Aksoy, Cevat Giray; Poutvaara, Panu

Working Paper

Refugees' and irregular migrants' self-selection into Europe: Who migrates where?

SOEPpapers on Multidisciplinary Panel Data Research, No. 1069

Provided in Cooperation with:
German Institute for Economic Research (DIW Berlin)

Suggested Citation: Aksoy, Cevat Giray; Poutvaara, Panu (2020) : Refugees' and irregular migrants' self-selection into Europe: Who migrates where?, SOEPpapers on Multidisciplinary Panel Data Research, No. 1069, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:
https://hdl.handle.net/10419/214241

SOEP The Socio-Economic Panel

1069 2020

SOEPpapers
on Multidisciplinary Panel Data Research

SOEP – The German Socio-Economic Panel at DIW Berlin 1069-2020

Refugees' and Irregular Migrants' Self-selection into Europe: Who Migrates Where?

Cevat Giray Aksoy and Panu Poutvaara

DIW SOEP

SOEPpapers on Multidisciplinary Panel Data Research at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at **http://www.diw.de/soeppapers**

ISSN: 1864-6689 (online)

German Socio-Economic Panel (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: soeppapers@diw.de

Refugees' and Irregular Migrants' Self-selection into Europe: Who Migrates Where?

Cevat Giray Aksoy and Panu Poutvaara*

Abstract

We analyze self-selection of refugees and irregular migrants and test our theory in the context of the European refugee crisis. Using unique datasets from the International Organization for Migration and Gallup World Polls, we provide the first large-scale evidence on reasons to emigrate, and the self-selection and sorting of refugees and irregular migrants. Refugees and female irregular migrants are positively self-selected with respect to human capital, while male irregular migrants are negatively self-selected. These patterns are similar when analyzing individually stated main reason to emigrate, country-level conflict intensity, and sub-regional conflict intensity. Migrants respond to economic incentives and border policies.

JEL Codes: F22, J15, J16, J24, O15

Keywords: international migration, refugees, irregular migrants, self-selection, human capital, gender differences in migration

*Aksoy is a Principal Economist at the European Bank for Reconstruction and Development, part-time Assistant Professor of Economics at King's College London, and Research Associate at IZA Institute of Labor Economics and London School of Economics, aksoyc@ebrd.com. Poutvaara is a Professor at the University of Munich, Director of the ifo Center for International Institutional Comparisons and Migration Research at the ifo Institute and Research Fellow at CESifo, CReAM, and IZA Institute of Labor Economics, poutvaara@ifo.de. We are grateful to Nuno Nunes, Ivona Zakoska-Todorovska, and the International Organization for Migration (IOM) for kindly providing the Flow Monitoring Surveys. We also thank Giacomo Battiston, Michal Burzynski, Daniele Coen-Pirani, Ralph De Haas, Anthony Edo, Guido Friebel, Kai Gehring, Yvonne Giesing, Thomas Ginn, Sergei Guriev, Hubert Jayet, Nadzeya Laurentsyeva, Till Nikolka, Caglar Ozden, Helmut Rainer, Carla Rhode, Dominik Sachs, Orkun Saka and participants at the AFD-World Bank Migration and Development Conference, CEPR-LSE Migration Workshop, CEPS, CERDI, *CES*ifo Conference on Employment and Social Protection, CReAM/RWI Workshop on Migration, COMPAS and Queen Mary Forced Displacement Conference, EPCS, German Economic Association Conference, Institute for Employment Research (IAB), IZA Annual Migration Meeting, OECD Migration Conference, SOLE, Silvaplana Political Economy Workshop, University of Glasgow Migration and Mobility Workshop, University of Luxembourg, University of Pittsburgh Migration Workshop and University of Trento for their helpful comments. Zalina Alborova, Benedikt Behringer, Stefan Pasch and Michael Ganslmeier provided outstanding research assistance. Views presented are those of the authors and not necessarily of the EBRD, IOM, or any other organization. All interpretations, errors, and omissions are our own.

1. Introduction

Nearly 66 million people were forcibly displaced worldwide at the end of 2016 (UNHCR, 2017). The total number of people seeking safety across international borders as refugees reached 22.5 million, with more than half of all refugees worldwide coming from only three countries: Afghanistan, South Sudan, and Syria (UNHCR, 2017). The 1951 Refugee Convention and its extension in 1967 define a refugee as a person who is outside his or her country of nationality "owing to well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion," and is unable or unwilling to return there. The Convention forbids returning a refugee to a territory where his or her life or freedom would be threatened due to persecution (UNHCR, 1967). An irregular migrant, in contrast, is broadly defined as a person who travels abroad voluntarily in search of economic opportunities, but has no legal right to remain in the intended destination country. Distinguishing between refugees and irregular migrants is complicated: an irregular migrant has a strong incentive to claim to be a refugee to obtain permission to stay. Receiving countries, instead, may aim to reduce the number of refugees by tightening the burden of proof for persecution.

As policymakers try to make sense of a complex reality, it is important to understand the self-selection of refugees and irregular migrants in terms of their skills and demographic characteristics. This is the focus of our paper. We provide the first large-scale systematic evidence on the motivations, self-selection, and intended sorting of refugees and irregular migrants from multiple countries of origin.

We make four main contributions. We start by presenting a model of how refugee self-selection in the presence of risks related to staying in an unsafe country and risks related to unsafe passage differs from the self-selection of irregular economic migrants. We then provide evidence on the motivations and intended destinations of both refugees and irregular migrants who arrived in Europe in 2015 or 2016 using data from the International Organization for Migration's Flow Monitoring Surveys (FMS). Next, we test our theory by combining FMS with Gallup World Polls to understand how migrant groups are self-selected from the origin population in terms of education and predicted income, and how this self-

selection differs between refugees and irregular migrants.[1] As a last step, we examine how the country characteristics (such as the duration of the asylum procedure and social spending) and border policies are associated with the sorting of migrants into different intended destination countries.

The number of refugees and irregular migrants is at its highest level ever, even surpassing the post-World War II numbers (UNHCR, 2019). In 2016, the ten main source countries of refugees were in Asia and Africa, as well as nine of the ten largest refugee-hosting countries (see Figure 1). Although most refugees stay in developing countries (Chin and Cortes 2015; Hatton 2016), about 1.6 million refugees and irregular migrants arrived in Europe between January 2015 and December 2017 (see Figure 2). Refugee migration is also at the top of the political agenda in the United States, Canada, and Australia, which have resettled, respectively, about 3 million, 0.7 million and 0.5 million refugees since 1980.[2] Moreover, self-selection of irregular migrants is a much-debated topic in the United States. The estimated total number of irregular migrants in the United States was 10.5 million in 2017, representing 3.2 percent of the total U.S. population that year (Pew Research Center, 2019).

Understanding the motivations of migrants, their self-selection in terms of socio-demographic characteristics, and their intended destination countries is important for several reasons. First, knowing the motivations of migrants helps to distinguish between the refugee crisis and challenges associated with irregular migration.[3] Second, refugees' self-selection has implications for rebuilding their home countries. The more skilled refugees are, the more difficult it is to fill the gap they leave once the country enters the reconstruction stage. Third, knowing the skill distribution and intended destinations of refugees who make it to the transit

[1] We use the term "refugees and irregular migrants" to refer to individuals surveyed in Flow Monitoring Surveys. We will sometimes refer to each group separately. We will also sometimes refer to both groups as migrants. When analyzing self-selection with respect to education and predicted income, we restrict the attention to 9,828 respondents aged 25 to 64, in order to focus on migrants who have likely already completed their education and are not yet retired.

[2] While most refugees in European countries first enter and apply for asylum, the United States, Canada and Australia admit refugees mostly through resettlement programs, based on referrals by UNHCR and other organizations. From 1980 to 2016, the United States resettled in each year more refugees than the rest of the world together, but the Trump administration cut the number of admissions from the peak of 97,000 in 2016 to 23,000 in 2018. In 2018, Canada overtook the United States as the country resettling the largest absolute number of refugees (Pew Research Center, 2019).

[3] Irregular migrants do not have a legal right to stay in their intended destination country, but returning them to their country of origin is often costly and complicated. Irregular migrants often work in the shadow economy and pose a security concern as those who have not applied for asylum have not been subject to normal security screening, and may resort to crime. They also face a higher risk of becoming a victim of crime, and their human rights are infringed in many transit and destination countries (UNHCR, 2008).

countries is helpful in planning integration policies, and thereby contributes to social stability in host countries and intended destination countries.

Building on Roy (1951), Borjas (1987) showed that if skills are sufficiently transferable across countries, theory suggests that migrants from a less egalitarian to a more egalitarian country should come from the lower end of the skill distribution, while migrants from a more egalitarian to a less egalitarian country should come from the upper end of the skill distribution. Given that most African and Middle Eastern countries have much wider income differences than European countries, the Roy-Borjas model suggests that economic migrants from these regions to European welfare states should come from the lower end of the skill distribution. Subsequently, Grogger and Hanson (2011) presented a model that also builds on the Roy model but assumes linear utility and absolute cost of migration, instead of time-equivalent costs as in Borjas (1987). The model by Grogger and Hanson (2011) predicts that migrants are also generally positively self-selected from less egalitarian countries and that the sorting of migrants depends on relative skill prices in different destinations. Therefore, more highly educated migrants are more likely to settle in countries that offer high rewards for skill.

Our theoretical predictions show that a severe conflict can reverse the self-selection pattern from an inegalitarian country. The key mechanism is that a sufficiently high risk of staying in an unsafe country lowers expected returns to human capital below that available in more egalitarian destination countries. This novel result contrasts with the previous assumption that self-selection in terms of productivity is less intense for migrants who leave their home country based on non-economic motives (Chiswick, 1999, p. 181). This positive self-selection from countries plagued by major conflict arises without assuming borrowing constraints. Borrowing constraints would further strengthen the result as those with more education are likely to have more resources that allow them to leave.

We find important differences in how refugees and irregular migrants are self-selected relative to the country of origin population in terms of their education. Both male and female migrants whose main reason to emigrate was conflict or persecution are positively self-selected with respect to secondary and tertiary education. Female irregular migrants (those motivated by other reasons than conflict or persecution) are also positively self-selected in terms of their education, although not as strongly as female refugees. Male irregular migrants, instead, are weakly negatively self-selected in terms of their education. Negative self-selection of male irregular migrants is in line with Borjas (1987) prediction that migrants from countries with wider income differences should be negatively self-selected in terms of

their productivity. Positive self-selection of women, instead, can be rationalized by very low female labor force participation rates and severe gender discrimination in almost all origin countries as these imply a substantially lower rate of return for each year of education for women as compared to men.[4]

We also analyze self-selection according to conflict intensity, irrespective of migrants' individually stated reason to emigrate. We compare how self-selection of migrants from major conflict countries with more than 1000 battle-related deaths (that is, major conflict countries) in a single year differs from self-selection of migrants from countries with fewer or no battle-related deaths (that is, minor or no conflict countries). Results are similar to those found using individually stated main reason to migrate: both male and female migrants from major conflict countries are positively self-selected in terms of their education. When it comes to migrants from minor or no conflict countries, women are also positively self-selected, while the negative self-selection pattern is found for men.

A major concern in any cross-country analysis is the omitted variable bias. In a seminal contribution, Acemoglu et al. (2001) suggest that European colonial powers adopted more extractive institutions in countries with higher settler mortality and that these institutions have a long-lasting detrimental effect on economic performance. In our setting, some omitted factors, such as extractive institutions, could both increase the risk of conflict and depress opportunities available to those with higher education, pushing them to emigrate with a higher probability. This could generate a spurious correlation between self-selection at the upper part of the skill distribution and conflict. To address this concern, we exploit variation in conflict intensity at the sub-regional level.[5] We scrape geo-coded conflict data from Uppsala Conflict Database for all countries with major or minor conflict in the 2009-2014 period and use these data to identify high conflict and low conflict intensity sub-regions for each country, based on whether the number of battle-related deaths is above or below the country median. We then analyze how migrants from each sub-region are self-selected from

[4] According to the United Nations' Gender Inequality Index (2017), all of the origin countries we have in our sample score poorly and are clearly below Western European countries.

[5] Glaeser et al. (2004) argue that European colonists brought along both their human capital and institutions, and that better human capital tends to result in better institutions and higher growth. We do not take stance on the relative strength of the mechanisms highlighted by Acemoglu et al. (2001) and Glaeser et al. (2004). What is crucial to motivate our sub-regional analysis is that there is a strong correlation between human capital and institutions in present-day world, and that bad institutions increase the risk of conflict (Besley and Persson 2011). For example, Hanushek and Woessmann (2015) find that the effect of cognitive skills on economic growth is higher in countries that have been open to international trade, as well as in countries with better protection against expropriation. Focusing on sub-regional differences in self-selection allows us to control for all country-specific factors, independently of the direction of causality.

the origin population living in that sub-region in terms of their education and other demographic characteristics.[6] Consistent with our main results, we find that migrants who left from high conflict intensity sub-regions are positively self-selected in terms of education, as are female migrants from sub-regions with less intensive conflict, although to a lesser extent. Male migrants from sub-regions with less intensive conflict are, instead, negatively self-selected in terms of their education.[7]

Taken together, the self-selection patterns are similar when analyzing: (i) individually stated main reason to emigrate; (ii) country-level conflict intensity; and (iii) sub-regional conflict intensity. Importantly, our sub-regional level findings suggest that the effect of conflict on self-selection of migrants goes beyond mere correlation. The reason is that when analyzing self-selection at the sub-regional level we can include country fixed effects to account for omitted variables that could influence both the probability of conflict and self-selection. Moreover, having consistent results when using alternative conflict definitions strongly implies that migrants do not systematically misreport while answering the survey (also see Figure 6). They also did not have any incentive to lie as the survey was anonymous, no contact or other identifying information was collected, and the respondents were informed that their answers did not influence their legal status.

To further investigate the self-selection, we also compare the predicted earnings of refugees and irregular migrants and non-migrants. As GWP collects earnings at the household level, this analysis is restricted to singles. We find that refugees and female irregular migrants are strongly positively self-selected in terms of their predicted earnings, while male irregular migrants are negatively self-selected.

Using unique datasets, we further complement our analysis by investigating the self-selection patterns into Germany and Turkey. Consistent with our main results, we find a strong positive self-selection of refugees with respect to education in Germany. When it comes to Turkey, we find that refugees are strongly positively self-selected and irregular migrants strongly negatively self-selected. Unlike in Europe, female irregular migrants are not positively self-selected.

[6] With the exception of Iran, each major or minor conflict country included in our sample was a colony (or part of a colony) before its independence. Country fixed effects capture the effects of colonial past.

[7] We also find that main motivations of migrants are in line with our sub-regional conflict measure. In all countries, more migrants emigrated because of conflict or persecution from high conflict intensity sub-regions. For example, in Afghanistan, Iraq, and Syria, the share of men (women) emigrating from high conflict intensity sub-regions because of conflict or persecution is 31 to 53 (16 to 33) percentage points higher than among those emigrating from sub-regions with less intensive conflict.

The remainder of the paper is organized as follows: Section 2 reviews the related literature. Section 3 discusses push and pull factors for international migration. Section 4 outlines a theoretical framework for understanding the self-selection mechanisms. Section 5 documents the data sources. Section 6 shows descriptive statistics. Section 7 describes the estimation strategy. Section 8 presents the results concerning refugee self-selection into Europe, section 9 on sorting, and section 10 on border policies. Section 11 presents results on self-selection into Turkey, after which section 12 concludes.

2. Related Literature

Although there is a large literature on migrant self-selection, its focus has been on economic migrants (see Borjas 1987; Chiquiar and Hanson 2005; Hatton and Williamson 1998; Fernandez-Huertas Moraga 2011; Grogger and Hanson 2011; Abramitzky et al. 2012; Belot and Hatton 2012; Parey et al. 2017; Hendricks and Schoellman 2018; Borjas et al. 2019). However, refugee migration has been significant throughout history, with the United States being the most prominent destination. Religious minorities fleeing persecution played an important role in early colonies. In the late 19th and early 20th century, the United States maintained open borders to European immigrants, whether economic migrants or refugees and these flows decisively shaped America (Bandiera et al. 2013; Abramitzky and Boustan 2017; Sequeira et al., forthcoming).

Research related to refugee migration has focused on assimilation, and the socio-economic and political effects on countries of origin or host countries, instead of self-selection. Most studies use refugee flows as an exogenous source of variation to study questions where identification is otherwise difficult. Cortes (2004) compares human capital investments and earnings growth between refugees and economic migrants who arrived in the United States from 1975 to 1980. She finds that refugees invested more in human capital and made greater earnings gains, which can be explained by refugees having longer expected time horizon in the host country as they lack the option of returning to their country of origin. Waldinger (2010) uses the exogenous drop in faculty quality following Nazis dismissing all Jewish professors, as well as professors they considered politically unreliable to show that faculty quality has a major impact on Ph.D. student outcomes. Moser et al. (2014) show that refugee scientists from Nazi Germany helped to boost US science not only through their own direct contributions but also by attracting new researchers to their fields. Hornung (2014)

goes even further back in history and analyzes the impact of the 1685 settlement of French Protestants (Huguenots) who were persecuted because of their religion in Prussia. He finds that Huguenot settlement boosted the long-run productivity of the Prussian textile industry. Focusing on asylum seekers in Switzerland, Couttenier et al. (forthcoming) examine whether the past exposure to conflict in origin countries makes migrants more prone to violence in their host country. They show that cohorts exposed to civil conflicts during childhood are more likely to commit violent crimes than their co-nationals born after the conflict.

There are a few studies that analyze the self-selection of refugees from some individual countries. Cohen (2007) compares the educational levels and earnings assimilation of Jewish immigrants from the former Soviet Union in the United States and Israel and Birgier et al. (2018) examines the self-selection patterns in relation to Argentinian and Chilean refugees' economic assimilation in the United States, Sweden, and Israel. Brücker et al. (2016) show that recently arrived refugees in Germany have a relatively low level of education compared with the German population. Guichard (2017) finds that refugees from Afghanistan, Iraq, and Syria to Germany are positively self-selected in terms of education relative to the population at the origin, which is in line with what our theory predicts. However, there is no systematic analysis on the self-selection of refugees from multiple countries of origin into multiple destinations.

The European refugee crisis in 2015 and 2016 differs from that in the 1990s in three crucial respects, as summarized by Dustmann et al. (2017). First, the current crisis has had a much stronger political impact on Europe, which was already divided by populist and separatist national movements and weakened by the Great Recession and the Eurozone crisis. Second, the current crisis involves multiple actors over which Western nations have much weaker influence than in conflicts in the 1990s. Third, the refugees arriving in Europe are perceived to be culturally more distant than those from former Yugoslavia in the early 1990s. In 2015 and 2016, concerns about refugees and irregular migrants resulted in the re-introduction of border controls inside the Schengen area, disrupting the central principle of intra-EU free mobility and intra-European trade and supply chains. The refugee crisis also played a central role in the Leave campaign against the United Kingdom’s membership of the European Union (despite the United Kingdom being outside the Schengen area and, therefore, relatively unaffected by the refugee crisis). Worries about immigration have powered the rise of populist parties and candidates in Austria (Halla et al. 2017), Denmark (Dustmann et al., 2019), France (Edo et al., 2019), Germany (Otto and Steinhardt 2014), and Greece (Dinas et al. 2019). This makes it all the more important that we understand the socio-

demographic characteristics of such refugees and migrants, many of whom are likely to apply for asylum in their destination countries.

Finally, our paper is also related to the literature on irregular migration. In an early theoretical contribution, Ethier (1986) shows that using a mixture of border controls and internal enforcement helps to reduce the costs of deterring illegal immigration. Friebel and Guriev (2006) present a model in which illegal immigrants may enter into temporary servitude contracts to repay smuggler services and show that in that case, stricter deportation policies may backfire as they make enforcing servitude for illegal immigrants in illegal sector easier. Most of the empirical research on irregular migration has focused on migration from Mexico to the United States (Hanson and Spilimbergo 1999; Orrenius and Zavodny 2005). We contribute to this literature by analyzing how irregular migrants arriving in Europe from multiple countries of origin are self-selected, and how this self-selection differs from the self-selection of refugees.

3. Push and Pull Factors for International Migration

Most of the population growth throughout human history took place in the 20th century. According to the United Nations, the world population reached one billion in 1804, two billion in 1927, and three billion in 1960.[8] Subsequently, it reached four billion in 1974 and seven billion in 2011. Most of the growth has taken place in Africa and Asia. From 1950 to 2015, the share of Africa of the world population increased from 9.0 percent to 16.0 percent and that of Asia from 55.4 percent to 60.1 percent. At the same time, the population share of Europe declined from 21.7 percent to 10.1 percent.[9]

Figure 3 depicts the change in the population living in Africa, Asia, Europe, and the rest of the world from 1950 to 2019, as well as the United Nations projection (medium variant) until the end of the 21st century (grey background). Notably, the population in Africa has grown rapidly over the last 40 years and more than half of global population growth is expected to take place in Africa from 2029 onwards. In contrast, fertility rates have been below the replacement level (2.1 children per woman) in most of Europe since the 1970s (UN, 2015). As suggested by Malthus (1798), a major concern is that fast population growth will increase demand on shared resources and combined with poverty, it could result in

[8] UN (2014), p. 22.

[9] UN World Population Prospects (2019).

conflict. Indeed, recent papers by Brückner (2010) and Acemoglu et al. (forthcoming) estimate the causal effect of population growth on conflict and find evidence for Malthusian dynamics.[10]

The dramatic demographic divergence around the world goes together with vast earnings gaps. Figure 4 shows the gross domestic product per capita (GDPPC) at purchasing power parity in 10 major origin countries from which refugees and irregular migrants arrived in Europe via Mediterranean routes in 2015 and 2016, as well as in Germany, Sweden, France, and Italy as the European countries that received the largest number of asylum applications in 2015 and 2016 and Turkey as the country hosting the largest number of refugees and irregular migrants. The GDPPC at purchasing power parity in Germany and Sweden is 8 to 35 times higher than the sub-Saharan African countries, and 2.7 times that in Iran, which has the highest GDPPC at purchasing power parity among major countries of origin.

Liquidity constraints and immigration restrictions can mean that the poorest simply cannot migrate, and could help to explain why the highly-skilled have been found to be more mobile (Docquier et al., 2009). Research has also shown that low growth, high (youth) unemployment, and environmental problems are powerful push factors for international migration in poor countries of origin (Beine and Parsons 2015; Cattaneo and Peri 2016; Hatton 2016; Mayda 2010).

4. Theoretical Framework

We analyze emigration from several countries of origin into one destination. We denote countries of origin with index *k*, and the potential destination with index *d*.[11] Our model builds on Borjas (1987) but adds into it gender-specific risks related to conflict or persecution in the home country and to migrating, and gender-specific returns to human capital. As Grogger and Hanson (2011), we simplify the Borjas (1987) framework by leaving out

[10] Brückner (2010) instruments population in sub-Saharan Africa using randomly occurring droughts, while Acemoglu et al. (forthcoming) rely on international epidemiological transition in the 1940s that increased population growth especially in countries that had initially higher mortality from infectious diseases.

