I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

The Labor Market Integration of Refugees to the United States: Do Entrepreneurs in the Network Help?

Olivier Dagnelie
Anna Maria Mayda
Jean-François Maystadt

DISCUSSION PAPER SERIES

IZA DP No. 11343

The Labor Market Integration of Refugees to the United States: Do Entrepreneurs in the Network Help?

Olivier Dagnelie
CREM, Universite de Caen Normandie

Anna Maria Mayda
SFS, Georgetown University and IZA

Jean-François Maystadt
Lancaster University Management School and LICOS

FEBRUARY 2018

ABSTRACT

The Labor Market Integration of Refugees to the United States: Do Entrepreneurs in the Network Help?*

We investigate whether entrepreneurs in the network of refugees - from the same country of origin - help refugees' labor-market integration by hiring them in their businesses. We analyze the *universe* of refugee cases without U.S. ties who were resettled in the United States between 2005 and 2010. We address threats to identification due to sorting of refugees into specific labor markets and to strategic placement by resettlement agencies. We find that the probability that refugees are employed 90 days after arrival is positively affected by the number of business owners in their network, but negatively affected by the number of those who are employees. This suggests that network members who are entrepreneurs hire refugees in their business, while network members working as employees compete with them, consistent with refugees complementing the former and substituting for the latter.

Corresponding author:
Anna Maria Mayda
Georgetown University
Department of Economics
37th and O Streets, NW
Washington DC, 20057
USA

E-mail: amm223@georgetown.edu

"An immigrant himself, Chobani yogurt founder [Hamdi Ulukaya] becomes icon for refugees. ... Despite warnings against hiring refugees, Ulukaya has made executive decisions to offer employment to people who have fled from hunger, persecution and fear."(Al Monitor, October 5, 2015)

1   Introduction

Since its creation with the 1980 Refugee Act, the U.S. Refugee Admissions Program (USRAP) has resettled in the United States more than 3 million individuals fleeing persecution, war or violence in their countries of origin. One of the main goals of the program is the successful labor market integration of refugees into the local communities where they are placed.[1] The drivers of refugees' (and in general immigrants') labor market integration are also the focus of academic interest. The existing literature shows that, among the many factors affecting the labor market assimilation of foreign-born workers, social networks are especially important. Social networks are broadly defined as the group of migrants from the same country of origin or community as foreign workers. They are believed to provide information on labor market conditions and opportunities to recent refugees and immigrants, as well as job referrals to firms about foreign workers (Munshi, 2003; Beaman, 2012).[2]

In this paper we investigate a different (but related) channel through which network members might help foreign-born workers in the labor market. Refugees and migrants may face discrimination in the labor market. Hence, even with additional information from network members, they may not be able to find a job. Consistent with recent research and motivated by anecdotal evidence, we explore the role of entrepreneurs within the network in facilitating the labor-market integration of refugees. First we note that, according to recent work, entrepreneurship is high among foreign-born workers: Kerr and Kerr (2017) finds that about 27 percent of immigrants in the U.S. were entrepreneurs

---

[1]U.S. government agencies closely monitor the economic assimilation of other types of immigrants as well. See for example https://www.dhs.gov/blog/2015/12/16/keeping-american-dream-alive.

[2]The migration literature has analyzed other mechanisms, besides information transmission, through which networks impact newly arrived foreign-born individuals, as discussed in Dolphin and Genicot (2010), Munshi (2014a), and Munshi (2014b). One of them is the role of networks to provide credit. Networks also mitigate the impact of shocks in the countries of origin by facilitating migration flows and integration in the U.S.(Clemens, 2017; Mahajan and Yang, 2017).

in 2008 (24 percent during 1995-2008)). Anecdotal evidence also suggests a link between migrant entrepreneurs and access to the labor market by foreign workers. According to a recent piece on National Public Radio (NPR), Belgian Turks fare better in the Belgian labor market than Belgian Moroccans because they are helped (hired) by entrepreneurs in their network (NPR 2016). A similar story made newspapers headlines lately – that the CEO of Greek-yogurt Chobani "fills" his plants with refugees. This is the channel we explore in this paper. Specifically, we investigate whether network members who are entrepreneurs help refugees' labor-market integration by hiring them in their businesses.[3] To our knowledge, this channel has not been previously analyzed in the network and migration literature.[4]

U.S. State Department is responsible for the initial placement and resettlement of refugees in the United States. In practice, U.S. State Department enters into agreements with various resettlement agencies, which provide reception and placement services for the first three months of refugees' stay. Broadly speaking, refugees resettled in the United States are of two types: refugees "with U.S. ties" and those "without U.S. ties." Refugees who report to have a U.S. tie are likely placed in the same geographic location as their family or friends living in the U.S.. Cases without U.S. ties are those of refugees with no family or friends in the United States. In this paper we focus on the latter group and analyze the *universe* of refugee cases without U.S. ties who were resettled in the United States between 2005 and 2010. Importantly, refugees without U.S. ties do not decide where they go upon arrival to the United States. Their placement decisions are taken by resettlement agencies, as a function of individual characteristics we observe in the data and control for in the analysis. Importantly, note that no employee of the resettlement agencies meets the refugee before the placement decision has been made. This implies that our results cannot be driven by refugees sorting into specific labor markets nor by resettlement agencies placing them according to unobserved individual characteristics. At the same time, the resettlement agencies do take into account the characteristics of the location where they might place refugees – for example, the availability of local programs and communities

---

[3]There may be different reasons why business owners in the network want to hire immigrants and refugees from the same country of origin, one of them being greater information about these workers.

[4]We should note that, in the broader literature, there is evidence consistent with our hypothesis. For example, as discussed in Section 2, some papers show that migrant managers are more likely to hire workers from the same ethnic background.

able to meet the refugees' needs. This might be an issue for identification since unobserved (by the econometrician) characteristics of the community of placement could be correlated with both our main regressors (the overall size and number of business owners in the network) and the dependent variable (the labor market integration of refugees). For example, if the returns to the skills of Iranian workers are especially high in Los Angeles, both our regressors and the dependent variable are likely to be high as well. We address these threats to identification by including a full set of fixed effects and by exploiting the practical features of the U.S. refugee resettlement program (as in Beaman, 2012). Hence our results are not driven by strategic placement of resettlement agencies as a function of location characteristics.

