencing the effect of age at immigration in the estimated sample.

7.1 Critical periods

It is called a structural break when a time series abruptly changes at a point in time. A test for structural breaks is performed in order to determine whether there is a significant change in the data. The change in the case of this empirical analysis involves a change in means of the dependent variable of interest for a specific value of the main explanatory variable, age at immigration. In order to determine whether and when there is a structural break, we conduct a Wald test, where we test the following null hypothesis:
H_0: There is no structural break

The Wald test is applied on an unrestricted regression, which means that there are

no restrictions on the estimates and the data finds the breakpoints. In other words, the unrestricted regression considers the model that reflects the alternative hypothesis allowing the parameter estimates to take on any values. The Wald test is based on the parameter estimates and their covariance derived using maximum likelihood ratios or least squares (Wald, 1943). The test statistic is calculated in line with the following equation:

$$\Xi_W = \frac{(\hat{\beta}-\beta_0)^2}{Var(X)}$$

8 Results

To begin with, the results from the estimation of the baseline regression with immigration age dummies for second and first-generation, are presented. Secondly, the Wald-statistic[7] results using the full sibling sample for two different groups (childhood immigrants and the full sample, up to age 26) are presented. The results from the Wald tests are used in order to estimate a spline regression with slope changes. The aim for that is to facilitate the interpretation from our estimates. We divide the sample by gender to see heterogeneity in the estimated sample. Lastly, we analyse potential effects from geographical distance.

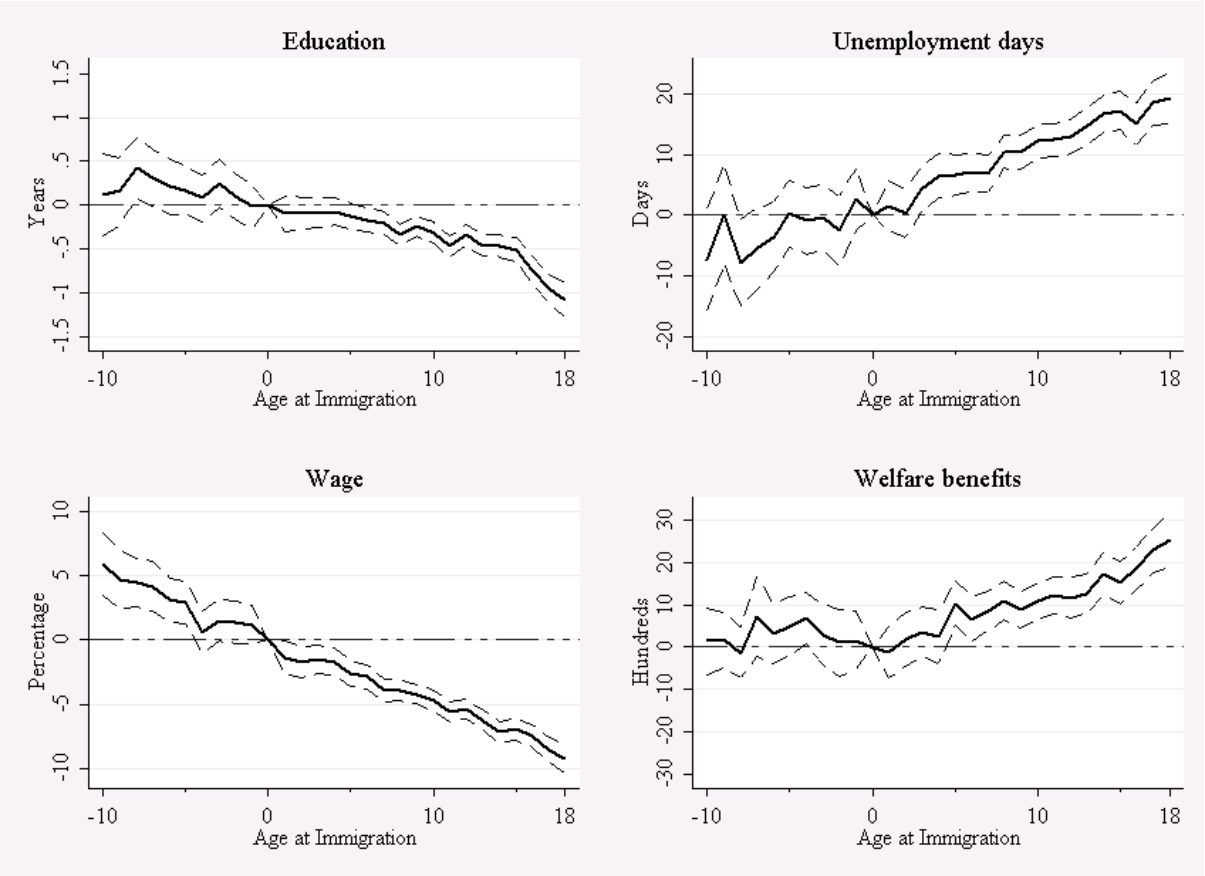
8.1 Immigration age as a binary variable

Figure 4 illustrates the estimated coefficients at each arrival age with a 95% confidence interval between ages -10 to 18. The horizontal line indicates the age-zero immigrants (the reference group). The impact from arrival age compared to the age-zero immigrant is visualized.

The effects for the different dependent variables of interest differ in magnitude, and significance between the two generations. Starting with the upper panel in the first graph, we see that the impact on attained education of the age at immigration have a a downward sloping trend. The results for the second-generation are slightly above zero, but with large standard errors and lack of significance. The negative effect of later arrival to Sweden becomes significant if the individual from the first-generation group arrives after the age 6, just before the beginning of the primary school in Sweden (regression results using the dummies are found in the appendix E.1). When plotting the effect on educational attainment for the full sample up to age 26, it is clear that after age 18 the estimate changes trend and standard errors increase (the education graph with dummies up to age 26 is found in appendix E.1). This makes the results for higher ages than 18 difficult to interpret, which is consistent with the fact that the immigrant may have finished her education before immigration. Next, the trend for the number of employment days is similar to the trend for attained education. The estimated coefficient indicates that higher age at immigration results in more days of unemployment in adult life. The effects for second-generation are not as strong as for the first-generation and have higher standard errors. The effect on unemployment becomes significant already if the individual arrives to Sweden at age 4. After that point, every extra year of the delay in arrival means more years of unemployment in Sweden. The average reaches a large magnitude; when the individual arrives at age 11, on average she can expect to have 13 days more in unemployment than the reference group at ages 28-30. The lower panel starts with the

[7]See appendix E.2 for additional explanation

Figure 4: Binary variable - refugee sample



Note: Age at immigration as binary variable - The estimates and 95% confidence intervals from family fixed effects specifications. The reference group is the age-zero immigrant. All outcomes are observed at age 28 to 30. Beta coefficients are found in appendix E.1.

wage estimate in percentage of the reference group. We find a strong negative trend both in the effect on the economic outcomes for the second and first-generation immigrants. There is clear evidence that the children born in Sweden, with parents that arrived to Sweden at least 5 years before the birth, benefit by the earlier arrival of the parents to Sweden. In comparison with the age-zero immigrant, they earn more, the earlier their parents have arrived. The negative impact of later arrival becomes significant, already for the age-zero immigrants that are born in Sweden. Thereafter, the effect increases. The last economic outcome, received welfare benefits in hundreds of Swedish kronor, follows a positive trend. Later arrival means more welfare benefits received. In line with the initial expectation, the first-generation immigrants receive more welfare benefits than the reference group and the amount is increasing with the age at immigration. The age at immigration starts to matter after the age 7. Further, the received welfare benefits expand substantially from later arrivals than age 18 (see appendix E.1).

All in all, we observe following effects in the regression with immigration age as the dependent variable:

- Negative relationship between immigration age and educational attainment when the individual arrives after the age 6.
- Positive relationship between immigration age and unemployment days if the individual arrives after the age 4.

- Negative relationship between immigration age and logarithm of wage both for the first- and second-generation immigrants.
- Positive relationship between immigration age and welfare benefit if the individual arrives after the age 7.

When we exclude the family fixed effects and include additional controls, the results change in magnitude for some of our dependent variables (see appendix E.1). The average negative effects from later arrival are stronger for the second-generation immigrants, except the income. The impact on education is negative both for first- and second-generation, however the effect for first-generation immigrants lacks significance before age 9. Interestingly, the coefficients for welfare benefits for second-generation immigrants, whose parents arrived to Sweden more than 8 years ago, become significant and negative, if fixed effects are excluded. This seems reasonable, since the integration of parents before the birth of immigrants is important for the integration of children into the labour market. Wage have a similar trend, but lower magnitude in comparison with the regression including fixed effects.

When analysed for men and women separately, the trend is similar (see appendix E.1). The magnitude of coefficients differ, with slightly stronger effect of age at immigration for the female part of the sample.

To conclude, we find significant negative impact of age at immigration on all the dependent variables, in comparison with the the reference group. We also see indications that the analysed effects are stronger for the first-generation immigrants than for the second-generation immigrants, which is consistent with the initial expectations.

8.2 Structural breaks

In order to analyse if age at immigration has a specific breakpoint, we estimate two Wald tests for an unknown structural break in the data. We begin with analysing the critical threshold solely for the sample of childhood immigrants below the age 18. The first table presents the results. Next, we analyse the full sample, up to age 26. We do not analyse the effect on education for the older sample (aged 18-26), since there is risk of misinterpretation of the results due to the fact that we do not know whether the education has been acquired in the host country or in the source country, as seen in the section above. The estimate finds the breakpoints at the age at immigration of 11-12 years for the sample of children up to age 18. In other words, the effect of age at immigration changes significantly depending on if the immigrant arrives before or after the puberty. For the full sample, we see that it matters if the individual completes the secondary school in the home country or not, since the breakpoints are around age 18. We find some indications of a second breakpoint at age zero when plotting the Wald statistic for all ages, the results of the Wald test are illustrated graphically in appendix E.2.

Table 5: Wald test

	Attained education	Unempl. days	Wage (log)	Welfare benefits
Up to age 18				
Estimated Break age	11	12	12	13
Statistics	55.1952	44.1952	7.4669	24.4435
p-value	0.0000	0.0000	0.0839	0.0000
Total ages	-10 to 18	-10 to 18	-10 to 18	-10 to 18
Trimmed ages	-5 to 17	-5 to 17	-5 to 17	-5 to 17
Number of observations	66,578	70,947	66,673	71,555
Up to age 26				
Estimated Break age	-	17	18	19
Statistics	-	100.2828	189.2309	65.3146
p-value	-	0.0000	0.0000	0.0000
Total ages	-	-10 to 26	-10 to 26	-10 to 26
Trimmed ages	-	-4 to 21	-4 to 21	-4 to 21
Number of observations	-	108,430	86,124	111,615

Note: The estimated use family dummies to allow for the time series of immigration age. No difference in the estimated coefficient compared to the family fixed effect estimate.

8.3 Spline regression

To make the estimates more precise, we conduct a linear regression with the possibility to change slope in two places. We decided to set the first split at prior to and the second post birth immigration. The threshold is interesting due to the fact that the first and second-generation immigrants may have totally different characteristics even if the dummy graph illustrating the estimations showed little difference in slope. The zero age split is used for all dependent variables. Further, we analyse if it matters to come before or after completing the secondary education, at age 18. Variables *wage* and *education* are estimated using the transformation explained in the section about the empirical strategy. For welfare benefits and unemployment days, we use *Business Cycle Barometer* as a control for time trends.

In the table below, it can be seen that a significant negative effect of coming later to Sweden for all the generations of immigrants studied and all the dependent variables of interest is found. The effect of late arrival on educational attainment is larger in magnitude for first-generation immigrants. When it comes to the effect on unemployment days, the older the immigrant, the higher the number of unemployment days due to one year later arrival to Sweden. Interestingly, the second-generation immigrants are more affected, in terms of lower earnings, than first-generation immigrants. Although, the difference in slope is minimal. The effect of coming one year later on adjusted logarithm of wage is negative. The effect is largest for the oldest sample. The estimated relationship on welfare benefits is negative for the sample of younger immigrants, which means that later arrival results in less welfare benefits received in comparison with age-zero immigrants. However, the effect becomes positive for adult immigrants. The estimated coefficient for the variable *Business Cycle Barometer* indicates that the higher value of the indicator implies lower number of unemployment days (see appendix E.3 for results with controls). In other words, the improved economic state results in lower number of unemployment days among immigrants.

Furthermore, the performed F-tests result in the following remark: the sensitive thresholds that are assumed to be at age 18 and age 0 exist. The F-test shows that we cannot reject the null hypothesis at 1% level that there is no structural break at age 0. We see that the difference in slope in the estimated effect between the second-generation and the younger part of first-generation is minimal. The most important conclusion is that the differences in the slope between the younger and older part of first-generation immigrants are significant and it is an evidence of an existence of a structural break.

Table 6: Linear spline estimate - refugee sample

VARIABLE	Attained education	Unempl. days	Wage (log)	Welfare benefits
Immage [-10;0]	-0.0319***	0.3442	-0.496***	-0.256
	(0.0123)	(0.260)	(0.0633)	(0.381)
Immage [0;18]	-0.0633***	0.8567***	-0.449***	-0.653***
	(0.00562)	(0.091)	(0.0241)	(0.136)
Immage [18;26]	-	1.447***	-0.752***	10.27***
	-	(0.201)	(0.0571)	(0.295)
Firstborn	0.0999***	7.492***	0.161	16.28***
	(0.0316)	(0.564)	(0.154)	(0.868)
Boom (at age18)	-	0.999*	-0.618***	5.481***
	-	(0.578)	(0.152)	(0.874)
Business Cycle Barometer	-	-0.540***	-	0.0623
	-	(0.0498)	-	(0.0741)
Female dummy	0.769***	-9.810***	-5.133***	-10.54***
	(0.0266)	(0.535)	(0.143)	(0.816)
Observations	66,578	108,43	86,124	111,615
R-squared	0.676	0.601	0.679	0.667
Family Fixed Effect	33,226	52,171	36,956	54,605
Df	33,347	56,251	49,161	57,002
Same slope p-value [-10;-1] [0;18] [18;26]	-	0.0000	0.0000	0.0000
Same slope p-value [-10;0][0;18]	0.0211	0.1388	0.5334	0.3747
Same slope p-value [0;18] [18;26]	-	0.0007	0.0000	0.0000
Family FE	YES	YES	YES	YES

Note: Immage is the age at immigration interval for the linear estimate. Boom (at age 18) is a dummy for recession or boom in the business cycle when turning 18 years old. Business Cycle Barometer indicate economic state at observation year, see appendix G.1. Standard errors in parentheses ***p< 0.01, **p <0.05, *p<0.1

8.3.1 Heterogenous effects: gender, region

In this section we investigate the gender differences in the estimated effect, illustrated in the table below. In the performed F-tests, we look if the slopes of the trends for women versus men are the same.

Prima facie, while performing all the estimations women have always better outcomes than men in terms of higher educational level, less unemployment days, less welfare benefits received. In spite of that, they face substantial disadvantage in terms of lower earnings.

To begin with, the second-generation females are more negatively impacted by later arrival of the parents than the male part of the population. Regarding the younger part of the first-generation, the negative effect on education and income is higher for female immigrants than male. However, the effect on unemployment is lower for women than for men. The same patterns are seen within the sample of adult immigrants. The positive effect on welfare benefits is stronger for the adult female part of the sample than for other groups.

The null hypothesis that the slope of the trend for earnings is the same for both males and females for all age groups can be rejected. Regarding the rest of the sample, we can reject the null hypothesis on 5% level for both the younger and older first-generation of the sample for all the outcomes, except the welfare benefits.

Table 7: Linear spline estimate - refugee sample gender

VARIABLE	Attained Education	Unempl. days	Wage (log)	Welfare benefits
Female (F)				
Immage [-10;0]	-0.0114	1.346***	-0.597***	-0.672
	(0.0216)	(0.409)	(0.108)	(0.579)
Immage [0;18]	-0.0729***	1.797***	-0.408***	-0.00892
	(0.0102)	(0.164)	(0.0452)	(0.234)
Immage [18;26]	-	1.828***	-1.188***	10.24***
	-	(0.387)	(0.118)	(0.546)
Observations	31,482	54,827	40,867	56,197
R-squared	0.826	0.807	0.840	0.843
Male (M)				
Immage [-10;0]	-0.0187	0.149	-0.497***	-0.401
	(0.0209)	(0.463)	(0.113)	(0.749)
Immage [0;18]	-0.0568***	1.912***	-0.361***	0.197
	(0.00959)	(0.195)	(0.0485)	(0.321)
Immage [18;26]	-	3.812***	-0.577***	8.004***
	-	(0.541)	(0.136)	(0.870)
Observations	35,096	53,603	45,257	55,418
R-squared	0.795	0.835	0.845	0.838
Same slope M/F p-value [-10;0]	0.6561	0.0981	0.0000	0.0231
Same slope M/F p-value [0;18]	0.0000	0.0154	0.0000	0.1721
Same slope M/F p-value [18;26]	-	0.0000	0.0001	0.0000
Controls	YES	YES	YES	YES
Family FE	YES	YES	YES	YES

Note: The same slope test is against the null hypothesis no difference between Mate(M) and female (F) at different age intervals. Control estimates and slope difference within gender, see appendix E.3. Standard errors in parentheses ***p<0.01, **p<0.05, *p<0.1

All in all, the results have strong trends in line with the outlined hypotheses for all the dependent variables of interest. The results follow the same patterns if the estimations are separated by gender. In the sample consisting of men we find a stronger effect on employment, while in the sample consisting of women, we find stronger effects on education and earnings. Further, the breakpoints have been found in early puberty and after completing the secondary education. The breakpoints have been confirmed in the F-tests.

When the impact due to differences in geographical distance is analysed, the linear model is estimated within the sample of first-generation immigrants from different regions. It has been assumed that there are not only differences in levels of the difficulty to acculturate between the siblings, but also differences in the overall trend. Therefore, even though variation in family characteristics is absorbed, the effects due to cultural influences can be seen.

To begin with, the trends are very similar for the differing regions. The region with strongest effects of age on labour market performance is Asia. The studied relationship is weakest for immigrants from South America, especially in attained education. The later arrival of geographically close Yugoslavian immigrants is not as determining for the integration in labour market as for the Asian and African refugees. However, the European subgroup has a strong negative trend in attained education.