[11] We model migration decision between staying in the home country or migrating into one destination, rather than a specific choice between various European countries, as refugees and irregular migrants may face considerable uncertainty on whether they can make it to their preferred destination country. In our empirical analysis, destination *d* may be interpreted as a composite of various European countries, in which different countries receive a weight corresponding to the conditional probability that a migrant from country *k* ends up there in case of reaching Europe.

unobservable skill components in origin and in destination as our data do not allow testing hypotheses related to it. We denote the human capital of individual i beyond primary education by h_i. We denote gender by index g, $g \in \{f, m\}$, in which f denotes females and m denotes males. The wage individual i of gender g would receive in home country k is given by

$$w_i^k = \exp(\alpha_{k,g} + r_{k,g} h_i),$$

in which $\exp(\alpha_{k,g})$ captures the wage available to those of gender g with primary education, and $r_{k,g}$ is the gender-specific return to human capital above primary education in country k. To capture the risks associated with conflict or persecution, we assume that there is a country-specific and gender-specific risk $q_{k,g}$, $0 \leq q_{k,g} < 1$ of losing the wage income and suffering an additional utility loss $L_{k,g}$, $L_{k,g} \geq 0$. The utility is logarithmic in terms of consumption in case of working and linear in terms of the loss in case the risk is realized. It is reasonable to expect $q_{k,g}$ and $L_{k,g}$ to be highest in countries suffering war or civil war, intermediate in countries with a low-intensity conflict or persecution, while it can be viewed as zero, or close to zero, at least for men in relatively safe countries from which migrants are motivated by the lack of job opportunities, rather than conflict or persecution. For women, the risk term could include also gender-based violence and repression, like forced marriage, which could be a problem even in countries that are not suffering war or civil war. The expected utility in case of no migration is therefore

$$EU_i^k = (1 - q_{k,g}) \log(w_i^k) - q_{k,g} L_{k,g}.$$

In case of a successful migration, individual i receives wage $w_i^d = \exp(\alpha_{d,g} + r_{d,g} h_i - \pi_{k,g})$, in which $\exp(\alpha_{d,g})$ depicts the wage available to natives of gender g with primary education in country d, $r_{d,g}$ is the gender-specific return to human capital above primary education in country d and $\pi_{k,g}$ measures the loss of potential productivity due to the imperfect applicability of migrant's human capital and lack of language skills. It plays an equivalent role to time-equivalent migration costs in Borjas (1987), and can differ across countries of origin, for example reflecting linguistic distance and quality of schooling.

To capture the idea that refugees and irregular migrants face various risks also on their way to the destination countries, we assume that migrants of gender g from country k face risk $s_{k,g}$, $0 \leq s_{k,g} < 1$ of not making it to their intended destination in case of trying to

migrate. This risk could include, at its extreme, dying on the way, as well as being caught in a transit country the migrant would prefer not to stay in because of imprisonment, lack of funds, or unexpected border closures. We denote the expected loss in case the risk is realized by $L_{M,g}, L_{M,g} \geq 0$. Gender-specific migration cost is denoted by $c_{k,g}$. We expect lower female migration rates as many traditional societies have social norms that may make traveling alone more difficult for women. Lastly, an idiosyncratic component ε_i depicts various costs and benefits related to migration that are not captured by other terms, including the valuation of different cultural norms and social networks. The expected utility in case of migration is therefore

$$EU_i^d = (1 - s_{k,g}) \log(w_i^d) - s_{k,g} L_{M,g} - c_{k,g} + \varepsilon_i.$$

We assume that ε_i is independently and identically distributed and denote its density function by f and its cumulative distribution function by F, which we assume to be differentiable. We assume that the distribution is sufficiently wide so that under all relevant circumstances, some people migrate and some stay.

It is ex-ante rational to migrate if $EU_i^d > EU_i^k$. This simplifies into the condition

(1) $\varepsilon_i > \varepsilon_i^* = \left[(1 - q_{k,g}) r_{k,g} - (1 - s_{k,g}) r_{d,g}\right] h_i + (1 - q_{k,g}) \alpha_{k,g} - (1 - s_{k,g}) \alpha_{d,g} + \pi_{k,g} - q_{k,g} L_{k,g} + s_{k,g} L_{M,g} + c_{k,g}$.

The probability that individual *i* migrates is given by

(2) $p_i = 1 - F(\varepsilon_i^*)$.

Our main testable prediction relates to migrant self-selection:

Proposition 1. Migrants are positively self-selected in terms of their human capital if $(1 - q_{k,g}) r_{k,g} < (1 - s_{k,g}) r_{d,g}$ and negatively self-selected in terms of their human capital if $(1 - q_{k,g}) r_{k,g} > (1 - s_{k,g}) r_{d,g}$.

Proof. Inserting ε_i^* from equation (1) into equation (2) and partially differentiating gives $\frac{\partial p_i}{\partial h_i} = -\left[(1 - q_{k,g}) r_{k,g} - (1 - s_{k,g}) r_{d,g}\right] f(\varepsilon_i^*)$.

Proposition 1 shows that the self-selection of migrants with respect to their human capital depends not just on gender-specific returns to human capital, but also gender-specific risks related to conflict or persecution if staying in the home country and risks related to

potential migration. Migrants are positively self-selected if $(1-q_{k,g})r_{k,g} < (1-s_{k,g})r_{d,g}$ and negatively self-selected if $(1-q_{k,g})r_{k,g} > (1-s_{k,g})r_{d,g}$. This condition refers to gender-specific risk-adjusted expected returns to human capital and it implies that if returns to skills are higher in the country of origin and the country of origin is relatively safe (a low risk term $q_{k,g}$) migrants are negatively self-selected, in line with Borjas (1987). However, if the country of origin suffers from a sufficiently severe conflict, given by $q_{k,g} > 1 - \frac{(1-s_{k,g})r_{d,g}}{r_{k,g}}$, the self-selection is reversed, and migrants tend to come from the upper part of the skill distribution. If gender-specific return to human capital is lower in the country of origin than the risk-adjusted return $(1-s_{k,g})r_{d,g}$ in the potential destination, migrants are positively self-selected even in the absence of risks of staying in the origin country.

Our model allows self-selection to differ between men and women. Given that female labor force participation rates are typically much lower and gender discrimination much more severe in African and Asian countries than in European destination countries, it is plausible to expect that even though average returns to human capital among men are higher in origin countries than in Europe, the reverse is the case for women once non-participation is accounted for. Therefore, we expect that the self-selection of men is negative from countries with low risk of conflict or persecution and positive from countries with a high risk of conflict or persecution, while self-selection of women should be positive from both low-risk and high-risk countries.

Refugee self-selection has been previously studied also by Borjas (1987) and Chin and Cortes (2015), but with opposite predictions. Refugee self-selection in Borjas (1987) refers to the particular case in which skills which are rewarded in the destination country are penalized in the country of origin. When motivating this, Borjas (1987) refers to "countries that have recently experienced a Communist takeover." We analyze the case in which human capital is valued both in the countries of origin and in the destination, introducing risks related to migration and staying. According to Chin and Cortes, "refugees will be less selected on characteristics associated with labor market success in the destination country compared to other migrants." According to our model, refugees' self-selection depends on the relative risks faced in the country of origin and during the migration process. Even if everyone in the country of origin faced the same risk of losing one's job due to persecution, an increase in such a risk could result in stronger self-selection in terms of human capital.

The different prediction arises as Chin and Cortes (2015) model persecution as a disamenity which does not influence the wage income in the country of origin, while we assume that people who are subject to persecution both lose their wage income and are subject to an additional utility loss, for example, due to maltreatment if being imprisoned.

The comparative statics of equation (2) also show that increases in the basic and skill-related component of the wage in the home country (destination country) reduce (increase) migration, in line with Borjas (1987). Higher risks related to staying in the country of origin (migration) increase (decrease) migration. Also, policies related to border closures or openings directly influence incentives to migrate as they influence migration costs, and higher migration costs decrease migration. A safer passage (corresponding to a lower risk term s_k) has a similar impact on how migrants are self-selected with respect to their human capital as a higher return to human capital in the potential destination, making it more likely that refugees would be positively self-selected. Lastly, higher gender-specific risks when traveling alone for female migrants (or higher risks from conscription for males if staying in the country of origin that could be captured by $q_{k,m}$ and $L_{k,m}$) reduce the share of female migrants; note that the model does not impose that the share of female migrants would be lower but that this is an empirical question.

Although our empirical application focuses on migration from African and Asian origin countries into Europe, our theoretical model can be applied to other contexts. For example, if the aim would be to analyze the self-selection of migrants from Latin America to the United States, risk terms $q_{k,g}$ and $L_{k,g}$ could capture risks related to gang violence and drug wars, as well as conflict or persecution in countries such as Nicaragua and Venezuela. In the historical context, our model would generally predict that refugees are positively self-selected in case the risk term is sufficiently high.

5. Data

The data used in this paper come from the Flow Monitoring Surveys (IOM), Gallup World Polls, IAB-BAMF-SOEP Survey of Refugees in Germany, the World Bank's World Development Indicators (WDI), and the Uppsala Conflict Data Program (UCDP). The level of analysis is the individual level, and the details on how the dataset was constructed are provided below.

5.1. Flow Monitoring Surveys

Our analysis is based on the Flow Monitoring Survey (FMS) obtained from the International Organization for Migration (IOM). The FMS provides in-depth insight into the profile, motivations, experiences, and intentions of migrants. It aims to derive quantitative estimates of the flow of (non-European) third-country nationals who are migrating towards Europe through the Central and the Eastern Mediterranean routes. The surveys were conducted in 11 languages (Arabic, Dari, English, Farsi, French, Italian, Kurdish, Pashto, Somali, Tigrinya, and Urdu) and administered by trained (male and female) data collectors with a mix of cultural and linguistic backgrounds (IOM, 2017).[12] FMS only gathers information from migrants aged 14 and older.

FMS provides rich information on migrants' country of origin, demographic characteristics (age, gender, educational attainment, and marital status), employment status before migration, reasons for leaving the place of residence, and intended destination. We use three waves of FMS in our main analyses. The first wave (October 2015 to December 2015) conducted interviews in Croatia, Greece, Slovenia, and the Former Yugoslav Republic of Macedonia[13]. The second wave of FMS (January 2016 to November 2016) covers Bulgaria, Croatia, Greece, Hungary, the Former Yugoslav Republic of Macedonia, Serbia, and Slovenia. The third wave of the survey only covers Italy, spanning from June to November 2016. In the raw data, the three waves of FMS included 21,497 respondents. Furthermore, we have data for 13,469 respondents who were interviewed in Turkey as the fourth wave from November 2016 to August 2018. We report corresponding summary statistics for observations with non-missing responses to age, gender, and educational attainment questions in Tables 2a, 2b and in the Appendix Table A.2.

The surveys aim to be representative of migrants arriving in Europe (IOM, 2017). The sample is large enough to gather evidence about migrants with different demographic profiles, and it is broadly reflective of the migrant population (IOM, 2017). Figure 5 illustrates the main migratory routes across the Mediterranean: (i) the Central Mediterranean Route refers to the sea journey from sub-Saharan Africa to Italy, with Libya being the main point of departure; (ii) the Eastern Mediterranean Route refers to the sea crossing from

[12] In the case of large groups, the surveys were conducted on a sample of the population. In case of small groups, the entire population was surveyed.

[13] The Former Yugoslav Republic of Macedonia has officially adopted its new name "the Republic of North Macedonia" in February 2019. However, throughout the paper we use the Former Yugoslav Republic of Macedonia or FYR Macedonia, which were originally used in the surveys.

Turkey to Greece.[14] Migrants who entered Greece then travel through Western Balkan countries with the aim of reaching the western Europe.[15]; (iii) the Western Mediterranean Route refers to sea crossing from Morocco to mainland Spain, which is not part of the analysis. More than 99 percent of migrants who arrived in Europe in 2015 and 2016 using the Mediterranean routes did so through the routes we analyze. More specifically, 1,030,173 migrants arrived in Europe using the Eastern Mediterranean Route; 335,278 migrants used the Central Mediterranean Route; 13,400 migrants used the Western Mediterranean Route (ECFR, 2017; UNHCR, 2017).

The surveys are conducted at the transit points.[16] They are fully anonymous and voluntary. Respondents are approached by IOM field staff and are informed of the purposes of the research and of the fact that participation does not influence their legal status in the country of the interview. The rest of the questions are posed only to those migrants who give their consent to proceed. IOM also often uses data collector from the same nationality as migrants or at least speaking their language. The response rates were very high, which alleviates concerns about selection bias in participating in the surveys.[17] Of those approached in the first two waves, only 2 percent declined to participate. The share declining to participate was somewhat higher in the third wave (4 percent) and fourth wave (5 percent).[18]

It is important to note that, despite the fast-changing conditions in the field, FMS provides a good representation of migrant groups. More specifically, in Table 1, we evaluate the representativeness of FMS data from the first three waves by comparing it with official

[14] Ajzenman et al. (2019) study the impact of the recent migration crisis on entrepreneurship in transit countries by exploiting cross-locality variation in distance to main transit migrant routes. They show that the entrepreneurial activity of natives falls considerably in localities that are more exposed to mass migration, compared to those located further away.

[15] In Greece, all surveyed locations were on the mainland and not on the islands, and represent a mix of official sites or camps managed by the national authorities and IOM premises in Athens. In the case of Balkan countries, the interviews were conducted both in spontaneous and organized transit points, including in official reception centers run by national authorities (IOM, 2017). As for Italy, the FMS was conducted in 36 entry and transit points in Sicily, Calabria and Apulia in 2016 including 3 of the 4 hotspots operating in the country (Trapani, Pozzallo and Taranto). The hotspots are first reception facilities to identify and register migrants soon upon arrival (IOM, 2016).

[16] The share of interviews by survey country in the first three waves is as follows: 24 percent in Italy, 20 percent in Greece, 19 percent in FYR Macedonia, 18 percent in Croatia, 6 percent in Bulgaria, 5 percent in Hungary, 5 percent in Slovenia, and 3 percent in Serbia.

[17] There were no incentives provided for participation. It was on a voluntary basis and migrants could terminate the interview at any time. However, data collectors provided a clear explanation of the purpose of the survey before asking for consent and emphasized the fact that it was anonymous. IOM also selected locations for the survey where migrants stay longer and that provide the right environment in terms of space and confidentiality. These helped migrants to feel comfortable and respond to the survey.

[18] In the first two waves (Eastern Med.), interviewers approached 15,016 migrants on transit and only 286 of them did not participate the survey; in the third wave (Central Med.), 7,028 migrants were approached and only 261 of them gave no consent; and in the last wave (Turkey) only 730 migrants out of 14,199 did not want to take part in the survey.

Eurostat data on asylum applications in 2015 and 2016. The table includes only those countries that were named as a country of nationality by at least 100 respondents in the FMS data. Most origin countries in our sample are predominantly Muslim and low or lower-middle income nations. Syrians are by far the largest group (25.6 percent in FMS data and 25.5 percent in Eurostat data), Afghans the second largest group (18.5 percent in FMS data and 12.8 percent in Eurostat data) and Iraqis the third-largest group (9.3 percent in FMS data and 8.5 percent in Eurostat data) in both data sets. Overall, the correlation between the shares of the nationalities in FMS and Eurostat data is remarkably high: 0.98. The share of males is somewhat higher in FMS data, but the differences are small for most countries. To sum up, comparisons with Eurostat data do not raise any major concerns about the representativeness of FMS data, with the caveat that women may have been somewhat less likely to answer the FMS survey, as suggested by IOM experts based on their field experience (UNICEF, 2017).

5.2. IAB-BAMF-SOEP Survey of Refugees

To investigate self-selection into Germany as the main destination country, we use the IAB-BAMF-SOEP Survey of Refugees (2016), which is collected as part of the German Socio-Economic Panel. The dataset provides information on refugees' country of origin, demographic characteristics, and employment histories and covers about 4,000 adult refugees in Germany. We focus on refugees who arrived in Germany in 2015 and 2016 and migrated from major conflict countries with more than 100 respondents (Afghanistan, Iraq, Pakistan, Syria) to make the analysis comparable with the rest of the paper.

5.3. Gallup World Poll and Country Level Characteristics

Our primary data on the origin country population come from the 2009-2014 Gallup World Polls (GWP).[19] These nationally representative surveys are fielded every year in over 120 countries and interview approximately 1,000 individuals in each country on a range of topics. The GWP provides detailed information on individuals' demographic characteristics (age, gender, educational attainment, and marital status), labor market outcomes, and income.

[19] Specifically, we observe the data for Syria between 2009 and 2013, for Algeria between 2010 and 2014, for Iran between 2011 and 2014 and for Morocco between 2010 and 2013. For the rest of the source countries, we have data spanning from 2009 to 2014. In Appendix (Tables A.6.-A.9.), we also show that our results remain qualitatively the same when we use the data on origin population between 2009 and 2011 (i.e. pre-Syrian conflict period; descriptive characteristics are shown in Appendix Table A.1).

The GWP's main advantage for our purposes is that the poll allows us to combine the FMS data with data on the country of origin (henceforth, non-migrant) population for a broad spectrum of countries.[20] Specifically, we pool two unique datasets based on 13 origin countries reported by at least 100 respondents aged 25 to 64 in the first three waves of FMS.[21] We then create a "*migrant*" indicator variable, which takes a value of 1 for respondents surveyed in the FMS and zero otherwise. Using this pseudo-cross-sectional sample, we investigate how refugees and irregular migrants are self-selected from the origin population. Importantly, each variable in GWP is harmonized with the definitions used in FMS to ensure comparability.

To further understand the characteristics of refugees and economic migrants, we use Uppsala Conflict Data Program (UCDP) battle-related deaths dataset and classify the origin countries based on their conflict intensity (Pettersson and Eck, 2018).[22] Following the definitions provided by UCDP, we define: i) *major conflict* category as countries with 1000 or more battle-related deaths in any of the years between 2009 and 2014 (this category includes Afghanistan, Iraq, Nigeria, Pakistan, Sudan, Syria); ii) *minor conflict* category as countries with 25 to 999 battle-related deaths in any of the years between 2009 and 2014 (this category includes Algeria and Iran); iii) *no conflict* category as countries that did not experience a major war or minor conflict in any of the years between 2009 and 2014 (this category includes Bangladesh, Cameroon, Côte d'Ivoire, Morocco, and Senegal). It is important to note that: (i) Political Terror Scale (Gibney et al. 2018) gives us the same classification using the data between 2009 and 2014; (ii) there is some movement of countries between the categories across years; (iii) using a continuous measure of the conflict intensity produces qualitatively similar results.[23] We also use geo-coded conflict data from UCDP to classify each origin country's sub-regions by their conflict intensity.

We also use several destination country characteristics in our analysis. We obtained country-level unemployment rates and the GINI Index (0-100 range) from the World Bank's World Development Indicators database. We use the *migrant integration policy index* variable from the MIPEX, which measures migrants' opportunities to participate in society.

[20] Of course, it is possible that respondents surveyed in GWP might migrate later on. This means that if the scale of emigration is large, we actually underestimate differences between migrants and non-migrants.

[21] The data come from the following thirteen countries: Afghanistan, Algeria, Bangladesh, Cameroon, Côte d'Ivoire, Iran, Iraq, Morocco, Nigeria, Pakistan, Senegal, Sudan, and Syria.

[22] We use the UCDP's best estimate for battle-related deaths to classify the countries based on their conflict intensity.

[23] These results are not presented here but available upon request.

The index uses 167 policy indicators covering 8 policy areas (labor market mobility, family reunion, education, political participation, long-term residence, access to nationality, anti-discrimination, and health) to rate countries from 0 to 100, with 100 being the top score.[24] We obtained data on the *average duration of the asylum procedure* from Eurofound (2016). This variable reports the average number of months passed between the submission of the asylum claim and the first decision. For ease of interpretation, we rescale this variable from 0 to 1, with 1 being the longest duration. Data on waiting duration for labor market access come from the Organisation for Economic Co-operation and Development (OECD, 2016). This variable measures the waiting period, in months, that it takes to obtain a work permit after successfully claiming asylum. Again, we rescale this variable from 0 to 1, with 1 being the longest duration.

6. Descriptive Statistics

Tables 2.a and 2.b present descriptive statistics of FMS data from Waves 1-3, with respect to when and where the interviews took place and basic demographic and socioeconomic characteristics of survey respondents. In Table 2.a, we show the descriptive statistics for the full sample (i.e., with no age restrictions, and non-missing responses to age, gender, and educational attainment questions). In Table 2.b, we report the characteristics of the sample of migrants aged 25 to 64. We focus on this age group when analyzing self-selection according to education and predicted earnings as those younger than 25 have often not yet completed their education, while those older than 64 may have already retired. We are not able to include Eritrea in our main analysis as it is not surveyed by GWP. We also exclude nationalities with less than 100 respondents. Our final sample consists of 9,828 observations provided by migrants of 13 different nationalities (i.e., origin countries) with at least 100 respondents.

Table 2.a shows descriptive statistics for the main FMS sample without age restrictions. An overwhelming majority of survey respondents are male (82 percent) with an average age of 26. 30 percent of the respondents are married, and 18 percent have tertiary education.[25] About half of the respondents also report having been employed before

[24] For details of the compilation of the MIPEX, see Huddleston et al. (2015).

[25] These figures are in line with those found in Brücker et al. (2016) who show that in Germany, 13 percent of refugees aged 18 or more have a university degree and 6 percent have a vocational qualification.

migration. 77 percent cited conflict or persecution and 18 percent economic reasons as the main reason why they had emigrated. Limited access to basic services (like school and health care) or lack of food or accommodation was named as the main reason by only 2 percent of respondents. 2 percent referred to natural disasters or other reasons.

When we focus on individuals aged 25 to 64 in Table 2.b, we again find that a vast majority of survey respondents are male (80 percent). There are also notable differences between the two samples: migrants in this age band are more likely than those in the full sample to be married (57 percent), to have tertiary education (25 percent), and to have been employed before migration (63 percent). These differences in education levels and employment status before migration also highlight the importance of focusing on individuals aged 25 and older when testing our theory. By doing so, we avoid the share of young respondents mechanically driving results for self-selection with respect to tertiary education and employment. When it comes to the reasons for leaving their home country, migrants cite "conflict or persecution" (79 percent) and "economic reasons" (17 percent) as the main causes. There is also a major gender difference in marital and past employment status: women are more likely to be married and much less likely to have been employed. The share of Syrians is also considerably higher among women.

There are some noteworthy differences between the survey waves. First, Syrians, Afghans, and Iraqis are the most common nationalities in Wave 1, Wave 2, and Wave 4, but almost absent in Wave 3. In Wave 3, most migrants come from Africa, with Nigerians, Eritreans, and Guineans being the largest groups. This is expected given that Wave 3 was fielded in Italy, which is the main arrival point for people fleeing conflict and poverty in Africa.[26] In Wave 4 (see Appendix Table A.2.), Syrians, Iraqis, Iranians, and Afghans constitute more than 90 percent of the sample. Second, although most migrants are motivated by conflict or persecution in all waves, this motivation dominates most clearly in Waves 1 and 2. Third, there is a huge variation in the level of education across survey waves: the share of those with tertiary education in the full sample is 36 percent in Wave 1, 20 percent in Wave 2, 7 percent in Wave 3, and 14 percent in Wave 4. Corresponding patterns prevail if the analysis is restricted to those aged 25 to 64.