We measure labor market integration with a variable indicating whether the refugee is employed 90 days after arrival. We define the social network as the community of refugees from the same country of origin as the incoming refugee and living in the same commuting zone (CZ) where the refugee is placed upon arrival and is currently observed. In addition, using the American Community Survey, we can measure the fraction of, respectively, business owners (self-employed) and employees in each network at the beginning of the period of analysis. We find that, the greater the number of business owners in the network of the refugee, the higher the probability that the refugee is employed 90 days after arrival. At the same time, the greater the number within the network of those who are employees, the lower the probability that the refugee is employed 90 days after arrival. These results are consistent with the hypothesis that network members owning their business hire refugees, while network members working as employees compete with refugees. In other words, refugees "complement" network members who own businesses and "substitute" for network members who work as employees. Our findings indicate that at the mean, doubling the number of business owners in the refugee's network raises the probability that the refugee is employed by about 1.3 percentage points; similarly, at the mean, doubling the number of employees in the refugee's network decreases the probability that the refugee is employed by about 4.9 percentage points. To put these marginal effects into perspective, the greater number of business owners in the network implies that a refugee from Vietnam (who finds on average 244 entrepreneurs in the network in the CZ of first placement) has 4 additional percentage points probability to be employed compared to a refugee from

Liberia (who finds on average only 1 entrepreneur). At the same time, the higher number of employees in the network implies that a refugee from Vietnam (who finds on average 1,465 employees in the network in the CZ of first placement) has 10 percentage points lower probability to be employed compared to a refugee from Liberia (who finds on average only 210 employees). The results are both statistically and economically significant and are robust to using alternative specifications. We discuss and rule out alternative interpretations of our findings. An important policy implication can be derived from our results. Policymakers will be able to achieve two goals at once: by providing business incentives and opportunities to tenured refugees and migrants, they can help the latter as well as just-arrived refugees.

Our paper contributes to the literature for two main reasons. First, as discussed at length above, it provides robust evidence about the causal impact of networks – on the labor market success of foreign-born workers – through a new channel. Second, very few papers in the academic literature analyze the labor market outcomes of refugees resettled within the United States, most likely due to unavailability of data.[5] We are the first to observe directly and analyze a very large number of refugee cases resettled in the United States (indeed, as mentioned above, the *universe* of refugees without U.S. ties who arrived between 2005 and 2010). Using these data we provide the first large-scale evidence on the fraction of refugees, resettled in the United States, who find a job in the short run and on the individual-level characteristics that are correlated with finding a job. We show that almost a third of refugees (without U.S. ties) are employed at 90 days after arrival. We also find that the probability that a refugee is employed after three months from arrival is positively correlated with the refugee's level of education, negatively correlated with most "support" variables – denoting for example if the applicant has a government cash, medical, etc. assistance source of support – and follows an inverse U-shaped pattern over the life cycle. Controlling for the level of education, age and other individual-level controls, we also find that the greater the number of household members who accompany the refugee, the lower the probability that the refugee is employed after 90 days from

---

[5]To our knowledge, the only two academic papers which shed light on this topic are Cortes (2004) and Beaman (2012).

arrival.[6]

The remainder of the paper is organized as follows. Section 2 explains how our paper fits in the existing literature. Section 3 gives information on the data and shows the summary statistics. We explain the identification strategy in Section 4 and introduce the empirical specification in Section 5. In Section 6 we present the main empirical results while in Section 7 we discuss a number of extensions and robustness tests. Finally Section 8 concludes.

2 Literature Review

Our contribution lies at the intersection of three strands of the literature which analyze, respectively: the impact of social networks on the labor market outcomes of foreign workers, the role of migrant entrepreneurs in the economy, and the labor market integration of migrants.

As mentioned in the Introduction, social networks can play an informational role and facilitate the labor market assimilation of migrants and refugees. One channel through which networks reduce asymmetric information problems is by providing information on the quality of potential workers through job referrals, which enables managers to attain a better match between workers and firms (Montgomery, 1991). In addition the new hires will have an incentive to work hard as to avoid social sanctions from the network in case of misbehaviour (Munshi, 2014b). Networks can also provide information on labor market opportunities to agents in search of a job (Calvó-Armengol and Jackson, 2004). The informational role of social networks for both sides of the market has been abundantly investigated in the migration literature, for example by Munshi (2003) and Beaman (2012).[7] In his seminal paper, Munshi (2003) shows that the labor market outcomes of Mexican migrants improve the greater the size of the network – from their community of origin back in Mexico – and suggests that networks help migrants by providing job referrals in a situation in which information about jobs

---

[6]The relationship between household size and employment is theoretically ambiguous but, in this context, a negative sign is expected since it is likely to be a proxy for poverty.

[7]In this Section we focus on papers that analyze U.S. data. However note that similar results are found for other countries. Edin et al. (2003) and Damm (2009) find that migrant networks, defined as ethnic enclaves, improve the labor market outcomes of immigrants in Sweden (for the less skilled) and Denmark. The authors of both papers provide an interpretation of their results based on the informational role of networks.

is not perfect. The estimates in the paper control for individual fixed effects and show that Mexican immigrants to the United States are more likely to be employed and earn a higher wage when the size of their network is exogenously larger.

Beaman (2012) refines Munshi's idea and further explores the informational role of networks. She builds a dynamic model of social networks with multiple cohorts and finds that the vintage of the network is a key determinant of the economic assimilation of refugees. Beaman (2012) focuses on the labor market integration of newly-arrived refugees who just came to the United States (90 days before). She analyzes the role in this process played by respectively "recent" refugees, who came to the United States in the previous two years, and "tenured" refugees, who have been in the country for more than two years. She finds that an increase in the number of recent refugees worsens the labor market outcomes of newly-arrived refugees while an increase in the number of tenured refugees improves them. Beaman interprets the results as consistent with a job information story according to which tenured refugees provide information to newly-arrived refugees while recent refugees compete with them for information. Beaman takes her model to the data and she is the first in the literature to take advantage of the exogenous variation in the allocation of refugees without U.S. ties in the United States. Our identification strategy is very closely related to Beaman's one. However, we analyze the entire universe of refugees without U.S. ties while Beaman looks at cases resettled by only one of the resettlement agencies. In addition, the channels of investigation are clearly different although related: contrary to the common belief that networks are usually beneficial for new arrivals, both our paper and Beaman's analysis point out that the effect of networks is highly non linear, often positive but other times negative.