Table 8: Linear estimate - region specific sample

REGION	obs.	Attained Education		Unempl. days		Wage (log)		Welfare benefits	
European	12,170	-0.0994***	(0.0154)	0.977**	(0.386)	-0.313***	(0.0881)	0.842	(0.582)
West Asia	30,333	-0.0574***	(0.00773)	1.808***	(0.188)	-0.602***	(0.0547)	1.683***	(0.299)
Central Asia	6,004	-0.147***	(0.0201)	2.289***	(0.434)	-0.341***	(0.124)	1.909***	(0.507)
Africa	4,715	-0.0845***	(0.0204)	1.214**	(0.495)	-0.320**	(0.128)	0.236	(1.047)
South America	10,362	-0.00656	(0.0137)	1.797***	(0.326)	-0.231***	(0.0865)	0.816	(0.553)

Note: Estimates use first-generation immigrants aged 0 to 18 years. Family fixed effects specifications included. Standard errors in parentheses ***p<0.01, **p<0.05, *p<0.1

Looking more specifically at the country level, we see that the negative effect of education is strongest for Vietnam and Somalia, countries that are both culturally and geographically distant from Sweden. The Peruvian refugees are mostly affected when it comes to the unemployment. The refugees from Uganda and Lebanon suffer from lower earnings due to later arrival more than any other group. The effect on welfare benefits is strongest for Pakistanian refugees. However, there is bias due to differences in sample sizes between the regions and the results should be interpreted with precaution. Despite the fact that a large part of the regional variation is removed when we absorb fixed effects, the pattern is still clear: the more culturally distant the region of origin is, the stronger the effect for the individual.

Table 9: Linear estimate - country specific sample

	COUNTRY	obs.	Attained Education	Unempl. days	Wage (log)	Welfare benefits
Europe	Yugoslavia	6,001	-0.106*** (0.0199)	0.974* (0.517)	-0.414*** (0.119)	1.418* (0.804)
	Bosnia	6,169	-0.0821*** (0.0247)	1.062* (0.590)	-0.192 (0.134)	-0.220 (0.852)
West Asia	Syria	2,247	-0.0780*** (0.0278)	1.912*** (0.671)	-0.710*** (0.181)	0.797 (0.888)
	Turkey	8,795	-0.0962*** (0.0133)	1.123*** (0.317)	-0.512*** (0.0968)	1.334*** (0.349)
	Iraq	6,406	-0.0706*** (0.0217)	1.435** (0.566)	-0.635*** (0.147)	2.791*** (0.990)
	Iran	8,408	-0.0334** (0.0163)	2.326*** (0.342)	-0.548*** (0.111)	1.913*** (0.652)
	Lebanon	4,477	-0.0275 (0.0182)	2.446*** (0.493)	-0.814*** (0.139)	1.431 (0.883)
Central Asia	Vietnam	2,515	-0.148*** (0.0250)	2.691*** (0.612)	-0.317* (0.168)	2.324*** (0.524)
	Afghanistan	820	-0.0576 (0.0667)	1.246 (1.566)	-0.317 (0.434)	-0.433 (2.528)
	Bangladesh	373	-0.0876 (0.0654)	3.799** (1.553)	-0.689 (0.480)	3.467* (1.827)
	Pakistan	199	-0.150 (0.0973)	4.496** (2.147)	-0.655 (0.648)	7.667*** (2.755)
	Sri Lanka	2,237	-0.0971 (0.0662)	0.0268 (1.062)	-0.417 (0.341)	0.195 (1.419)
Africa	Ethiopia	1,126	-0.0464 (0.0395)	2.476** (0.967)	-0.218 (0.272)	-0.891 (1.893)
	Somaila	1,997	-0.160*** (0.0398)	1.200 (0.961)	-0.685*** (0.239)	-0.929 (1.893)
	Gambia	357	0.0394 (0.0789)	0.808 (1.945)	0.0283 (0.673)	-1.277 (4.923)
	Uganda	495	-0.0308 (0.0567)	0.244 (1.244)	-0.805** (0.395)	3.067 (2.853)
	Erithrea	740	-0.0684 (0.0420)	1.079 (1.005)	-0.252 (0.250)	1.650 (1.577)
South America	Chile	7,512	-0.00375 (0.0154)	1.383*** (0.376)	-0.196* (0.100)	0.395 (0.617)
	Bolivia	386	-0.0358 (0.0618)	4.875** (2.036)	-0.377 (0.424)	4.460** (2.157)
	Peru	629	-0.0201 (0.0520)	5.570*** (1.286)	-0.701** (0.347)	2.621 (2.511)
	Colombia	1,835	0.0527 (0.0606)	0.743 (1.230)	-0.171 (0.352)	1.180 (2.579)

Note: Estimates use first-generation immigrants aged 0 to 18 years. Family fixed effects specifications included. Standard errors in parentheses ***p<0.01, **p<0.05, *p<0.1

9 Sensitivity analysis

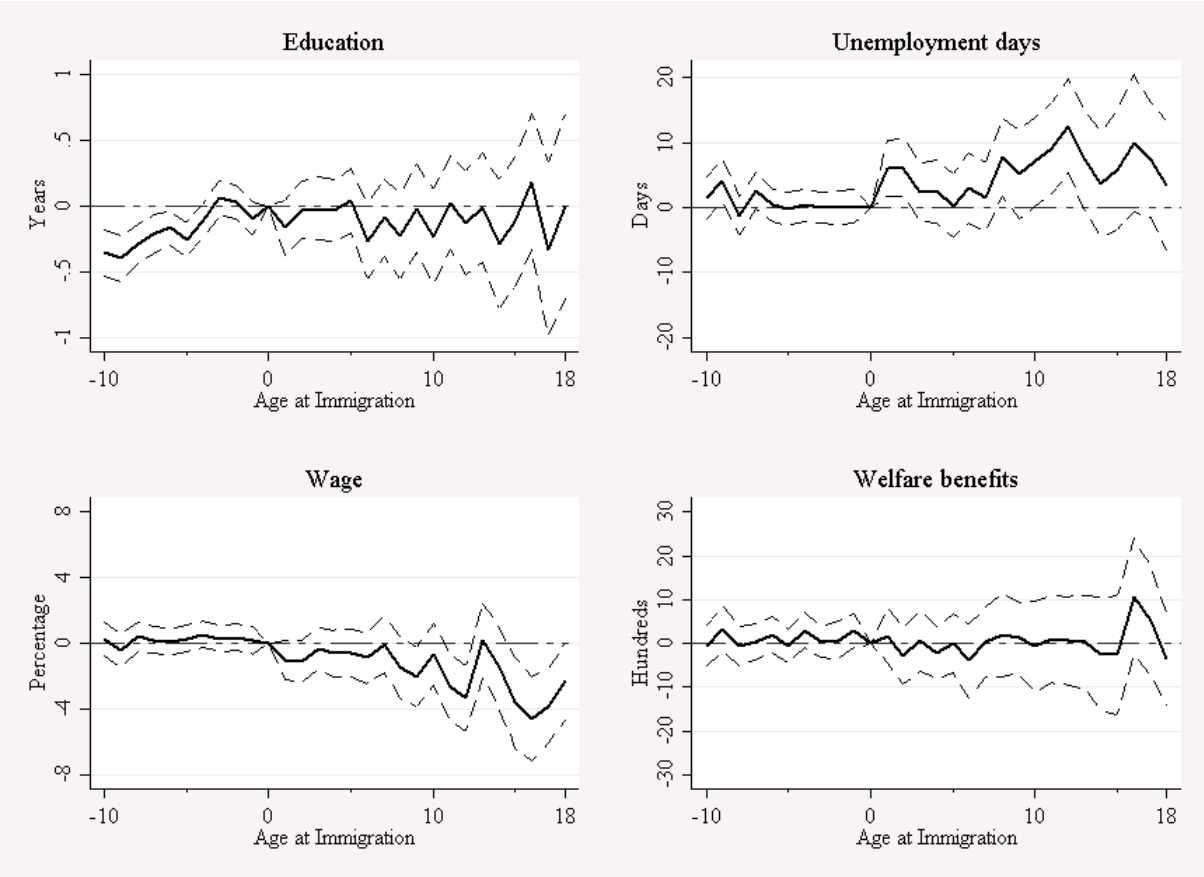
In this section we verify the stability and robustness of our results. In addition, we compare and check the consistency of the results for the refugee group with the samples of immigrants from geographically closer countries and the parentless immigrants. The analysis of the differences between the refugees and labour migrants highlights the difference between the results on migration due to exogenous shocks and migration due to economic reasons. By measuring the outcomes of the orphans, compared to other same aged children, we test for the negative effect of arriving without parents. Secondly, we test main regression equation using the same framework, but later in life, within the 34 to 36 age interval. When measuring the outcomes at a later stage in the immigrant working career gives us a good indication if the results hold in long-term and if the outcomes in ages 28-30 are a good approximation of the life-time outcomes.

9.1 Other immigrants

The subsample consisting of other immigrants include the immigrants classified as "labour migrants" from geographically closer countries. Thus, this sample contains mainly the immigrants that planned their migration and had the opportunity to influence the age at immigration. In comparison with the refugee sample, the analysed effect for the other

immigrants that has not experienced an exogenous shock is visibly weaker. In the figure below, it can be seen that the pattern for the second-generation migrants lacks significant difference from the reference group, except for a positive trend in education up to 4 years before birth. The first-generation migrants have weak effects from later arrival. The standard errors are high for the first-generation and the estimated values fluctuate significantly.

Figure 5: Binary variable - other immigrants



Note: Age at immigration as binary variable - The estimates and 95% confidence intervals from family fixed effects specifications. The reference group is the age-zero immigrant. All outcomes are observed at age 28 to 30. Beta coefficients are found in appendix F.

Further, the table below presents the estimates from the linear model. The coefficients for the other immigrants are lower and lack significance, but the trend goes in the same direction as for the refugee sample. For the second-generation immigrants, the only significant relationship is found for the level of attained education. However, the estimated relationship is positive, which is counterintuitive. For the younger part of the first-generation subgroup, the results are in line with the initial hypotheses. Nonetheless, we cannot reject the null hypothesis that there is no effect on welfare benefits and education. Within the last, oldest group of adult immigrants, no significant effects are found, despite the negative effect of later arrival on welfare benefits. In comparison with the refugee sample, we find significant differences for all the outcomes of interest, except welfare benefits. Overall, all the slopes for the estimated trends are steeper for the refugee sample.

Table 10: Linear spline estimate - other immigrants

VARIABLE	Attained education	Unemp. days	Wage (log)	Welfare benefits
Immage [-10;0]	0.3852***	-0.1355	-0.0442	-0.141
	(0.00663)	(0.149)	(0.0413)	(0.211)
Immage [0;18]	-0.0098	0.294***	-0.121***	-0.0614
	(0.00815)	(0.0902)	(0.0237)	(0.121)
Immage [18;26]	-	0.2665	-0.0894	0.749*
	-	(0.297)	(0.0812)	(0.389)
Firstborn	0.199***	3.011***	0.2633*	0.278
	(0.0363)	(0.574)	(0.1587)	(0.810)
Boom (at age 18)	-	-0.239	0.135	1.438*
	-	(0.589)	(0.162)	(0.830)
Business Cycle Barometer	-	-0.291***	-	0.0153
	-	(0.0514)	-	(0.0693)
Female dummy	0.641***	-3.782***	-6.195***	-3.580***
	(0.0321)	(0.568)	(0.158)	(0.803)
Observations	49,109	74,205	70,237	80,313
R-squared	0.743	0.688	0.715	0.631
Family Fixed Effect	20,044	29,792	25,802	32,469
Df	29,060	44,405	44,428	47,836
Same slope p-value [-10;-1] [0;18] [18;26]	-	0,0101	0.3114	0.1120
Same slope p-value [-10;0][0;18]	0.0029	0,019	0.1381	0.7645
Same slope p-value [0;18] [18;26]	-	0,0057	0.7424	0.0862
Family FE	YES	YES	YES	YES
Same slope Other/Refugee p-value [-10;0]	0.0000	0.0526	0.0000	0.2654
Same slope Other/Refugee p-value [0;18]	0.0000	0.0000	0.0000	0.0063
Same slope Other/Refugee p-value [18;26]	-	0.0000	0.0000	0.0000

Note: Immage is the age at immigration interval for the linear estimate. Boom (at age 18) is a dummy for recession or boom when turning 18 years old. Standard errors in parentheses ***p<0.01, **p<0.05, *p<0.1

We see that the reason for immigration matters for the analyzed effect. The effect becomes significantly stronger with not only geographical distance, but also the fact that refugees cannot plan and choose the best time to arrive to Sweden. Despite the fact that we use family fixed effects, there are differences in trends among the other immigrants and refuge siblings that have an impact on the results. Therefore, the effect on the sample of other migrants is lower in magnitude and less significant. However, it is still in line with the initial hypotheses. This is the explanation for why this study finds stronger effects than previous research on the topic.

9.2 Orphan sample

This is a very interesting sample of refugees that encounter an exogenous shock in terms of that they are forced to immigrate, but in the majority of cases not solely due to political conflicts. The orphans come to Sweden as potential adoptees. This part of the sample has no connections with the culture of the country of origin, since they arrive

without the parents. Orphan siblings are here a binary variable equal to one if parentless, compared to the children arriving with parents. Therefore, instead of using family fixed effects, region or year fixed effect estimations are performed, in order to control for calendar effects and regional variations. In the table below we see the effect of coming to Sweden as a child 0-5 years old without the parents on the performance on labour market compared to same aged children with parents. It is important to emphasize that the parentless children in the sample immigrate from three countries: Chile, Colombia and Sri Lanka. This is in line with the statistical reports by Migrationsverket (2016).

Table 11: Binary variable - orphans

VARIABLE	Attained Education		Unempl. days		Wage (log)		Welfare Benefits	
Orphans dummy	0.0745	-0.211**	-3.185**	-0.170	1.005**	0.581	-6.207**	-2.312
	(0.0769)	(0.0970)	(1.373)	(1.742)	(0.393)	(0.499)	(2.769)	(3.515)
Boom (at age18)	-	-	-1.095	-1.283	0.0223	-0.269	-5.090*	-2.258
	-	-	(1.504)	(1.230)	(0.432)	(0.353)	(3.033)	(2.483)
Firstborn	0.414***	0.413***	-2.268	0.119	0.366	0.312	-1.628	-0.00758
	(0.0781)	(0.0736)	(1.394)	(1.319)	(0.396)	(0.376)	(2.813)	(2.662)
Female dummy	0.827***	0.807***	-7.899***	-7.147***	-3.524***	-3.593***	-7.064***	-6.200**
	(0.0687)	(0.0688)	(1.226)	(1.227)	(0.351)	(0.352)	(2.474)	(2.476)
Business Cycle Barometer	-	-	-	-0.147	-	-	-	-0.298
	-	-	-	(0.122)	-	-	-	(0.246)
Observations	6,613	6,613	6,624	6,624	5,891	5,891	6,624	6,624
R-squared	0.025	0.025	0.008	0.006	0.018	0.018	0.003	0.001
Number of years	17	17	18	18	18	18	18	18
Number of countries	3	3	3	3	3	3	3	3
Year FE	YES	NO	YES	NO	YES	NO	YES	NO
Country FE	NO	YES	NO	YES	NO	YES	NO	YES

Note: Immage is the age at immigration interval for the linear estimate. Boom (at age 18) is a dummy for recession or boom when turning 18 years old. Standard errors in parentheses ***p<0.01, **p<0.05, *p<0.1

To begin with, the effect on educational level is ambiguous; it changes with the two different estimation models applied. There is a negative effect on the number of unemployment days for coming to Sweden without parents in comparison with coming to Sweden with the entire family. It can be interpreted that the parentless children perform better in the labour market in adult life, compared to those who arrive with their parents. In general, the parentless children receive less welfare benefits in adulthood. These remarks indicate that parentless children integrate better in the Swedish society than those who arrive to Sweden with the entire families. In this section the binary and spline estimates are presented for the same refugee sample as used in the result section, in order to extend the validity and see how the results differ if observed at another point in time.

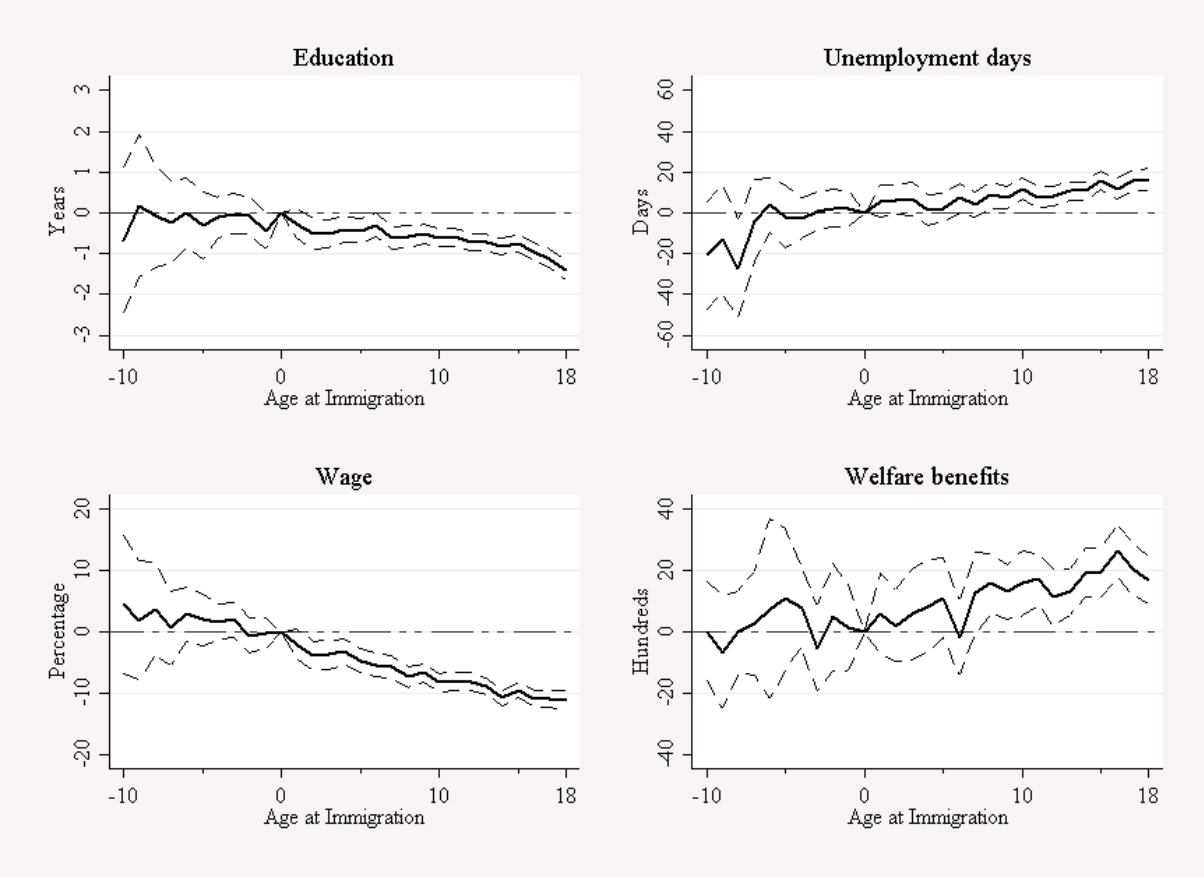
9.3 Long-term validity

The binary and spline estimates are presented for the same refugee sample as used in the result section, in order to extend the validity and see how the results differ if observed at another point in time.