Appendix Table A.3. shows descriptive statistics for the IAB-BAMF-SOEP Survey of Refugees. As this survey included only adults, it is not surprising that all respondents are

[26] The collapse of the Gaddafi regime in 2011 opened the sea route from Libya to Italy. Friebel et al. (2018) show that this significantly increased migration intentions from African countries from which migration route to Europe was shortened most.

somewhat older and more educated than in the FMS. If restricting the attention to those aged 25 to 64, the main difference between IAB-BAMF-SOEP Survey of Refugees and the FMS is that the share of women is clearly higher in Germany, and respondents are on average older. In terms of education, differences compared with the average if pooling Waves 1 and 2 are small.[27]

In Table 3, we present descriptive characteristics of the origin population from the GWP.[28] Unlike in Tables 2a and 2b, the gender ratio is balanced. People in origin countries (compared with refugees and irregular migrants) are also older on average, more likely to be married than those in the FMSs, and less likely to have (completed) tertiary education. When differentiating the analysis by gender, employment rates in the age group 25 to 64 are quite similar in FMSs and GWP, and somewhat higher in the IAB-BAMF-SOEP Survey. These patterns remain qualitatively similar when we use the origin population data between 2009 and 2011. Women are much less likely to be employed than men in all data sets, and somewhat less educated than men in GWP.

Table 4 presents the differences in educational attainment for the main origin countries. Apart from Nigeria, Bangladesh, and Senegal, migrants are better educated than non-migrants, whether analyzing the full population or restricting the attention to those aged 25 to 64. Of the major countries of origin, Syrian migrants are especially highly educated. In terms of our model, this can be expected to reflect higher risks of staying in Syria due to high conflict intensity, and higher recognition rates of Syrian asylum applicants.

Figure 6 illustrates the reasons for leaving by nationality. We find that more than 90 percent of respondents from Eritrea, Iraq, Somalia, Sudan, and Syria report leaving their country due to conflict or persecution. At the other end of the scale, a vast majority of respondents from Morocco and Algeria cite economic conditions as the main reason for leaving their home country. Additional analysis suggests that there are no noteworthy gender- and age-specific differences in reported reasons. Figure A.1 in the Appendix illustrates reasons for leaving by nationality among respondents in Turkey. Compared with Figure 6, the most striking difference is that only 45 percent of respondents from Afghanistan living in Turkey reported that conflict or persecution was their main reason for emigrating, compared

[27] Brücker et al. (2016) allowed multiple responses for reason to migrate and report that 70 percent of respondents listed war and conflict as reason to migrate, followed by persecution (44 percent), poor personal living conditions (39 percent), discrimination (38 percent), and fear of forced conscription (36 percent).

[28] Table 3 reports unweighted descriptive statistics from the Gallup World Polls. Importantly, we find that using sampling weights does not substantially change the shares we report in Table 3. These estimates are not reported here but available upon request.

with 89 percent among Afghans in the first three waves in Europe. The share emigrating due to conflict or persecution is lower in Turkey also among Iraqis, Somalians, and Syrians, but somewhat higher among Iranians.

The respondents were also asked their intended destination country, which we report by survey waves. As shown in Figure 7, Germany was the main destination in the first two waves, chosen by 61 percent of respondents, followed by France (6 percent), Italy, and Sweden (5 percent each). In wave 3 that was collected in Italy, 61 percent reported an intention to stay in Italy, followed by 8 percent aiming to Germany and 6 percent aiming to France. In wave 4, 63 percent intended to stay in Turkey, 7 percent aimed to Germany, and 18 percent did not know where to aim. It is plausible that this reflects the effective closing of the route from Turkey to the European Union following the March 2016 agreement that allows the European Union to send migrants arriving from Turkey to Greece illegally back in exchange for taking recognized refugees from Turkey.

Figure 8 shows the reasons for leaving in the first three waves of the FMS, according to the intended destination country. We find that more than 80 percent of those aiming for Germany, Denmark, Finland, Sweden, Norway, and Austria had left their country of origin because of conflict or persecution. This share was below 60 percent among those aiming for France, Belgium, and Italy, which had, correspondingly, the highest shares of economic migrants. In Figure A.2 in the Appendix, among respondents in Turkey, we find that more than 80 percent of those aiming for Germany, Sweden, Switzerland, Austria, the United States, and the United Kingdom had left their country of origin due to conflict or persecution.

Finally, respondents were asked about the total cost of the journey and the total time spent traveling. In the Eastern (Central) Mediterranean Route, 7 (15) percent reported total cost below USD 1000, 74 (44) percent USD 1000-5000, and 16 (12) percent above USD 5000, with 3 (19) percent not being able or willing to give a cost estimate. In Turkey, costs were considerably lower, with 52 percent reporting that they had spent less than USD 1000, 18 percent USD 1000-5000, and only 4 percent more than USD 5000. When it comes to the duration of travel, we find that 59 (65) percent of survey respondents in the Eastern (Central) Mediterranean Route spent less than 30 days in transit. 18 (20) percent of the respondents spent between 30 and 59 days, and 10 (13) percent of the respondents spent between 60 and 119 days.

7. Empirical Approach

To assess the self-selection of migrants, we estimate a series of multivariate regression models relating to how refugees and irregular migrants in the FMS differ from the overall population in Gallup World Polls (GWP). We restrict our attention to those aged 25 to 64 to focus on individuals most likely to have completed their education, and not yet retired. Our main variables of interest are age, gender, and level of education. In some of the analyses, we also use predicted incomes to study how these are related to self-selection into emigration. We proceed by estimating linear probability models for our outcomes for ease of interpretation, though logistic regression models returned similar patterns. We also estimate most models separately by gender. Our models of self-selection take the form:

$$(1) \qquad Migrant_{ic} = \alpha + \beta_1 X_i + \beta_2 O_c + \varepsilon_i$$

where $Migrant_{ic}$ takes a value of 1 if individual *i* from country *c* is in the FMS sample and 0 otherwise. X_i is a vector of demographic variables that include: age group dummies (25-34, 35-44, 45-54); education dummies (tertiary education, secondary education); dummy variable indicating the labor market status before migration (employed); a dummy variable indicating the marital status (married, divorced, widowed). To account for other unobservable characteristics, we include a full set of origin country (O_c) dummies, which control for all time-invariant variation caused by factors that vary cross-nationally.

We use data for the origin population from GWP between 2009 and 2014. However, the results are robust to using data between 2009 and 2011 (that is, pre-Syrian conflict period). Note that in equation (1) the relevant reference category is composed of individuals who are in the GWP. We estimate standard errors robust to heteroscedasticity.

We also use GWP to investigate the self-selection of refugees and irregular migrants with respect to predicted income. Specifically, we first estimate Mincer regressions for each origin country controlling for the level of education, employment status, gender, and age. We then use these estimates to predict individual income for those surveyed in the FMS to gain further insights into how migrants from different country groups are self-selected with respect to their earnings potential.

After analyzing the self-selection of refugees and irregular migrants, we investigate sorting into different intended destinations by using FMS data only. Our models of sorting take the form:

(2) $OUTCOME_c = \alpha + \beta_1 X_i + \beta_2 O_c + \varepsilon_i$

where $OUTCOME_c$ is one of the country-level indicators: (i) Gini coefficient; (ii) migrant integration policy index; (iii) average duration of asylum procedure; (iv) average waiting duration for labor market access; (v) social expenditure as a percentage of GDP; (vi) country-level unemployment rate (log); (vii) log GDP per capita. All other variables are as described above.

8. Self-selection of Refugees and Irregular Migrants into Europe

This section presents two sets of results. We first show self-selection of refugees and irregular migrants with respect to education and other demographic characteristics. Second, we analyze the self-selection of refugees and irregular migrants according to their predicted earnings.

8.1. Self-selection of Refugees and Irregular Migrants with respect to Demographic Characteristics

In this subsection, we test our theoretical predictions on how the distribution of education among refugees and irregular migrants compares with that of non-migrants. We conduct this analysis in three different ways. First, we use individuals' responses to the "main reason to migrate" question in FMS. Second, we classify countries based on their country-level conflict intensity following the classifications provided by the UCDP. Third, we match migrants' origin provinces to the sub-regions included in the GWP and define each sub-region's conflict intensity using geo-coded conflict data.

Using individuals' main reason for migration, Table 5 presents evidence for males and females together, Table 6 for males only, and Table 7 for females only, restricting the analysis of migrants to the first three waves. The columns are structured as follows: *migrant due to all reasons* is the outcome variable in Column 1; *refugee (which refers to migrants who left their country of origin due to conflict or persecution)* in Column 2; *irregular migrant (which refers to migrants who left their country of origin due to economic or other reasons, apart from conflict or persecution)* in Column 3. As mentioned above, we restrict attention to

individuals aged 25 to 64 to focus on individuals most likely to have completed their education and not yet retired.

In column 1 of Table 5, we find that educated people are significantly more likely to migrate. Columns 2 and 3 show that the pattern is mainly driven by migrants escaping conflict or persecution.[29] Irregular migrants are also somewhat more likely to have completed tertiary education than non-migrants but to a lesser degree.[30] Stronger positive self-selection of refugees is in line with our theoretical predictions suggesting that if the risk of being a victim of conflict or persecution increases, the probability of emigration becomes eventually increasing in human capital even if returns to human capital would be higher in the country of origin in the absence of conflict or persecution. An additional mechanism outside our model is that better-educated individuals would find it easier to finance their trip, while liquidity constraints would prevent the least educated people from migrating. However, if liquidity constraints would explain different self-selection patterns among refugees and irregular migrants, we would expect these groups to differ in their total spending to finance their trip. Additional analysis suggests that the distribution of total costs is almost the same among refugees and irregular migrants along both Eastern Mediterranean Route and Central Mediterranean Route (see Appendix Table A.5.). Therefore, we find it unlikely that liquidity constraints would explain differences in self-selection patterns between these groups. Turning to other covariates, we find that the probability of emigration is higher for men, younger people, and singles. Young people being most likely to emigrate is intuitive as they have the longest time horizon to reap returns to migration investment (Sjaastad 1962).

Tables 6 and 7 present corresponding analyses separately for men and women.[31] Refugees are positively self-selected in terms of their education among both genders. Men who left due to economic or other reasons are negatively self-selected in terms of their education. This is intuitive and in line with Borjas (1987), as African and Asian countries have much wider income differences than European countries.

[29] We also find similar results when we analyze migrants arriving from the Eastern Mediterranean and the Central Mediterranean routes separately.

[30] The results hold also if excluding Iran, which has by far the highest share of tertiary educated emigrants.

[31] We find that among those who left their country due to conflict or persecution (economic or other reasons) 35 (50) percent of men and 11 (31) percent of women were travelling alone. Among female refugees (irregular migrants) who travel with others, 57 (55) percent report to have travelled with their spouses and 41 (42) percent with their children, parents or siblings. Among male migrants who travel with others from major (minor or no) conflict countries, 80 (78) percent of them report to have travelled with their spouses and 17 (20) percent with their children, parents or siblings. About 2 percent of migrants also report to have travelled with their extended family members or friends.

Women who left due to economic or other reasons are, instead, positively self-selected in terms of their education. In terms of our model, this can be explained in two ways. First, gender discrimination is a much bigger problem in Africa and Asia than in Europe. Together with very low female labor force participation rates in most, but not all, origin countries this can lower expected returns to education for women below expected returns in Europe, even when the opposite holds for men.[32] An alternative explanation would be that a non-negligible subset of women could be escaping gender-based violence and repression even from countries that are conventionally considered "safe". This would produce a self-selection pattern that is similar to refugees who escape conflict and persecution.

The decision of whether or not to migrate is a complicated one. It is often difficult to distinguish between refugees and those migrating for other reasons but claiming to be refugees. Economic incentives can be expected to play an important role even for those leaving countries suffering from civil war. Table 8 sheds light on this issue and presents the results by the conflict intensity classification at the country level. The columns are structured as follows: both genders and major conflict countries (column 1); both genders and minor or no conflict countries (column 2); males and major conflict countries (column 3); males and minor or no conflict countries (column 4); females and major conflict countries (column 5); and females and minor or no conflict countries (column 6). We again restrict attention to individuals aged 25 to 64.

The results in column 1 and 2 suggest that the probability of emigration increases in tertiary education in both country groups, but decreases in secondary education in countries with a minor or no conflict. In line with our results in Table 5, we find that refugees and irregular migrants escaping major conflicts tend to be highly educated relative to the national average in their country of origin.

Columns 3 to 6 present corresponding analyses separately for men and women. The probability of emigration increases in both secondary and tertiary education for both men and women in countries suffering from a major conflict. In countries with minor or no conflict, men with secondary education are somewhat less likely to emigrate than men with less than secondary education, while the effect of tertiary education is statistically insignificant.

[32] In the age group 25 or older, female labor force participation rate varies between 15 percent and 33 percent and male labor force participation rate varies between 75 percent and 92 percent in Afghanistan, Algeria, Bangladesh, Iran, Iraq, Morocco, Pakistan, Sudan, and Syria (ILO 2019), with the gender gap in labor force participation rates varying between 53 and 68 percentage points. The gender gap in labor force participation rates is lower in Cameroon, Côte d'Ivoire, Nigeria, and Senegal, varying between 12 and 29 percentage points.

In countries with minor or no conflict, women with tertiary education are significantly more likely to emigrate than women with less than secondary education, while women with a secondary education do not differ from women with less than a secondary education in their probability of emigration.

A major concern in interpreting our results in a causal way is that the presence of some unobserved reasons could affect both the probability for a country to be in conflict and the skill composition of emigrants. For example, poor institutional quality could both increase the risk of conflict and generate high-educated emigration due to bad job market prospects. Such unobserved heterogeneity might lead us to find that migrants from major conflict countries are positively self-selected. Furthermore, one can always ask whether the Uppsala Conflict Database cut-off of 1000 or more battle-related deaths within a year is an appropriate measure for high level country risks. To address these concerns, we make use of the fact that we observe the origin province information for a subset of migrants surveyed in the FMS. More specifically, we match each of the migrants' origin provinces to the sub-regions included in the GWP to investigate whether our results hold when we conduct our analysis at the sub-regional level.[33] Formally, we proceed as follows: (i) we scrape geo-coded conflict data from Uppsala Conflict Database for the 2009-2014 period; (ii) we assign the number of battle-related deaths for each year to the corresponding sub-regions in the GWP based on geographical coordinates; (iii) we then classify each GWP sub-region as *more intensive conflict* or *less intensive conflict*. *More intensive conflict* refers to the sub-regions that experience equal to or more battle-related deaths than the country median (over the 2009-2014 period), and *less intensive conflict* refers to the sub-regions that experience less battle-related deaths than the country median (over the 2009-2014 period).[34] Importantly, we do not include countries that did not experience any conflict, given that it is practically impossible to classify their sub-regions based on the conflict intensity.

We present the results at the sub-regional level in Table 9, in which we again restrict attention to individuals aged 25 to 64. The samples across the columns are as follows: both genders and more intensive conflict sub-regions (column 1); both genders and less intensive conflict sub-regions (column 2); males only and more intensive conflict sub-regions (column

[33] Gallup defines sub-regions based on official boundaries. To address potential concern that GWP might be unrepresentative for conflict-affected sub-regions, we calculated GWP respondent shares and population shares based on censuses in those sub-regions covered by GWP in 2010 for both major and minor conflict countries. Average correlation across all countries is 0.941. Correlation is not only remarkably high for major conflict countries (0.983 for Afghanistan, 0.966 for Iraq and 0.987 for Syria) but also across high conflict intensive and low conflict intensive sub-regions. The full list of sub-regional population shares is available upon request.

[34] The yearly numbers of observations by sub-regions are stable in all high conflict and minor conflict countries.

3); males only and less intensive conflict sub-regions (column 4); females only and more intensive conflict sub-regions (column 5); and females only and less intensive conflict sub-regions (column 6).

The results in Table 9 are in line with findings from Tables 5, 6, 7 and 8: (i) refugees and irregular migrants escaping from more intensive conflict sub-regions tend to be highly educated relative to the average in their sub-region of origin; (ii) men from more intensive conflict sub-regions are positively self-selected, whereas men from less intensive sub-regions are negatively self-selected; (iii) women are positively self-selected from both less and more intensive conflict sub-regions, but more so from more intensive conflict sub-regions.

We also analyzed the share of those who emigrated due to conflict or persecution among migrants from sub-regions with more intensive and less intensive conflict, relative to the country median. Appendix Table A.4. reports this separately for men and women from all major and minor conflict countries. With the exception of women from Algeria, Nigeria, and Pakistan, the share emigrating due to conflict or persecution is always higher from regions with more intensive conflict than from regions with less intensive conflict. For most countries, the difference is substantial among both men and women, suggesting that the major difference in self-selection pattern goes together with major differences in main motivations to emigrate, even when controlling for country fixed effects. Note that the overall share having emigrated due to conflict or persecution is lower in Appendix Table A.4. than among respondents included in Table 6. The discrepancy arises as sub-regions are observed only for a subset of respondents, and somewhat less common for those who emigrated due to conflict or persecution.

We establish a link between self-selection among refugees and irregular migrants and their main motivation to emigrate in Table 10. More specifically, our outcome variable is *reason to migrate: conflict/persecution*, which is equal to 1 for respondents who cite conflict or persecution as the main reason to migrate and 0 for respondents who cite other reasons (economic reasons, limited access to amenities, or natural disasters and other reasons) in the Flow Monitoring Surveys. The columns are structured as follows: all respondents (column 1); males (column 2); females (column 3). We again restrict our attention to individuals aged 25 to 64. We find that both male and female refugees are significantly more likely to have secondary and tertiary education compared with migrants who cite other reasons for leaving their countries.

In Table 11, we complement our analysis by investigating the self-selection of refugees (who escaped major conflict countries) into Germany, which is the main destination

country. Consistent with our main results, we find strong positive self-selection with respect to education. Unlike in FMS, the gender composition of refugees who have made it into Germany is relatively balanced.

8.2. Self-selection with respect to Predicted Income

As the FMS did not ask respondents about their income, we can only test self-selection with respect to predicted incomes. We use predicted incomes for both migrants and non-migrants to focus on both groups' income arising from observable characteristics. In Table 12, we use predicted individual pre-migration income as the only explanatory variable to predict the emigration of single people aged 25 to 64, in addition to country fixed effects. We only focus on single individuals as income is reported at the household level in GWP. Specifically, we obtain predicted individual income by estimating within-country Mincer regressions which control for the level of education, employment status, gender, and age separately for each country of origin. We then use the estimated coefficients to calculate predicted income for both migrants and non-migrants.

Note that our models in Table 12 do not include individual covariates again as these are all strongly correlated with predicted income. We find that refugees are strongly positively self-selected with respect to predicted income, as are female irregular migrants. Male irregular migrants, instead, are negatively self-selected with respect to predicted income. We also find that predicted log income strongly increases the probability of emigration from countries suffering a major conflict, in line with our theoretical predictions. Male migrants from countries with minor or no conflict are negatively self-selected with respect to their predicted earnings, and female migrants positively self-selected.

A potential concern related to our findings is that self-selection, according to predicted earnings, could reflect borrowing constraints, with migrants from countries with higher income being more likely to be able to emigrate. Although we do not exclude this possibility, negative self-selection among male migrants from countries with minor or no conflict supports the idea that borrowing constraints cannot be the main explanatory factor. In other words, if it would be about borrowing constraints and these being less binding for highly educated, then we would observe positive self-selection also for men from minor or no conflict countries.

Borjas, Kauppinen, and Poutvaara (2019) showed that the Roy model has more precise predictions about the self-selection of migrants than previously realized. The same

conditions that result in positive or negative selection in terms of expected earnings also imply a stochastic dominance relationship between the earnings distributions of migrants and non-migrants.

Figure 9 compares cumulative distribution functions (CDFs) of predicted income (in the country of origin) of refugees who migrate due to conflict or persecution with non-migrants' CDF, and Figure 10 makes a corresponding comparison between irregular migrants whose main reason to emigrate was not conflict or persecution and non-migrants. These figures suggest that male and female refugees and female irregular migrants are positively self-selected with respect to their income. Male irregular migrants, instead, do not differ much from non-migrants.

Figures 11 and 12 show migrants' and non-migrants' cumulative distribution functions (CDFs) of predicted income for major conflict countries and for countries with minor or no conflict. We again find that both male and female migrants from major conflict countries are positively self-selected in terms of their predicted income. For countries with minor or no conflict, migrant and non-migrant men do not differ in terms of their income distribution. Migrating women, instead, are positively self-selected in terms of their income.

9. Sorting of Refugees and Irregular Migrants

We next analyze the sorting of refugees and irregular migrants into different intended destination countries. Our main question is whether refugees and irregular migrants with tertiary level education (or more) sort themselves into more inegalitarian countries and whether the choice of the destination country for refugees and irregular migrants is shaped by macro-level characteristics. Table 13 investigates how migrants' sorting according to their education differs between major conflict countries and countries without a major conflict. In this table, we only focus on migrants aged 25 to 64 in waves 1, 2, and 3 in the FMS sample, and our outcome variable is the Gini coefficient of the intended destination country. When studying all destination countries, those with tertiary education are more likely to choose more unequal countries, and those with secondary education are more likely to choose more equal countries than those below secondary education. This pattern also holds for those who emigrate from major conflict countries. When it comes to countries with minor or no conflict, the sorting pattern of those with tertiary education is strongly in line with the Roy-Borjas model: they are much more likely to aim for more unequal countries. An important caveat is

that these models implicitly assume that skills are sufficiently transferable across countries. These findings also do not imply causality, as other factors (such as culture and institutions) may affect the choice of destination and be linked to a person's education. Nonetheless, our findings indicate that education may play an important role when refugees and irregular migrants choose their destination.

Table 14 provides suggestive evidence on how other country-specific factors might influence migrants' choice of destination. As with the previous analysis, the sample is restricted to respondents covered by waves 1, 2, and 3 in the FMS surveys (ages 25 to 64). Outcome variables represent various characteristics of the intended destination country of each respondent. In line with previous tables, we report the estimates using migrants with *primary education or less* as a reference category.

Migrants who are educated to secondary and tertiary levels are less likely to head for countries that have more comprehensive migrant integration policies than migrants with primary education or less. Migrants who are educated to secondary and tertiary levels are more likely to choose destination countries where asylum applications are considered slower and where work permit applications, once asylum has been granted, take more time to process than migrants with primary education or less. More generous social safety nets make a destination country less attractive for migrants with secondary and tertiary education. Migrants who are educated to secondary and tertiary levels are also more likely to prefer countries that have higher unemployment rates and GDP per capita. In other words, refugees and irregular migrants coming to Europe respond to incentives at all stages of the migration process. This is in line with the Borjas (1999) finding that immigrants who receive welfare benefits in the United States tend to be more concentrated in states with generous benefits than natives or immigrants who do not receive welfare. Recently, Agersnap et al. (2019) have tested the effects of large changes in welfare benefits for immigrants in Denmark, finding evidence for the "welfare magnet" hypothesis. Their analysis is based on granted residence permits. Our paper provides first evidence on how the generosity of the welfare state affects the preferred destination country for migrants from multiple origin countries.