There is evidence in the literature that social networks not only affect the integration of foreign workers in the labor market but also the type of job they end up in. Patel and Vella (2012) uncover strong evidence of network effects in the occupational choice of immigrants to the United States. They find that newly arrived immigrants follow their compatriots, from previous waves of immigration, and tend to find employment in the same occupations, with a stronger effect for low-skilled migrants. Their analysis suggests that this result is not driven by skill-based sorting, rather by sharing of information about job opportunities through networks. Patel and Vella also highlight that the migrants following

the occupational choice of their predecessors benefit from a large wage premium. Similar to Patel and Vella (2012), occupational concentration of migrants by nationality is observed by Andersson et al. (2014) who find that immigrants are particularly likely to work with compatriots and are also more likely, than natives, to work with immigrants from other countries – although only a small share works in immigrant-only workplaces. Again the interpretation is that social networks play a role in shaping the observed occupational distribution.

Another strand of the literature emphasizes the role of migrant entrepreneurs in the economy. Occupational concentration among migrants seems particularly prevalent in self-employed activities – which are at the core of our hypothesis. Fairlie and Lofstrom (2013) document that immigrant entrepreneurship has grown steadily over the last decades. The authors show that migrant business-owners are concentrated in a few states – California, Florida, NY and Texas representing approximately 50% of the total – and industry sectors – Construction, Professional Services, Other Services, Retail Trade, Health Care and Social Assistance, Accommodation, Recreation and Entertainment. Kerr and Mandorff (2015) seek to explain the high prevalence of self-employment among migrants. They develop a theoretical model showing how social interactions lead to occupational stratification along ethnic lines. They conjecture that members of small social networks develop business-specific skills through informal exchanges of information on their business activities and posit that these social interactions are complementary to production.[8]

Consistent with our findings, Åslund et al. (2014) point out that, in the Swedish labor market, migrant managers are more likely to hire workers from the same ethnic background[9] and to offer them higher wages. This earnings premium disappears once unobserved worker characteristics are taken into account, suggesting that managers are able to detect or recruit better workers when they come from the same country of origin as these workers. Åslund et al. (2014) offer suggestive evidence that this is the result of profit maximizing behavior and not discrimination or preferences. They propose several explanations for this observation: sharing a language or business culture could

---

[8]We should note that Kerr and Mandorff (2015) broadly explain how social networks affect occupational specialization among ethnic groups. However, their two-sector model is applicable to a choice between self-employment and employment among migrants.

[9]Note that Bandiera et al. (2009) and Giuliano, Levine, and Leonard (2009, 2011) come to similar conclusions, in causal studies.

enhance productivity, lowering transaction and communication costs. Managers may also experience less noise in productivity signals from workers with a similar background. A related reason is that job-search networks may provide useful information lowering the cost of information acquisition.

Finally, our analysis touches upon the literature seeking to identify the main factors of assimilation among migrants. Several studies have stressed the importance of language skills and education for the labor market assimilation of migrants (Chiswick, 1991; Borjas, 1994; Chiswick and Miller, 1995). Similar results have been found for Europe (Dustmann and van Soet, 2012; Dustmann and Fabbri, 2003).

3 Data and Summary Statistics

We exploit variation in individual refugees' labor-market integration as a function of individual-level covariates as well as aggregate variables which vary across U.S. commuting zones (CZs), years of arrival and nationalities of origin of refugees. Our analysis is based on data from several sources. We use highly-confidential individual-level administrative data from the Worldwide Refugee Admissions Processing System (WRAPS) data set, housed at the Refugee Processing Center (RPC) – RPC is part of the Bureau of Population, Refugees, and Migration (PRM) at the U.S. State Department.[10] WRAPS contains detailed individual-level information about the universe of refugees resettled to the United States from 1990 to the present. Given that we observe the universe of refugees, there is no sampling error in our data. For every refugee resettled to the United States we know the year of arrival, the city and state of placement as well as individual characteristics such as the age, gender, marital status, level of education, number of household members in the application, country of origin and type of refugee – i.e., whether the refugee is with or without U.S. ties. For the analysis in this paper, for methodological reasons, we restrict the attention to refugees without U.S. ties. Importantly, together with arrival information, WRAPS also includes information on the labor-market outcomes of refugees (specifically their employment status) at 90 days after arrival. This information comes from follow-up interviews of the refugees which we can access only for refugees resettled from 2005

---

[10]Refugee records in WRAPS are protected under Section 222(f) of the Immigration and Nationality Act, 8 U.S.C. §1202(f), and may be subject to the Privacy Act of 1974, as amended, 5 U.S.C. §552a.

on, hence the period of our analysis is 2005-2010. Given the timing of the follow-up interviews, at 90 days after arrival, our focus is on refugees' integration in the labor market *in the short run*.

Finally, using the WRAPS data, we can also measure each refugee's social network which we define as the community of refugees without U.S. ties, arrived up to one year before (since 1990), from the same origin country as the refugee, living in the same CZ where the refugee is placed upon arrival and is currently observed. We merge these data with the U.S. American Community Survey (ACS). We use the ACS to measure the fraction of business owners (self-employed) and of employees for each nationality and CZ at the beginning of the period of analysis. Finally, we utilize data from the 1990 and 2000 U.S. Census to measure other characteristics of each nationality.