The binary estimates do not change significantly, when the outcomes are studied within the age interval 34-36. The figure below shows that there is no statistical difference from the age-zero immigrant, but the standard errors are large. Within the sample of

first-generation refugees that arrive before primary school, the effects of later arrival are significant in comparison with the reference group.

Figure 6: Binary variable - refugee sample aged 34 to 36



Note: Age at immigration as binary variable - The estimates and 95% confidence intervals from family fixed effects specifications. The reference group is the age-zero immigrant. All outcomes are observed at age 34 to 36. Beta coefficients are found in appendix F.

In the linear model the coefficients estimated have the same slope as in the previous estimations. Regarding the first outcome, education, the impact is still negative, yet not significant for the second-generation estimates. In comparison with the estimates at ages 28-30, the effect on unemployment and earnings increases for the second-generation immigrants, and decreases for the younger part of first-generation immigrants. This is a puzzling result that indicates that the assimilation of parents matters more for the economic integration than assimilation of the individual. Although, it is important to highlight that the sample size of the second-generation immigrants is now small and the standard errors are too large to draw any reliable conclusions. The effect on welfare benefits is weak in comparison with the estimates in the main regression, but goes in the positive direction: the older the individual when immigrating, the higher amount of welfare benefits received.

Interestingly, when looking at the control variable indicating if the individual i is born first among the siblings or not, we see that it has a significant positive impact on the performance of the individual in all the previous estimations. Although, while looking at this variable when the outcomes are observed later in life (ages 34-36), the coefficient is lower in magnitude and changes sign. This means that birth order matters in the beginning of the work-life and loses significance the older the individual becomes.

Table 12: Linear spline estimate - refugee sample (age 34 to 36)

VARIABLE	Attained education	Unempl. days	Wage (log)	Welfare benefits
Immage [-10;0]	-0.0223	1.391**	-0.833***	0.195
	(0.0324)	(0.617)	(0.165)	(0.800)
Immage [0;18]	-0.0773***	0.395***	-0.383***	0.0132
	(0.00946)	(0.142)	(0.0387)	(0.183)
Immage [18;26]	-	2.219***	-0.497***	1.877***
	-	(0.247)	(0.0702)	(0.347)
Firstborn	0.206***	2.234***	-0.403*	4.455***
	(0.0510)	(0.768)	(0.210)	(1.128)
Boom (at age18)	-	-0.886	-0.184	-2.338**
	-	(0.813)	(0.218)	(1.177)
Business Cycle Barometer		-0.121*		0.0929
		(0.0733)		(0.104)
Female dummy	0.611***	-1.889***	-4.195***	-5.375***
	(0.0424)	(0.724)	(0.198)	(1.055)
Observations	30,989	70,772	53,359	71,884
R-squared	0.745	0.682	0.771	0.709
Family Fixed Effect	11,935	26,097	16,352	26,949
Df	19,049	44,667	37,000	44,927
Same slope p-value [-10;-1] [0;18] [18;26]	-	0.0000	0.0419	0.0000
Same slope p-value [-10;0][0;18]	0.1127	0.1272	0.0133	0.8402
Same slope p-value [0;18] [18;26]	-	0.0000	0.1970	0.0000
Family FE	YES	YES	YES	YES

Note: Immage is the age at immigration interval for the linear estimate. Boom (at age 18) is a dummy for recession or boom when turning 18 years old. Standard errors in parentheses ***p<0.01, **p<0.05, *p<0.1

The results are robust to the changes in observation year and the patterns are similar, independently of when the outcome is studied. Thus we can reject the null hypothesis that there are no differences due to the age at arrival once again. Therefore, we can conclude that the chosen age when the results are observed (28-30 years) can be used as an approximation of the lifetime outcomes.

10 Discussion

This study finds evidence that labour market integration declines appreciably with age at immigration. The general effect might be due to the fact that it is more difficult to acquire Sweden-specific skills, such as the language, the higher the age upon arrival. Older immigrants may be less able to adjust to the cultural and linguistic challenges. At the same time less acculturation may lead to discrimination and lower returns to education. Ease of acculturation is inversely related to the age at immigration, and this is the main driver of our findings. In addition, we see that there are periods when the effect of age at immigration is amplified.

The effect is inherited through different generations, with the second-generation im-

migrants affected negatively by later arrival of the parents before their birth. This shows that parental assimilation has implications for the children in adult life. Plausibly, the relationship between age at immigration and the economic outcomes for second-generation immigrants is not as strong, as for the first-generation immigrants. It indicates that in long term the negative effect of age at immigration diminishes as the family becomes fully assimilated into the Swedish society.

Specifically, the results are in line with the initial hypotheses. When we conduct the Wald test, we see that the critical threshold for education is in line with the findings of previous studies - at age 11. This result goes hand in hand with CPH and the theory developed by Scovel (1988). According to the second language theories, the breakpoint is around the age 12. This indicates that implementation of efficient policies in order to enhance acquisition of Swedish language is crucial for the integration of refugees into the labour market. The strongest negative impact on education is found in the regression with binary variables at ages 16-18, which means that it is better to rather come before that period or complete the education in the home country. The proposal to prolong the compulsory school attendance in Sweden to the age 20 for immigrants will increase the educational attainment and give the possibility to acquire the Sweden-specific skills. It has potential to be an efficient tool in order to diminish the negative effects of coming late to Sweden on labour market performance. On the other side, we see that there are significant differences depending on when the immigrant arrives to Sweden and on the source country. Therefore, making it compulsory to attend school to age 20 will be unnecessary waste of resources, resulting in deadweight loss.

The unemployment days increase with later immigration to Sweden. The threshold is around age 12 for the sample of childhood immigrants, consistent with the theories about second language acquisition. The negative effect on unemployment days due to later arrival is twice as strong for male in comparison with female. Indirectly, the later the refugee arrives to Sweden, the harder to find a job and the trend is more negative for men than for women.

The negative impact on logarithm of wage in comparison with the reference group of age-zero immigrants is significant in all the regressions. In addition, the negative relationship is stronger for women than for men. Considering the fact that the effect is reversed when looking at employment it can be interpreted as the age at immigration does not matter for female unemployment, but matter significantly for male. Although, the incoming women receive the jobs of lower quality the later they arrive and the income decreases due to later arrival. Previous studies find that the negative effect on labour market performance is stronger for men than women due to the in general lower earnings of women, and therefore less room for discrimination at the lower end-scale of wage distribution. However, we see that there is a stronger negative effect of age at arrival within the sample of female immigrants on earnings and education, which means that there are mechanisms related to acculturation that influence the fact that women on average earn less than men.

Further, the critical age for earnings has been found at the age of 12, when analysing the below age 18 sample; and at the age of 18, when analysing the full sample. Therefore, in addition to the linguist aspects, it matters whether the individual arrives before or after the graduation from secondary school. This an additional evidence that educational attainment leaves a permanent scar on lifetime labour market performance. It seems to be beneficial to complete the education in the source country, however there is risk that the education received will not be recognized in Sweden.

Regarding the welfare benefits, the topic about the contribution of refugees to the public finance and whether they are stuck with welfare is on the top of agenda of politicians right now. We see that there is a negative pattern with age at immigration increasing the amount of money received from the state. This can be interpreted in fact as if the time in Sweden increases, the welfare benefits the refugee receives decrease, which is the opposite of the allegation that there are no incentives to work for the immigrants. The time is necessary to acquire the language and cultural knowledge. The younger the refugees come, the lower the amount of welfare benefits received. The second-generation immigrants receive less welfare benefits than age-zero immigrants, which means that the effect is not inherited through generations. To conclude, we do not see any clear evidence on that immigrants arriving to Sweden are stuck with welfare. In contrary, we see that the unemployment rate decreases, the longer they stay in Sweden.

We see that all the effects are amplified by cultural and geographical distance. We also observe that it matters whether there is room for planning of the time of immigration or if the immigrants need to flee due to an exogenous shock. Therefore, the new waves of refugees encounter more challenges and find it more difficult to integrate into the labour market than the previous waves of labour migration. It should be considered in the new integration policies and programs. Since the refugees come to Sweden due to an exogenous shock, there is little possibility to influence the age at immigration. Therefore, it is crucial for policymakers to consider the differences in various samples of immigrants and address the problems of distinctive levels and types of problems for the immigrants in different cohorts. The "one size fits all" approach is not feasible to integrate efficiently the refugees; since the differences within this group are considerable.

Furthermore, the new law implemented that limits the family reunification means that immigrants arriving without their families will become more common than they currently are. It also means that this type of study will be more difficult to conduct in the future, since sibling pairs immigrating to the country will become more infrequent. The analysis of the orphan sample arriving independently and individually is important for the future, since more people will arrive without the families. Overall, there is a positive effect of coming to Sweden without parents on labour market performance. The potential explanation for that could be the fact that parentless children are forced to integrate into the Swedish society faster, independently of their age. They cannot rely on the help from the biological parents and therefore have no choice, but acquire the language and adapt to the culture as efficiently as possible. The cultural ties and the connections with the country of origin are not as strong as for the children arriving with parents and the effect of later arrival is not as significant.

This is a broad study of the relationship between age at immigration and outcomes that mirror the integration into the labour market. The results are in line with the outcomes from the studies of the relationship in other countries; this shows that the conclusions can be generalized and applicable to other contexts. The results are robust to changes in the assumptions, in the selection of the sample and limitations that has been made throughout the whole study. The unique register dataset used in this study made it possible to isolate the effect of within family variations. This enabled us to eliminate a large proportion of the bias; something that the studies on immigration in other countries did not manage to do. A main drawback of this study is that some samples analysed are too small to make any reliable conclusions (specifically, the sample of 34-36 ages immigrants used in the sensitivity analysis). Unlike the existing studies, this study does not have the problem of attrition bias, since none of the individuals included in the

studied samples have emigrated during the period of analysis. We manage to not only minimize the selection bias in the analysis, but also the bias due to differences between different cohorts. Since we adjust for the calendar effects by comparing with the age-zero immigrants, comparing only the differences within families and the economic state when the individual completes the secondary education, the bias due to cohort differences is significantly reduced. However, there is still risk that the individuals differ due differences in how skillful a specific cohort is due to varying quality of institutions, differences in trends in second language acquisitions and other aspects that may influence the differences in integration among siblings. This would be a threat to the rather strong assumption that the siblings would perform equally well in the absence of immigration. Nevertheless, previous studies show that cohort groups of immigrants are not very different from each other, (Borjas, 1985, 1995), so this bias should not threaten the validity of this study.

In addition to that, we see that there are differences dependent on whether we apply family fixed effects or not. This means that the use of exogenous shocks does not reduce the selection bias entirely and that there is still a lot of unobserved heterogeneity within the sample of refugees. Therefore, the conclusion is that family fixed effects is a more feasible method for these types of studies. There are differences in the estimates for the refugee sample and other immigrants sample, as well as for the sample used in study by Åslund et al.(2015). This means that both the reason for immigration, geographical distance and cultural/linguistic differences influence the estimated effect. In other words, there are differences in not only levels, but also trends among the siblings. Every year more in the country of source will result in longer process of acculturation, which means that there are not only differences in the level of economic outcomes between the younger and older sibling, but also there are differences in the slope of the estimated trends. The trend of the difficulty to adapt to the new culture is positively related to the age at immigration. The family fixed effects method absorbs the differences in levels of difficulty to acculturate, but does not consider the differences in trends. Therefore, we see differences between the immigrants of different nationalities. If the language cultural/distance has a similar effect on both siblings, then it is absorbed when the family fixed effect regression is conducted. However, since we see that there are still differences between the regions in the estimated coefficients, it implies that the cultural distance alters the rate of convergence and the slope changes with the time to converge. The results imply that there is a differential trend in the acculturation of siblings, not only differential level. This possible source of bias and feasible ways to control for it, should be studied in depth in future research.

Consequently, the observed differences in country characteristics could be to a large extent driven by differences in the trends of the effect between the siblings. In the light of the results of the Wald test, the existence of different thresholds means that the result is two-folded: the younger sibling is not only negatively impacted by the later arrival in comparison with the older sibling, but there are also differences in the slope of the trend. Looking at the example of South America, the very mild effect of age at immigration could be related to the immigrant profile that is characterized by on average low age at immigration. If both siblings arrive before the threshold found in the Wald test, the estimated magnitude of the coefficients will automatically be lower, compared with the other regions when the siblings arrived at higher ages, above or overlapping the threshold. In other words, the sibling pair arriving to Sweden when the slope of the trend is flatter will have a weaker effect, in comparison with the sibling pair that arrive when the slope is steeper.

Unlike the majority of previous research, the chosen comparison group does not consist of natives, but of the age-zero immigrant, born in Sweden. Therefore, all the conclusions are drawn based on the comparison within the sample of immigrants. This has both advantages and drawbacks, in the sense that the effect of age at immigration is isolated, but sometimes hard to interpret.

There is a lot of room for future research that should focus more specifically on analysing the sample of orphans that is a very unique group and has been studied briefly in this paper. There is also demand for more studies on the contribution of refugees to public finance and the amount of social assistance received. In addition to that, there is a lot of research about the existing crowding out effects of wages of natives due to immigration, however it would be valuable to see these studies conducted on the refugees from recent immigration waves, since we see that the cultural and geographical distance matters.

To conclude, we see that there are substantial differences between the immigrants due to the reason for migration, culture, gender and age at immigration. The diversity of the estimated effect means that any integration and anti-discrimination policies must take into account complex interactions between timing of arrival, type of immigration, culture, gender and the educational background. It should be taken into consideration that every new influx of immigration has different characteristics and acculturate in different ways. However, the general pattern is: a day late, a Krona short.

10.1 Policy implications

Prima facie, looking at the results of the analysis, we propose to implement specially tailored integrational programs for immigrants arriving at different ages. The focus needs to be on extra education for the young individuals arriving at ages 12-18. The quality of language education should be improved and adjusted for differing needs and abilities of different age groups.

Since the economic outcomes are significantly impaired by later age at arrival, we see that acculturation is key for the integration of immigrants in the labour market. In order to facilitate the integration of immigrants, the interpretation of grades and recognition of foreign qualifications should be improved. Additionally, the supply of jobs without requirements on Sweden-specific skills, such as language, should increase. It is very important to implement rules in order to promote fair treatment of immigrants. Since we see that assimilation is enhanced when the individual arrives independently, without family, it indicates that the new law limiting family immigration should have positive effects on labour market integration of immigrants. On the other hand, the family structure is not the most important determinant of labour market performance. The age at immigration has a larger influence and the focus should be to reduce the negative impact for teenagers. We should also have in mind, that since the age effects are stronger for refugees than other migrants, the unsuccessful integration of teenagers will have severe consequences for the society in the future.

11 Concluding remarks

This is the first study that finds significant effects of age at immigration on the economic outcomes for the recent waves of refugees. We find two critical thresholds: i) around the puberty for childhood/teen immigrants and ii) around the age 18 when including a wider age span. The existence of breakpoints implies that it is crucial to focus on efficient language programs, especially for teenage refugees. Furthermore, we find that parental assimilation is important for the performance of refugees in the labour market, since the age effect exists among children of refugees born in the host country (second-generation immigrants). The robustness checks confirm that the effects are seen even in later career life of the refugee. It seems to be more beneficial to come to Sweden after the graduation from secondary school than just before. Arriving near the transition out of high school is associated with a permanent reduction in educational attainment, and this reduces the future earnings. In line with previous research, we find that the cultural and linguistic distance matters for the studied effect. This implies that cultural differences are an important factor in the integration of refugees into the labour market and should be considered in the immigration politics. Furthermore, we find that there are substantial differences in the studied effect between men and women, as well as between the refugees and labour migrants. All things considered, the heterogeneity of the effect means not only that the integration of refugees is more challenging than of the labour migrants, but also that their characteristics are more contrasting and influence the integration to a larger degree.

References

[1] Acemoglu, D., and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of Labor Economics*, 4, 1043-1171.

[2] Aguíiar, R., and Gusafsson, B. (1991). The earnings assimilation of immigrants. *Labour*, 5(2), 37-58.

[3] Asher, J. J., and García, R. (1969). The optimal age to learn a foreign language. *Modern Language Journal*, 53(5), 334-341.

[4] Åslund, O. Bohlmark, A., and Nordstrom Skans, O. N. (2009). Age at migration and social integration. (IZA Discussion Paper No. 4263). Available at SSRN: `https://ssrn.com/abstract=1434577`

[5] ——.(2015). Childhood and family experiences and the social integration of young migrants. *Labour Economics*, 35, 135-144.

[6] Åslund, O. Hensvik, L., and Skans, O. N. (2014). Seeking similarity: How immigrants and natives manage in the labor market. *Journal of Labor Economics,* 32(3), 405-441.

[7] Bantekas, A. (1992). *Three essays on discrimination and factor demand.* (Doctoral dissertation). Uppsala: Department of Economics, Uppsala University.

[8] Baron R., and Kenny D.A. (1986) The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*, 51(6), 1173–1182.

[9] Basu, S. (2010). Age-of-entry effects on education of immigrant children: a sibling study. (Working Paper). Aailable at SSRN 1720: `https://ssrn.com/abstract=1720573`.

[10] Beckley, A. L. (2015). Age at immigration and crime in Stockholm using sibling comparisons. *Social Science Research,* 53, 239-251.

[11] Bohlmark, A. (2008). Age at immigration and school performance: A siblings analysis using Swedish register data. *Labour Economics,* 15(6), 1366-138.

[12] ——.(2009). Integration of childhood immigrants in the short and long run–Swedish evidence. *International Migration Review,* 43(2), 387-409.

[13] Borjas, G. J.(1985). Assimilation, change in cohort quality, and the earnings of immigrants. *Journal of Labor Economics*, 3, 463-89.

[14] ——.(1987). Self-selection and the earnings of immigrants. *American Economic Review*, 77, 531-53.

[15] ——.(1991). Immigration policy, national origin, and immigrant skills: A comparison of Canada and the United States. in D. Card and R. B. Freeman (Ed.), *Small Differences that Matter* (pp. 21-43). Chicago: University of Chicago Press.

[16] ——.(1995). Assimilation and changes in cohort quality revisited: what happened to immigrant earnings in the 1980s?. *Journal of Labor Economics*, 13(2), 201-245.

[17] ——.(1999). The economic analysis of immigration. *Handbook of Labor Economics*, 3, 1697-1760.

[18] Brady, D. S., and Friedman, R. D. (1947). Savings and the income distribution. In *Studies in Income and Wealth* (pp. 247-265). New York: NBER.

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

[20] Brown, J. N. (1989). Why do wages increase with tenure? On-the-job training and life-cycle wage growth observed within firms. *American Economic Review*, 79(5,6), 971-991.

[21] Byman, D. (2005). *Deadly Connections: States that Sponsor Terrorism.* New York: Cambridge University Press.