10. The Role of Border Controls

The refugee crisis was catapulted to the top of the European policy agenda following Chancellor Merkel's decision to open Germany's borders to Syrian refugees in August 2015.

Subsequently, several other policy changes, mostly tightening border controls, significantly reduced or redirected the numbers of refugees and irregular migrants to other countries. In this section, we analyze the effect of such policies on FMS respondents' intended destination country, using the information on the interview date. Did the policies change the intended destinations, and, if so, how? Alternatively, it could be that the policy changes would have relatively little effect on what the intended destinations were.

We focus on the major border policy changes over the sample period of FMS (October 2015-November 2016). In particular, we identified five important policy changes: (i) *Austria imposes quota* refers to a dummy variable that is equal to one, if interviews were conducted after 19th February 2016, when Austria imposed a quota of a maximum of 80 asylum applications, and a maximum of 3,200 people to travel through Austria per day; (ii) the *Austria quota announcement* is equal to one for interview dates after Austria announced this quota on 20th January 2016; (iii) *Hungary border closing* is equal to one if the interview took place after Hungary closed its border on 16th October 2015; (iv) *Slovenia and FYR Macedonia border tightening* refers to the date on which FYR Macedonia closed its border with Greece and Slovenia set stricter border controls, and it is equal to one if the interview was conducted after 9th March 2016. *Sweden border control* is equal to one if interviews took place after 11th November 2015. In this analysis, we use the first two waves of FMS since the waves 3 and 4 were conducted after the latest border policy change we analyze. Unlike in the analysis of self-selection, we do not impose any age restrictions and include all countries with more than 100 respondents in waves 1 and 2.

Table 15 reports the coefficients of border policy dummies listed above while controlling for origin and survey country fixed effects as well as individual characteristics. We focus on the top six intended destination countries reported by migrants in the first two waves of FMS (see the top panel of Figure 7). The outcomes across the columns are as follows: the intended destination is Germany (column 1); the intended destination is France (column 2); the intended destination is Italy (column 3); the intended destination is Sweden (column 4), intended destination is Austria (column 5), and the Netherlands (column 6).

The results suggest that border policies significantly affect the intended destinations of refugees and irregular migrants. Results in Column 1 suggest that while Austria quota announcement and Sweden border controls increase the likelihood of stating "Germany" as an intended destination country, Slovenia and FYR Macedonia border tightening significantly reduce migration intentions to Germany. Focusing on France (Column 2), we find that Slovenia and FYR Macedonia border tightening is the only policy that positively affects the

sorting patterns of refugees and irregular migrants. Moving to Column 3, we find that both the introduction of the refugee quota by Austria and Slovenia and FYR Macedonia border tightening significantly increase migration intentions to Italy.

Columns 4 and 6 presents results for Sweden and the Netherlands. We find that Sweden's border controls and Austria's quota policies significantly lower the migration intentions to both countries. Lastly, in Column 5, we focus on sorting to Austria. The results suggest that migrants are significantly more likely to aim for Austria after the quota announcement, but before the quota was implemented.

In results not reported, we also considered heterogeneity across several other dimensions and found weak evidence of heterogeneity in the role of the border controls by age, education, or origin country conflict intensity. Notably, we find stronger effects for men, which may be explained by traditional gender roles. Overall, these results provide suggestive evidence that refugees and irregular migrants are mostly well-informed about the border policy changes, and there is a strong need for deeper cooperation between destination countries (Hatton 2015).

11. Self-selection of Refugees and Irregular Migrants into Turkey

In Table 16, we investigate the self-selection of refugees and irregular migrants with respect to education and other demographic characteristics in Turkey. We restrict attention to individuals aged 25 to 64 and to origin countries with more than 100 respondents in this age group in Turkey. The results for men and women together are, if anything, even more pronounced than in Table 5: refugees are strongly positively self-selected with respect to their education, and irregular migrants strongly negatively self-selected. The results for men are similar to Table 6: male refugees are positively self-selected, and male irregular migrants negatively self-selected. We also find that female refugees are positively self-selected, in line with Table 7. Unlike in Europe, female irregular migrants in Turkey are not positively self-selected. In terms of our model, this can be explained by worse gender discrimination and lower female labor force participation in Turkey.

An additional analysis reported in Appendix Table A.5. suggests that, just as among migrants in European transit countries, refugees and irregular migrants do not differ much in terms of the total cost of their journey, despite differences in their education and predicted

earnings. A similar cost distribution makes it unlikely that liquidity constraints would explain different self-selection.

12. Conclusion

Using Flow Monitoring Surveys and Gallup World Polls, we analyzed the main reason to emigrate and self-selection of refugees and irregular migrants arriving in Europe through the Central and Eastern Mediterranean Routes in 2015 and 2016. We found that 77 percent of respondents were refugees in the sense of fleeing war, conflict or persecution, 18 percent were motivated by economic reasons, 2 percent by the lack of basic services, and 2 percent had left because of natural disasters or other reasons. Most of the respondents were from Syria, Afghanistan, and Iraq, followed by Nigeria, Pakistan, Eritrea, Morocco, and Iran. Refugees and irregular migrants are more likely to be single, male, and young, with vast cross-country variation in the main motivation to migrate.

We found that refugees and female irregular migrants who arrived in Europe are positively self-selected in terms of their education and expected earnings, while male irregular migrants are negatively self-selected. This pattern holds both when using individually stated main reason to emigrate, and when comparing migrants from major conflict countries (most plausibly refugees) with migrants from minor or no conflict countries. In order to address the concern that both emigration and conflict could reflect unobserved country-specific factors, we also used geo-coded conflict data to classify sub-regions in each conflict country according to their conflict intensity. Our findings on different self-selection patterns hold also in this sub-regional analysis, alleviating concerns that the patterns would be driven by cross-country differences in colonial history, institutions, or the presence of oil and other natural resources. We also complemented our analysis by investigating the self-selection of refugees into Germany, which is the main destination country, and self-selection of refugees and irregular migrants into Turkey, which currently hosts the highest number of refugees in the world. Consistent with our main results, we find strong positive self-selection with respect to the education of refugees into both Germany and Turkey; irregular migrants into Turkey, instead, are negatively self-selected. Our explanation for positive self-selection of refugees is that a sufficiently severe risk of being killed or imprisoned reduces also expected returns to human capital, reversing self-selection pattern

compared with irregular migrants. Negative self-selection of male irregular migrants is in line with the Roy-Borjas model. Positive self-selection of female irregular migrants can also be explained by extending Roy-Borjas model to include gender-specific returns to human capital. Low female labor force participation rates and gender discrimination can depress expected returns to women's human capital in African and Asian countries of origin, even if the reverse would hold for men.

Positive self-selection of refugees is good news for receiving countries if they keep taking steps to ease their path to labor market integration. For example, successful integration of refugees can boost subsequent exports to their country of origin (Parsons and Vezina 2018). At the same time, the positive self-selection of refugees is bad news for countries of origin as it makes rebuilding more difficult, once the conflict is over. To the extent that human capital is driving economic growth, such a loss may have severe long-term effects.

We also analyzed migrants' sorting across different destinations. We found that economic incentives, migrant integration policies, and border policies all play an important role. Most importantly, migrants with tertiary education are more likely to prefer a destination with high income differences and GDP per capita than their compatriots with primary or secondary education. Migrants with secondary or tertiary education are less likely to prefer a country that spends a large share of its GDP on social expenditures than those with primary education or less. Unemployment and average duration of the asylum process, instead, deter more those with less education.

Our detailed analysis of socio-demographic characteristics and background of refugees and irregular migrants points to several policy implications. For example, a significant number of migrants from Africa desires to seek out better economic opportunities abroad. While most of them may ultimately be denied asylum, they can slow down asylum application procedures. This may, in turn, undermine popular support for a well-managed and fair asylum system (Hatton, 2017). Aging European economies could consider tackling this problem by increasing legal employment opportunities for African citizens on a selective basis, depending on local needs. Such initiatives could form part of a broader strategy aimed at containing illegal migration to Europe (MEDAM, 2018).

Moreover, policies that support refugees' and migrants' integration into the labor market need to be tailored to their skills (World Bank, 2018). Migrants escaping major

conflicts (such as the fighting in Syria) may well benefit from receiving early access to language courses and other basic training while waiting for decisions on their asylum applications. Battisti et al. (2019) show that the labor market outcomes of certain groups of refugees can be improved through a basic job search assistance program. Offering labor market access to asylum seekers and fostering social integration are able to mitigate the detrimental effect of past conflict exposure on criminality Couttenier et al. (forthcoming). All of these aspects are particularly important in terms of promoting the integration of migrants and helping a country to make society more inclusive for everyone.

References

Abramitzky, R. and Boustan, L. (2017). Immigration in American Economic History. *Journal of Economic Literature* 55(4): 1311-45.

Abramitzky, R., Boustan, L., and Eriksson, K. (2012). Europe's Tired, Poor, Huddled Masses: Self-selection and Economic Outcomes in the Age of Mass Migration. *American Economic Review* 102(5): 1832–56.

Acemoglu, D., Fergusson, L., and Johnson, S. (forthcoming). Population and Conflict. *Review of Economic Studies.*

Acemoglu, D., Johnson, S., and Robinson, J.A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review* 91: 1369-1401.

Agersnap, O., Jensen, A.S., and Kleven, H. (2019). The Welfare Magnet Hypothesis: Evidence From an Immigrant Welfare Scheme in Denmark. NBER Working Paper No. 26454.

Ajzenman, N., Aksoy, C. G., & Guriev, S. (2019). Exposure to Transit Migration, Public Attitudes, and Entrepreneurship among the Native Population. EBRD Working Paper No. 236.

Bandiera, O., Rasul, I., and Viarengo, M. (2013). The Making of Modern America: Migratory Flows in the Age of Mass Migration. *Journal of Development Economics* 102: 23-47.

Battisti, M., Giesing, Y., and Laurentsyeva N. (2019). Can Job Search Assistance Improve the Labour Market Integration of Refugees? *Labour Economics* 61: No. 101745.

Beine, M. and Parsons, C. (2015). Climatic Factors as Determinants of International Migration. *Scandinavian Journal of Economics* 117: 723-767.

Belot, M. V., and Hatton, T. J. (2012). Immigrant Selection in the OECD. *The Scandinavian Journal of Economics* 114(4): 1105-1128.

Besley, T. and Persson, T. (2011). The Logic of Political Violence. *Quarterly Journal of Economics* 126 (3): 1411-1445.

Birgier, D. P. Lundh, C., Haberfeld, Y., and Elldér, E. (2018). Self-Selection and Host Country Context in the Economic Assimilation of Political Refugees in the United States, Sweden, and Israel. *International Migration Review* 52(2): 524–558.

Borjas, G.J. (1987). Self-selection and the Earnings of Immigrants. *American Economic Review* 77: 531-53.
Borjas, G.J. (1999). Immigration and Welfare Magnets. *Journal of Labor Economics* 17(4): 607-637.

Borjas, G.J., Kauppinen, I., and Poutvaara, P. (2019). Self-Selection of Emigrants: Theory and Evidence on Stochastic Dominance in Observable and Unobservable Characteristics. *Economic Journal* 129: 143–171.

Brücker, H., Rother, N., Schupp, J., von Gostomski, C.B., Böhm, A., Fendel, T., Friedrich, M., Giesselmann, M., Kosyakova, Y., Kroh, M., Kühne, S., Liebau, E., Richter, D., Romiti, A., Schacht, D., Scheible, J.A., Schmelzer, P., Siegert, M., Sirries, S., Trübswetter, P., and Vallizadeh, E. (2016). Forced Migration, Arrival in Germany, and First Steps toward Integration. In Britzke, J., and Schupp, J. (Eds.), SOEP Wave Report 2016, p. 149-164.

Brückner, M. (2010). Population Size and Civil Conflict Risk: Is There a Causal Link? *Economic Journal* 120: 535–550.

Cattaneo, C. and Peri, G. (2016). The Migration Response to Increasing Temperatures. *Journal of Development Economics* 122: 127-146.

Chin, A. and Cortes, K.E. (2015). The Refugee/Asylum Seeker. In Handbook of the Economics of International Migration, Volume 1, Edited by Chiswick, B.R. and Miller, P.W. Pages 585-658.

Chiquiar, D. and Hanson, G.H. (2005). International Migration, Self-selection, and the Distribution of Wages: Evidence from Mexico and the United States. *Journal of Political Economy* 113(2): 239–281.

Chiswick, B. (1999). Are Immigrants Favorably Self-selected? *American Economic Review* 89(2): 181-185.

Cohen, Y. (2007). Self-Selection and Earnings Assimilation: Immigrants from the Former Soviet Union in Israel and the United States. *Demography* 44(3): 649–668.

Cortes, K. E. (2004). Are Refugees Different from Economic Immigrants? Some Empirical Evidence on the Heterogeneity of Immigrant Groups in the United States. *Review of Economics and Statistics*, 86(2), 465-480.

Couttenier, M., Preotu, V., Rohner, D., and Thoenig, M. (forthcoming). The Violent Legacy of Victimization: Post-conflict Evidence on Asylum Seekers, Crimes and Public Policy in Switzerland. *American Economic Review.*

Dinas, E., Matakos, K., Xefteris, D., and Hangartner, D. (2019). Waking Up the Golden Dawn: Does Exposure to the Refugee Crisis Increase Support for Extreme-Right Parties? *Political Analysis* 27: 244-254.

Docquier, F., Lowell, B.L., and Marfouk, A. (2009). A Gendered Assessment of Highly Skilled Emigration. *Population and Development Review* 35: 297–321.

Dustmann, C., Fasani, F., Frattini, T., Minale, L., and Schönberg, U. (2017). On the Economics and Politics of Refugee Migration. *Economic Policy* 32(91): 497-550.

Dustmann, C., Vasiljeva, K., and Damm, A.P. (2019). Refugee Migration and Electoral Outcomes. *The Review of Economic Studies* 86(5): 2035–2091.

ECFR (2017). The Mediterranean and Migration: Postcards from a Crisis. Available here: https://www.ecfr.eu/specials/mapping_migration# (last accessed 17 November 2019)

Edo, A., Giesing, Y., Öztunc, J., and Poutvaara, P. (2019). Immigration and Electoral Support for the Far-Left and the Far-Right. *European Economic Review* 115: 99-143.

Ethier, W.J. (1986). Illegal Immigration: The Host-Country Problem. *American Economic Review* 76: 56-71.

Eurofound (2016). Approaches to the Labour Market Integration of Refugees and Asylum Seekers. Publications Office of the European Union, Luxembourg.

Fernandez-Huertas Moraga, J. (2011). New Evidence on Emigrant Selection. *Review of Economics and Statistics* 93(1): 72–96.

Friebel, G. and Guriev, S. (2006). Smuggling Humans: A Theory of Debt-financed Migration. *Journal of the European Economic Association* 4: 1085-1111.

Friebel, G., Manchin, M., Mendola, M., and Prarolo, G. (2018). International Migration Intentions and Illegal Costs: Evidence from Africa-to-Europe Smuggling Routes. IZA DP No. 11978.

Gibney, M., Cornett L., Wood R., Haschke P., Arnon D., and Pisanò A. (2018). *The Political Terror Scale 1976-2017*. Available here: http://www.politicalterrorscale.org (last accessed 17 November 2019)

Glaeser, E.L., La Porta, R., Lopez-de-Silanes, F., and Shleifer, A. (2004). Do Institutions Cause Growth? *Journal of Economic Growth* 9(3): 271-303.

Grogger, J., and Hanson, G.H. (2011). Income Maximization and the Selection and Sorting of International Migrants. *Journal of Development Economics* 95: 42-57.

Guichard, L. (2017). Self-selection of Refugees. Manuscript, IAB.

Halla, M., Wagner, A. F., and Zweimüller, J. (2017). Immigration and Voting for the Far Right. *Journal of the European Economic Association* 15: 1341–1385.

Hanson G.H. and Spilimbergo A. (1999). Illegal Immigration, Border Enforcement, and Relative Wages: Evidence from Apprehensions at the US-Mexico border. *American Economic Review* 89(5): 1337–1357.

Hanushek, E.A. and Woessmann, L. (2015). The Knowledge Capital of Nations: Education and the Economics of Growth. Cambridge, MA: MIT Press, 2015

Hatton, T. J. (2015). Asylum Policy in the EU: the Case for Deeper Integration. *CESifo Economic Studies* 61(3-4): 605-637.

Hatton, T. J. (2016). Refugees, Asylum Seekers, and Policy in OECD Countries. *American Economic Review* 106(5): 441-45.

Hatton, T.J. (2017). Refugees and Asylum Seekers, the Crisis in Europe and the Future of Policy. *Economic Policy* 32(91): 447-496.

Hatton, T. J., and Williamson, J. G. (1998). The Age of Mass Migration: Causes and Economic Impact. Oxford University Press.

Hendricks, L. and Schoellman, T. (2018). Human Capital and Development Accounting: New Evidence from Wage Gains at Migration. *Quarterly Journal of Economics* 133(2): 665–700.

Hornung, E. (2014). Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia. *American Economic Review* 104(1): 84-122.

Huddleston, T., Bilgili, Ö., Joki, A.L., and Vankova, Z. (2015). Migrant Integration Policy Index 2015. Barcelona Centre for International Affairs, Migration Policy Group.

ILO (2019). Labour Market Statistics. Available here: https://www.ilo.org/ilostat/ (last accessed 17 November 2019)

IOM (2016). Mixed Migration Flows in the Mediterranean and Beyond. Available here: http://migration.iom.int/docs/Analysis_Flow_Monitoring_and_Human_Trafficking_Surveys_in_the_Mediterranean_and_Beyond_8_DECEMBER_2016.pdf (last accessed 17 November 2019)

IOM (2017). Mixed Migration Flows in the Eastern Mediterranean and Beyond. Available here: http://migration.iom.int/docs/Analysis_Flow_Monitoring_Surveys_in_the_Mediterranean_and_Beyond_11_April_2017.pdf (last accessed 17 November 2019)

Malthus, T. R. (1798). *An Essay on the Principle of Population, As It Affects the Future Improvement of Society: With Remarks on the Speculations of Mr. Godwin, M. Condorcet and Other Writers* (London: Printed for J. John).

Mayda, A. (2010). International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows. *Journal of Population Economics* 23: 1249–1274.

MEDAM (2018). Flexible Solidarity: A Comprehensive Strategy for Asylum and Immigration in the EU – 2018 MEDAM Assessment Report on Asylum and Migration Policies in Europe, Mercator Dialogue on Asylum and Migration, Kiel.

Moser, P., Voena, A., and Waldinger, F. (2014). German-Jewish Émigrés and U.S. Invention. *American Economic Review* 104(10): 3222-3255.

OECD (2016). Making Integration Work: Refugees and others in need of protection, OECD Publishing, Paris.

Orrenius P.M. and Zavodny M. (2005). Self-Selection among Undocumented Immigrants from Mexico. *Journal of Development Economics* 78(1): 215–40.

Otto, A.H. and Steinhardt, M.F. (2014). Immigration and Election Outcomes: Evidence from City Districts in Hamburg. *Regional Science and Urban Economics* 45: 67–79.

Parey, M., Ruhose, J., Waldinger, F., and Netz, N. (2017). The Selection of High-skilled Emigrants. *Review of Economics and Statistics* 99(5): 776-792.

Parsons, C., and Vézina, P. L. (2018). Migrant networks and trade: The Vietnamese boat people as a natural experiment. The Economic Journal 128(612): F210-F234.

Pettersson, T. and Eck, K. (2018). Organized Violence, 1989-2017. *Journal of Peace Research* 55(4): 535-547.

Pew Research Center (2019). 5 facts about illegal immigration in the U.S. Available here: https://www.pewresearch.org/fact-tank/2019/06/12/5-facts-about-illegal-immigration-in-the-u-s/ (last accessed 17 November 2019)

Roy, A.D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers* 3(2): 135-146.

Sequeira, S., Nunn, N., & Qian, N. (forthcoming). Immigrants and the Making of America. *Review of Economic Studies.*

Sjaastad, L.A. (1962). The Costs and Returns of Human Migration. *Journal of Political Economy* 70(5, Part 2): 80-93.

UN (2014). World Urbanization Prospects: The 2014 Revision. Available here: https://population.un.org/wup/Publications/Files/WUP2014-Report.pdf (last accessed 17 November 2019)

UN (2017). Gender Inequality Index. Available here:http://hdr.undp.org/en/composite/GII (last accessed 17 November 2019)

UN (2019). World Population Prospects. Available here: https://population.un.org/wpp/ (last accessed 17 November 2019)

UNHCR (1967). Convention and Protocol Relating to the Status of Refugees. Available here: http://www.unhcr.org/3b66c2aa10 (last accessed 17 November 2019)

UNHCR (2008). Challenges of Irregular Migration: Addressing Mixed Migration Flows. Available here: www.unhcr.org/49e479c911.pdf (last accessed 17 November 2019)

UNHCR (2017). Global Trends: Forced Displacement in 2016. Available here: http://www.unhcr.org/5943e8a34.pdf (last accessed 17 November 2019)

UNHCR (2019). Global Trends: Forced Displacement in 2018. Available here: https://www.unhcr.org/5d08d7ee7.pdf (last accessed 17 November 2019)

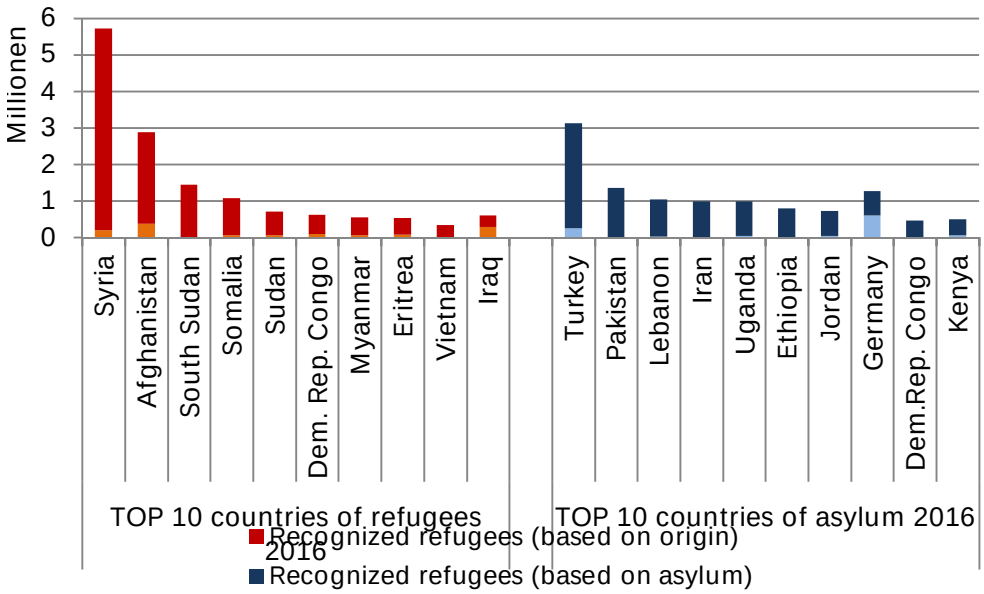
UNICEF (2017). Harrowing Journeys. Available here: https://data.unicef.org/resources/harrowing-journeys/ (last accessed 17 November 2019)

UNHCR (2018). See and Land Arrivals Monthly. Available here: https://data2.unhcr.org/en/situations/mediterranean (last accessed 17 November 2019)

Waldinger, F. (2010). Quality Matters: The Expulsion of Professors and the Consequences for Ph.D. Student Outcomes in Nazi Germany. *Journal of Political Economy* 118 (4): 787-831.