Figure 1: Number of refugees' arrivals between 2005 and 2010, by refugee type (refugees with or without U.S. ties)
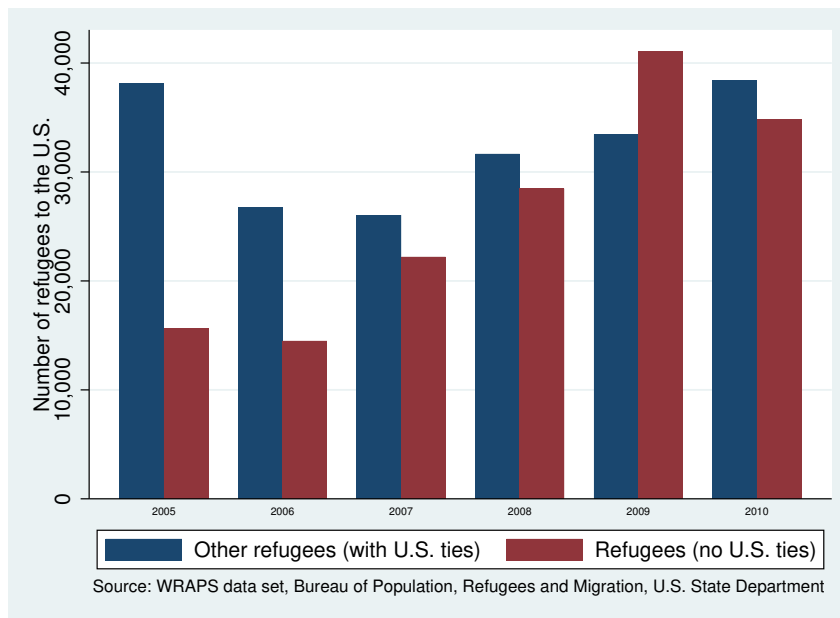
Note: Number of refugees between 2005 and 2010.

Source: Worldwide Refugee Admissions Processing System (WRAPS).

Figure 2: Size of networks of refugees (without U.S. ties) by nationality and U.S. state, in 2010
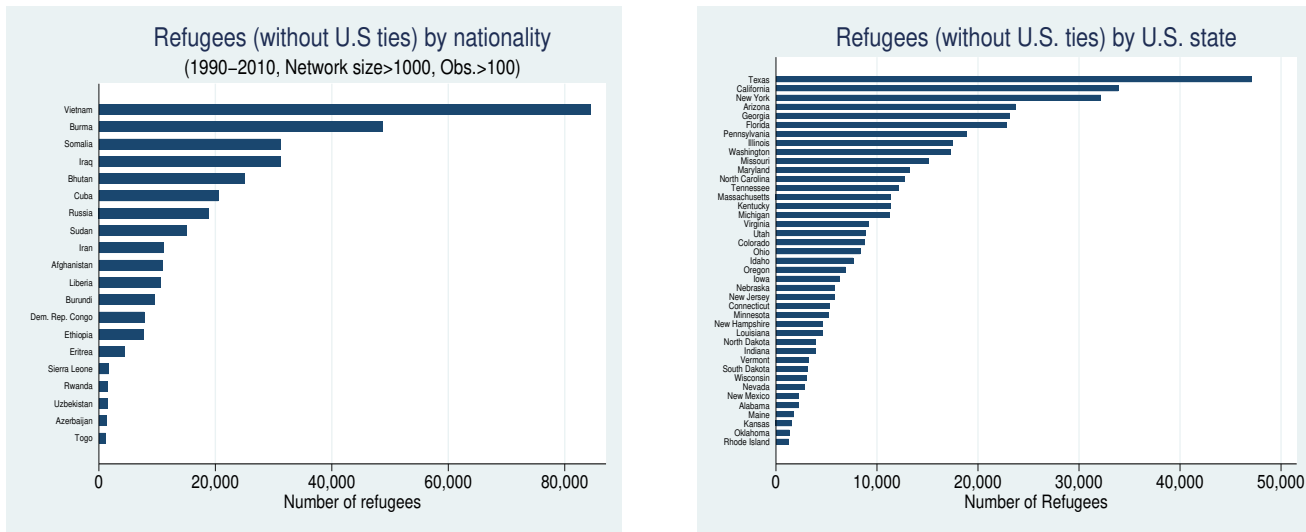
Note: Networks are constructed by summing the refugees arrivals between 1990 and 2010. For presentation purposes, only networks with size greater than 1,000 are considered. Only nationalities with more than 100 individual observations are shown.

Source: WRAPS data set, Bureau of Population, Refugees and Migration, U.S. State Department.

Figure 3: Size of networks of refugees (without U.S. ties) across commuting zones within the United States, in 2010
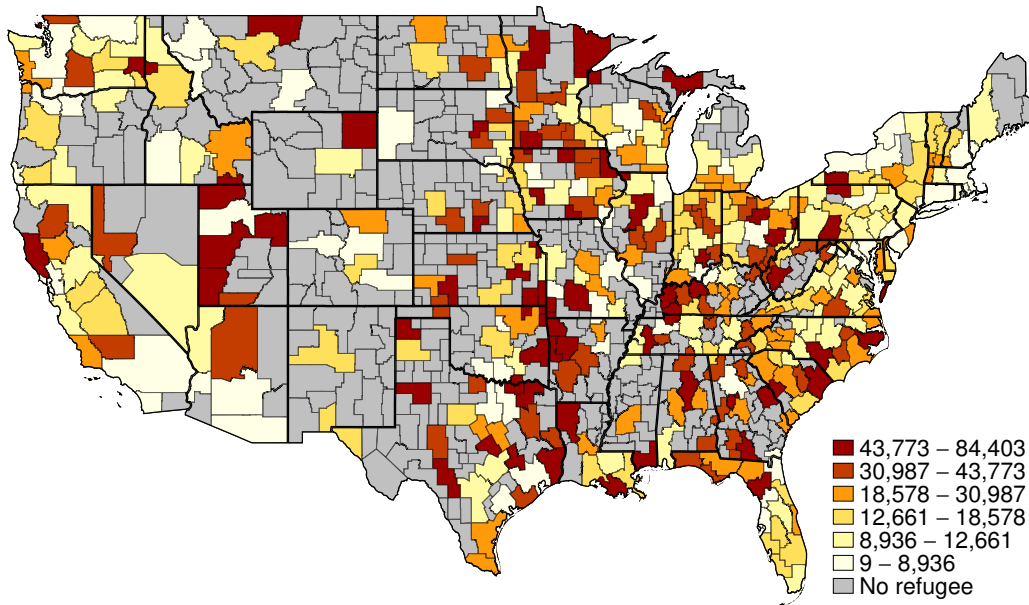
Source: Worldwide Refugee Admissions Processing System (WRAPS).