[22] Coppieters, R. (1987). Competence differences between native and near-native speakers. *Language*, 63, 544-573.

[23] Chiswick, B. R. (1978). The effect of Americanization on the earnings of foreign-born men. *Journal of Political Economy*, 86(5), 897-921.

[24] ——.(1991). Speaking, reading, and earnings among low-skilled immigrants.*Journal of Labor Economics*, 9(2), 149-170.

[25] Cohen-Goldner, S., and Paserman, M. D. (2011). The dynamic impact of immigration on natives' labor market outcomes: Evidence from Israel. *European Economic Review,* 55(8), 1027-1045.

[26] Dagens Nyheter (2016, July 9). De nya hårda moderaterna. Retrieved from `http://www.dn.se/nyheter/politik/de-nya-harda-moderaterna/`.

[27] Dagens Nyheter (2016, June 20). Så vill regeringen skärpa asyllagarna. Retrieved from `http://www.dn.se/nyheter/politik/sa-vill-regeringen-skarpa-asyllagarna/`.

[28] Dagens Samhälle (2016, April 13). Svårt att förstå regeringens ändring om kvotflyktningar. Retreived from `http://www.dagenssamhalle.se/debatt/svart-foersta-regeringens-aendring-om-kvotflyktingar-24177`.

[29] D'Amuri, F., Ottaviano, G. I., and Peri, G. (2010). The labor market impact of immigration in Western Germany in the 1990s. *European Economic Review*, 54(4), 550-570.

[30] DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(04), 499-533.

[31] Edin, P. A., LaLonde, R. J., and Åslund, O. (2000). Emigration of immigrants and measures of immigrant assimilation: Evidence from Sweden. *Swedish Economic Policy Review,* 7(2), 163-204.

[32] Global Jewish Advocacy (2016, September 19). The UN, the Two Types of Refugees, and The Arab-Israeli Conflict. Retrieved from http://www.ajc.org/site/apps/nlnet/content3.aspx?c=7oJILSPwFfJSG&b=9077539&ct=14924723.

[33] Government Offices of Sweden (2015). Swedish Integration Policy. Retreived from http://www.government.se/contentassets/b055a941e7a2\47348f1acf6ade2fd876/swedish-integration-policy-fact-sheet-in-english.

[34] Gustafsson, B., and Zheng, J. (2006). Earnings of Immigrants in Sweden, 1978 to 1999. *International Migration*, 44(2), 79-117.

[35] Haider, S., and Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review*, 96(4), 1308-1320.

[36] Harley, B., and Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition*, 19(03), 379-400.

[37] Johnson, J. S., and Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology,* 21(1), 60-99.

[38] Lagakos, D., Moll, B., Porzio, T., Qian, N., and Schoellman, T. (2015). Life-cycle wage growth across countries. (Unpublished Manuscript). Princeton University.

[39] Le Grand, C., Szulkin, R., and Tåhlin, M. (2004). Overeducation or skill shortage? Matching in the Swedish labor market 1974-2000. In M. Bygren, M. Gähler, and M. Nermo (Ed.), *Familj och arbete. Vardagsliv i förändring* (pp.282-321). Stockholm: SNS Förlag.

[40] Le Grand, C., and Szulkin, R. (1999). Invandrarnas löner i Sverige. *Arbetsliv och Arbetsmarknad.* Stockholm: Arbetslivsinstitutet.

[41] Liebig, T., and Widmaier, S. (2009). Children of Immigrants in the Labour Markets of EU and OECD Countries. (OECD Social Employment and Migration Working Papers 97). Paris: OECD Publishing.

[42] Lillard, L. A. (1977). An Essay on Human Wealth. In *Explorations in Economic Research* (pp. 108-158). NBER.

[43] Lundborg, P. (2013). Refugees' employment integration in Sweden: Cultural distance and labor market performance. *Review of International Economics,* 21(2), 219-232.

[44] Malm F. (2016, October 25). Stärk individen och motverka parallelsamhällen. Retrieved from http://www.fredrikmalm.nu/2016/10/liberalernas-nya-integrationsprogram.html.

[45] Martinoia, M. (2011). European integration, labour market dynamics and migration flows. *European Journal of Comparative Economics*, 8(1), 97.

[46] Migrationsinfo (2016, July 13). Anhöriginvandring. Retrieved from https://www.migrationsinfo.se/migration/sverige/anhoriginvandring/.

[47] Migrationsverket (2014, April 14). Fakta om migration. Retrieved from http://www.migrationsverket.se/Om-Migrationsverket/Fakta-om-migration.html.

[48] Migrationsverket (2016, June 23). Historik. Retrieved from www.migrationsverket.se/Om-Migrationsverket/Fakta-om-migration/Historik.html.

[49] Mincer, J. (1975). Education, Experience, and the Distribution of Earnings and Employment: An Overview. In *Education, Income, and Human Behavior* (pp. 71-94). NBER.

[50] Moav, O. (2002). Income distribution and macroeconomics: the persistence of inequality in a convex technology framework. *Economics Letters*, 75(2), 187-192.

[51] Modigliani, F., and Brumberg, R. (1954). Utility analysis and the consumption function: An interpretation of cross-section data. In K.K. Kurihara (Ed.), *Post-Keynesian Economics* (pp.388-436). New Brunswick: Rutgers University Press.

[52] Munro, M. J., Flege, J. E., and MacKay, I. R. (1996). The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics*, 17(03), 313-334.

[53] Nekby, L., and Rödin, M. (2010). Acculturation identity and employment among second and middle generation immigrants. *Journal of Economic Psychology*, 31(1), 35-50.

[54] Nordin, M., and Rooth, D. O. (2007). Income gap between natives and second generation immigrants in Sweden: Is skill the explanation?. (IZA Discussion Paper No. 2759). Available at SSRN: https://ssrn.com/abstract=984423.

[55] OECD (2016). International Migration Outlook 2016, OECD Publishing, Paris. Retrieved from http://dx.doi.org/10.1787/migr_outlook-2016.

[56] Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, 5(3), 261-283.

[57] Patkowski, M. S. (1980). The sensitive period for the acquisition of syntax in a second language 1. *Language Learning*, 30(2), 449-468.

[58] Peterman, W. B. (2016). The effect of endogenous human capital accumulation on optimal taxation. *Review of Economic Dynamics*, 21, 46-71.

[59] Rashid, S. (2004). Immigrant earnings, assimilation and heterogeneity. (The S-WoPEc Working Paper Collection). Umeå: Umeå Economic Studies, Umeå University.

[60] Rooth, D. O., and Ekberg, J. (2003). Unemployment and earnings for second generation immigrants in Sweden. Ethnic background and parent composition. *Journal of Population Economics*, 16(4), 787-814.

[61] Ruist, J. (2015). Refugee immigration and public finances in Sweden. (Working Papers in Economics no. 613). Göteborg: Göteborg University.

[62] Rödin, M., and Özcan, G. (2011). Is it how you look or speak that matters?: An experimental study exploring the mechanisms of ethnic discrimination. (SULCIS Working Paper Series 3). Available at: http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A821626&dswid=article

[63] Schaafsma, J., and Sweetman, A. (2001). Immigrant earnings: age at immigration matters. *Canadian Journal of Economics*, 34(4), 1066-1099.

[64] Scovel, T. (1988). *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech.* New York: Newbury House Publishers.

[65] ——. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics*, 20, 213-223.

[66] Socialstyrelsen (2016). Försörjningsstöd. Retrieved from https://www.socialstyrelsen.se/hittarattmyndighet/ekonomisktbistand/\forsorjningsstod.

[67] Skolverket (2004). *Elever med utländsk bakgrund-rapport till regeringen.* Stockholm: Swedish National Agency for Education.

[68] Smit, F., and Driessen, G. (2009). Creating effective family–school partnerships in highly diverse contexts. In R. Deslandes (Ed.), *International Perspectives on Contexts, Communities and Evaluated Innovative Practices: Family-School-Community Partnerships* (pp. 64-81). Oxford: Routledge.

[69] Smith, C. D., Helgertz, J., and Scott, K. (2016). Parents' years in Sweden and children's educational performance. *IZA Journal of Migration*, 5(1), 6.

[70] Sorace, A. (2005). Selective optionality in language development. In Cornips, L. and Corrigan, K.P. (Ed.), *Syntax and Variation: Reconciling the Biological and the Social* (pp. 55-80). Amsterdam/Philadelphia: John Benjamins.

[71] Szulkin, R., and Jonsson, J. O. (2007). Ethnic segregation and educational outcomes in Swedish comprehensive schools. (SULCIS reports and working papers). Available at: http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A176854&dswid=1217.

[72] Sverige Statistiska centralbyrån. (1977). *Invandrarnas Levnadsförhållanden 1975.* LiberFörlag.

[73] Tahta, S., Wood, M., and Loewenthal, K. (1981). Foreign accents: Factors relating to transfer of accent from the first language to a second language. *Language and Speech,* 24(3), 265-272.

[74] Three Classical Tests; Wald, LM (Score), and LR tesfts. (2016). Retrieved from http://faculty.ndhu.edu.tw/~jlin/files/3tests.pdf.

[75] The Local (2016, January 4). Why do refugees prefer Sweden to Denmark? Retrieved from http://www.thelocal.se/20160104/why-do-refugees-prefer-sweden-to-denmark.

[76] Van den Berg, G. J., Lundborg, P., Nystedt, P., and Rooth, D. O. (2009). Critical periods during childhood and adolescence: a study of adult height among immigrant siblings. (IZA Discussion Paper 4140). Available at SSRN: `https://ssrn.com/abstract=1395247`.

[77] Van Ours, J. C., and Veenman, J. (2006). Age at immigration and educational attainment of young immigrants. *Economics Letters*, 90(3), 310-316.

[78] Vanhove, J. (2013). The critical period hypothesis in second language acquisition: A statistical critique and a reanalysis. *PloS ONE,* 8(7), e69172.

[79] Vilhelmson, B. (2002). Rörlighet och förankring. Geografiska aspekter på människors välfärd. (CHOROS 2002:1). Göteborg: Kulturgeografiska institutionen, Göteborg University.

Appendices

A Variables

A.1 Variables

Variable Name	*Explanation*
Immage	Immigration age
LoneInk	Annual taxable income (100SEK)
SocBidrPerF04	Individual welfare benefits (100SEK)
Sun2000NivaOld	Educational attainment

A.2 Control Variables

Variable Name	*Explanation*
Business Cycle Barometer[8]	Confidence indicator in: manufacturing, building and retail
Boom	Business Cycle Barometer > 100
Firstborn	Indicator for being first-born among siblings
$\text{Education}_{Parent}$	Highest level of parental educational attainment

A.3 Sources

Variable Name[9]	*Source*	*Explanation*
GDP	World Bank	Gross Domestic Produin ct per capita
Corruption measure	World Development Index	Max=100 (highly clean); Min=0 (highly corrupt)
English proficiency	Education First	English tests completed in 2014 (100 is maximum)
Gender Equality	Gender Inequality Index	The higher value, the more unequal the country

A.4 The SUN2000 classification

SUN-code	Level	Edyrs
63	Doctoral(PhD)	21
62	Licentiate	19
60	Other advanced degree	18
55	Post-secondary 5 years or more	17
54	Post-secondary 4 years	16
53	Post-secondary 3 years	15
52	Post-secondary 2 years	14
41	Post-secondary less than years	13
33	Upper secondary 3 years	12
32	Upper secondary 2 years	11
31	Upper secondary less than 2 years	10
20	Compulsory school 9 (10) years	9
10	Compulsory school less than 9 years	8
00	Pre-School	

[9]Based on index from year 2014

B Wald test

The operating principle of the performed test is to fit the regression without the restrictions. Then we assess whether the results appear to agree with the hypothesis within sampling variability. The Wald distance of a coefficient estimate from the hypothesized value is the linear distance measured in standard deviation units. The following null hypothesis is being tested: $H_0 = \hat{\beta} = \beta_0$, where $\beta_0 =$ the hypothesized value of the coefficient. Next, the following density is assumed: $l(y;\beta)$. Let $L(\beta)$ be the log likelihood function of the model and $\hat{\beta}$ be the Maximum Likelihood Estimation of β. Wald test is based on the intuition that the null hypothesis is being accepted when $\hat{\beta}$ is close to β_0. The observed variables in the estimations are assumed to be partitioned into the endogenous variables X and exogenous variables Y. Then following all the regularity conditions for existence, consistency and asymptotic normality of Maximum Likelihood Estimation, the hypotheses of interest can be rewritten as:

$$H_0 : g(\beta_0) = 0$$
$$H_A : g(\beta_0) \neq 0$$

where g(.); $\mathrm{R}^P \rightarrow \mathrm{R}^r$ and the rank is $\frac{dg}{d\theta}$ is r.

Proposition 1
$\Xi_n^W = ng'(\hat{\theta}_n)(\frac{dg'(\hat{\theta}_n)}{d\theta'} I^{-1}(\hat{\theta}_n) \frac{dg'(\hat{\theta}_n)}{d\theta)})^{-1} g(\hat{\theta}_n) \sim \chi^2(r)$ under H_0
where $I = E_X E_\theta(-\frac{2logf}{d\theta \dot{\theta'}})$ and $I^-1(\hat{\theta}_n)$ is the inverse of I evaluated at $\theta = \hat{\theta_n}$
The assymptotic characteristics imply the following:

$\sqrt{n}(\hat{\theta_n} - \theta_0) \rightarrow N(0, I^{-1}(\theta_0))$
The first order Taylor series of $g'(\hat{\theta_n})$ around the true value of θ_0 results in:

$$\sqrt{n}(g(\hat{\theta}_n) - g(\theta_0)) = \frac{dg(\theta_0)}{d\theta'} \sqrt{n}(\hat{\theta}_n - \theta_0) + o_p(1)$$

By forming the quadratic form of the normal random variables and evaluating at MLE, it can be concluded that:

$$\Xi_n^W = ng'(\hat{\theta}_n)(\frac{dg'(\hat{\theta}_n)}{d\theta'} I^{-1}(\hat{\theta}_n) \frac{dg'(\hat{\theta}_n)}{d\theta)})^{-1} g(\hat{\theta}_n) \sim \chi^2(r)$$

C Country data

Table 13: Country data - refugee sibling sample

	COUNTRY	Obs.	Immage.	Female ==1	Attained Education	Unempl. days	Wage (log)	Welfare Benefits	Parents education	GDP of Sweden	CPI 2014	EPI 2014	GTI 2014
Europe	Yugoslavia	8,626	15.29	0.49	12.27	41.20	7.079	45.82	10.14	0.181	39		0.76
	Bosnia	9,328	15.99	0.51	13.35	32.75	7.225	39.00	11.04	0.082	39	56.17	0.76
West Asia	Lebanon	7,651	14.42	0.47	12.16	60.58	6.567	68.93	9.50	0.014	27		6.4
	Syria	4,082	15.91	0.56	12.25	52.01	6.621	48.43	8.17	0.042	20		8.12
	Turkey	18,492	8.31	0.50	12.17	41.09	6.683	22.32	9.41	0.175	45	47.62	5.98
	Iraq	14,159	18.28	0.52	12.02	53.58	6.647	88.81	9.14	0.108	16	38.02	10
	Iran	14,786	14.95	0.50	13.95	48.59	6.739	47.19	10.84	0.092	27	46.59	4.9
Central Asia	Vietnam	4,012	11.51	0.47	13.01	41.37	6.936	35.37	10.41	0.034	31	53.81	0
	Afghanistan	1,726	18.42	0.50	11.86	41.33	6.810	69.59	9.49	0.011	12		9.39
	Bangladesh	1,128	16.72	0.56	13.36	42.90	6.684	48.24	11.13	0.018	25		5.25
	Pakistan	782	20.35	0.53	13.41	36.49	6.730	33.80	8.94	0.022	30		9.37
	Sri Lanka	2,529	3.75	0.55	13.58	22.63	7.227	17.00	10.84	0.065	37	46.37	4.01
Africa	Gambia	855	17.45	0.52	11.38	55.89	6.557	81.14	10.60	0.007	28		3.7
	Ethiopia	2,249	16.45	0.41	12.76	63.63	6.641	76.45	8.51	0.009	33		0
	Somaila	4,030	18.04	0.50	10.23	53.15	6.595	110.13	10.05	0.009	8		7.41
	Uganda	1,468	14.76	0.59	13.37	42.51	6.942	50.58	9.30	0.007	18		2.45
	Erithrea	1,808	15.82	0.54	12.36	48.35	6.870	49.28	9.29	0.012	25		2.93
South America	Chile	10,516	8.79	0.50	12.88	38.67	6.891	29.98	11.21	0.247	73	51.88	2.59
	Bolivia	512	7.20	0.49	13.57	53.23	6.790	31.43	10.12	0.053	35		0.24
	Peru	954	13.70	0.49	13.61	35.69	6.902	31.99	10.61	0.111	38	52.46	2.96
	Colombia	2,027	3.89	0.48	13.25	28.68	7.141	28.20	10.31	0.134	37	48.54	6.24

Note: The refugee sibling sample by country used in the main analysis. The time span of observation lies between 1990 to 2014, the earliest possible immigration is 1950. Immigration year refer to the mother immigration year if the individual is second-generation. Outcomes measured in age 28 to 30. The average country characteristics are divided by number of country in the refugee sample, source of variable, see appendix A. Orphans divided by country and observations - Sri Lanka 1,777 - Chile 1,232 - Colombia 1,459.

Table 14: Country data - other immigrant sample

	COUNTRY	Obs.	Immage.	Female ==1	Attained Education	Unempl. days	Wage (log)	Welfare Benefits	Parents education	GDP of Sweden	CPI 2014	EPI 2014	GTI 2014
Northern Europe	Danmark	6,272	5.453	.510	13.16	28.810	6.98	16.54	13.14	1.04	91		0.19
	Finland	43,652	1.69	.50	13.30	28.60	7.068	16.827	12.95	0,84	89	65,32	
	Norway	7,371	5.516	0.51	13.48	25.95	7.046	13.44	12.23	1,65	86	67.83	3,57
	Poland	11,109	7.984	0.548	13.629	29.17	7.074	24.47	11.96	0,24	61	62,95	
	Germany	4,224	4.135	.493	13.80	23.52	7.072	10.57	12.54	0,81	79	61,83	1,02
	Great Britain	2,937	6.123	.4576	14.19	19.49	7.135	10.89	13.12	0,78	81	100	5.17
South Europe	Greece	4,035	5.043	.466	12.70	25.388	6.84	8.35	11.54	0,37	46		4.73
	Italy	742	6.83	.459	13.05	34.66	6.792	26.51	11.95	0,59	44	52.80	2.55

Note: The other immigrant sibling sample by country used in the sensitivity analysis.