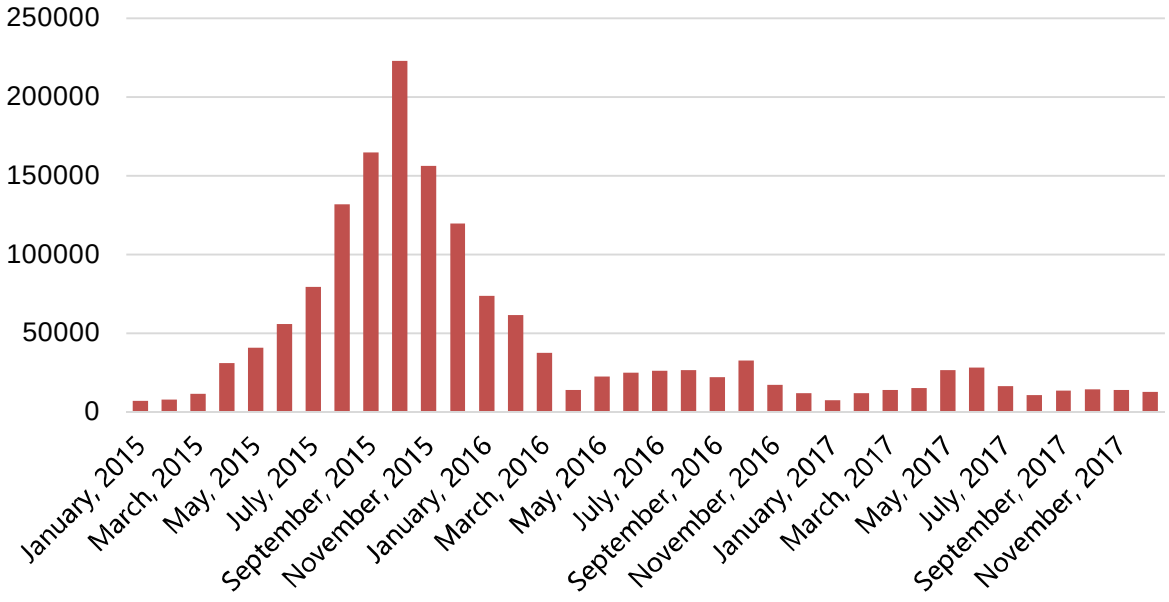
World Bank (2018). Asylum Seekers in the European Union: Building Evidence to Inform Policy Making, Washington, D.C.

Figure 1: Main countries of origin of refugees and main countries of asylum in 2016



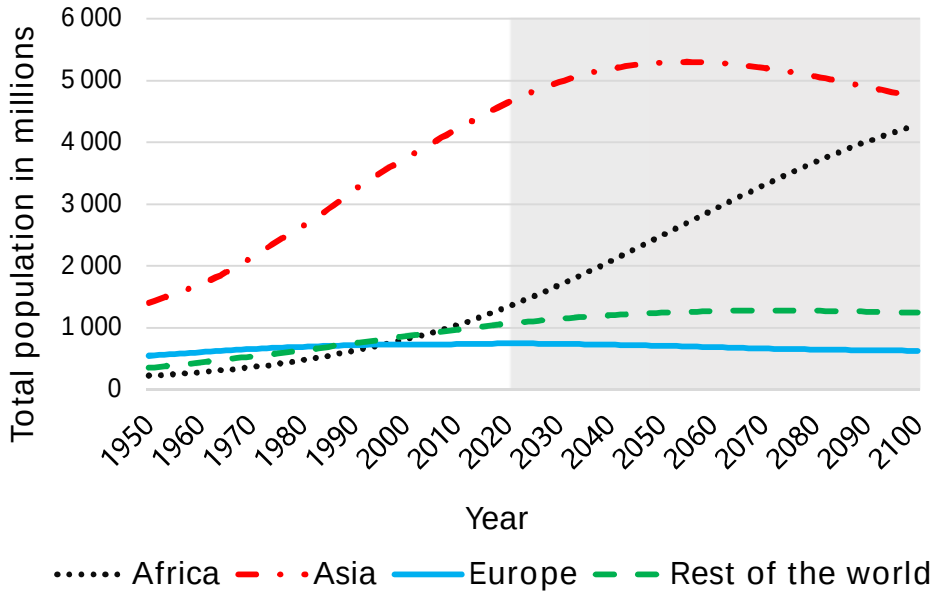
Source: Eurostat, UNHCR and authors' calculations.

Figure 2: Sea and land arrivals of refugees and irregular migrants to Europe, monthly



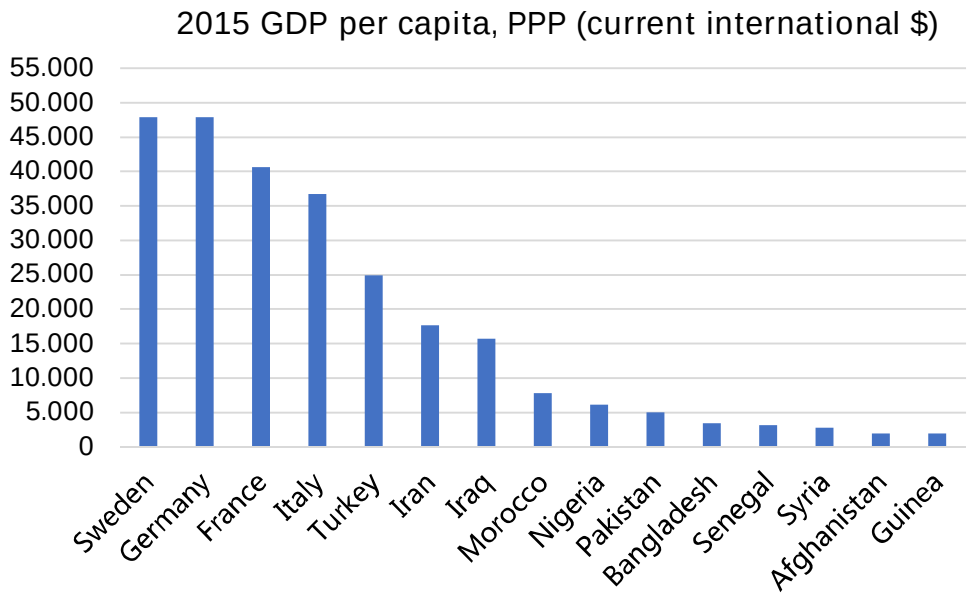
Source: UNHCR and authors' calculations.

Figure 3: Actual and predicted population development (medium-fertility variant)



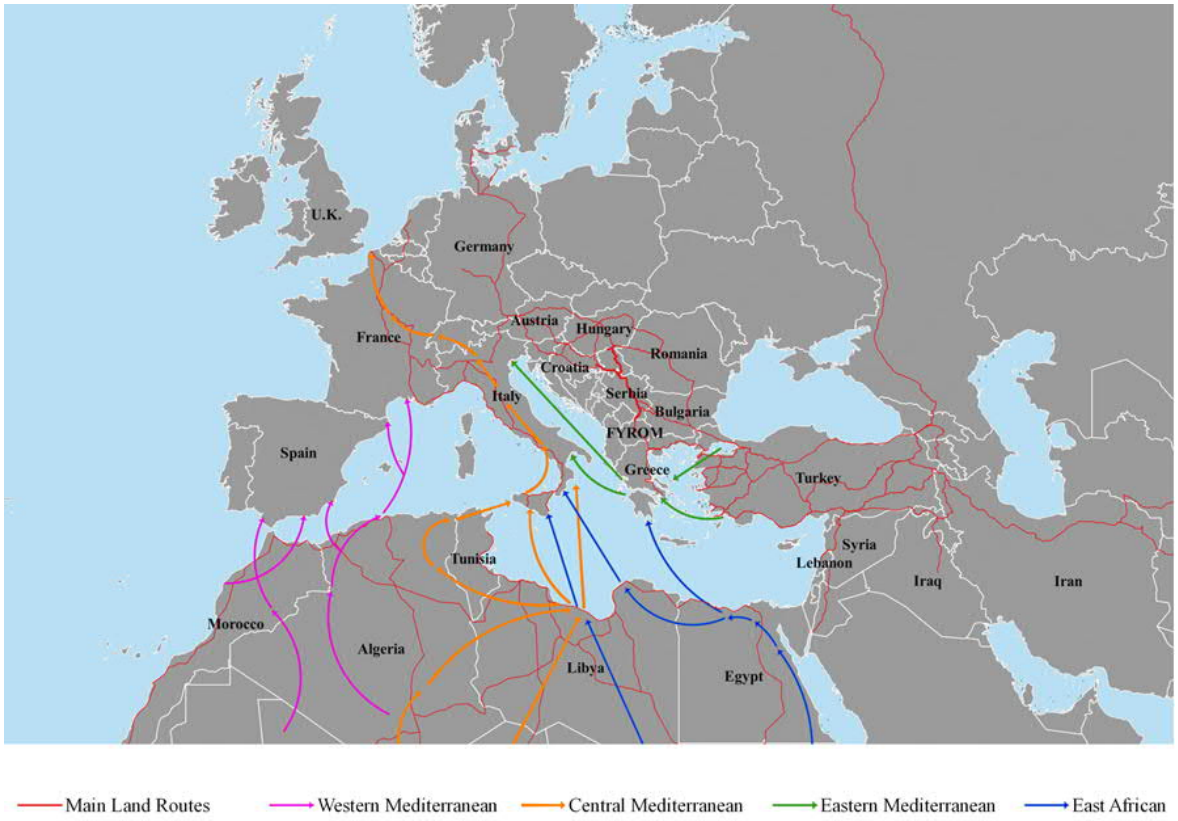
Source: United Nations World Population Prospects, the 2019 Revision.

Figure 4: GDP per capita at purchasing power parity in main origin and destination countries



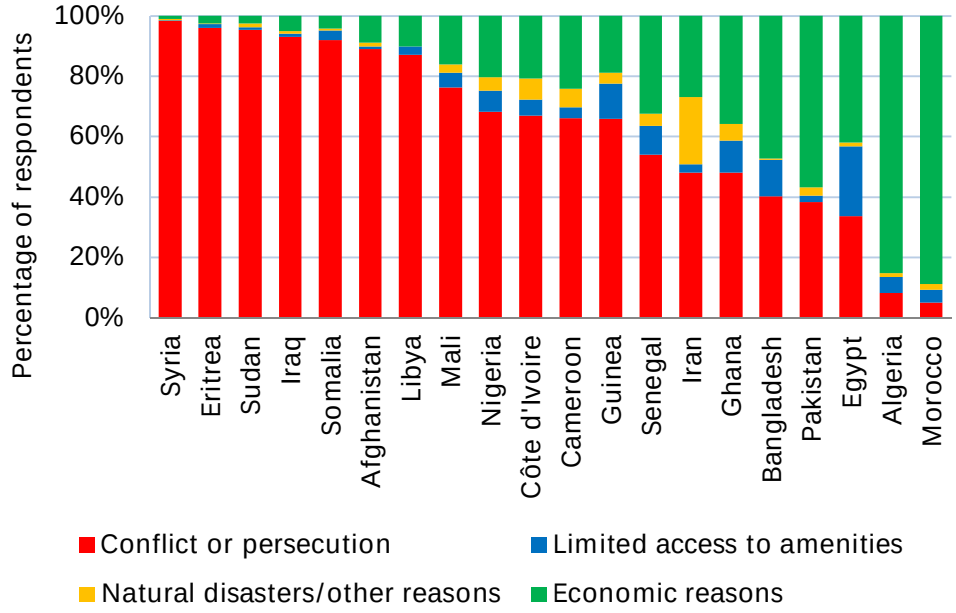
Notes: GDP per capita is depicted in purchasing power parity in 2015 for 10 largest origin countries in the FMS, and for France, Germany, Italy, Sweden, and Turkey. Source: The World Fact Book.

Figure 5: Mediterranean Sea routes and main land routes



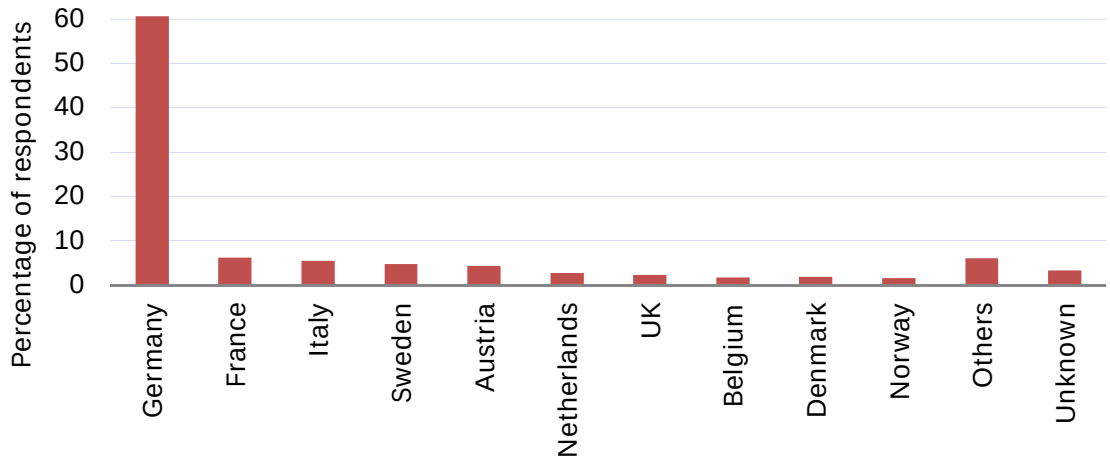
Notes: The map is for illustration purposes only. Names and boundaries do not imply official endorsement or acceptance by EBRD or IOM. Source: IOM and authors' calculations.

Figure 6: Reasons to emigrate by origin country, age 14+

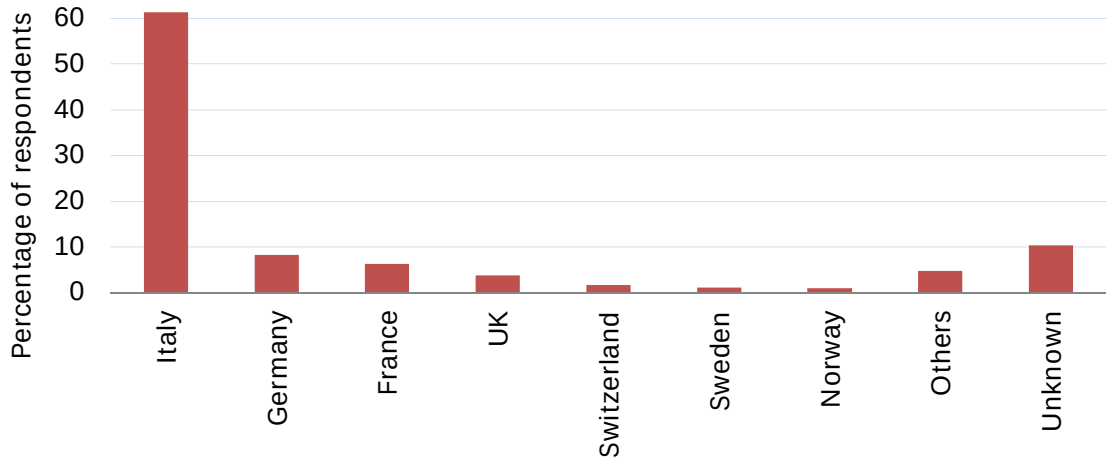


Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3) and authors' calculations.

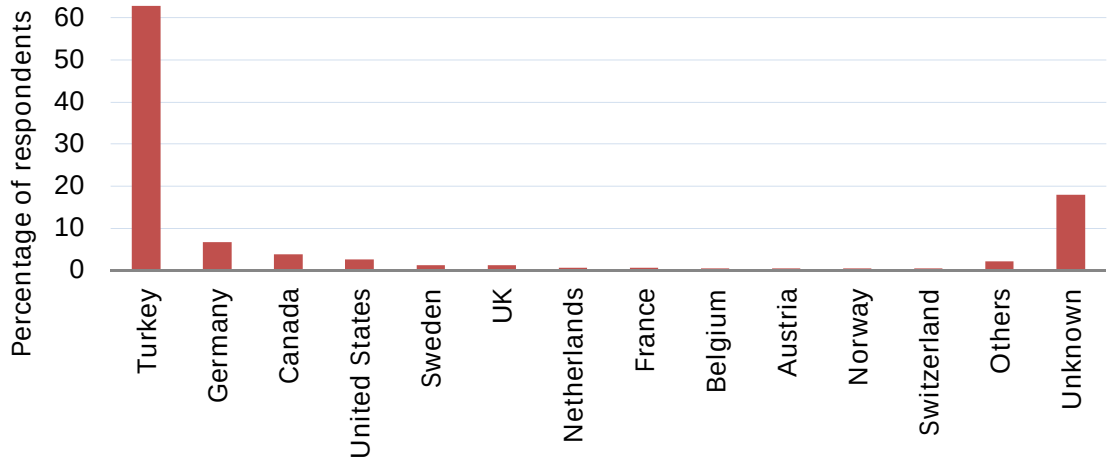
Figure 7: Intended destination country, age 14+
Waves 1 and 2



Wave 3 (Italy)

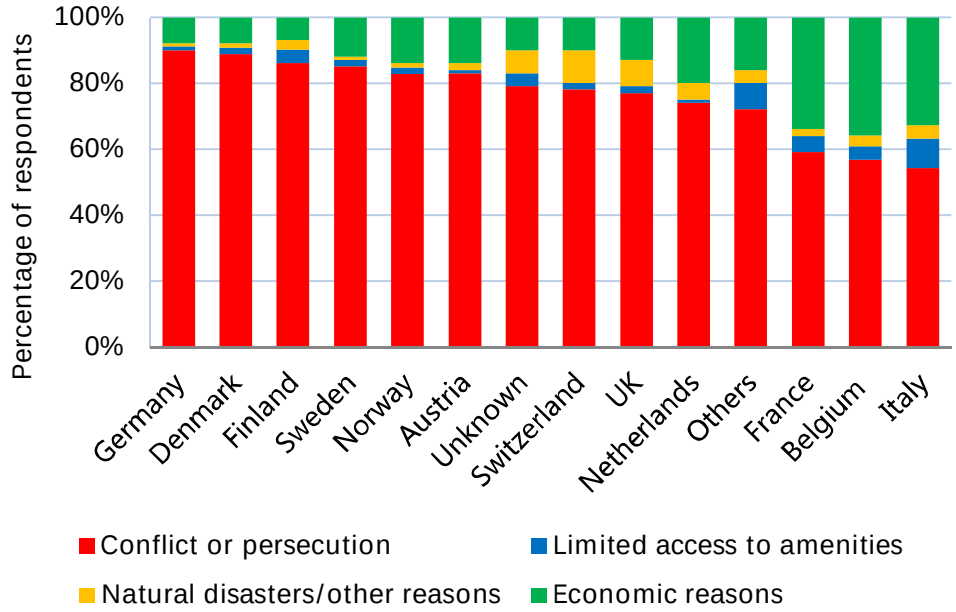


Wave 4 (Turkey)



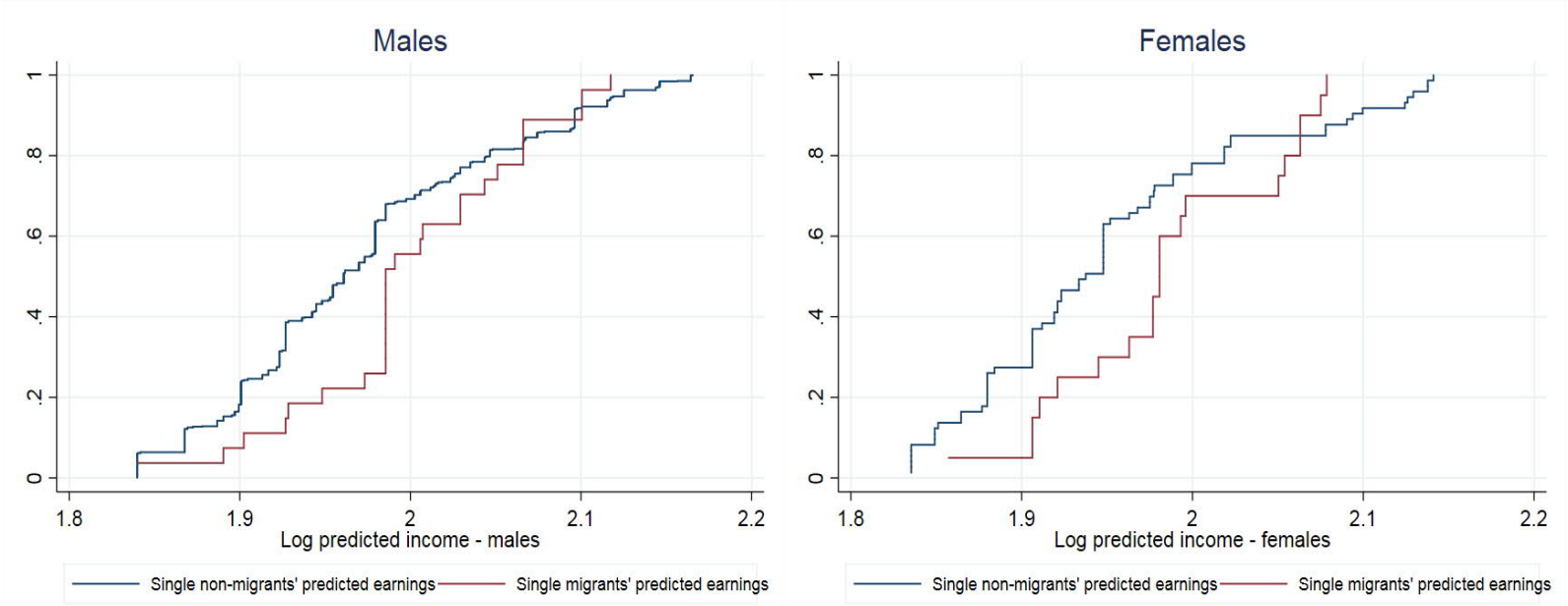
Source: Flow Monitoring Surveys (2015-18) and authors' calculations.