The U.S. welcomed about 15,000 refugees without U.S. ties in, respectively, 2005 and 2006. In

later years, the number increased up to around 41,000 in 2009 and 35,000 in 2010. Importantly, refugees without U.S. ties represent a substantial fraction of the overall number of refugees resettled to the United States (see Figure 1, which shows the number of refugees, broken down by refugee type, i.e. with or without U.S. ties).[11] In 2005, refugees without U.S. ties represented around 30 percent of total refugees but in every other year within our period of analysis this fraction was higher, reaching 55% in 2009.[12] Figure 2 shows the size of refugees' networks (without U.S. ties) in 2010 by, respectively, country of origin and U.S. state.[13] These numbers are constructed by taking the sum of refugees' arrivals in each year between 1990 and 2010. The top 5 nationalities of refugees (without U.S. ties) are, respectively, Vietnam, Burma, Somalia, Iraq and Bhutan. Interestingly, the U.S. states with the largest networks of refugees without U.S. ties in 2010 are, respectively, Texas, California, New York state, Arizona, Georgia and Florida. Finally, Figure 3 shows the geographical distribution of refugees (without U.S. ties) across CZs within the United States in 2010.[14]

Figure 4: Employment rates of refugees (without U.S. ties) by nationality and U.S state, in 2005-2010
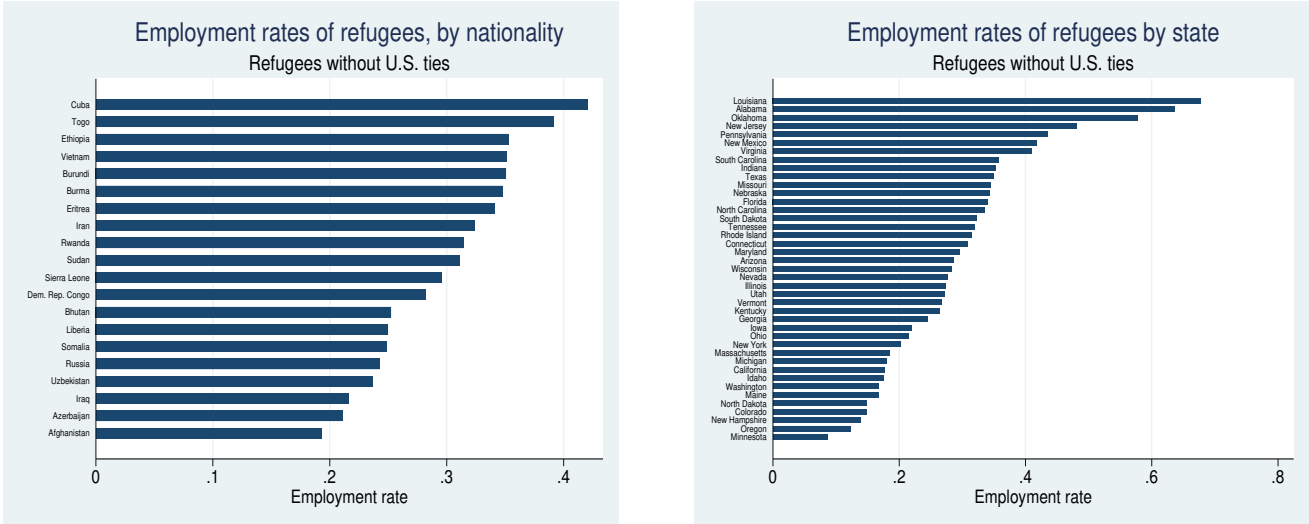
Note: Employment rates are defined as the share of employed people among refugees without U.S. ties aged between 18 and 64. For presentation purposes, only nationalities with more than 100 individual observations and a network size above 1,000 members are considered.

Source: WRAPS data set, Bureau of Population, Refugees and Migration, U.S. State Department.

---

[11]Figures 1 through 4 are based on the universe of refugees without U.S. ties.

[12]The fractions were: 29% in 2005, 35% in 2006, 46% in 2007, 47% in 2008, 55% in 2009 and, finally, 47% in 2010.

[13]From this point on, we always refer to refugees *without U.S. ties*, unless explicitly noted.

[14]Figure B.1 provides the same geographical distribution across CZs but is restricted to CZs for which there were arrivals in 2005-2010, i.e. CZs corresponding to the sample we analyze in the regressions.

Figure 5: Employment rates by U.S. state, in 2005-2010
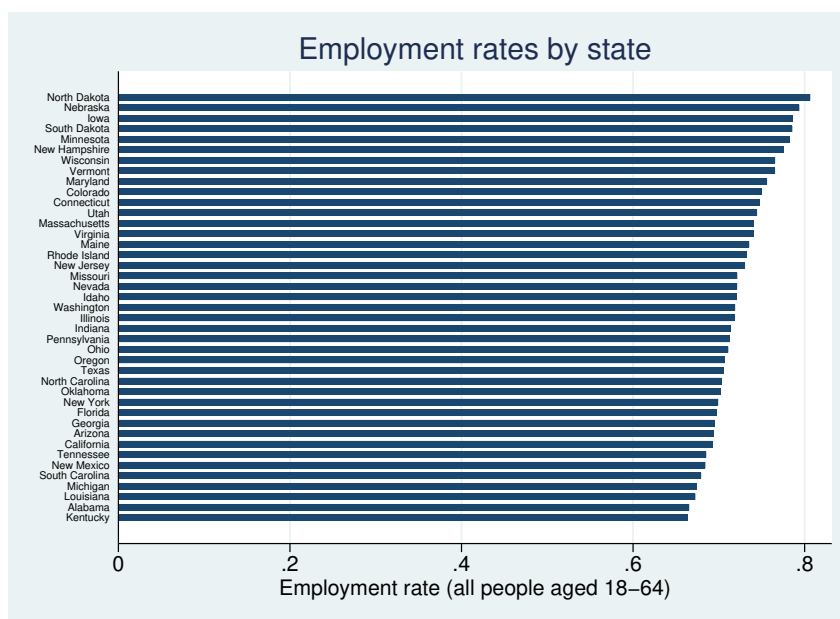
Note: Employment rates are defined as the share of employed people among individuals aged between 18 and 64. The rates are averaged between 2005 and 2010.