Table 15: Distribution - other immigrant sample

REGION	Siblings Immigrants	Distribution
Baltic region	75,565	0.90
Other european	4,777	0.10
Total	80,342	1.00

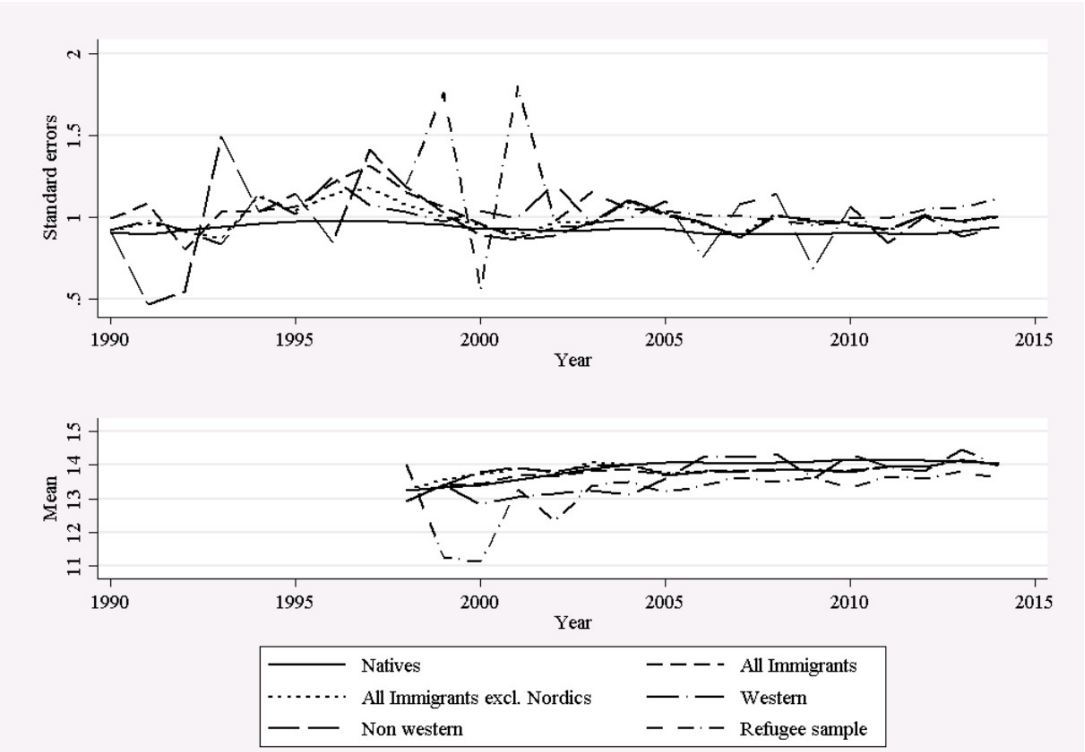
D Transformation

Table 16: Mean and standard errors - natives and age-zero immigrants

	obs.	Attained Education		Unempl. days		Wage (log)		Welfare Benefits	
GROUP		Mean	se	Mean	se	Mean	se	Mean	se
Natives	2,687,071	13,92	2,67	21,32	49,16	7,19	0,97	8,71	57,96
All Immigrants	11,211	13,83	2,74	22,99	49,74	7,20	1,01	13,70	75,58
Immigrants excl. Nordic	9,160	13,89	2,75	20,44	45,91	7,26	1,00	13,36	76,71
Western	2,818	13,53	2,72	33,21	62,12	7,01	1,04	16,54	77,83
Non-Western	8,393	13,88	2,74	19,86	44,81	7,27	1,00	12,75	74,78
Refugee Sample	4,086	13.55	2,92	21.59	49,29	7,28	1,05	16.88	94,23

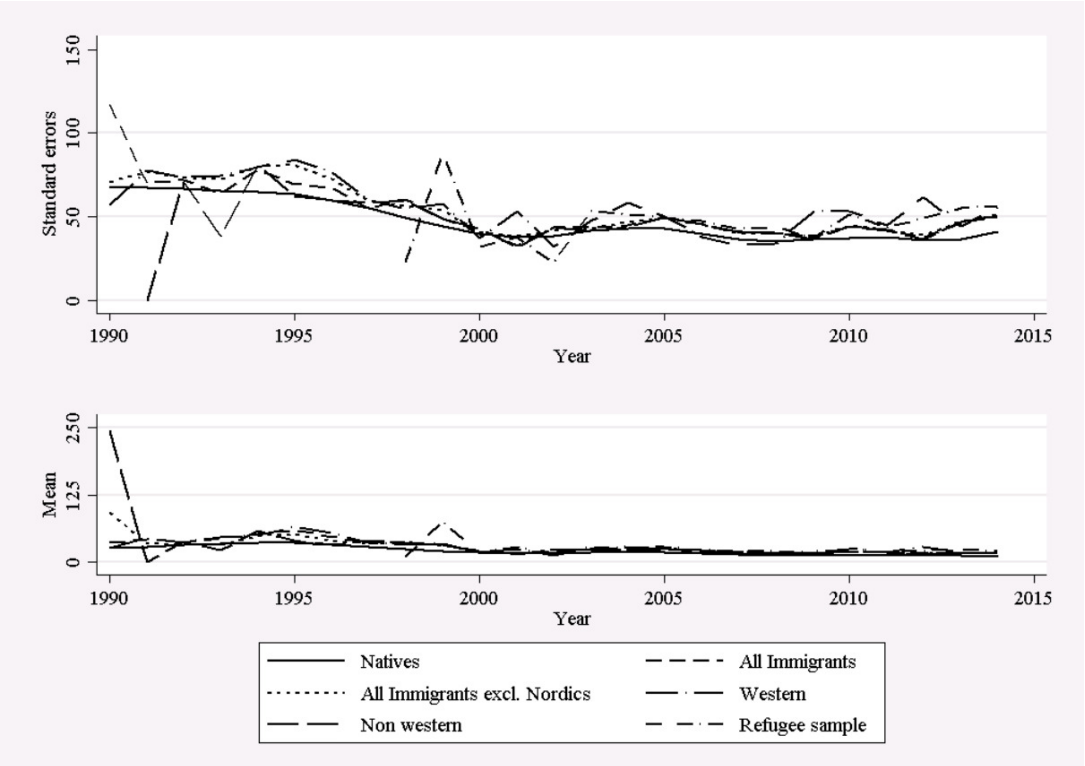
Note: The mean and standard error for each dependent variable use age-zero natives or immigrants. Time span of observation between 1990 to 2014.

Figure 7: Mean and standard errors - education



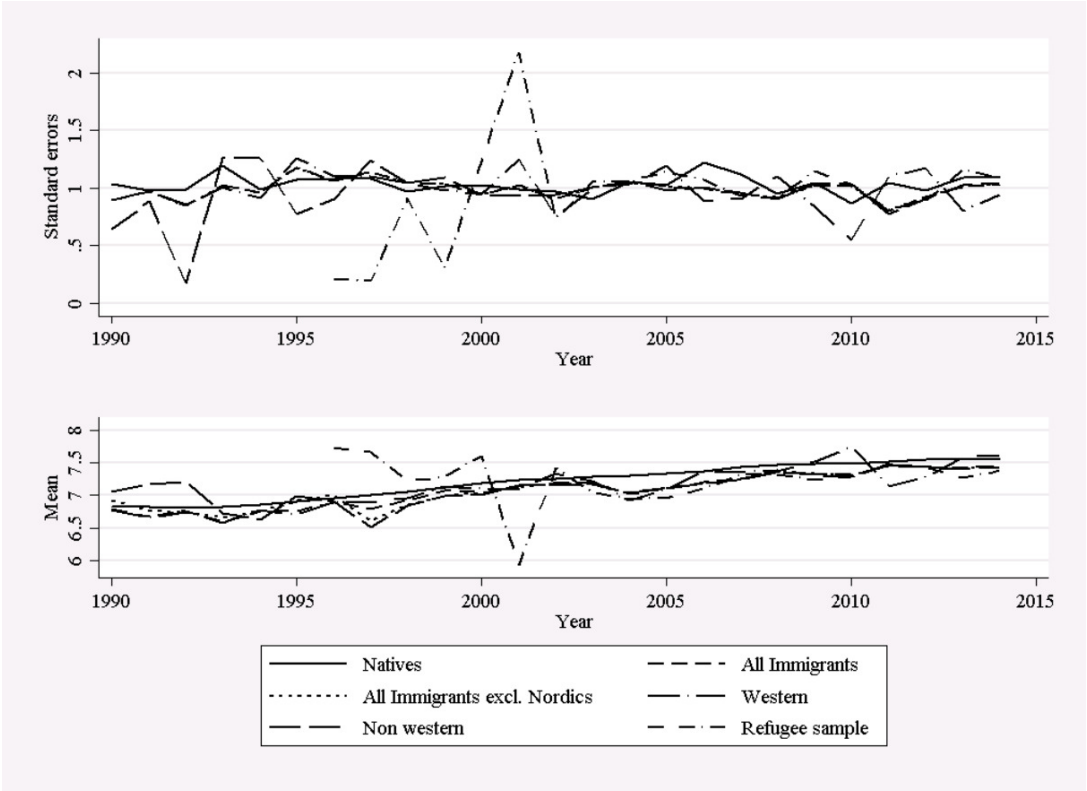
Note: The mean and standard error of age-zero natives or immigrants.

Figure 8: Mean and standard errors - unemployment



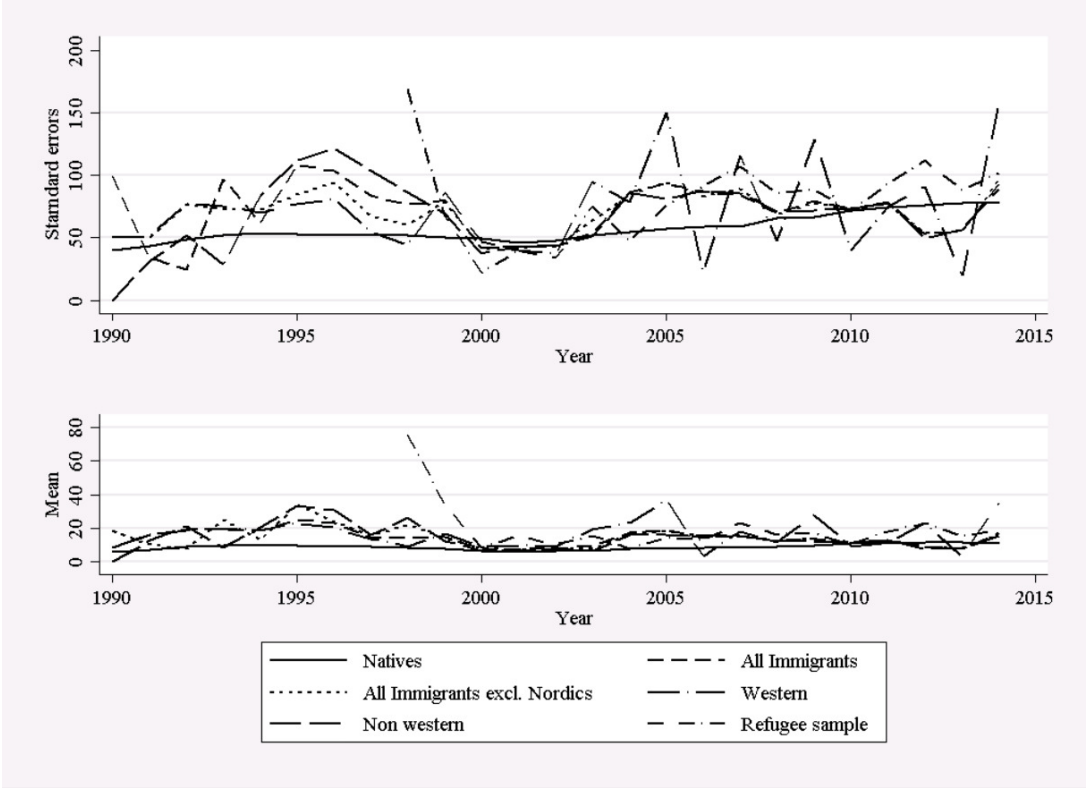
Note: The mean and standard error of age-zero natives or immigrants.

Figure 9: Mean and standard errors - wage



Note: The mean and standard error of age-zero natives or immigrants.

Figure 10: Mean and standard errors - welfare benefits

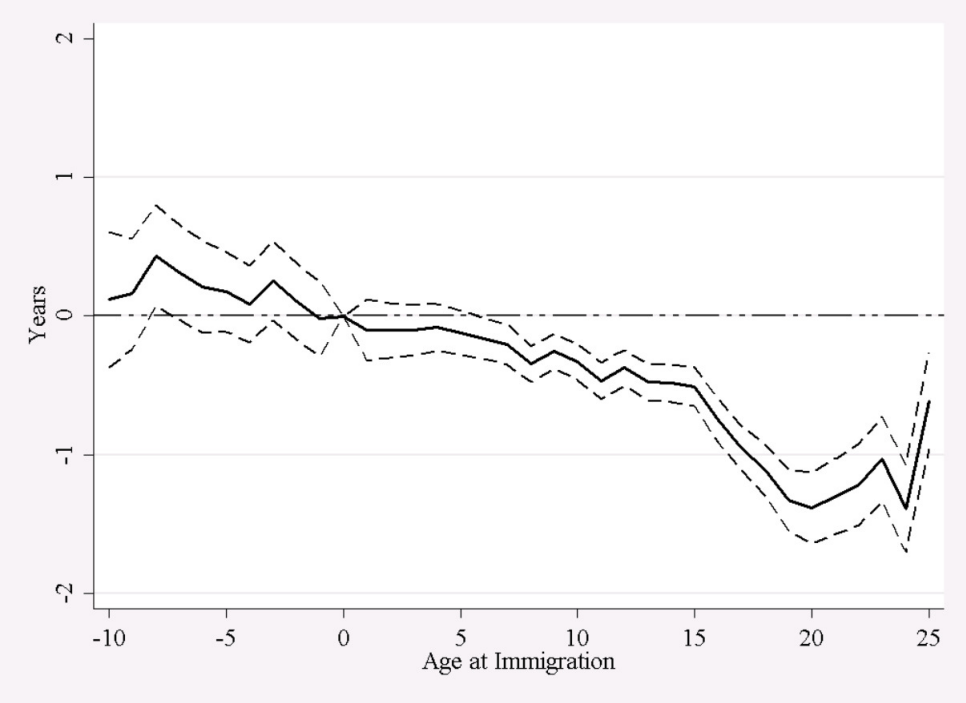


Note: The mean and standard error of age-zero natives or immigrants.

E Results

E.1 Binary variable

Figure 11: Binary variable - education up to age 26



Note: Age at immigration as binary variable - The estimates and 95% confidence intervals from family fixed effects specifications. The reference group is the age-zero immigrant.

Figure 12: Binary variable - refugee sample fixed effect/without fixed effect

VARIABLE (age)	Attained Education				Unemployment days				Wage(log)				Welfare Benefits			
Immage (-10)	0.191	(0.214)	-0.405**	(0.161)	-8.090*	(4.484)	-3.934	(3.272)	5.965***	(1.188)	2.527***	(0.863)	1.452	(6.904)	-13.35**	(5.443)
Immage (-9)	0.205	(0.190)	-0.289**	(0.146)	-0.282	(4.061)	3.142	(2.967)	4.742***	(1.081)	0.816	(0.794)	1.758	(6.303)	-10.36**	(4.935)
Immage (-8)	0.485***	(0.172)	-0.103	(0.132)	-8.179**	(3.738)	-0.844	(2.682)	4.594***	(0.990)	0.791	(0.709)	-1.203	(5.800)	-12.52***	(4.461)
Immage (-7)	0.366**	(0.170)	-0.216*	(0.129)	-5.900	(3.682)	-1.871	(2.628)	4.223***	(0.976)	0.688	(0.693)	7.303	(5.751)	-4.829	(4.371)
Immage (-6)	0.260*	(0.154)	-0.136	(0.120)	-4.120	(3.409)	0.268	(2.444)	3.208***	(0.909)	0.781	(0.647)	3.211	(5.347)	-9.202**	(4.065)
Immage (-5)	0.213	(0.148)	-0.171	(0.117)	-0.192	(3.317)	0.164	(2.369)	2.966***	(0.880)	0.596	(0.626)	5.071	(5.228)	-6.618*	(3.941)
Immage (-4)	0.116	(0.141)	-0.212*	(0.111)	-1.212	(3.191)	3.734*	(2.266)	0.654	(0.860)	-1.441**	(0.602)	7.000	(5.041)	-5.128	(3.769)
Immage (-3)	0.268*	(0.141)	-0.172	(0.114)	-0.646	(3.218)	0.940	(2.312)	1.547*	(0.866)	-1.116*	(0.613)	3.042	(5.089)	-6.840*	(3.846)
Immage (-2)	0.120	(0.140)	-0.154	(0.114)	-2.744	(3.207)	1.861	(2.309)	1.369	(0.864)	-0.374	(0.617)	0.903	(5.076)	-7.252*	(3.841)
Immage (-1)	-0.0183	(0.134)	-0.197*	(0.106)	2.340	(3.095)	3.719*	(2.144)	1.158	(0.837)	0.0586	(0.571)	1.503	(4.908)	-4.187	(3.567)
Reference group (0)	0	(0.296)	0	(0.254)	0	(6.507)	0	(4.495)	0	(0.560)	0	(0.295)	0	(10.21)	0	(7.336)
Immage (1)	-0.114	(0.106)	-0.0129	(0.0789)	1.762	(2.459)	1.634	(1.603)	-1.389**	(0.664)	-0.407	(0.424)	-1.253	(3.908)	0.889	(2.666)
Immage (2)	-0.114	(0.117)	-0.121	(0.0866)	0.546	(2.702)	1.979	(1.757)	-1.781**	(0.736)	-0.762	(0.469)	1.950	(4.293)	0.214	(2.923)
Immage (3)	-0.113	(0.116)	0.131	(0.0840)	4.516*	(2.680)	2.753	(1.702)	-1.548**	(0.728)	-0.360	(0.455)	3.747	(4.256)	1.643	(2.832)
Immage (4)	-0.109	(0.115)	0.0644	(0.0815)	6.746**	(2.638)	6.247***	(1.648)	-1.754**	(0.721)	-1.062**	(0.442)	2.639	(4.185)	6.292**	(2.741)
Immage (5)	-0.176	(0.114)	-0.0270	(0.0794)	7.141***	(2.605)	6.296***	(1.604)	-2.694***	(0.713)	-1.289***	(0.430)	10.36**	(4.132)	6.928***	(2.669)
Immage (6)	-0.226**	(0.115)	-0.0277	(0.0779)	7.710***	(2.604)	7.444***	(1.572)	-2.989***	(0.706)	-1.574***	(0.420)	6.536	(4.127)	7.469***	(2.614)
Immage (7)	-0.275**	(0.115)	-0.00563	(0.0753)	7.465***	(2.585)	6.366***	(1.514)	-4.066***	(0.701)	-1.446***	(0.404)	8.557**	(4.094)	4.897*	(2.518)
Immage (8)	-0.411***	(0.114)	-0.119	(0.0727)	11.10***	(2.549)	6.914***	(1.458)	-4.069***	(0.693)	-0.798**	(0.388)	11.37***	(4.033)	7.210***	(2.425)
Immage (9)	-0.323***	(0.115)	-0.188***	(0.0725)	11.02***	(2.563)	8.781***	(1.454)	-4.399***	(0.695)	-0.912**	(0.387)	8.790**	(4.052)	7.061***	(2.417)
Immage (10)	-0.414***	(0.117)	-0.259***	(0.0733)	13.12***	(2.591)	9.216***	(1.469)	-4.893***	(0.703)	-1.174***	(0.392)	11.04***	(4.095)	9.401***	(2.440)
Immage (11)	-0.589***	(0.118)	-0.428***	(0.0735)	13.50***	(2.606)	9.471***	(1.468)	-5.832***	(0.707)	-1.546***	(0.392)	12.27***	(4.114)	9.482***	(2.439)
Immage (12)	-0.471***	(0.120)	-0.433***	(0.0740)	13.88***	(2.627)	11.80***	(1.475)	-5.639***	(0.713)	-1.802***	(0.394)	11.17***	(4.142)	11.49***	(2.449)
Immage (13)	-0.606***	(0.122)	-0.532***	(0.0737)	15.78***	(2.649)	14.16***	(1.470)	-6.541***	(0.720)	-2.012***	(0.394)	12.29***	(4.173)	13.86***	(2.441)
Immage (14)	-0.615***	(0.124)	-0.616***	(0.0747)	18.07***	(2.695)	15.86***	(1.487)	-7.568***	(0.734)	-2.899***	(0.398)	17.13***	(4.244)	18.17***	(2.467)
Immage (15)	-0.686***	(0.127)	-0.848***	(0.0750)	18.55***	(2.735)	16.96***	(1.490)	-7.383***	(0.747)	-2.527***	(0.400)	14.95***	(4.296)	18.46***	(2.468)
Immage (16)	-0.929***	(0.130)	-1.086***	(0.0755)	16.22***	(2.789)	17.63***	(1.498)	-7.925***	(0.757)	-3.391***	(0.401)	18.35***	(4.365)	22.64***	(2.475)
Immage (17)	-1.178***	(0.136)	-1.440***	(0.0767)	19.92***	(2.878)	21.21***	(1.507)	-9.168***	(0.786)	-4.160***	(0.406)	21.94***	(4.499)	24.70***	(2.490)
Immage (18)	-1.326***	(0.146)	-1.906***	(0.0780)	21.10***	(2.974)	22.26***	(1.505)	-9.869***	(0.817)	-6.292***	(0.409)	25.02***	(4.629)	27.85***	(2.485)
Firstborn	0.131***	(0.0326)	0.436***	(0.0273)	-0.801	(0.701)	-2.339***	(0.527)	0.396**	(0.194)	-0.134	(0.143)	0.863	(1.107)	-2.215**	(0.871)
BCB	-	-	-	-	0.129**	(0.0583)	-0.0164	(0.0417)	-	-	-	-	0.0894	(0.0913)	0.213***	(0.0679)
Parent Education	-	-	0.034	(0.0523)	-	-	0.242	(0.551)	-	-	0.015	(0.142)	-	-	0.405	(0.445)
Observations	59,921		59,921		70,947		70,947		66,673		66,673		71,555		71,555	
R-squared	0.617		0.011		0.546		0.011		0.668		0.026		0.582		0.009	
Family FE	YES		NO		YES		NO		YES		NO		YES		NO	
Controls	YES		YES		YES		YES		YES		YES		YES		YES	