Figure 8: Reasons to emigrate by intended destination country, age 14+



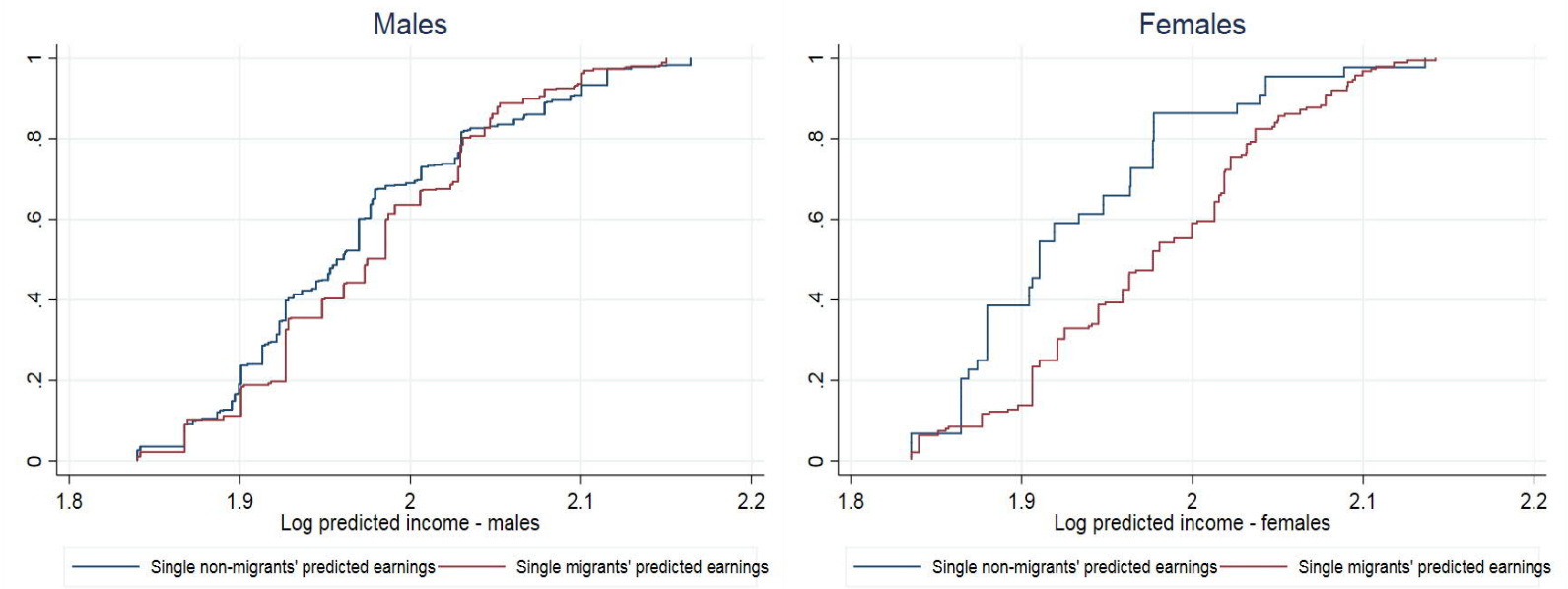
Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3) and authors' calculations.

Figure 9: CDF for single migrants' and single non-migrants' predicted income, reason to migrate: conflict/persecution



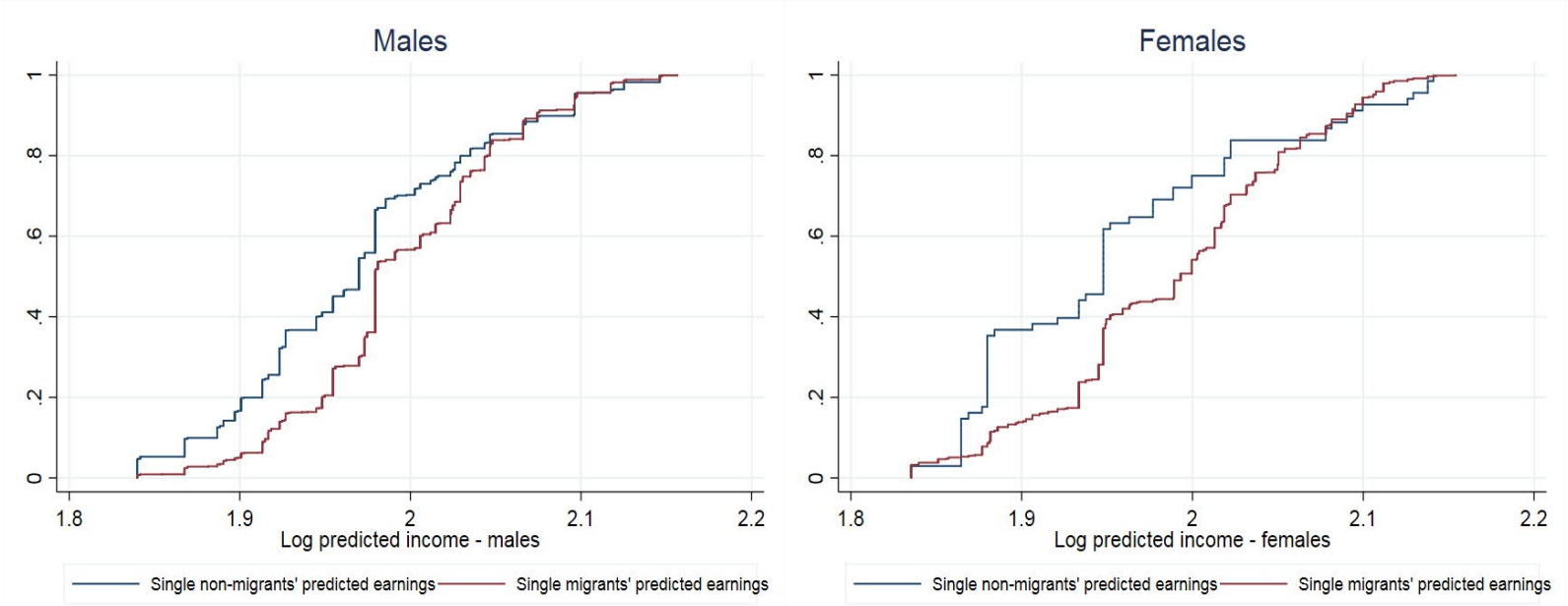
Notes: These charts present cumulative distribution functions (CDFs) of predicted income (in the country of origin) of refugees who migrate due to conflict or persecution with non-migrants' CDF. Income variables are calculated using the latest purchasing power parity estimates available and estimated to the 2015 US dollar by Gallup World Polls. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3), Gallup World Polls and authors' calculations.

Figure 10: CDF for single migrants' and single non-migrants' predicted income, reason to migrate: others (i.e. all but conflict/persecution)



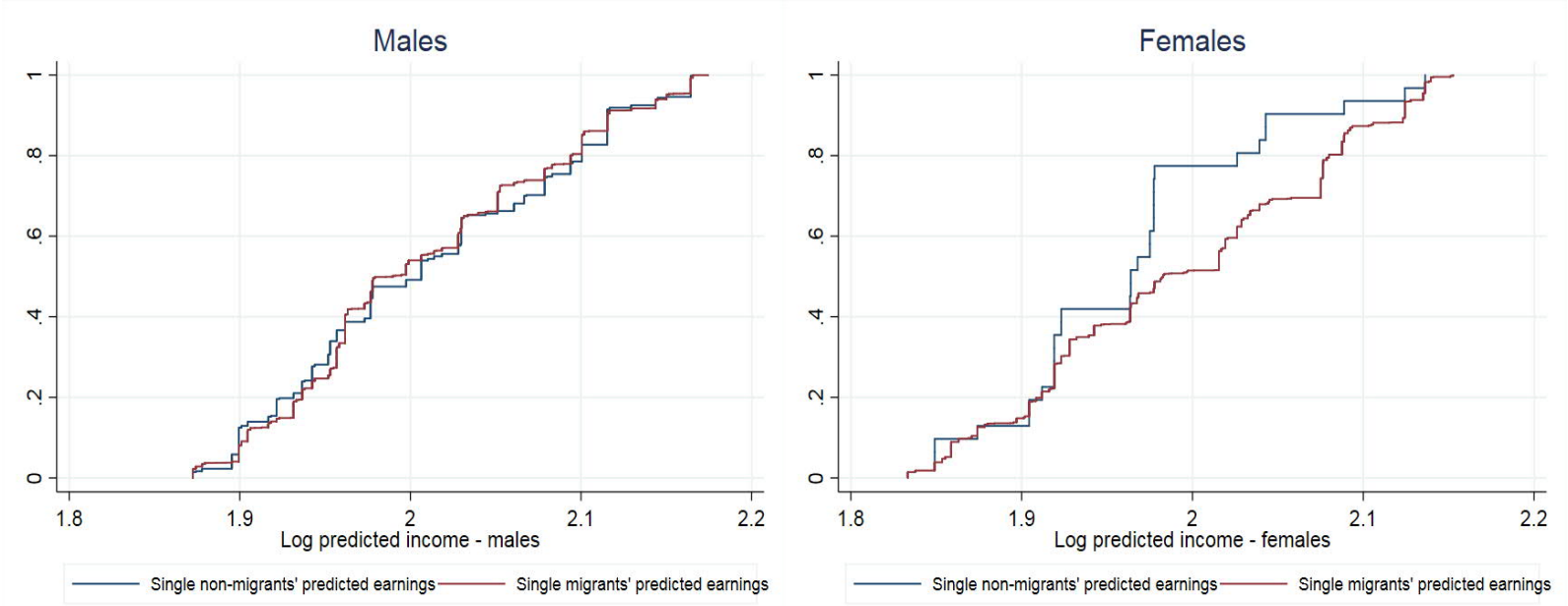
Notes: These charts present cumulative distribution functions (CDFs) of predicted income (in the country of origin) of irregular migrants whose main reason to emigrate was not conflict or persecution with non-migrants' CDF. Income variables are calculated using the latest purchasing power parity estimates available and estimated to the 2015 US dollar by Gallup World Polls. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3), Gallup World Polls and authors' calculations.

Figure 11: CDF for single migrants' and single non-migrants' predicted income, major conflict countries and by gender



Notes: These charts present migrants' and non-migrants' cumulative distribution functions (CDFs) of predicted income for major conflict countries. Income variables are calculated using the latest purchasing power parity estimates available and estimated to the 2015 US dollar by Gallup World Polls. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3), Gallup World Polls, Uppsala Conflict Database and authors' calculations.

Figure 12: CDF for single migrants' and single non-migrants' predicted income, minor or no conflict countries and by gender



Notes: These charts present migrants' and non-migrants' cumulative distribution functions (CDFs) of predicted income for minor or no conflict countries. Income variables are calculated using the latest purchasing power parity estimates available and estimated to the 2015 US dollar by Gallup World Polls. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3), Gallup World Polls, Uppsala Conflict Database and authors' calculations.

Table 1: Comparison of Flow Monitoring Surveys and Eurostat data by nationality

Nationality	Number of observations from FMS	Share		Share of Males	
		FMS	Eurostat	FMS	Eurostat
Syria	5562	0.256	0.255	0.719	0.674
Afghanistan	4026	0.185	0.128	0.866	0.765
Iraq	2028	0.093	0.085	0.803	0.681
Nigeria	1223	0.056	0.031	0.729	0.698
Pakistan	1056	0.049	0.036	0.985	0.933
Eritrea	971	0.045	0.040	0.778	0.705
Morocco	817	0.038	0.007	0.902	0.896
Iran	721	0.033	0.024	0.770	0.711
Guinea	553	0.025	0.008	0.955	0.790
Bangladesh	418	0.019	0.014	0.995	0.940
Senegal	400	0.018	0.008	0.983	0.939
Mali	323	0.015	0.009	0.954	0.936
Sudan	317	0.015	0.009	0.962	0.916
Somalia	285	0.013	0.019	0.677	0.678
Côte d'Ivoire	284	0.013	0.006	0.852	0.806
Algeria	246	0.011	0.008	0.967	0.888
Egypt	237	0.011	0.004	0.987	0.830
Ghana	215	0.010	0.006	0.986	0.882
Cameroon	165	0.008	0.003	0.812	0.675
Libya	108	0.005	0.004	0.935	0.758
		Correlation: 0.978***		Correlation: 0.875***	

Notes: Eurostat data set contains information on the number of asylum applications by nationality. The table only includes countries, which at least 100 respondents in the FMS named their nationality. Eritrea is not included in the analysis as it is not surveyed by Gallup World Polls. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3). Eurostat, 2015 and 2016.
*** p < .01, ** p < .05, * p < .1.

Table 2.a: Descriptive characteristics from Flow Monitoring Surveys Waves 1-3, full sample with no restrictions (age 14+)

Variables	(1) Full sample	(2) Males	(3) Females	(4) Wave 1 (10/2015 - 12/2015)	(5) Wave 2 (01/2016 - 11/2016)	(6) Wave 3 (06/2016 - 11/2016)
Age	26.41 (8.43)	26.09 (8.24)	27.93 (9.15)	27.40 (8.83)	27.77 (8.71)	22.55 (6.01)
Male	0.82 (0.38)	--	--	0.79 (0.40)	0.81 (0.39)	0.88 (0.33)
Married	0.30 (0.46)	0.25 (0.43)	0.55 (0.50)	--	0.42 (0.49)	0.17 (0.38)
Divorced	0.01 (0.01)	0.00 (0.06)	0.03 (0.18)	--	0.01 (0.11)	0.00 (0.08)
Widowed	0.01 (0.01)	0.00 (0.06)	0.03 (0.18)	--	0.01 (0.10)	0.00 (0.07)
Secondary education	0.50 (0.50)	0.50 (0.50)	0.48 (0.50)	0.40 (0.49)	0.53 (0.50)	0.46 (0.50)
Tertiary education	0.18 (0.39)	0.17 (0.38)	0.22 (0.41)	0.36 (0.48)	0.20 (0.40)	0.07 (0.25)
Employed	0.47 (0.50)	0.51 (0.50)	0.25 (0.43)	--	0.47 (0.50)	0.47 (0.50)
Reasons for leaving:						
Conflict or persecution	0.77 (0.42)	0.76 (0.43)	0.84 (0.36)	0.86 (0.35)	0.79 (0.40)	0.68 (0.47)
Economic reasons	0.18 (0.38)	0.19 (0.40)	0.09 (0.29)	0.11 (0.31)	0.17 (0.38)	0.21 (0.41)
Limited access to amenities	0.02 (0.15)	0.02 (0.15)	0.02 (0.14)	0.01 (0.08)	0.01 (0.07)	0.07 (0.26)
Other reasons	0.02 (0.15)	0.02 (0.14)	0.03 (0.18)	0.02 (0.10)	0.02 (0.14)	0.03 (0.17)
Nationalities:						
Syria	0.28 (0.45)	0.25 (0.43)	0.46 (0.50)	0.45 (0.50)	0.36 (0.48)	0.01 (0.11)
Afghanistan	0.20 (0.40)	0.22 (0.41)	0.16 (0.36)	0.23 (0.42)	0.28 (0.45)	0.01 (0.09)
Iraq	0.10 (0.30)	0.10 (0.30)	0.12 (0.32)	0.14 (0.34)	0.14 (0.34)	0.01 (0.09)
Nigeria	0.06 (0.25)	0.06 (0.25)	0.10 (0.30)	0.01 (0.07)	0.00 (0.05)	0.22 (0.41)
Pakistan	0.05 (0.22)	0.06 (0.25)	0.00 (0.07)	0.06 (0.23)	0.07 (0.25)	0.02 (0.13)
Eritrea	0.04 (0.21)	0.05 (0.23)	0.05 (0.10)	0.01 (0.04)	0.00 (0.03)	0.14 (0.35)
Morocco	0.04 (0.20)	0.05 (0.21)	0.02 (0.15)	0.02 (0.15)	0.05 (0.22)	0.03 (0.17)
Iran	0.04 (0.20)	0.03 (0.18)	0.05 (0.21)	0.05 (0.22)	0.05 (0.21)	0.01 (0.05)
Guinea	0.03 (0.18)	0.03 (0.17)	0.01 (0.09)	0.01 (0.02)	0.01 (0.03)	0.10 (0.31)
Bangladesh	0.02 (0.14)	0.03 (0.16)	0.00 (0.02)	0.01 (0.08)	0.00 (0.06)	0.07 (0.25)
N	20,465	16,781	3,684	1,667	13,063	5,735

Notes: Means (standard deviations). The sample sizes for some variables are different either due to missing data or because they were not asked in each wave. Secondary and tertiary education refer to the highest completed education. We report shares for the top 10 nationalities in the full sample. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3).

Table 2.b: Descriptive characteristics from Flow Monitoring Surveys Waves 1-3, restricted sample with adults aged 25 to 64

Variables	(1) Full sample	(2) Males	(3) Females	(4) Wave 1 (10/2015 - 12/2015)	(5) Wave 2 (01/2016 - 11/2016)	(6) Wave 3 (06/2016 - 11/2016)
Age	32.42 (7.60)	32.20 (7.39)	33.29 (8.36)	32.70 (8.32)	32.92 (7.80)	29.60 (5.03)
Male	0.80 (0.40)	--	--	0.80 (0.39)	0.79 (0.41)	0.87 (0.33)
Married	0.57 (0.49)	0.52 (0.50)	0.78 (0.41)	--	0.65 (0.47)	0.40 (0.50)
Divorced	0.01 (0.13)	0.01 (0.10)	0.05 (0.21)	--	0.02 (0.14)	0.01 (0.12)
Widowed	0.01 (0.16)	0.00 (0.08)	0.05 (0.21)	--	0.01 (0.13)	0.01 (0.11)
Secondary education	0.48 (0.50)	0.49 (0.50)	0.45 (0.50)	0.35 (0.47)	0.50 (0.49)	0.44 (0.50)
Tertiary education	0.25 (0.43)	0.25 (0.43)	0.28 (0.45)	0.41 (0.49)	0.26 (0.44)	0.12 (0.32)
Employed	0.63 (0.48)	0.73 (0.44)	0.28 (0.45)	--	0.61 (0.49)	0.69 (0.46)
Reasons for leaving:						
Conflict or persecution	0.79 (0.41)	0.77 (0.42)	0.86 (0.35)	0.86 (0.34)	0.79 (0.41)	0.69 (0.46)
Economic reasons	0.17 (0.37)	0.19 (0.39)	0.09 (0.28)	0.10 (0.30)	0.17 (0.37)	0.22 (0.41)
Limited access to amenities	0.02 (0.13)	0.01 (0.11)	0.01 (0.10)	0.02 (0.08)	0.01 (0.07)	0.06 (0.23)
Other reasons	0.02 (0.16)	0.02 (0.14)	0.04 (0.19)	0.02 (0.08)	0.03 (0.16)	0.02 (0.15)
Nationalities:						
Syria	0.37 (0.48)	0.33 (0.47)	0.52 (0.50)	0.51 (0.50)	0.41 (0.49)	0.03 (0.07)
Afghanistan	0.17 (0.37)	0.18 (0.38)	0.15 (0.36)	0.15 (0.36)	0.20 (0.40)	0.01 (0.06)
Iraq	0.14 (0.35)	0.14 (0.35)	0.13 (0.34)	0.15 (0.36)	0.16 (0.37)	0.02 (0.15)
Nigeria	0.05 (0.21)	0.05 (0.22)	0.04 (0.20)	0.01 (0.08)	0.00 (0.05)	0.30 (0.46)
Pakistan	0.06 (0.23)	0.07 (0.25)	0.01 (0.08)	0.06 (0.24)	0.06 (0.23)	0.03 (0.17)
Morocco	0.05 (0.20)	0.05 (0.22)	0.03 (0.17)	0.02 (0.13)	0.05 (0.21)	0.18 (0.38)
Iran	0.06 (0.23)	0.05 (0.23)	0.07 (0.25)	0.07 (0.24)	0.06 (0.25)	0.01 (0.07)
Bangladesh	0.01 (0.12)	0.02 (0.13)	0.00 (0.02)	0.01 (0.07)	0.01 (0.05)	0.07 (0.26)
Senegal	0.01 (0.10)	0.01 (0.12)	0.00 (0.06)	0.01 (0.05)	0.01 (0.06)	0.06 (0.24)
Sudan	0.01 (0.11)	0.01 (0.09)	0.00 (0.07)	0.01 (0.06)	0.01 (0.05)	0.05 (0.20)
N	9,828	7,865	1,963	925	7,475	1,428

Notes: Means (standard deviations). Respondents from Eritrea and countries with less than 100 respondents are excluded. The sample sizes for some variables are different either due to missing data or because they were not asked in each wave. Secondary and tertiary education refer to the highest completed education. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3).

Table 3: Descriptive characteristics from Gallup World Polls, 2009-2014

	(1) Full sample (all ages)	(2) Full sample (all ages) males	(3) Full sample (all ages) females
Age	35.22	36.14	34.23
	(14.68)	(15.23)	(14.00)
Male	0.51	--	--
	(0.50)		
Married	0.61	0.60	0.63
	(0.49)	(0.48)	(0.48)
Divorced	0.02	0.02	0.03
	(0.15)	(0.13)	(0.18)
Widowed	0.04	0.02	0.08
	(0.21)	(0.14)	(0.27)
Secondary education	0.42	0.45	0.40
	(0.49)	(0.50)	(0.50)
Tertiary education	0.08	0.10	0.06
	(0.27)	(0.30)	(0.24)
Employed	0.48	0.65	0.29
	(0.50)	(0.48)	(0.45)
N	129,431	67,167	62,264

	Restricted sample (ages 25-64)	Restricted sample (ages 25-64) males	Restricted sample (ages 25-64) females
Age	39.31	39.95	38.63
	(10.90)	(11.13)	(10.61)
Male	0.52	--	--
	(0.50)		
Married	0.76	0.77	0.76
	(0.42)	(0.42)	(0.43)
Divorced	0.03	0.02	0.04
	(0.17)	(0.15)	(0.20)
Widowed	0.04	0.02	0.08
	(0.21)	(0.15)	(0.71)
Secondary education	0.37	0.40	0.35
	(0.48)	(0.49)	(0.48)
Tertiary education	0.10	0.12	0.07
	(0.29)	(0.32)	(0.26)
Employed	0.55	0.75	0.33
	(0.50)	(0.43)	(0.47)
N	89,484	46,493	42,991

Notes: Means (standard deviations). This table presents summary statistics for origin countries included in the analysis. Secondary and tertiary education refer to the highest completed education. Source: Gallup World Polls, 2009-2014.