Source: American Community Survey.

Between 2005 and 2010, slightly less than one third of refugees without U.S. ties found a job within 90 days after arrival (the precise percentage is 30.2%)[15] However, there is substantial variation across countries of origin in the "employment rate of refugees" – we define the latter as the share of employed people among refugees without U.S. ties aged 18-64.[16] The employment rate ranges between more than 40% for Cuban refugees and less than 20% for Afghani refugees (see left panel of Figure 4). There is also considerable variation across U.S. states (see right panel of Figure 4). Interestingly, a refugee resettled in Louisiana or Alabama has more than 60 percent chance to find a job within 3 months, as opposed to a probability below 10 percent in Minnesota. At first sight, there is no strong link across U.S. states between the employment rate of refugees (without U.S. ties) and the employment rate of natives (see Figure 5). For example, Louisiana and Alabama rank very low in terms of the

---

[15]Note that the percentage which appears in Table A.1 is slightly different (30.7%) because it is based on the sample we use for the regressions.

[16]In other words, in the denominator of the "employment rate of refugees", we include both individuals in and out of the labor force. This is because we cannot observe whether an individual is outside of the labor force in the refugees' data. We define the employment rates in Figure 5 in a similar way.

probability that overall workers are employed, while these two states rank very high in terms of the probability that refugees without U.S. ties are employed. Clearly the differences in employment rates of refugees across, respectively, U.S. states and countries of origin are the outcome of a combination of factors, such as the strength of the local labor markets, the composition of refugees in different places according to individual characteristics and nationalities, etc. The variation in our data at the individual, CZ, nationality and year levels will allow us to control for many of these factors. Finally, there is also considerable variation in self-employment rates, both across nationalities (see left panel of Figure 6) and U.S. states (see right panel of Figure 6). Self-employment rates are high for immigrants from Burma, Iran, Vietnam and Azerbaijan as well as for immigrants living in Georgia, Tennessee and Oklahoma.

Figure 6: Self-Employment rates of migrants by nationality and U.S. state, in 2004
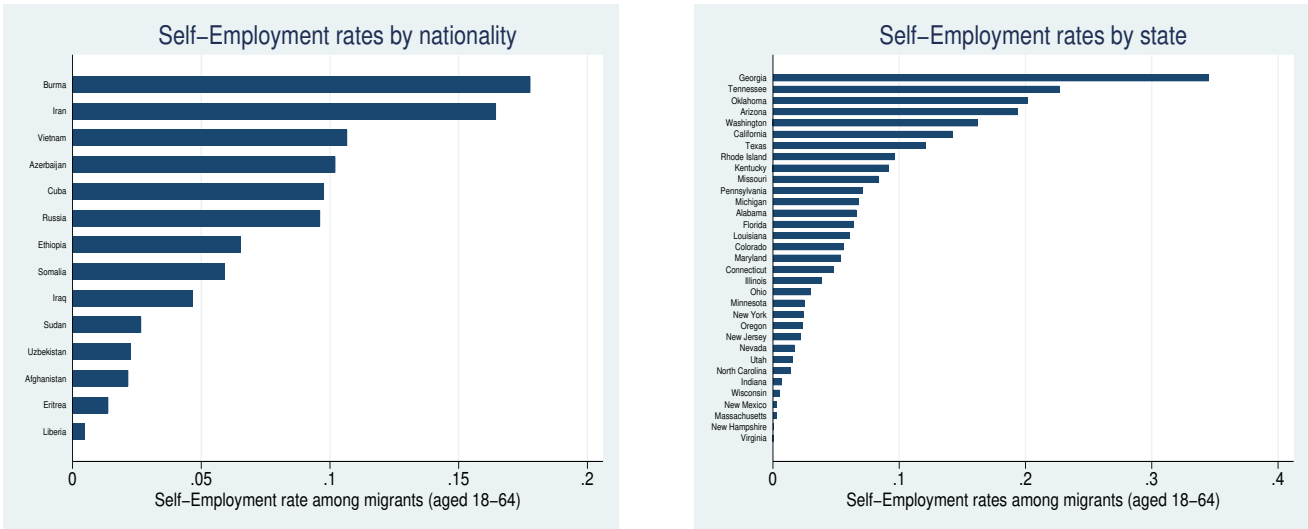
Note: Self-Employment rates are defined as the share of self-employed people among migrants aged between 18 and 64. For presentation purposes, only nationalities with more than 100 individual observations, non-zero self-employment rates and a network size above 1,000 members are considered.

Source: American Community Survey.

Finally, Table A.1 in the Appendix shows the summary statistics of the variables used in the empirical analysis. Refugees without U.S ties resettled in 2005-2010 tend to be in their early 30's, have on average two family members accompanying them, are more likely to be males (with a probability of around 58%) and to be married (with a probability of around 60%). Almost 40 percent of refugees without U.S. ties have no formal education. Around 17 percent of refugees without U.S. ties have a

primary education while 29 percent of them have a secondary education and approximately 7 percent of them went to university/college. As already mentioned above, around 30 percent of refugees in our sample are employed at 90 days after arrival. They are placed in commuting zones where the network from the same country of origin has on average 492 members, of whom around 78 are predicted to be business owners, 298 are predicted to be employees, while the rest are either unemployed or outside of the labor force.

4 Identification Strategy

We take advantage of the institutional features of the U.S. Refugee Admissions Program in order to estimate the causal effect of our variables of interest. First note that, to assess their case for admission to the United States, all refugee applicants are interviewed overseas by an officer from the Department of Homeland Security. This makes the initial screening process of potential refugees independent from the subsequent resettlement and allocation process within the United States. The latter is administered by resettlement agencies which work with the U.S. Department of State. Aside from this, there are two main threats to identification of a causal effect which we need to address: first, individual sorting of refugees into commuting zones (CZs) and, second, the non-random placement of refugees across CZs by resettlement agencies. In the following paragraphs we explain how we address each of these threats.