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Note: Age at immigration as binary variable - The reference group is the age-zero immigrant.

Figure 13: Binary variable - refugee sample without fixed effect divided gender

	Attained Education				Unemployment days				Wage(log)				Welfare benefits			
VARIABLE (age)	Female		Male		Female		Male		Female		Male		Female		Male	
Immage (-10)	-0.409*	(0.240)	-0.291	(0.214)	-7.111	(4.622)	-1.953	(4.613)	1.289	(1.286)	2.788**	(1.147)	-11.38	(7.141)	-15.98**	(8.044)
Immage (-9)	-0.0540	(0.212)	-0.413**	(0.198)	7.264*	(4.086)	-0.801	(4.270)	0.474	(1.155)	0.560	(1.075)	-4.388	(6.314)	-16.34**	(7.446)
Immage (-8)	-0.166	(0.181)	-0.0702	(0.190)	2.533	(3.482)	-4.652	(4.091)	0.276	(0.962)	1.470	(1.026)	-11.51**	(5.380)	-13.53*	(7.134)
Immage (-7)	0.0376	(0.181)	-0.465**	(0.182)	1.368	(3.487)	-5.390	(3.915)	0.0120	(0.964)	1.253	(0.978)	-1.933	(5.388)	-8.019	(6.827)
Immage (-6)	-0.160	(0.174)	-0.0437	(0.164)	1.082	(3.348)	-0.795	(3.534)	0.586	(0.921)	0.605	(0.893)	-11.19**	(5.172)	-8.169	(6.164)
Immage (-5)	-0.0398	(0.172)	-0.164	(0.156)	1.088	(3.322)	-1.073	(3.364)	0.232	(0.921)	0.284	(0.842)	-3.945	(5.132)	-9.724*	(5.867)
Immage (-4)	-0.134	(0.159)	-0.235	(0.154)	2.155	(3.066)	4.956	(3.313)	-1.611*	(0.856)	-1.621*	(0.832)	-7.580	(4.736)	-3.369	(5.778)
Immage (-3)	-0.106	(0.159)	-0.232	(0.160)	0.936	(3.070)	0.880	(3.442)	-0.873	(0.847)	-1.420	(0.870)	-7.416	(4.743)	-6.430	(6.002)
Immage (-2)	-0.115	(0.161)	-0.169	(0.158)	0.266	(3.095)	3.257	(3.404)	-0.239	(0.862)	-0.621	(0.867)	-10.15**	(4.782)	-4.672	(5.937)
Immage (-1)	-0.288*	(0.149)	-0.0847	(0.147)	3.550	(2.876)	3.766	(3.160)	0.391	(0.806)	-0.502	(0.794)	-6.520	(4.444)	-2.230	(5.512)
Reference group (0)	0	(0.0768)	0	(0.0771)	0	(6.117)	0	(6.537)	0	(0.409)	0	(0.418)	0	(9.290)	0	(11.17)
Immage (1)	0.0757	(0.111)	-0.101	(0.111)	1.360	(2.129)	1.924	(2.384)	0.126	(0.588)	-0.964	(0.600)	0.0452	(3.290)	1.721	(4.158)
Immage (2)	-0.232*	(0.123)	0.0272	(0.120)	1.747	(2.366)	1.955	(2.581)	0.00534	(0.665)	-1.754***	(0.651)	0.403	(3.655)	-0.380	(4.501)
Immage (3)	0.0612	(0.119)	0.244**	(0.116)	1.343	(2.292)	3.868	(2.502)	0.589	(0.640)	-1.531**	(0.634)	1.298	(3.540)	1.464	(4.363)
Immage (4)	0.0174	(0.117)	0.181	(0.112)	7.444***	(2.240)	4.852**	(2.402)	-0.888	(0.629)	-1.626***	(0.611)	7.785**	(3.461)	4.307	(4.189)
Immage (5)	-0.129	(0.114)	0.138	(0.109)	5.984***	(2.182)	6.218***	(2.338)	-1.529**	(0.610)	-1.491**	(0.595)	4.368	(3.370)	8.422**	(4.077)
Immage (6)	0.0143	(0.111)	0.0115	(0.107)	6.698***	(2.134)	7.735***	(2.293)	-1.185**	(0.597)	-2.309***	(0.581)	1.878	(3.297)	11.61***	(3.999)
Immage (7)	-0.0252	(0.106)	0.0566	(0.105)	8.138***	(2.028)	4.509**	(2.234)	-1.431**	(0.566)	-1.721***	(0.566)	5.538*	(3.133)	3.868	(3.896)
Immage (8)	-0.0936	(0.103)	-0.0838	(0.101)	6.331***	(1.960)	7.188***	(2.145)	-0.471	(0.545)	-1.393**	(0.543)	5.416*	(3.028)	8.319**	(3.741)
Immage (9)	-0.178*	(0.103)	-0.133	(0.100)	9.817***	(1.960)	7.566***	(2.134)	-0.602	(0.546)	-1.534***	(0.540)	5.958**	(3.026)	7.472**	(3.720)
Immage (10)	-0.242**	(0.105)	-0.194*	(0.101)	8.646***	(1.991)	9.298***	(2.146)	-1.945***	(0.556)	-0.897*	(0.543)	10.49***	(3.074)	7.837**	(3.738)
Immage (11)	-0.547***	(0.104)	-0.267***	(0.102)	9.451***	(1.977)	9.274***	(2.157)	-1.938***	(0.554)	-1.495***	(0.546)	8.939***	(3.051)	9.524**	(3.757)
Immage (12)	-0.450***	(0.106)	-0.348***	(0.102)	12.86***	(1.998)	10.54***	(2.157)	-1.895***	(0.558)	-2.025***	(0.547)	8.927***	(3.080)	13.18***	(3.755)
Immage (13)	-0.471***	(0.106)	-0.505***	(0.101)	14.10***	(1.994)	13.85***	(2.148)	-2.381***	(0.559)	-1.996***	(0.546)	13.30***	(3.074)	13.80***	(3.738)
Immage (14)	-0.609***	(0.107)	-0.550***	(0.103)	15.34***	(2.020)	15.95***	(2.168)	-3.366***	(0.566)	-2.805***	(0.551)	11.36***	(3.112)	23.41***	(3.775)
Immage (15)	-0.912***	(0.108)	-0.716***	(0.103)	17.57***	(2.030)	16.02***	(2.168)	-2.875***	(0.574)	-2.607***	(0.550)	15.51***	(3.122)	20.36***	(3.767)
Immage (16)	-1.094***	(0.108)	-1.007***	(0.104)	16.10***	(2.031)	18.65***	(2.188)	-4.242***	(0.570)	-2.939***	(0.555)	17.83***	(3.114)	26.39***	(3.792)
Immage (17)	-1.473***	(0.111)	-1.323***	(0.105)	20.01***	(2.049)	21.87***	(2.196)	-5.534***	(0.582)	-3.338***	(0.558)	25.63***	(3.141)	23.56***	(3.809)
Immage (18)	-2.257***	(0.108)	-1.558***	(0.111)	22.50***	(1.980)	21.83***	(2.268)	-8.220***	(0.571)	-4.349***	(0.575)	27.88***	(3.035)	28.42***	(3.925)
Firstborn	0.445***	(0.0396)	0.457***	(0.0371)	-4.436***	(0.728)	-0.896	(0.760)	0.392*	(0.206)	-1.024***	(0.195)	-1.746	(1.117)	-3.420***	(1.320)
BCB	-	-	-	-	-0.0506	(0.0568)	0.0156	(0.0606)	-	-	-	-	0.177**	(0.0861)	0.240**	(0.103)
Observations	31,529		35,144		33,652		37,295		28,254		31,667		33,955		37,6	
R-squared	0.034		0.021		0.013		0.011		0.019		0.008		0.011		0.008	
Family FE	NO		NO		NO		NO		NO		NO		NO		NO	
Controls	YES		YES		YES		YES		YES		YES		YES		YES	

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Note: Age at immigration as binary variable - The reference group is the age-zero immigrant.

Figure 14: Binary variable - refugee sample fixed effect divided gender

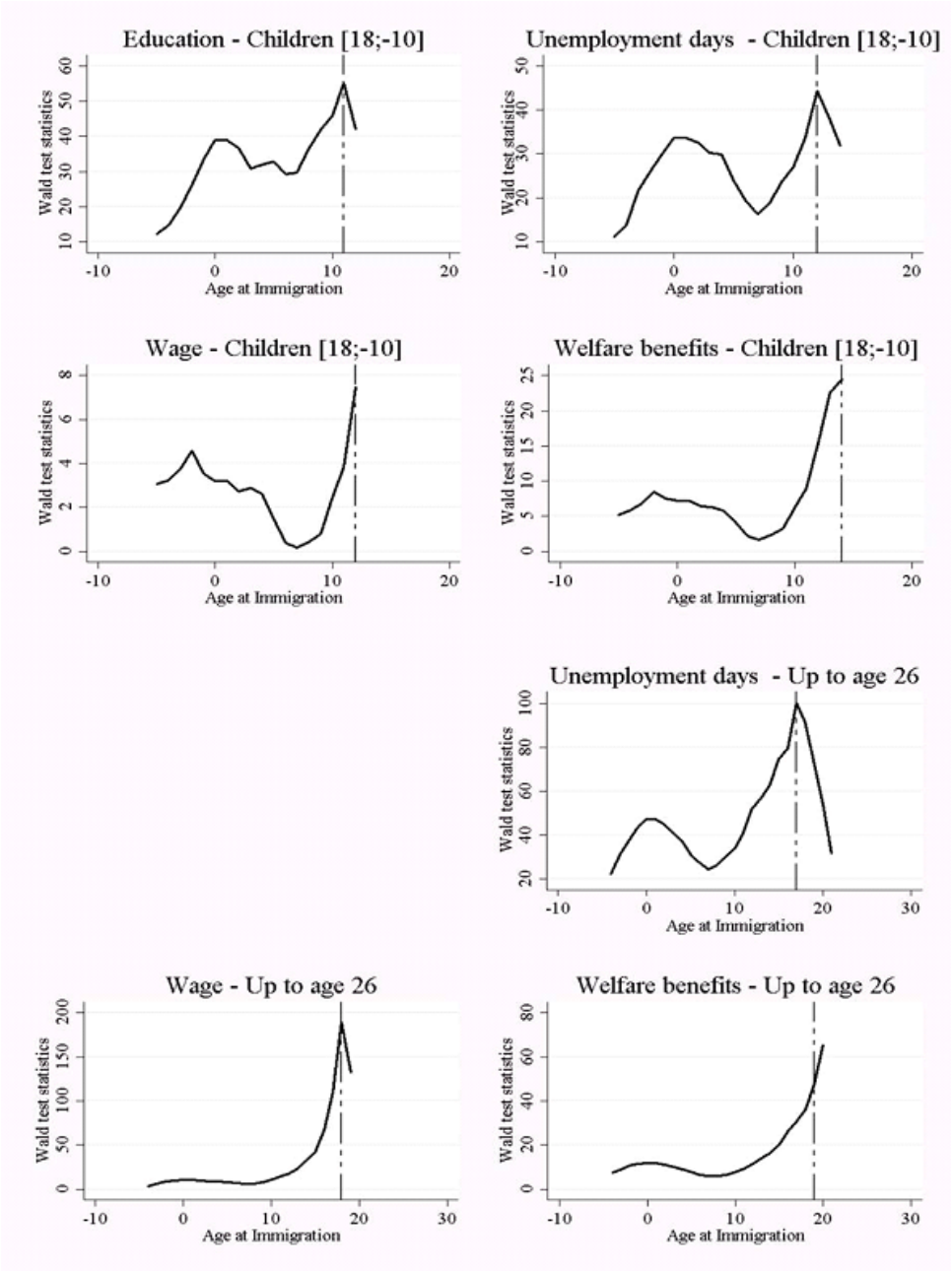
	Attained Education				Unemployment days				Wage(log)				Welfare benefits			
VARIABLE (age)	Female		Male		Female		Male		Female		Male		Female		Male	
Immage (-10)	-0.214	(0.400)	0.436	(0.353)	-19.20**	(7.821)	6.278	(8.026)	7.531***	(2.198)	5.612***	(2.002)	4.359	(10.90)	6.178	(13.55)
Immage (-9)	0.337	(0.338)	0.211	(0.322)	-1.291	(6.922)	4.006	(7.547)	5.669***	(1.942)	3.114*	(1.879)	7.189	(9.739)	12.66	(12.78)
Immage (-8)	0.520*	(0.294)	0.453	(0.316)	-17.30***	(6.182)	2.171	(7.421)	6.213***	(1.716)	3.849**	(1.850)	-1.710	(8.726)	10.75	(12.48)
Immage (-7)	0.251	(0.301)	0.393	(0.307)	-15.25**	(6.258)	3.763	(7.273)	6.091***	(1.712)	2.624	(1.848)	10.65	(8.812)	19.00	(12.31)
Immage (-6)	0.121	(0.275)	0.538**	(0.265)	-6.808	(5.905)	1.858	(6.341)	5.162***	(1.630)	2.441	(1.608)	-0.166	(8.357)	18.61*	(10.81)
Immage (-5)	0.475*	(0.267)	0.314	(0.254)	-4.873	(5.834)	1.609	(6.075)	4.482***	(1.600)	3.335**	(1.533)	7.624	(8.319)	9.540	(10.50)
Immage (-4)	0.0323	(0.257)	0.171	(0.245)	-12.72**	(5.661)	8.287	(5.936)	1.396	(1.569)	1.096	(1.511)	2.883	(8.106)	17.75*	(10.24)
Immage (-3)	0.463*	(0.254)	0.369	(0.249)	-10.33*	(5.690)	0.562	(6.031)	3.032*	(1.579)	-0.0285	(1.543)	0.556	(8.136)	18.81*	(10.47)
Immage (-2)	0.118	(0.247)	0.478*	(0.247)	-9.164	(5.628)	1.107	(6.009)	1.506	(1.542)	1.565	(1.544)	-0.930	(8.062)	12.79	(10.43)
Immage (-1)	-0.0670	(0.235)	0.367	(0.240)	1.818	(5.363)	4.392	(5.866)	2.563*	(1.512)	1.967	(1.496)	2.535	(7.729)	13.53	(10.20)
Reference group (0)	0	(0.168)	0	(0.166)	0	(11.86)	0	(12.02)	0	(1.044)	0	(1.024)	0	(16.89)	0	(20.69)
Immage (1)	-0.0298	(0.213)	0.130	(0.214)	0.318	(4.900)	-1.632	(5.243)	-0.826	(1.375)	-0.501	(1.326)	0.247	(7.083)	10.62	(9.122)
Immage (2)	-0.124	(0.217)	0.275	(0.215)	-4.153	(4.979)	3.937	(5.279)	-0.733	(1.419)	-0.836	(1.344)	2.586	(7.191)	13.33	(9.180)
Immage (3)	0.0320	(0.212)	-0.148	(0.209)	-0.136	(4.853)	9.863*	(5.117)	-1.237	(1.367)	-2.151*	(1.307)	3.122	(7.015)	19.59**	(8.890)
Immage (4)	-0.0521	(0.213)	-0.139	(0.205)	-0.195	(4.835)	4.364	(4.987)	-1.254	(1.369)	-0.783	(1.278)	2.870	(6.985)	12.18	(8.656)
Immage (5)	-0.104	(0.209)	0.0511	(0.201)	3.477	(4.728)	8.675*	(4.890)	-3.159**	(1.332)	-2.104*	(1.257)	6.086	(6.827)	23.89***	(8.485)
Immage (6)	-0.131	(0.210)	-0.117	(0.202)	0.654	(4.727)	8.178*	(4.896)	-0.982	(1.330)	-2.767**	(1.252)	0.0145	(6.821)	25.93***	(8.489)
Immage (7)	-0.165	(0.210)	-0.185	(0.204)	2.391	(4.687)	7.854	(4.903)	-2.888**	(1.314)	-4.023***	(1.249)	-0.390	(6.759)	30.97***	(8.494)
Immage (8)	-0.291	(0.209)	-0.290	(0.202)	4.033	(4.591)	13.07***	(4.847)	-1.608	(1.291)	-4.760***	(1.241)	8.603	(6.615)	30.32***	(8.387)
Immage (9)	-0.222	(0.210)	-0.252	(0.204)	4.031	(4.631)	12.68***	(4.867)	-2.952**	(1.297)	-4.612***	(1.241)	10.52	(6.666)	24.06***	(8.418)
Immage (10)	-0.347	(0.214)	-0.169	(0.206)	6.013	(4.700)	16.99***	(4.917)	-4.468***	(1.320)	-3.931***	(1.252)	13.05*	(6.766)	23.55***	(8.498)
Immage (11)	-0.531**	(0.214)	-0.381*	(0.209)	4.871	(4.670)	18.68***	(4.980)	-4.423***	(1.312)	-5.698***	(1.267)	9.966	(6.717)	30.80***	(8.592)
Immage (12)	-0.411*	(0.219)	-0.345	(0.211)	6.371	(4.725)	17.30***	(4.974)	-4.309***	(1.324)	-5.362***	(1.272)	5.683	(6.788)	31.26***	(8.581)
Immage (13)	-0.685***	(0.222)	-0.478**	(0.213)	6.897	(4.756)	18.34***	(5.030)	-5.235***	(1.335)	-6.019***	(1.291)	10.96	(6.827)	31.94***	(8.668)
Immage (14)	-0.549**	(0.225)	-0.476**	(0.218)	8.939*	(4.824)	21.82***	(5.136)	-5.532***	(1.356)	-7.312***	(1.317)	10.07	(6.916)	41.47***	(8.841)
Immage (15)	-0.696***	(0.231)	-0.542**	(0.220)	9.803**	(4.929)	20.94***	(5.181)	-6.112***	(1.387)	-5.829***	(1.330)	14.95**	(7.039)	32.17***	(8.900)
Immage (16)	-1.038***	(0.238)	-0.795***	(0.227)	4.989	(5.016)	21.64***	(5.344)	-6.505***	(1.400)	-7.013***	(1.364)	9.676	(7.123)	38.80***	(9.145)
Immage (17)	-1.414***	(0.248)	-0.926***	(0.237)	10.36**	(5.135)	23.14***	(5.552)	-9.140***	(1.446)	-7.794***	(1.427)	24.58***	(7.289)	31.76***	(9.498)
Immage (18)	-1.651***	(0.268)	-1.089***	(0.255)	12.41**	(5.411)	20.58***	(5.871)	-9.145***	(1.516)	-9.091***	(1.503)	25.45***	(7.654)	37.93***	(10.01)
Firstborn	0.103	(0.0628)	0.229***	(0.0581)	-0.645	(1.361)	2.818**	(1.371)	0.568	(0.388)	-0.0255	(0.355)	1.946	(1.961)	2.099	(2.364)
BCB	-	-	-	-	0.229**	(0.106)	0.00288	(0.107)	-	-	-	-	0.0271	(0.150)	0.236	(0.184)
Observations	31,529		35,144		33,652		37,295		28,254		31,667		33,955		37,6	
R-squared	0.826		0.795		0.736		0.725		0.791		0.776		0.765		0.724	
Family FE	YES		YES		YES		YES		YES		YES		YES		YES	
Controls	YES		YES		YES		YES		YES		YES		YES		YES	