Table 4: Educational attainment by origin country

	(1) Secondary education (Full sample)	(2) Tertiary education (Full sample)	(3) Secondary education (Ages 25-64)	(4) Tertiary education (Ages 25-64)
Flow Monitoring Surveys, Waves 1-3				
Syria	0.51 (0.50)	0.31 (0.46)	0.48 (0.50)	0.34 (0.47)
Afghanistan	0.45 (0.50)	0.13 (0.34)	0.39 (0.49)	0.19 (0.39)
Iraq	0.61 (0.49)	0.20 (0.40)	0.59 (0.49)	0.23 (0.42)
Nigeria	0.50 (0.50)	0.06 (0.23)	0.52 (0.50)	0.09 (0.29)
Pakistan	0.56 (0.50)	0.11 (0.31)	0.53 (0.50)	0.14 (0.35)
Morocco	0.66 (0.47)	0.09 (0.29)	0.63 (0.48)	0.11 (0.32)
Iran	0.44 (0.50)	0.40 (0.49)	0.41 (0.49)	0.45 (0.50)
Guinea	0.47 (0.50)	0.06 (0.24)	0.38 (0.48)	0.18 (0.39)
Bangladesh	0.39 (0.48)	0.02 (0.14)	0.44 (0.49)	0.05 (0.22)
Senegal	0.36 (0.48)	0.04 (0.20)	0.32 (0.47)	0.06 (0.23)
Gallup World Polls				
Syria	0.37 (0.48)	0.07 (0.25)	0.35 (0.48)	0.09 (0.28)
Afghanistan	0.25 (0.44)	0.04 (0.19)	0.21 (0.41)	0.05 (0.22)
Iraq	0.46 (0.50)	0.13 (0.33)	0.43 (0.50)	0.15 (0.35)
Nigeria	0.71 (0.45)	0.03 (0.17)	0.69 (0.46)	0.04 (0.20)
Pakistan	0.27 (0.44)	0.05 (0.22)	0.24 (0.42)	0.06 (0.23)
Morocco	0.34 (0.47)	0.06 (0.25)	0.26 (0.44)	0.08 (0.27)
Iran	0.58 (0.49)	0.25 (0.43)	0.54 (0.50)	0.28 (0.44)
Guinea	0.16 (0.37)	0.06 (0.24)	0.12 (0.32)	0.08 (0.27)
Bangladesh	0.47 (0.49)	0.03 (0.16)	0.41 (0.49)	0.03 (0.18)
Senegal	0.39 (0.48)	0.02 (0.15)	0.33 (0.47)	0.03 (0.18)

Notes: Means (standard deviations). Secondary and tertiary education refer to the highest completed education. Source: Flow Monitoring Surveys, 2015-2016 (waves 1-3). Gallup World Polls, 2009-2014.

Table 5: Self-selection of refugees and irregular migrants, adults aged 25-64

Outcome ➔	(1) Migrant due to any reason	(2) Refugee	(3) Irregular migrant
Secondary education	0.022***	0.021***	-0.007***
	(0.002)	(0.002)	(0.001)
Tertiary education	0.037***	0.032***	0.011***
	(0.004)	(0.003)	(0.002)
Employed	0.001	0.000	0.002**
	(0.002)	(0.002)	(0.001)
Male	0.057***	0.035***	0.014***
	(0.002)	(0.002)	(0.001)
Age 25-34	0.079***	0.051***	0.024***
	(0.002)	(0.002)	(0.001)
Age 35-44	0.035***	0.024***	0.010***
	(0.002)	(0.002)	(0.001)
Age 45-54	0.012***	0.008***	0.003***
	(0.002)	(0.002)	(0.001)
Married	-0.035***	-0.010***	-0.014***
	(0.003)	(0.003)	(0.002)
Divorced	-0.036***	-0.032***	0.005
	(0.006)	(0.004)	(0.004)
Widowed	-0.013***	-0.000	-0.006**
	(0.005)	(0.004)	(0.002)
Country FE	Yes	Yes	Yes
r2	0.081	0.077	0.024
N	62488	62488	62488

Notes: Robust standard errors in parentheses. Outcome variable, *migrant*, is equal to 1 for respondents in the Flow Monitoring Surveys and 0 for participants in Gallup World Polls. Reference categories are as follows: less than secondary education, unemployed or out of labor force, female, age 54+, and single. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3). Gallup World Polls, 2009-2014.
*** p < .01, ** p < .05, * p < .1.

Table 6: Self-selection of refugees and irregular migrants, adults aged 25-64, males

Outcome ➔	(1) Migrant due to any reason	(2) Refugee	(3) Irregular migrant
Secondary education	0.021***	0.019***	-0.013***
	(0.004)	(0.003)	(0.002)
Tertiary education	0.036***	0.035***	0.002
	(0.005)	(0.005)	(0.003)
Employed	0.012***	0.018***	-0.004**
	(0.004)	(0.003)	(0.002)
Age 25-34	0.107***	0.065***	0.035***
	(0.004)	(0.003)	(0.002)
Age 35-44	0.043***	0.025***	0.018***
	(0.003)	(0.003)	(0.002)
Age 45-54	0.014***	0.008***	0.006***
	(0.003)	(0.003)	(0.001)
Married	-0.062***	-0.028***	-0.019***
	(0.005)	(0.004)	(0.003)
Divorced	-0.081***	-0.069***	0.002
	(0.011)	(0.008)	(0.007)
Widowed	-0.070***	-0.051***	-0.008
	(0.011)	(0.010)	(0.005)
Country FE	Yes	Yes	Yes
r2	0.086	0.082	0.032
N	33253	33253	33253

Notes: Robust standard errors in parentheses. Outcome variable, *migrant*, is equal to 1 for respondents in the Flow Monitoring Surveys and 0 for participants in Gallup World Polls. Reference categories are as follows: less than secondary education, unemployed or out of labor force, female, age 54+, and single. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3). Gallup World Polls, 2009-2014.
*** p < .01, ** p < .05, * p < .1.

Table 7: Self-selection of refugees and irregular migrants, adults aged 25-64, females

Outcome ➔	(1) Migrant due to any reason	(2) Refugee	(3) Irregular migrant
Secondary education	0.025***	0.026***	-0.001
	(0.003)	(0.002)	(0.001)
Tertiary education	0.045***	0.035***	0.024***
	(0.005)	(0.004)	(0.003)
Employed	-0.012***	-0.016***	0.004***
	(0.002)	(0.002)	(0.001)
Age 25-34	0.035***	0.024***	0.012***
	(0.003)	(0.003)	(0.001)
Age 35-44	0.016***	0.013***	0.002**
	(0.003)	(0.002)	(0.001)
Age 45-54	0.003	0.003	-0.000
	(0.002)	(0.002)	(0.001)
Married	0.024***	0.024***	0.001
	(0.003)	(0.002)	(0.002)
Divorced	0.022***	0.008*	0.015***
	(0.006)	(0.005)	(0.005)
Widowed	0.031***	0.029***	0.004
	(0.005)	(0.004)	(0.002)
Country FE	Yes	Yes	Yes
r2	0.061	0.071	0.020
N	29235	29235	29235

Notes: Robust standard errors in parentheses. Outcome variable, *migrant*, is equal to 1 for respondents in the Flow Monitoring Surveys and 0 for participants in Gallup World Polls. Reference categories are as follows: less than secondary education, unemployed or out of labor force, female, age 54+, and single. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3). Gallup World Polls, 2009-2014.
*** p < .01, ** p < .05, * p < .1.

Table 8: Self-selection of refugees and irregular migrants, adults aged 25-64

Sample ➔	(1) Both genders & Major conflict countries	(2) Both genders & Minor/no conflict countries	(3) Males & Major conflict countries	(4) Males & Minor/no conflict countries	(5) Females & Major conflict countries	(6) Females & Minor/no conflict countries
Secondary education	0.045***	-0.006**	0.045***	-0.010**	0.046***	0.001
	(0.004)	(0.002)	(0.005)	(0.004)	(0.004)	(0.002)
Tertiary education	0.049***	0.016***	0.053***	0.008	0.057***	0.028***
	(0.005)	(0.005)	(0.007)	(0.008)	(0.007)	(0.006)
Employed	0.003	-0.001	0.029***	-0.011**	-0.034***	0.015***
	(0.003)	(0.003)	(0.005)	(0.005)	(0.003)	(0.003)
Male	0.066***	0.044***	--	--	--	--
	(0.003)	(0.003)				
Age 25-34	0.097***	0.061***	0.126***	0.095***	0.050***	0.017***
	(0.004)	(0.003)	(0.006)	(0.005)	(0.005)	(0.002)
Age 35-44	0.052***	0.016***	0.068***	0.021***	0.027***	0.001
	(0.004)	(0.002)	(0.005)	(0.003)	(0.005)	(0.001)
Age 45-54	0.021***	0.005***	0.028***	0.005**	0.009*	-0.001
	(0.003)	(0.001)	(0.005)	(0.002)	(0.005)	(0.001)
Married	-0.031***	-0.043***	-0.067***	-0.059***	0.039***	0.005
	(0.005)	(0.004)	(0.007)	(0.006)	(0.005)	(0.003)
Divorced	-0.066***	-0.003	-0.133***	-0.001	0.016	0.025***
	(0.008)	(0.008)	(0.013)	(0.018)	(0.010)	(0.008)
Widowed	-0.015**	-0.018***	-0.088***	-0.033**	0.049***	0.007*
	(0.007)	(0.005)	(0.016)	(0.013)	(0.008)	(0.004)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
r2	0.079	0.061	0.082	0.072	0.068	0.028
N	34405	28083	18869	14384	15536	13699

Notes: Robust standard errors in parentheses. Outcome variable, migrant, is equal to 1 for respondents in the Flow Monitoring Surveys and 0 for participants in Gallup World Polls. Countries are classified by the level of conflict following the definitions provided by Uppsala Conflict Data Program: major conflict category includes countries with 1000 or more battle-related deaths in any of the years over the sample period (Afghanistan, Iraq, Nigeria, Pakistan, Sudan, Syria); minor conflict category includes countries with 25 to 999 battle-related casualties in any of the years over the sample period (Algeria, Iran); no conflict category includes countries that did not experience a major conflict or minor conflict in any of the years over the sample period (Bangladesh, Cameroon, Côte d'Ivoire, Morocco, Senegal). Reference categories are as follows: less than secondary education, unemployed or out of labor force, female, age 54+, and single. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3). Gallup World Polls, 2009-2014.
*** p < .01, ** p < .05, * p < .1.

Table 9: Self-selection of refugees and irregular migrants at the sub-regional level, adults aged 25-64, major and minor conflict countries only

	(1)	(2)	(3)	(4)	(5)	(6)
Sample ➔	Both genders & More intensive conflict	Both genders & Less intensive conflict	Males & More intensive conflict	Males & Less intensive conflict	Females & More intensive conflict	Females & Less intensive conflict
Secondary education	0.055***	0.001	0.045***	-0.018**	0.077***	0.025***
	(0.007)	(0.006)	(0.009)	(0.008)	(0.010)	(0.008)
Tertiary education	0.091***	0.013	0.096***	-0.006	0.123***	0.056***
	(0.010)	(0.008)	(0.013)	(0.011)	(0.017)	(0.012)
Employed	0.067***	0.081***	0.167***	0.112***	-0.047***	0.040***
	(0.007)	(0.006)	(0.010)	(0.008)	(0.011)	(0.008)
Male	0.069***	0.056***	--	--	--	--
	(0.007)	(0.005)				
Age 25-34	0.197***	0.126***	0.183***	0.127***	0.176***	0.105***
	(0.012)	(0.008)	(0.015)	(0.011)	(0.017)	(0.011)
Age 35-44	0.112***	0.099***	0.085***	0.110***	0.121***	0.076***
	(0.012)	(0.008)	(0.015)	(0.011)	(0.017)	(0.011)
Age 45-54	0.064***	0.037***	0.032**	0.038***	0.081***	0.030***
	(0.012)	(0.008)	(0.016)	(0.011)	(0.018)	(0.011)
Married	0.033***	-0.032***	-0.047***	-0.089***	0.177***	0.066***
	(0.008)	(0.007)	(0.011)	(0.010)	(0.012)	(0.009)
Divorced	-0.050**	0.067***	-0.232***	-0.101***	0.151***	0.221***
	(0.020)	(0.018)	(0.032)	(0.028)	(0.026)	(0.022)
Widowed	0.060***	0.043***	-0.234***	-0.159***	0.254***	0.173***
	(0.018)	(0.016)	(0.036)	(0.029)	(0.021)	(0.018)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
r2	0.284	0.137	0.272	0.132	0.330	0.147
N	20450	24308	11542	13424	8908	10884

Notes: Robust standard errors in parentheses. Sample only includes major and minor conflict countries: Afghanistan, Algeria, Iran, Iraq, Nigeria, Pakistan, Sudan, and Syria. Outcome variable, migrant, is equal to 1 for respondents in the Flow Monitoring Surveys and 0 for participants in Gallup World Polls. Sub-regions in Gallup World Polls classified as *more intensive conflict* or *less intensive conflict* based on the number of battle-related deaths data obtained from the Uppsala Conflict Data Program. *More intensive conflict* refers to the sub-regions that experience equal to or more battle-related deaths than the country median (over 2009-2014 period) and *less intensive conflict* refers to the sub-regions that experience less battle-related deaths than the country median (over 2009-2014 period). Reference categories are as follows: less than secondary education, unemployed or out of labor force, female, age 54+, and single. Source: FMS, 2016 and 2016 (waves 1-3) and Gallup, 2009-2014.

*** p < .01, ** p < .05, * p < .1.

Table 10: Self-selection of refugees and irregular migrants, adults aged 25-64, FMS sample only

	(1)	(2)	(3)
Outcome: *reason to migrate: conflict or persecution*			
Sample ➔	All	Males	Females
Secondary education	0.067***	0.054***	0.121***
	(0.013)	(0.015)	(0.026)
Tertiary education	0.080***	0.087***	0.048*
	(0.017)	(0.019)	(0.023)
Employed	-0.027**	-0.020	-0.006
	(0.014)	(0.003)	(0.034)
Married	0.019	0.015	0.050
	(0.014)	(0.015)	(0.049)
Divorced	-0.009	0.026	0.013
	(0.042)	(0.057)	(0.072)
Widowed	0.040	0.068	0.031
	(0.043)	(0.069)	(0.070)
Country FE	Yes	Yes	Yes
r2	0.359	0.350	0.429
N	5473	4553	1920

Notes: Robust standard errors in parentheses. Outcome variable, *reason to migrate: conflict or persecution*, is equal to 1 for respondents who cite conflict or persecution as the main reason to migrate and 0 for other respondents who cite other reasons (economic reasons, limited access to amenities and natural disasters and other reasons) in the Flow Monitoring Surveys. Reference categories are as follows: less than secondary education, unemployed or out of labor force, and single. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3).
*** p < .01, ** p < .05, * p < .1.

Table 11: Self-selection of refugees and irregular migrants into Germany from major conflict countries, adults aged 25-64

	(1)	(2)	(3)
Sample ➔	Major conflict (both genders)	Major conflict (males only)	Major conflict (females only)
Secondary education	0.036***	0.033***	0.042***
	(0.003)	(0.003)	(0.004)
Tertiary education	0.089***	0.072***	0.126***
	(0.005)	(0.006)	(0.010)
Employed	0.023***	0.043***	0.001
	(0.003)	(0.003)	(0.005)
Male	-0.005*	--	--
	(0.003)		
Age 25-34	0.024***	0.014***	0.030***
	(0.004)	(0.005)	(0.005)
Age 35-44	0.028***	0.022***	0.030***
	(0.004)	(0.006)	(0.005)
Age 45-54	0.016***	0.013**	0.019***
	(0.004)	(0.006)	(0.005)
Married	0.019***	-0.004	0.053***
	(0.003)	(0.005)	(0.004)
Divorced	-0.010	-0.039***	0.032***
	(0.008)	(0.012)	(0.010)
Widowed	0.003	-0.069***	0.052***
	(0.006)	(0.008)	(0.007)
Country FE	Yes	Yes	Yes
r2	0.070	0.070	0.083
N	34180	18975	15205

Notes: Robust standard errors in parentheses. Outcome variable, *migrant*, is equal to 1 for respondents in the IAB-BAMF-SOEP Survey of Refugees and 0 for participants in Gallup World Polls. All specifications include origin country fixed effects. Countries are classified by the level of conflict following the definitions provided by the Uppsala Conflict Data Program: major conflict category includes countries with 1000 or more battle-related deaths in a given year over the sample period (Afghanistan, Iraq, Pakistan, Syria). Reference categories are as follows: less than secondary education, unemployed or out of labor force, female, age 54+, and single. Source: IAB-BAMF-SOEP Survey of Refugees in Germany, 2016 and Gallup World Polls, 2009-2014.
*** p < .01, ** p < .05, * p < .1.

Table 12: Self-selection of single migrants based on predicted income, adults aged 25-64

	(1)	(2)	(3)	(4)
Sample ➔	All countries	All countries	Major conflict countries	Minor or no conflict countries
Outcome ➔	Refugee	Irregular migrant	Migrant	Migrant
		Males		
Predicted log income	0.052***	-0.034**	0.019**	-0.015***
	(0.015)	(0.015)	(0.008)	(0.004)
r2	0.134	0.102	0.093	0.043
N	5845	5845	4648	3751
Country FE	Yes	Yes	Yes	Yes
		Females		
Predicted log income	0.043***	0.029***	0.044***	0.018***
	(0.015)	(0.012)	(0.008)	(0.003)
r2	0.087	0.050	0.036	0.014
N	2826	2826	2191	2036
Country FE	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. Outcome variable, *migrant*, is equal to 1 for respondents in the Flow Monitoring Surveys and 0 for participants in Gallup World Polls. Income variables are calculated using the latest purchasing power parity estimates available and estimated to the 2015 US dollar by Gallup World Polls. In columns 3 and 4, countries are classified by the level of conflict following the definitions provided by Uppsala Conflict Data Program: major conflict category includes countries with 1000 or more battle-related deaths in any of the years over the sample period (Afghanistan, Iraq, Nigeria, Pakistan, Sudan, Syria); minor conflict category includes countries with 25 to 999 battle-related casualties in any of the years over the sample period (Algeria, Iran); no conflict category includes countries that did not experience a major conflict or minor conflict in any of the years over the sample period (Bangladesh, Cameroon, Côte d'Ivoire, Morocco, Senegal). Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3). Gallup World Polls, 2009-2014.
*** p < .01, ** p < .05, * p < .1.

Table 13: Sorting of refugees and irregular migrants in respect to Gini index, adults aged 25-64

	(1)	(2)	(3)
Sample ➔	All	Major conflict	Minor or no conflict
Secondary education	-0.180**	-0.175**	-0.143
	(0.074)	(0.079)	(0.184)
Tertiary education	0.405***	0.081**	1.065***
	(0.124)	(0.033)	(0.267)
Employed	0.227***	0.118	0.557***
	(0.081)	(0.093)	(0.162)
Female	-0.014	0.021	-0.325
	(0.111)	(0.113)	(0.340)
Age 25-34	0.513	0.304	1.690
	(0.412)	(0.446)	(1.007)
Age 35-44	0.117	-0.037	1.231
	(0.413)	(0.446)	(1.024)
Age 45-54	0.151	-0.009	1.319
	(0.418)	(0.448)	(1.084)
Married	-0.142*	-0.200**	-0.046
	(0.084)	(0.092)	(0.187)
Divorced	-0.582*	-0.547	-0.104
	(0.316)	(0.370)	(0.737)
Widowed	-0.172	-0.226	-0.289
	(0.287)	(0.334)	(0.538)
Country FE	Yes	Yes	Yes
r2	0.306	0.394	0.113
N	3512	2650	862

Notes: Robust standard errors in parentheses. *Gini* coefficient of the intended destination country is the outcome variable and measured between 0 (no inequality), and 100 (perfect inequality). Countries are classified by the level of conflict following the definitions provided by Uppsala Conflict Data Program: major conflict category includes countries with 1000 or more battle-related deaths in any of the years over the sample period (Afghanistan, Iraq, Nigeria, Pakistan, Sudan, Syria); minor conflict category includes countries with 25 to 999 battle-related casualties in any of the years over the sample period (Algeria, Iran); no conflict category includes countries that did not experience a major conflict or minor conflict in any of the years over the sample period (Bangladesh, Cameroon, Côte d'Ivoire, Morocco, Senegal). Reference categories are as follows: less than secondary education, unemployed or out of labor force, female, age 54+, and single. Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3). World Development Indicators, 2016 or earliest available.
*** p < .01, ** p < .05, * p < .1.

Table 14: Sorting of refugees and irregular migrants in respect of characteristics of destination countries, adults aged 25-64

Outcome ➔	(1) Migrant integration policy index	(2) Average duration of the asylum procedure	(3) Waiting duration for labor market access	(4) Social expenditure (as a percentage of GDP)	(5) Unemployment rate (log)	(6) GDP per capita (log)
Secondary Education	-0.841***	0.011***	0.051***	-0.723***	0.084***	0.026**
	(0.265)	(0.003)	(0.013)	(0.137)	(0.018)	(0.012)
Tertiary Education	-0.844***	0.035***	0.072***	-0.687***	0.092***	0.044***
	(0.270)	(0.007)	(0.014)	(0.140)	(0.019)	(0.011)
Employed	0.020	0.002	0.004	-0.136	0.024*	-0.002
	(0.189)	(0.005)	(0.009)	(0.103)	(0.015)	(0.008)
Female	0.774***	0.020***	0.005	-0.384***	-0.042**	0.027**
	(0.270)	(0.007)	(0.012)	(0.132)	(0.017)	(0.011)
Age 25-34	-0.311	0.007	0.043	0.437	0.075*	-0.014
	(1.145)	(0.030)	(0.044)	(0.369)	(0.044)	(0.039)
Age 35-44	0.460	-0.002	0.011	0.290	0.005	-0.012
	(1.152)	(0.030)	(0.044)	(0.370)	(0.045)	(0.039)
Age 45-54	0.357	0.010	-0.008	-0.314	-0.031	-0.004
	(1.217)	(0.031)	(0.044)	(0.393)	(0.418)	(0.040)
Married	0.637***	-0.001	-0.016*	0.068	-0.012	0.003
	(0.190)	(0.005)	(0.009)	(0.104)	(0.014)	(0.008)
Divorced	0.221	-0.032*	0.004	0.389	0.037	0.006
	(0.676)	(0.019)	(0.029)	(0.373)	(0.045)	(0.025)
Widowed	0.867	-0.001	-0.017	-0.356	-0.050	0.037
	(1.008)	(0.027)	(0.034)	(0.452)	(0.056)	(0.037)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
r2	0.053	0.192	0.076	0.353	0.302	0.394
N	3509	3484	3509	3423	3492	3122

Notes: Robust standard errors in parentheses. Migrant integration policy index is a continuous variable (0-100, with 100 being the top score) and measures the country specific integration outcomes, integration policies, and other contextual factors for migrants' integration. Average duration of asylum procedure is the duration (number of months) between the submission of the asylum claim and the first decision (rescaled from 0 to 1, with 1 being the longest duration). Waiting duration for labor market access is the waiting period (number of months) for obtaining the work permit after claiming asylum (rescaled from 0 to 1, with 1 being the longest duration). Unemployment (log) is the unemployment rate in the intended destination country in the survey year. Social expenditure is measured as a percentage of GDP and comprises cash benefits, direct in-kind provision of goods and services, and tax breaks with social purposes. GDP per capita (log) is the gross domestic product per capita in current U.S. dollars in the intended destination country in the survey year. Reference categories are as follows: tertiary education, unemployed or our of labor force, female, age 54+, and single. Source: Eurofound, Flow Monitoring Surveys, 2015 and 2016 (waves 1-3), MIPEX, OECD, World Development Indicators.