Whenever one observes the labor-market outcomes of migrants in a given locality, as a function of the number of migrants, individual sorting into CZs might be a concern. Sorting affects both the explanatory variable and the sample of individuals analyzed. First, favorable labor-market conditions in a given location might both increase the number of migrants (and migrant entrepreneurs) and improve labor-market outcomes for a given pool of newly-arrived refugees. To get around this problem, it is necessary to make sure that the variation in the explanatory variable is exogenous. In addition, the pool of newly-arrived refugees can be affected by individual self-selection – based on unobserved (by the econometrician) characteristics. For example, if newly-arrived refugees are free to choose where to locate, those especially driven and smart might go where there are better labor-market conditions or more opportunities to open a business. In that case, we would observe a positive

correlation between the number of entrepreneurs in the network and the refugees' employment status but it would be driven by the selection of easily-employable refugees into a location with a large number of entrepreneurs. In other words the estimates might be biased due to newly-arrived refugees sorting into specific labor markets.[17] These issues do not arise in our empirical analysis since we analyze cases of refugees with no family members or friends already in the United States, the so called cases without U.S. ties. Importantly, we focus our analysis on the universe of refugees without U.S. ties not only to define our dependent variable for the pool of newly-arrived refugees but also to construct the network variables. The placement upon arrival of refugees without U.S. ties is decided by refugee resettlement agencies, not by the refugees.[18] Hence refugees' placement is exogenous with respect to their preferences.

The second threat to identification of a causal effect is the non-random placement of refugees across CZs by resettlement agencies, which could take place based on refugees' individual characteristics and/or locations' characteristics. However we observe all the individual characteristics of refugees known by the resettlement agencies at the time of the placement decision and can control for them in the empirical analysis. Importantly, no employee of the resettlement agencies meets the refugee before the placement decision has been made. This implies that our results cannot be driven by resettlement agencies placing refugees according to unobserved (by the econometrician) individual characteristics. At the same time, refugees without U.S. ties have to be placed close to the offices of resettlement agencies, which are likely to be located in places with non-random characteristics. In general, strategic placement by resettlement agencies may take place according to unobserved characteristics of local labor markets. For example, a given location may have higher returns to the skills owned by foreign workers from a given country of origin. Hence it might be that resettlement agencies place refugees from that country of origin in that location, i.e. that they carry out strategic placement. In our main specifications, we include CZs by nationality fixed effects, together with CZ by year and nationality by year fixed effects. In other words, we control for the fact that a given CZ may be a better match on average for refugees from a given country of origin as well as for

---

[17]Note that, if the variation in the explanatory variable is exogenous, self-selection of newly-arrived refugees will not bias the estimate of the causal effect but will only affect the interpretation of the causal effect.

[18]Strictly speaking, refugees with U.S. ties do not chose their location either but are placed close to their relatives.

time-varying conditions of each CZ and of refugees of each nationality. Given the full battery of fixed effects we include in the empirical model, our estimates are only based on variation over time within a commuting-zone by country of origin pair. Hence strategic placement by resettlement agencies can be an issue for identification only if it is time varying, i.e. if resettlement agencies are able to adjust their location decisions as a function of information which is specific for the time in which the refugee arrives to the United States. However, as pointed out by Beaman (2012), this is unlikely to be the case, due to uncertainty and to delays between the time of the location decision and the time of arrival of the refugee. In her paper, Beaman (2012) notes: "Overall, the IRC employee who is solely in charge of placement states that the effectiveness of strategic decision-making is limited since she never knows when a refugee who is assigned to the IRC by the State Department will actually be allowed to travel. To highlight the stochastic component, consider 2005: there were cases that were given refugee status in 2001 but who arrived in 2005 due to delays associated with heightened September 11, 2001 security requirements." (Beaman 2012, p.139). Note that IRC (International Rescue Committee) is one of the nine U.S. refugee resettlement agencies. Finally, the results we find suggest that time-varying strategic placement is not an issue in our analysis: If the CZ where the refugee is placed had higher returns to the skills owned by workers of her country of origin, at a given point in time, then we should observe a positive correlation between the probability of employment of newly arrived refugees (our dependent variable) and the number of employed refugees in the network from the same country of origin. As shown in section 6, we instead observe a negative correlation.

To conclude, we also worry that mobility *after first placement* of, respectively, newly-arrived refugees and refugees in the network, might affect our estimates. To the extent that newly-arrived refugees leave the initial location where they are placed, the pool of remaining individuals we observe to construct the dependent variable might be highly selected. We can calculate the fraction of refugees without U.S. ties who move to another location within the 3-months period. That fraction stands at a very low 7.4 percent (note that this fraction is very similar to what Beaman (2012) finds for refugees resettled by IRC). This is consistent with the fact that refugees should not be inclined to leave their initial location, given that this would imply losing the services provided by the resettlement agency. Another possibility is that mobility after first placement of tenured refugees –

which constitute the network – might result in a number of tenured refugees which is correlated with unobserved characteristics of the labor market. To address this issue, we exploit variation in the size of networks driven by the *initial* placement of refugees without U.S. ties in a given location.

Finally, if resettlement agencies were able to carry out strategic placement over time – for example as a function of the size of the communities from the same country of origin as the refugee – we should observe a positive correlation between the number of refugees without U.S. ties and the overall *migrant* network. As Figure 7 shows, this is not the case.

Figure 7: Migrants vs. Refugees (without U.S. ties), in 2005-2010
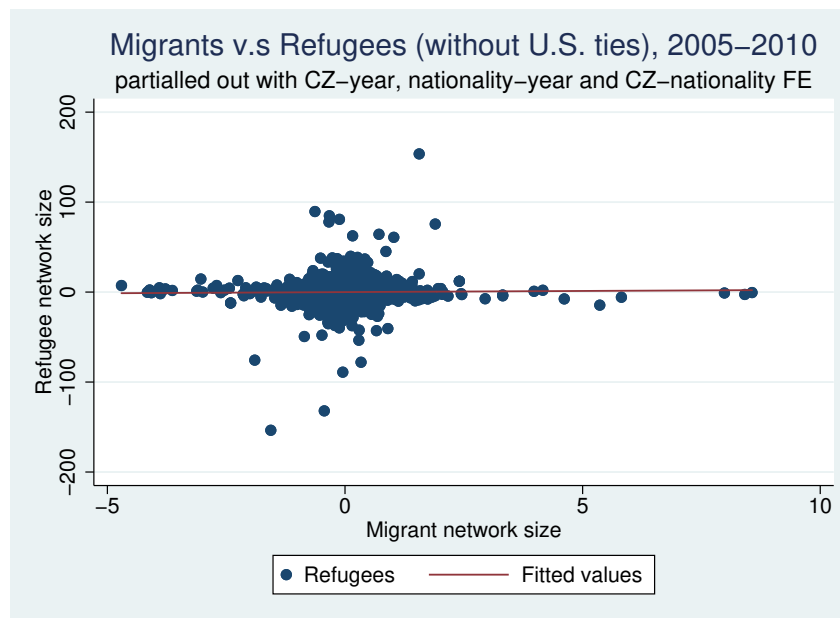
Note: Networks are constructed by summing the refugees arrivals up the previous year between 2005 and 2010. The network of refugees has been substracted from the network of migrants. All variables are partialled-out by commuting zone-year, nationality-year, and commuting-nationality fixed effects. The corresponding coefficient stands at 0.0012 with a standard error (clustered at the CZ and year levels) at 0.0022.