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Note: Age at immigration as binary variable - The reference group is the age-zero immigrant.

E.2 Structural break

Figure 15: Structual break - Wald estimates



Note: The figure illustrate the estimated Wald test using family dummies. No difference in the estimated coefficient compared to the family fixed effect estimate. presented numerical in 8.2.

E.3 Spline estimate

Figure 16: Linear spline estimate - refugee sample divided by gender controls displayed

	Female				Male			
VARIABLE	Attained Education	Unemp. days	Wage (log)	Welfare Benefits	Attained Education	Unemp. days	Wage (log)	Welfare Benefits
Immage [-10;0]	-0.0114	1.346***	-0.597***	-0.672	-0.0187	0.149	-0.497***	-0.401
	(0.0216)	(0.409)	(0.108)	(0.579)	(0.0209)	(0.463)	(0.113)	(0.749)
Immage [0;18]	-0.0729***	1.797***	-0.408***	-0.00892	-0.0568***	1.912***	-0.361***	0.197
	(0.0102)	(0.164)	(0.0452)	(0.234)	(0.00959)	(0.195)	(0.0485)	(0.321)
Immage [18;26]	-	1.828***	-1.188***	10.24***	-	3.812***	-0.577***	8.004***
	-	(0.387)	(0.118)	(0.546)	-	(0.541)	(0.136)	(0.870)
Business Cycle Barometer	-	-0.182*	-	0.0557	-	-0.247**	-	0.246
	-	(0.0934)	-	(0.135)	-	(0.101)	-	(0.171)
Boom (at age18)	-	1.075	-0.158	4.297***	-	2.737**	-0.710**	3.556*
	-	(1.083)	(0.298)	(1.576)	-	(1.097)	(0.276)	(1.869)
Firstborn	0.0114	2.268*	-0.382	11.08***	0.187***	4.636***	-0.633**	8.164***
	(0.0614)	(1.191)	(0.340)	(1.782)	(0.0567)	(1.238)	(0.317)	(2.121)
Observations	31,482	54,827	40,867	56,197	35,096	53,603	45,257	55,418
R-squared	0.826	0.807	0.840	0.843	0.795	0.835	0.845	0.838
Family Fixed Effect	9,006	12,659	9,105	12,345	10,913	13,078	9,913	13,442
Df	22,472	42,161	31,756	42,845	24,179	40,518	35,338	41,969
Same slope p-value [-1;-10] [18;0] [26;18]	-	0.5463	0.0000	0.0000	-	0.0000	0.3562	0.0000
Same slope p-value [0;-10][18;0]	0.0108	0.3561	0.1542	0.3481	0.1007	0.0011	0.3131	0.5082
Same slope p-value [18;0] [26;18]	-	0.9496	0.0000	0.0000	-	0.0027	0.1823	0.0000
Same slope M/F p-value [0;-10]	0.6561	0.0981	0.0000	0.0231	0.6561	0.0981	0.0000	0.0231
Same slope M/F p-value [18;0]	0.0000	0.0154	0.0000	0.1721	0.0000	0.0154	0.0000	0.1721
Same slope M/F p-value [26;18]	-	0.0000	0.001	0.0000	-	0.0000	0.001	0.0000
Family FE	YES	YES	YES	YES	YES	YES	YES	YES

Note: Immage is the age at immigration interval for the linear estimate. Boom (at age 18) is a dummy for recession or boom in the business cycle when turning 18 years old. Business Cycle Barometer indicate economic state at observation year, see appendix G.1. Standard errors in parentheses ***p< 0.01, **p <0.05, *p<0.1

F Sensitivity analysis

Table 17: Binary variable - other immigrants with controls

VARIABLE (age)	Attained Education		Unemployment days		Wage(log)		Welfare Benefits	
Immage (-10)	-0.360***	(0.108)	1.461	(2.027)	0.281	(0.597)	-0.461	(3.006)
Immage (-9)	-0.397***	(0.106)	4.099**	(2.002)	-0.433	(0.589)	3.336	(2.974)
Immage (-8)	-0.290***	(0.101)	-1.314	(1.909)	0.412	(0.565)	-0.609	(2.849)
Immage (-7)	-0.213**	(0.0979)	2.630	(1.845)	0.213	(0.546)	0.418	(2.752)
Immage (-6)	-0.164*	(0.0946)	0.483	(1.792)	0.0875	(0.531)	2.224	(2.678)
Immage (-5)	-0.254***	(0.0937)	-0.162	(1.773)	0.300	(0.527)	-0.596	(2.653)
Immage (-4)	-0.111	(0.0925)	0.362	(1.746)	0.550	(0.520)	3.082	(2.613)
Immage (-3)	0.0637	(0.0908)	-0.00954	(1.719)	0.278	(0.509)	0.340	(2.572)
Immage (-2)	0.0354	(0.0917)	0.0308	(1.735)	0.394	(0.515)	0.808	(2.598)
Immage (-1)	-0.0964	(0.0942)	0.175	(1.775)	0.186	(0.529)	2.909	(2.658)
Reference group (0)	0	(0.0735)	0	(1.371)	0	(0.408)	0	(2.052)
Immage (1)	-0.162	(0.117)	6.155***	(2.108)	-1.001	(0.628)	1.656	(3.133)
Immage (2)	-0.0290	(0.123)	6.237***	(2.182)	-1.067	(0.653)	-2.836	(3.243)
Immage (3)	-0.0169	(0.124)	2.430	(2.266)	-0.325	(0.677)	0.656	(3.355)
Immage (4)	-0.0362	(0.129)	2.602	(2.330)	-0.609	(0.687)	-2.142	(3.438)
Immage (5)	0.0402	(0.133)	0.659	(2.417)	-0.610	(0.711)	-0.0264	(3.546)
Immage (6)	-0.263*	(0.142)	3.195	(2.504)	-0.872	(0.747)	-3.933	(3.691)
Immage (7)	-0.0843	(0.149)	1.648	(2.633)	-0.0719	(0.783)	0.320	(3.853)
Immage (8)	-0.228	(0.160)	7.883***	(2.794)	-1.426*	(0.835)	1.911	(4.048)
Immage (9)	-0.0183	(0.167)	5.256*	(2.918)	-2.021**	(0.850)	1.262	(4.129)
Immage (10)	-0.233	(0.174)	7.203**	(3.050)	-0.670	(0.886)	-0.696	(4.325)
Immage (11)	0.0242	(0.178)	9.185***	(3.053)	-2.645***	(0.914)	0.990	(4.442)
Immage (12)	-0.123	(0.182)	12.71***	(3.140)	-3.337***	(0.927)	0.581	(4.521)
Immage (13)	-0.00918	(0.194)	7.346**	(3.235)	0.187	(0.961)	0.370	(4.614)
Immage (14)	-0.289	(0.199)	3.723	(3.419)	-1.471	(1.014)	-2.390	(4.796)
Immage (15)	-0.119	(0.218)	5.920	(3.737)	-3.594***	(1.099)	-2.547	(5.266)
Immage (16)	0.186	(0.234)	10.02***	(3.827)	-4.586***	(1.063)	10.59**	(5.112)
Immage (17)	-0.329	(0.255)	7.404**	(3.731)	-3.829***	(0.982)	5.326	(4.800)
Immage (18)	-0.00397	(0.269)	3.463	(3.634)	-2.321**	(0.938)	-3.592	(4.573)
Firstborn	0.202***	(0.0365)	-0.476	(0.658)	0.384**	(0.188)	0.482	(0.962)
Gender	0.639***	(0.0321)	-4.250***	(0.602)	-5.681***	(0.173)	-3.842***	(0.886)
BCB			-0.0357	(0.0556)			-0.0599	(0.0794)
Observations	49,129		63,869		60,302		67,964	
R-squared	0.737		0.684		0.704		0.655	
FamilyFE	YES		YES		YES		YES	

Note: Age at immigration as binary variable - The estimates from family fixed effects specifications. The reference group is the age-zero immigrant. Standard errors in parentheses ***p< 0.01, **p <0.05, *p<0.1

Table 18: Binary variable - refugee age 34 to 36 with controls

VARIABLE (age)	Attained Education		Unemployment days		Wage(log)		Welfare Benefits	
Immage (-10)	-0.646	(0.803)	-20.94	(14.90)	4.025	(4.010)	0.294	(26.21)
Immage (-9)	0.340	(0.723)	-12.66	(14.19)	1.554	(3.830)	-6.591	(24.98)
Immage (-8)	-0.0525	(0.531)	-27.40**	(10.68)	3.846	(2.937)	-0.133	(18.79)
Immage (-7)	-0.234	(0.414)	-3.658	(8.187)	0.742	(2.246)	2.941	(14.40)
Immage (-6)	0.136	(0.356)	3.948	(7.453)	2.238	(2.029)	7.579	(13.10)
Immage (-5)	-0.238	(0.324)	-1.679	(6.905)	2.104	(1.859)	10.86	(12.16)
Immage (-4)	-0.0723	(0.232)	-2.477	(5.177)	1.562	(1.369)	7.988	(9.102)
Immage (-3)	-0.00521	(0.209)	0.825	(4.776)	1.882	(1.252)	-5.402	(8.406)
Immage (-2)	-0.0587	(0.208)	2.460	(4.773)	-0.573	(1.230)	4.906	(8.403)
Immage (-1)	-0.447**	(0.200)	2.437	(4.660)	-0.247	(1.198)	1.360	(8.206)
Reference group (0)	0	(0.157)	0	(11.28)	0	(0.915)	0	(6.045)
Immage (1)	-0.283	(0.173)	5.935	(4.104)	-2.042*	(1.050)	5.932	(7.225)
Immage (2)	-0.545***	(0.185)	6.753	(4.372)	-3.897***	(1.139)	2.096	(7.696)
Immage (3)	-0.569***	(0.189)	7.055	(4.470)	-3.703***	(1.150)	5.820	(7.866)
Immage (4)	-0.473**	(0.188)	2.062	(4.400)	-3.099***	(1.149)	8.362	(7.728)
Immage (5)	-0.531***	(0.184)	2.932	(4.298)	-4.497***	(1.127)	11.23	(7.537)
Immage (6)	-0.400**	(0.188)	8.153*	(4.349)	-5.294***	(1.137)	-1.830	(7.621)
Immage (7)	-0.739***	(0.189)	4.869	(4.341)	-5.626***	(1.141)	13.24*	(7.596)
Immage (8)	-0.736***	(0.190)	9.701**	(4.324)	-7.160***	(1.135)	15.80**	(7.539)
Immage (9)	-0.700***	(0.190)	8.851**	(4.273)	-6.427***	(1.126)	13.22*	(7.435)
Immage (10)	-0.765***	(0.193)	13.11***	(4.294)	-8.281***	(1.131)	15.96**	(7.453)
Immage (11)	-0.802***	(0.194)	9.298**	(4.284)	-7.869***	(1.134)	17.18**	(7.409)
Immage (12)	-0.911***	(0.196)	9.942**	(4.289)	-8.113***	(1.143)	11.27	(7.398)
Immage (13)	-0.954***	(0.196)	12.36***	(4.266)	-8.598***	(1.133)	13.16*	(7.346)
Immage (14)	-1.069***	(0.198)	12.51***	(4.259)	-10.47***	(1.133)	19.45***	(7.305)
Immage (15)	-1.023***	(0.202)	18.20***	(4.322)	-9.307***	(1.153)	19.67***	(7.348)
Immage (16)	-1.277***	(0.207)	14.36***	(4.375)	-10.38***	(1.173)	26.31***	(7.378)
Immage (17)	-1.490***	(0.214)	18.79***	(4.479)	-10.55***	(1.204)	20.56***	(7.441)
Immage (18)	-1.816***	(0.218)	19.24***	(4.479)	-10.60***	(1.218)	16.89**	(7.451)
Firstborn	0.236***	(0.0515)	-2.143*	(1.134)	-0.276	(0.308)	3.043	(1.995)
Gender	0.623***	(0.0424)	-0.652	(0.979)	-3.456***	(0.262)	-9.511***	(1.721)
BCB	-	-	0.0546	(0.104)	-	-	-0.243	(0.183)
Observations	31,004		32,676		26,964		32,676	
R-squared	0.746		0.642		0.720		0.669	
Family FE	YES		YES		YES		YES	

Note: Age at immigration as binary variable - The estimates from family fixed effects specifications. The reference group is the age-zero immigrant. Standard errors in parentheses ***p< 0.01, **p <0.05, *p<0.1

G Descriptive statistics

Table 19: Descriptive statistics - orphans

VARIABLE	Obs.	Mean	Std. Dev	Min	Max
Birth year	4,397	1980	7.291	1969	1984
Immigration year	4,397	1981	7.414	1973	1991
Age at immigration	4,397	0.999	1.239	0	5
Female==1	4,397	0.51	0.499	0	1
Attained education	4,396	13.34	2.88	0	17
Unemployment days	4,397	22.31	47.89	0	365
Wage(log)	3,931	7.256	0.979	2.303	8.963
Welfare benefts	4,397	19.58	95.85	0	1083

Note: The sibling sample of refugee orphans used in the sensitivity analysis. The time span of observation lies between 1990 to 2014, the earliest possible immigration is 1960. Outcomes measured in age 28 to 30.

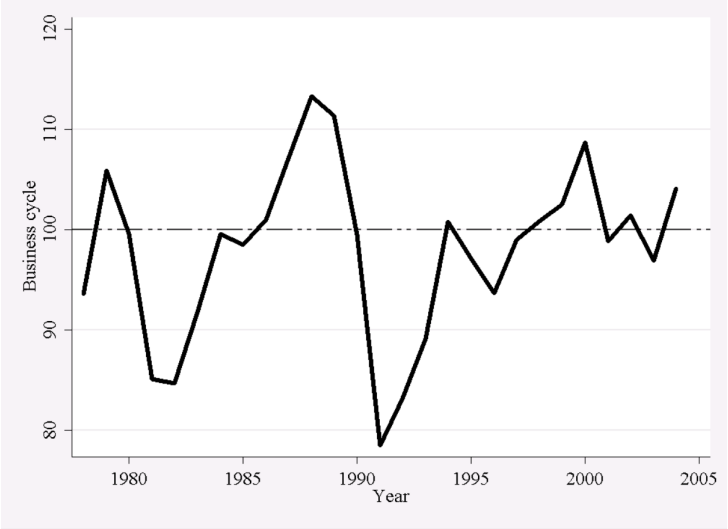
Table 20: Descriptive statistics - refugee sample aged 34 to 36

VARIABLE	Obs.	Mean	Std. Dev	Min	Max
Birth Year	71,903	1971	6.68	1960	1978
Immigration Year	71,903	1988	7.106	1970	2004
Age at Immigration	71,903	17.11	7.85	-10	26
Female==1	71,903	0.53	0.499	0	1
Attained Education	60,111	12.25	3.257	0	17
Unemployment days	70,791	37.86	66.67	0	365
Wage(log)	55,862	7.03	1.202	2.303	10.60
Welfare benefts	71,903	29.55	104.16	0	2740
Year first born	25,581	1967	6.21	1960	1978
Number of siblings	71,903	2.50	0.899	2	6
Parent highest Education	42,532	9,23	2,57	0	17

Note: The sibling sample of refugee orphans used in the sensitivity analysis. The time span of observation lies between 1990 to 2014, the earliest possible immigration is 1960. Outcomes measured in age 28 to 30.