*** p < .01, ** p < .05, * p < .1.

Table 15: Border policies and sorting of refugees and irregular migrants, all ages

Outcome ➔ *Intended destination country*	(1) **Germany**	(2) **France**	(3) **Italy**	(4) **Sweden**	(5) **Austria**	(6) **the Netherlands**
Austria quota announcement	0.160***	0.001	-0.003	-0.059***	0.016***	-0.024***
	(0.012)	(0.003)	(0.003)	(0.007)	(0.005)	(0.005)
Austria imposes quota	-0.004	0.006	0.012***	-0.015***	-0.025***	-0.006*
	(0.010)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
Hungary border closing	-0.114	0.005	0.015	-0.055	0.030***	0.015
	(0.102)	(0.007)	(0.016)	(0.085)	(0.009)	(0.042)
Slovenia and Macedonia border tightening	-0.165***	0.020***	0.012*	0.046***	0.010*	0.028***
	(0.012)	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)
Sweden border control	0.117***	-0.015**	0.005	-0.072***	-0.001	-0.017*
	(0.023)	(0.006)	(0.005)	(0.016)	(0.008)	(0.010)
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Employment status before migration	Yes	Yes	Yes	Yes	Yes	Yes
Origin Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey Country FE	Yes	Yes	Yes	Yes	Yes	Yes
r2	0.289	0.262	0.338	0.098	0.028	0.031
N	13983	13983	13983	13983	13983	13983

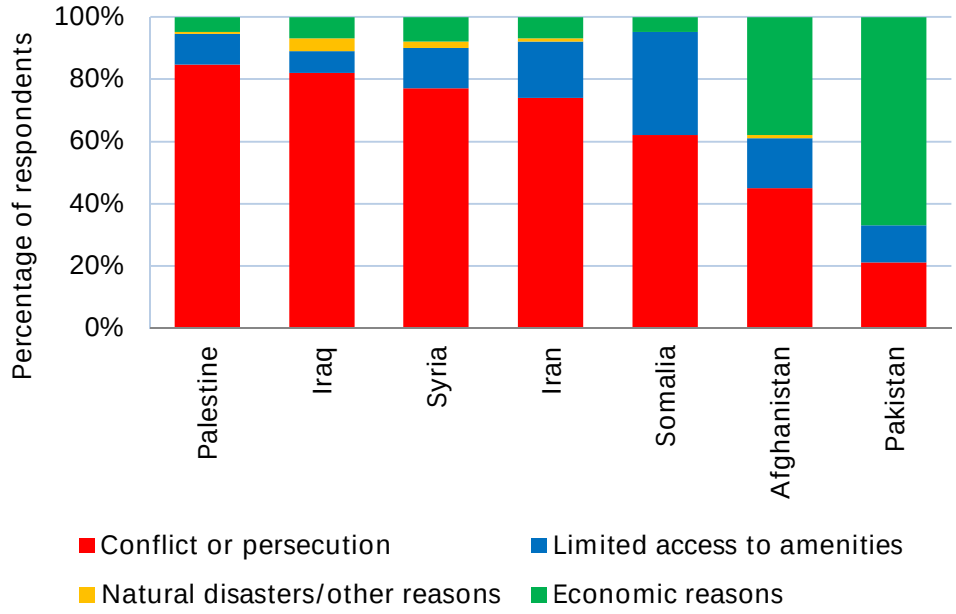
Notes: Robust standard errors in parentheses. *Intended destination country* is the outcome variable, which is equal to one if a refugee or irregular migrant names a particular country as their country of destination and zero otherwise. *Austria quota announcement* is equal to one for interview dates after Austria announced this quota on 20th January 2016. *Austria imposes quota* refers to a dummy variable that is equal to one, if interviews were conducted after 19th February 2016, when Austria imposed a quota of accepting maximum of 80 refugees or irregular migrants and a maximum of 3,200 people allowed traveling through Austria per day. *Hungary border closing* is equal to one if the interview took place after Hungary closed its border on 16th October 2015. *Slovenia and Macedonia border tightening* refers to the date on which Macedonia closed its border with Greece and Slovenia set stricter border controls and it is equal to one if the interview was conducted after 9th March 2016. *Sweden border control* is equal to one, if interviews took place after the 11th November 2015. For details on individual characteristics, see notes to Table 5. Source: Flow Monitoring Surveys, 2015 and 2016 – waves 1 and 2, migrants of different nationalities with at least 100 respondents.
*** p < .01, ** p < .05, * p < .1

Table 16: Self-selection of refugees and irregular migrants into Turkey, adults aged 25-64

Sample ➔	(1) Refugee	(2) Irregular migrant	(3) Refugee (males only)	(4) Irregular migrant (males only)	(5) Refugee (females only)	(6) Irregular migrant (females only)
Secondary Education	0.048***	-0.070***	0.061***	-0.023***	0.040***	-0.006
	(0.004)	(0.003)	(0.006)	(0.004)	(0.006)	(0.004)
Tertiary Education	0.032***	-0.014***	0.050***	-0.010**	0.037***	-0.002
	(0.006)	(0.005)	(0.008)	(0.003)	(0.009)	(0.006)
Employed	0.059***	0.015***	0.135***	0.033***	-0.026***	0.006
	(0.004)	(0.003)	(0.005)	(0.004)	(0.006)	(0.004)
Male	0.004	0.026***	--	--	--	--
	(0.004)	(0.003)				
Age 25-34	0.076***	0.045***	0.046***	0.053***	0.094***	0.042***
	(0.007)	(0.004)	(0.009)	(0.005)	(0.010)	(0.006)
Age 35-44	0.048***	0.039***	0.018**	0.030***	0.073***	0.036***
	(0.007)	(0.004)	(0.009)	(0.005)	(0.010)	(0.006)
Age 45-54	0.025***	0.017***	-0.001	0.013**	0.048***	0.023***
	(0.007)	(0.004)	(0.009)	(0.005)	(0.010)	(0.006)
Married	0.056***	-0.026***	0.010	-0.025***	0.110***	0.034***
	(0.006)	(0.004)	(0.008)	(0.006)	(0.008)	(0.005)
Divorced	0.013	0.074***	-0.119***	-0.032**	0.129***	0.081***
	(0.014)	(0.012)	(0.020)	(0.014)	(0.019)	(0.013)
Widowed	0.098***	0.024***	-0.125***	-0.045***	0.203***	0.081***
	(0.012)	(0.008)	(0.020)	(0.012)	(0.014)	(0.010)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
r2	0.111	0.063	0.126	0.036	0.125	0.028
N	36366	36366	19378	19378	16988	16988

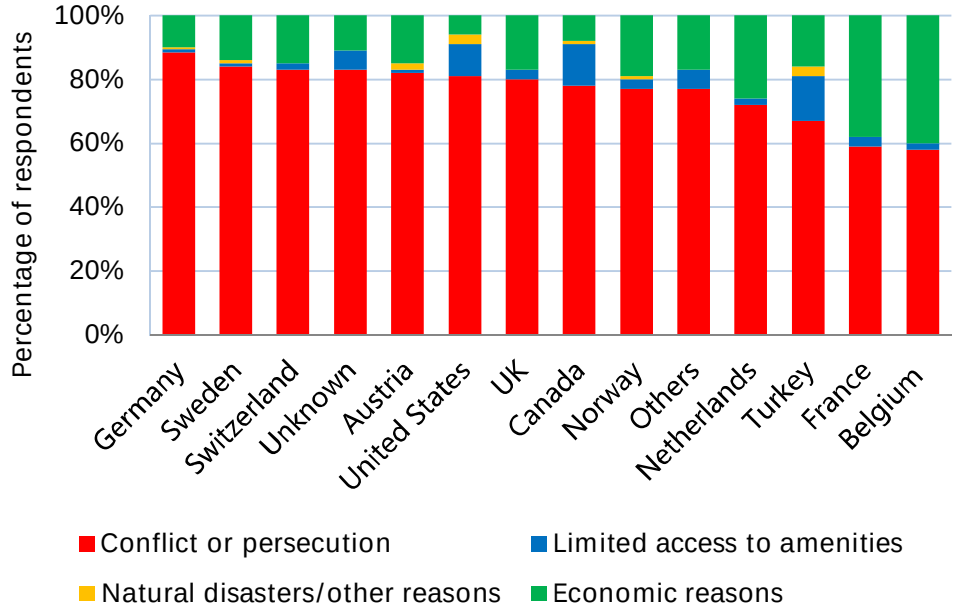
Notes: Robust standard errors in parentheses. Outcome variable, migrant, is equal to 1 for respondents in the Flow Monitoring Surveys and 0 for participants in Gallup World Polls. Reference categories are as follows: less than secondary education, unemployed or out of labor force, female, age 54+, and single. Source: Flow Monitoring Surveys (wave 4, Turkey only), 2016, 2017 and 2018. Gallup World Polls, 2009-2014.
*** p < .01, ** p < .05, * p < .1.

Figure A.1: Reasons to emigrate by origin country, among migrants who were interviewed in Turkey



Source: Flow Monitoring Surveys (wave 4, Turkey only, 2016-2018) and authors' calculations.

Figure A.2: Reasons to emigrate by intended destination country, among migrants who were interviewed in Turkey



Source: Flow Monitoring Surveys (wave 4, Turkey only, 2016-2018) and authors' calculations.

Table A.1: Descriptive characteristics from Gallup World Polls, 2009-2011

	(1) Full sample (all ages)	(1) Restricted sample (ages 25-64)
Age	33.20	38.87
	(12.49)	(10.51)
Male	0.51	0.51
	(0.50)	(0.50)
Married	0.58	0.73
	(0.49)	(0.44)
Divorced	0.02	0.03
	(0.14)	(0.18)
Widowed	0.03	0.04
	(0.19)	(0.20)
Secondary education	0.42	0.36
	(0.49)	(0.48)
Tertiary education	0.08	0.10
	(0.27)	(0.30)
Employed	0.49	0.56
	(0.50)	(0.50)
N	65688	46270

Source: Gallup World Polls, 2009-2011. Notes: Means (standard deviations). This table presents summary statistics for origin countries included in the analysis. Secondary and tertiary education refer to highest completed education.

Table A.2: Descriptive characteristics from Flow Monitoring Surveys, Turkey

Variables	(1) All ages	(2) Ages 25-64
Age	33.17 (11.61)	37.58 (10.23)
Male	0.61 (0.49)	0.59 (0.49)
Married	0.63 (0.48)	0.75 (0.43)
Divorced	0.03 (0.18)	0.04 (0.20)
Widowed	0.04 (0.20)	0.06 (0.23)
Secondary education	0.39 (0.48)	0.36 (0.48)
Tertiary education	0.14 (0.34)	0.17 (0.37)
Employed	0.51 (0.50)	0.59 (0.49)
Reasons for leaving:		
Conflict or persecution	0.69 (0.46)	0.68 (0.47)
Economic reasons	0.14 (0.34)	0.11 (0.32)
Limited access to amenities	0.11 (0.30)	0.10 (0.29)
Other reasons	0.06 (0.28)	0.11 (0.29)
Nationalities:		
Syria	0.33 (0.47)	0.33 (0.47)
Afghanistan	0.20 (0.40)	0.16 (0.36)
Iraq	0.27 (0.44)	0.29 (0.46)
Iran	0.16 (0.37)	0.18 (0.38)
Palestine	0.02 (0.12)	0.02 (0.08)
Somalia	0.02 (0.11)	0.02 (0.09)
Pakistan	0.01 (0.12)	0.01 (0.07)
N	12,203	9,047

Source: Flow Monitoring Surveys Turkey (wave 4), 11/2016 - 8/2018. Notes: Means (standard deviations). The sample sizes for some variables are different either due to missing data or because they were not asked in each wave. Secondary and tertiary education refer to highest completed education.

Table A.3: Descriptive characteristics from IAB-BAMF-SOEP Survey of Refugees, Germany

Variables	(1) Full sample (all ages)	(2) Full sample (all ages) males	(3) Full sample (all ages) females	(4) Restricted sample (ages 25-64)	(5) Restricted sample (ages 25-64) males	(6) Restricted sample (ages 25-64) females
Age	33.47 (10.37)	33.13 (10.59)	34.05 (9.97)	36.94 (8.56)	37.17 (8.67)	36.58 (8.37)
Male	0.63 (0.48)	--	--	0.60 (0.48)	--	--
Married	0.67 (0.47)	0.60 (0.49)	0.80 (0.40)	0.80 (0.40)	0.77 (0.42)	0.84 (0.36)
Divorced	0.02 (0.15)	0.02 (0.12)	0.04 (0.20)	0.03 (0.17)	0.02 (0.13)	0.05 (0.21)
Widowed	0.02 (0.15)	0.02 (0.04)	0.06 (0.23)	0.03 (0.16)	0.00 (0.05)	0.06 (0.24)
Secondary education	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.48 (0.50)	0.48 (0.50)	0.48 (0.50)
Tertiary education	0.27 (0.44)	0.26 (0.43)	0.30 (0.46)	0.28 (0.45)	0.27 (0.44)	0.30 (0.46)
Employed	0.62 (0.48)	0.79 (0.40)	0.35 (0.47)	0.68 (0.47)	0.87 (0.34)	0.39 (0.49)
Nationalities:						
Syria	0.58 (0.50)	0.59 (0.49)	0.55 (0.50)	0.58 (0.49)	0.60 (0.49)	0.56 (0.50)
Afghanistan	0.13 (0.34)	0.13 (0.34)	0.13 (0.34)	0.12 (0.33)	0.12 (0.33)	0.12 (0.32)
Iraq	0.16 (0.37)	0.16 (0.37)	0.16 (0.37)	0.16 (0.37)	0.16 (0.36)	0.16 (0.37)
Pakistan	0.02 (0.11)	0.02 (0.14)	0.00 (0.00)	0.02 (0.11)	0.02 (0.14)	0.00 (0.00)
N	2,657	1,675	982	2,043	1,244	799

Source: IAB-BAMF-SOEP Survey of Refugees, Germany, 2016. Notes: Means (standard deviations). Following the definition used in Gallup World Polls, secondary education refers to 9 to 15 years of education and tertiary education refers to completed four years of education beyond "secondary education".

Table A.4: Sub-regional conflict intensity and shares of emigration due to conflict or persecution

	Males		Females	
	(1) More intensive conflict sub-regions	(2) Less intensive conflict sub-regions	(3) More intensive conflict sub-regions	(4) Less intensive conflict sub-regions
Afghanistan	0.76	0.26	0.42	0.10
Algeria	0.09	0.05	0.02	0.03
Iran	0.61	0.08	0.40	0.07
Iraq	0.32	0.01	0.17	0.01
Nigeria	0.21	0.04	0.01	0.01
Pakistan	0.25	0.02	0.03	0.01
Sudan	0.81	0.21	0.72	0.19
Syria	0.61	0.08	0.40	0.07

Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3) and Uppsala Conflict Database.

Table A.5: Cost of migrants' journey

	(1) Wave 1 (10/2015 - 12/2015)	(2) Wave 2 (01/2016 - 11/2016)	(3) Wave 3 (06/2016 - 11/2016)	(4) Wave 4 (11/2016 - 8/2018)
		Refugees		
Less than 1000 Euros	0.09	0.07	0.20	0.57
Between 1000 and 5000 Euros	0.79	0.76	0.63	0.40
More than 5000 Euros	0.12	0.17	0.17	0.03
		Irregular migrants		
Less than 1000 Euros	0.09	0.05	0.26	0.51
Between 1000 and 5000 Euros	0.77	0.78	0.61	0.45
More than 5000 Euros	0.13	0.17	0.13	0.04

Source: Flow Monitoring Surveys (all waves). The first wave conducted interviews in Croatia, Greece, Slovenia, and the Former Yugoslav Republic of Macedonia. The second wave covers Bulgaria, Croatia, Greece, Hungary, the Former Yugoslav Republic of Macedonia, Serbia, and Slovenia. The third wave of the survey only covers Italy and the fourth wave only covers Turkey. Notes: Means. This table presents summary statistics by survey wave for the cost of migrants' journey between their origin country and country of interview.

Table A.6: Self-selection of refugees and irregular migrants, adults aged 25-64, GWP 2009-2011

Outcome ➔	(1) Migrant due to any reason	(2) Refugee	(3) Irregular migrant
Secondary education	0.044***	0.044***	-0.015***
	(0.004)	(0.003)	(0.002)
Tertiary education	0.088***	0.076***	0.027***
	(0.007)	(0.006)	(0.004)
Employed	-0.004	-0.006*	0.004**
	(0.004)	(0.003)	(0.002)
Male	0.106***	0.069***	0.024***
	(0.003)	(0.003)	(0.002)
Age 25-34	0.153***	0.010***	0.046***
	(0.004)	(0.004)	(0.002)
Age 35-44	0.069***	0.049***	0.020***
	(0.004)	(0.003)	(0.001)
Age 45-54	0.025***	0.018***	0.006***
	(0.004)	(0.003)	(0.001)
Married	-0.049***	-0.009*	-0.019***
	(0.005)	(0.004)	(0.003)
Divorced	-0.040***	-0.040***	0.018**
	(0.011)	(0.008)	(0.008)
Widowed	-0.008	0.008	-0.004
	(0.009)	(0.008)	(0.004)
Country FE	Yes	Yes	Yes
r2	0.160	0.130	0.068
N	29401	29401	29401

Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3). Gallup World Polls, 2009-2011. Notes: Robust standard errors in parentheses. *** p < .01, ** p < .05, * p < .1. Outcome variable, *migrant*, is equal to 1 for respondents in the Flow Monitoring Surveys and 0 for participants in Gallup World Polls. Reference categories are as follows: less than secondary education, unemployed or out of labor force, female, age 54+, and single.

Table A.7: Self-selection of refugees and irregular migrants, adults aged 25-64, males, GWP 2009-2011

Outcome ➔	(1) Migrant due to any reason	(2) Refugee	(3) Irregular migrant
Secondary education	0.039***	0.040***	-0.028***
	(0.006)	(0.005)	(0.003)
Tertiary education	0.081***	0.079***	0.006
	(0.009)	(0.009)	(0.005)
Employed	0.017**	0.026***	-0.006
	(0.006)	(0.006)	(0.004)
Age 25-34	0.209***	0.132***	0.063***
	(0.007)	(0.006)	(0.003)
Age 35-44	0.095***	0.061***	0.037***
	(0.006)	(0.006)	(0.003)
Age 45-54	0.033***	0.023***	0.012***
	(0.006)	(0.005)	(0.002)
Married	-0.088***	-0.036***	-0.025**
	(0.008)	(0.007)	(0.004)
Divorced	-0.104***	-0.093***	0.013
	(0.019)	(0.015)	(0.013)
Widowed	-0.104***	-0.074***	-0.009
	(0.020)	(0.019)	(0.008)
Country FE	Yes	Yes	Yes
r2	0.170	0.139	0.032
N	16152	16152	16152

Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3). Gallup World Polls, 2009-2011. Notes: Robust standard errors in parentheses. *** p < .01, ** p < .05, * p < .1. Outcome variable, *migrant*, is equal to 1 for respondents in the Flow Monitoring Surveys and 0 for participants in Gallup World Polls. Reference categories are as follows: less than secondary education, unemployed or out of labor force, female, age 54+, and single.

Table A.8: Self-selection of refugees and irregular migrants, adults aged 25-64, females, GWP 2009-2011

Outcome ➔	(1) Migrant due to any reason	(2) Refugee	(3) Irregular migrant
Secondary education	0.048***	0.052***	-0.004*
	(0.005)	(0.004)	(0.002)
Tertiary education	0.108***	0.080***	0.060***
	(0.010)	(0.008)	(0.007)
Employed	-0.029***	-0.034***	0.005*
	(0.004)	(0.003)	(0.002)
Age 25-34	0.072***	0.051***	0.025***
	(0.005)	(0.005)	(0.002)
Age 35-44	0.030***	0.025***	0.002
	(0.005)	(0.005)	(0.002)
Age 45-54	0.008	0.006	-0.000
	(0.005)	(0.005)	(0.001)
Married	0.048***	0.048***	0.002
	(0.005)	(0.004)	(0.003)
Divorced	0.056***	0.026	0.033***
	(0.013)	(0.010)	(0.010)
Widowed	0.066***	0.059***	0.008*
	(0.009)	(0.008)	(0.004)
Country FE	Yes	Yes	Yes
r2	0.121	0.112	0.020
N	13249	13249	13249

Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3). Gallup World Polls, 2009-2011. Notes: Robust standard errors in parentheses. *** p < .01, ** p < .05, * p < .1. Outcome variable, *migrant*, is equal to 1 for respondents in the Flow Monitoring Surveys and 0 for participants in Gallup World Polls. Reference categories are as follows: less than secondary education, unemployed or out of labor force, female, age 54+, and single.

Table A.9: Self-selection of refugees and irregular migrants, GWP 2009-2011

	(1)	(2)	(3)
Sample ➔	All	Major conflict	Minor or no conflict
Secondary education	0.044***	0.076***	-0.015***
	(0.004)	(0.006)	(0.005)
Tertiary education	0.089***	0.096***	0.077***
	(0.007)	(0.009)	(0.014)
Employed	-0.004	-0.008	0.003
	(0.004)	(0.005)	(0.007)
Male	0.107***	0.113***	0.093***
	(0.004)	(0.005)	(0.006)
Age 25-34	0.153***	0.166***	0.130***
	(0.005)	(0.007)	(0.006)
Age 35-44	0.069***	0.088***	0.032***
	(0.004)	(0.006)	(0.005)
Age 45-54	0.025***	0.036***	0.009**
	(0.004)	(0.006)	(0.004)
Married	-0.049***	-0.037***	-0.080***
	(0.006)	(0.007)	(0.008)
Divorced	-0.041***	-0.076***	0.021
	(0.011)	(0.014)	(0.019)
Widowed	-0.008	-0.010	-0.017*
	(0.009)	(0.013)	(0.010)
Country FE	Yes	Yes	Yes
r2	0.161	0.134	0.234
N	29401	19023	10378

Source: Flow Monitoring Surveys, 2015 and 2016 (waves 1-3). Gallup World Polls, 2009-2011. Notes: Robust standard errors in parentheses. *** p < .01, ** p < .05, * p < .1. Outcome variable, *refugee/migrant*, is equal to 1 for respondents in the Flow Monitoring Surveys and 0 for participants in Gallup World Polls (adults aged 25-64). All specifications include source country fixed effects. Countries are classified by the level of conflict following the definitions provided by Uppsala Conflict Data Program: major conflict category includes countries with 1000 or more battle-related deaths in a given year over the sample period (Afghanistan, Iraq, Nigeria, Pakistan, Sudan, Syria); minor conflict category includes countries with 25 to 999 battle-related casualties in a given year over the sample period (Algeria, Iran); no conflict category includes countries that did not experience a major conflict or minor conflict over the sample period (Bangladesh, Cameroon, Côte d'Ivoire, Morocco, Senegal). Reference categories are as follows: less than secondary education, unemployed or out of labor force, female, age 54+, and single.