G.1 Business Cycle Barometer (BCB)

Figure 17: Business Cycle Barometer



Note: The figure illustrate when the economy is booming or in a recession.

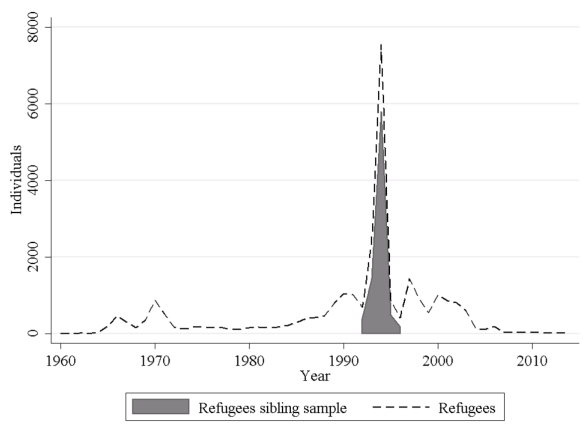
H Countries and exogenous shocks

Country	Period
Yugoslavia	1990 - 1998
Bosnia	1992 - 1998
Lebanon	1984 - 1994
Syria	1986 - 2009
Turkey	1970 - 1973
Iraq	1984 - 2009
Iran	1984 - 1995
Vietnam	1978 - 1993
Afghanistan	1988 - 2010
Bangladesh	1978 - 2010
Pakistan	1986 - 2010
Sri Lanka	1974 - 1994
Gambia	1980 - 2010
Ethiopia	1985 - 1992
Somalia	1989 - 2006
Uganda	1980 - 1998
Erithrea	1978 - 1993
Chile	1973 - 1993
Bolivia	1980 - 1986
Peru	1980 - 2003
Colombia	1973 -1994

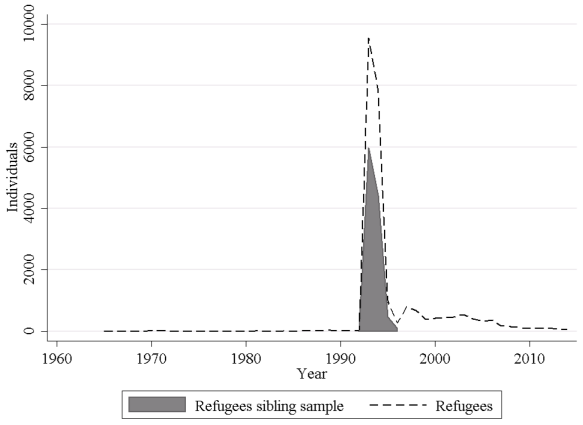
Note: Period indicate the year span to collect the refugee sibling sample.

The following figures illustrate the last immigration year of first generation refugees between age 0 to 26. The observation is made if the refugee turn 28 to 30 during the time span 1990 to 2014 in Sweden.

H.1 Europe

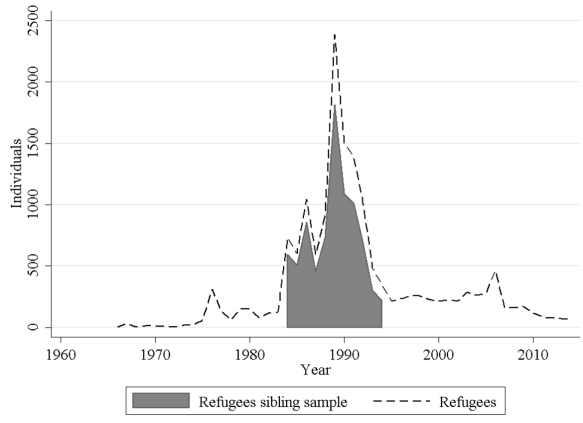


(a) Yugoslavia

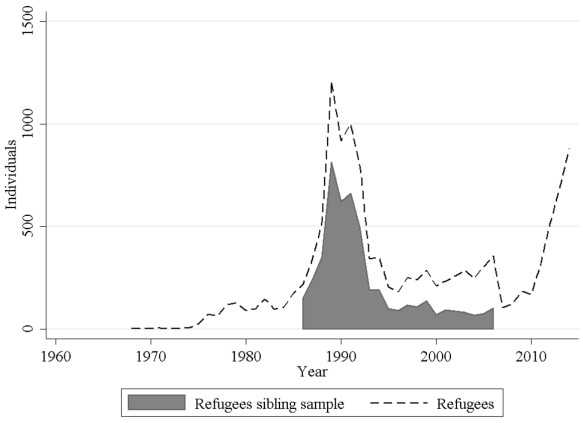


(b) Bosnia

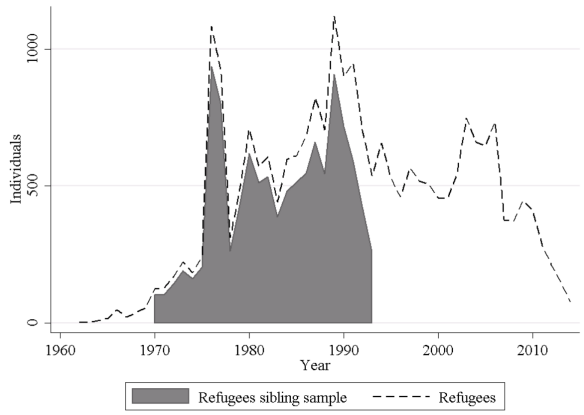
H.2 West Asia



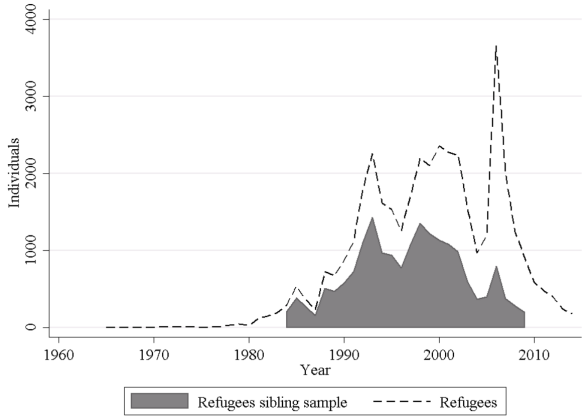
(a) Lebanon



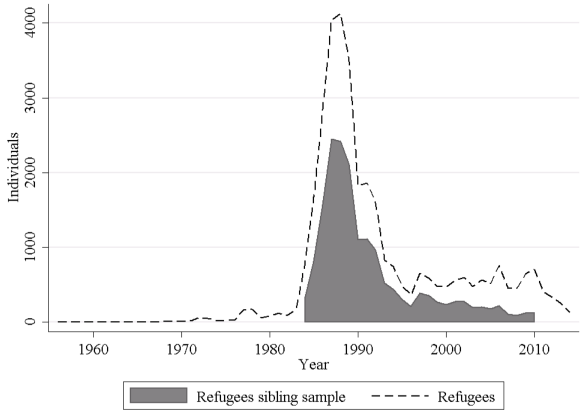
(b) Syria



(a) Turkey

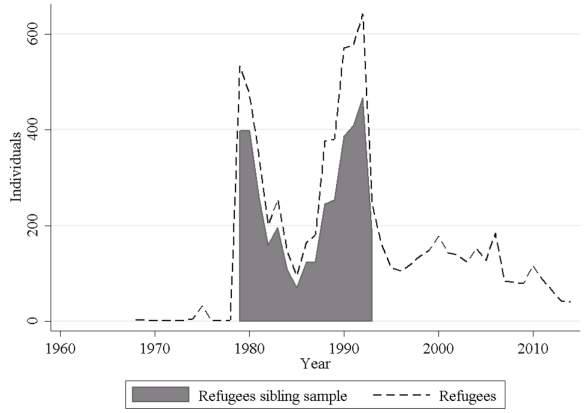


(b) Iraq

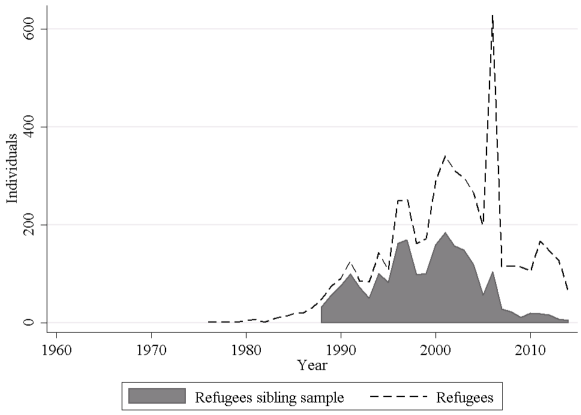


(a) Iran

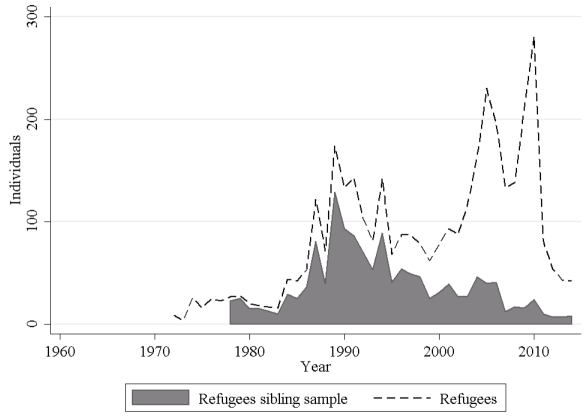
H.3 Central Asia



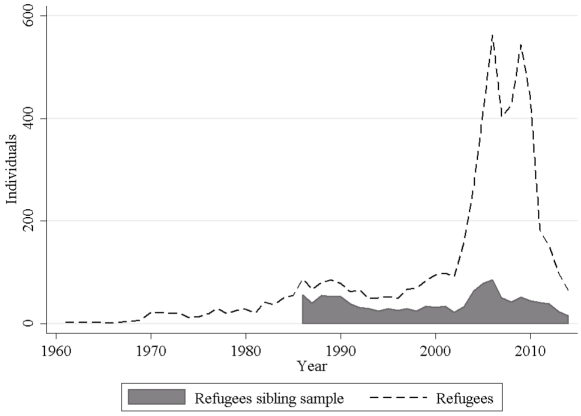
(a) Vietnam



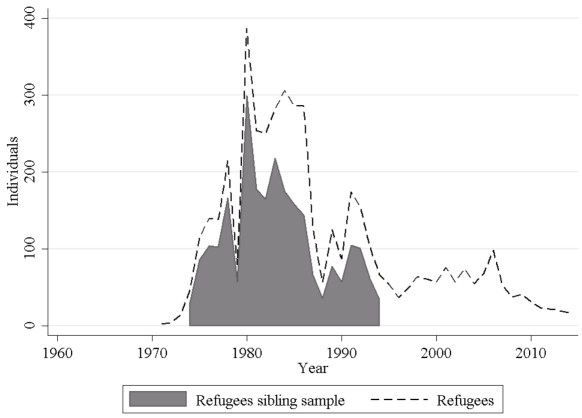
(b) Afghanistan



(a) Bangladesh

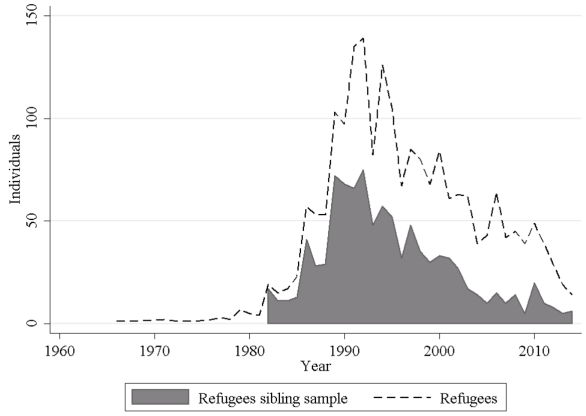


(b) Pakistan

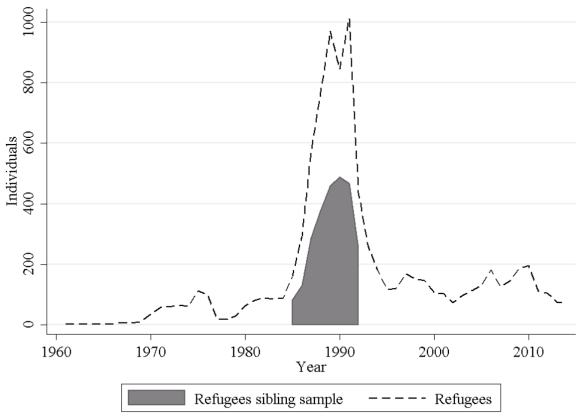


(a) Sri Lanka

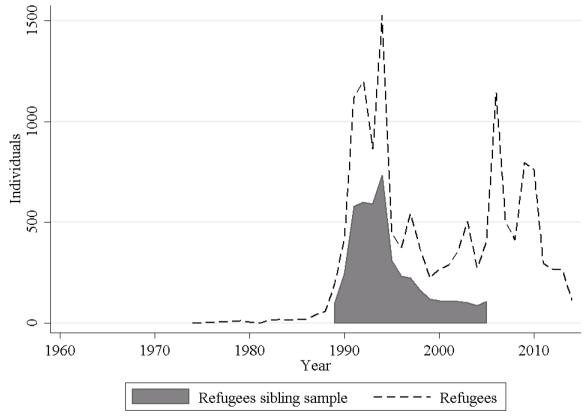
H.4 Africa



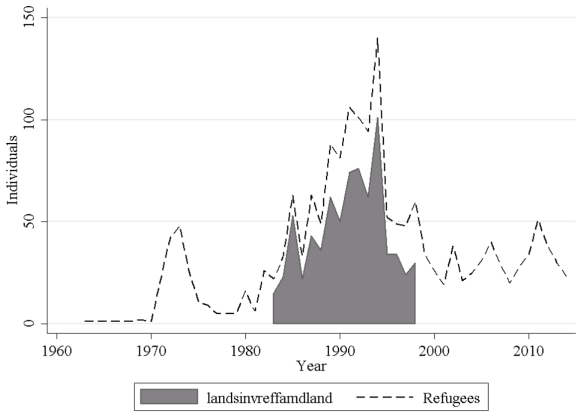
(a) Gambia



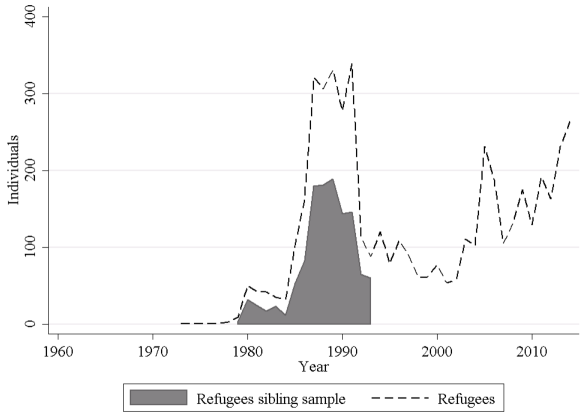
(b) Ethiopia



(a) Somalia

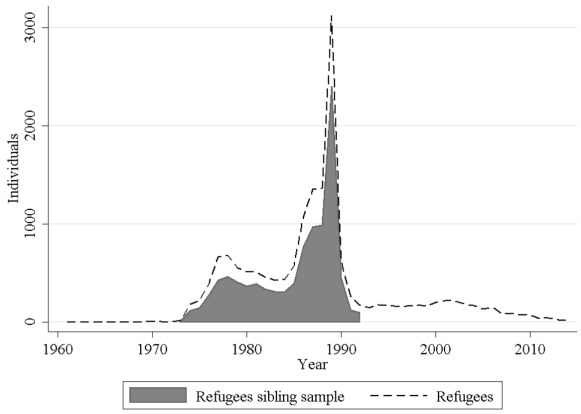


(b) Uganda

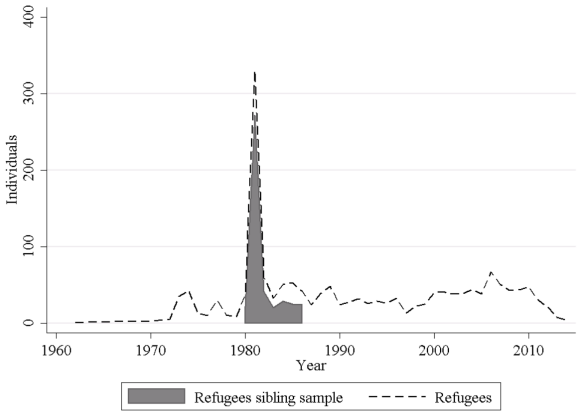


(a) Eritrea

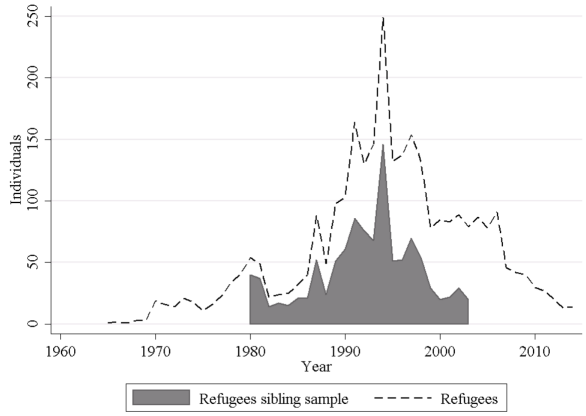
H.5 South America



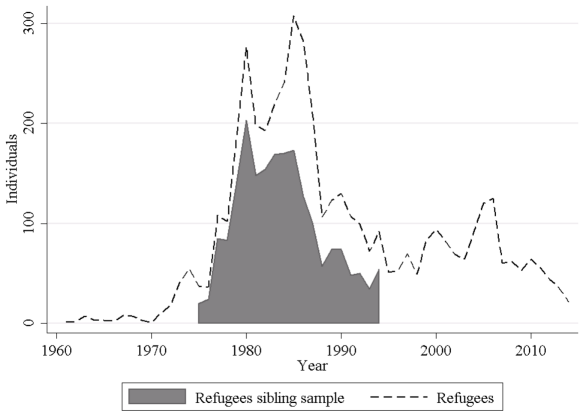
(a) Chile



(b) Bolivia



(a) Peru



(b) Colombia

I Other immigrants

Country	Period
Denmark	1950 - 2010
Finland	1950 - 2010
Norway	1950 - 2010
Germany	1950 -2010
Poland	1950 - 2010
Great Britain	1950 - 2010
Greece	1950 - 1988
Italy	1950 - 1980