Source: WRAPS data set, Bureau of Population, Refugees and Migration, U.S. State Department and American Community Survey.

5   Empirical Specification

We formulate and test our main hypothesis in two steps. First, we show that using our data the size of the refugee's network does *not* have a (positive and) significant impact on the labor market

integration of the refugee (see Table 1). This result is consistent with what Beaman (2012) finds. Next, we investigate the relationship between the size of the network and our dependent variable separately by nationality. We find evidence of heterogeneity as the relationship is positive for some countries of origin but negative for others. Among those with a positive relationship (Figure B.2), countries like Iran, Russia, Vietnam, or Sudan show a relatively high level of self-employment. Our next step is motivated by both research in the literature and anecdotal evidence which together suggest a potential important role for entrepreneurs in the network of migrants. Indeed Kerr and Mandorff (2015) show that entrepreneurship is high among foreign-born workers. Moreover, as mentioned in the Introduction, anecdotal evidence suggests a link between migrant entrepreneurs and access to the labor market by foreign workers. Hence, in Table 2, panel A, we explore whether the impact of the refugee's network varies according to the business ownership rate in the network. Pushing further this story, in Table 2, Panel B, we also investigate whether the impact of the network varies according to the share of employees among immigrants from the same country of origin and in the same CZ as the refugee. In other words we assess the role of networks, in shaping refugees' labor market outcomes 90 days upon arrival, by estimating the following linear probability model:

$$
\begin{aligned}
Y_{ijkt} &= \beta_0 + \beta_1 Network_{jk(t-1)} + \beta_2(Network_{jk(t-1)} * SelfEmployedShare_{jk(2004)}) \\
&+ \beta_3(Network_{jk(t-1)} * EmployeeShare_{jk(2004)}) + \delta^{'} X_{ijkt} + +\gamma^{'} Z_{ijkt} + \alpha_{kt} + \alpha_{jt} + \alpha_{jk} + \epsilon_{ijkt}
\end{aligned}
$$

(1)

for individual i from nationality group j in CZ k at year t. In our main specifications, the sample includes around 76,000 individuals, aged between 18 and 65, coming from 58 origin countries and arriving in 120 U.S. commuting zones (CZs) between 2005 and 2010.[19] CZs have been recognized as the most coherent unit of analysis to investigate labor market dynamics in the United States (Autor and Dorn, 2013).[20] The variable Y_{ijkt} is the employment status 90 days after arrival of individual i

---

[19]In comparison, Beaman (2012) exploits information on 1,600 refugees resettled by IRC in 16 metropolitan areas between 2001 and 2005.

[20]"Commuting Zones are clusters of U.S. counties that are characterized by strong within-cluster and weak between-cluster commuting ties." (David Dorn: $http://www.ddorn.net/data.htm$). We use the crosswalks provided by David Dorn to match PUMAs from the American Community Survey to 1990 CZs. We also assign individuals from public Use Microdata Areas to CZs, by multiplying the ACS sampling weights (corresponding to the inverse of the probability

from nationality group j in CZ k at year t. The main regressors of interest are the network variables. We first construct the stock of refugees (without U.S. ties) of nationality j resettled in CZ k up to the year before individual i's placement, $Network_{jk(t-1)}$. Importantly, as pointed out above, we measure the stock of refugees by summing the number of refugees without U.S. ties *initially* placed in a given location every year since 1990 until $(t-1)$. In addition, based on the American Community Survey, we construct the share of self-employed in the stock of migrants from the same country of origin and in the same CZ as the refugee, using data prior to the period of investigation (2004), $SelfEmployedShare_{jk(2004)}$. The interaction term between the refugee network and the share of self-employed in the migrant network is the first key variable of the empirical analysis. It represents the predicted number of business owners in the network of the refugee. Our hypothesis is that migrant entrepreneurs facilitate the labor market integration of refugees from the same country of origin by hiring them in their businesses. In that case we would expect to estimate $\beta_2 > 0$. Similarly, we construct the share of employees in the stock of migrants from the same origin country and in the same CZ as the refugee, $EmployeeShare_{jk(2004)}$.[21] The interaction term between the refugee network and the share of employees in the migrant network is the second key variable of the empirical analysis. It represents the predicted number, within the network of the refugee, of people working as employees. Our hypothesis is that a greater number of tenured employees in the network makes it more difficult for newly arrived refugees from the same country of origin to find a job, because of greater competition. Hence we would expect to estimate $\beta_3 < 0$.

We augment the specification with individual characteristics known by the resettlement agency at the time of placement (age, gender, marital status, level of education, number of household members in the application). We denote this set of variables X_{ijkt} in equation (1) above and Individual Controls (1) and (2) in the tables. In additional specifications, we also include information at the individual level from follow-up surveys about the refugee's participation in support programs, that might affect the probability of employment. We denote this set of variables Z_{ijkt} in equation (1) above and

to be selected in the surveyed sample) by the individual weights provided by David Dorn to map PUMAs to CZs. In addition, note that the WRAPS data set includes information on the city and state of placement which we used to determine the corresponding commuting zone.

[21]Moreover, in additional results, we also use the share of *unemployed* migrants.

Individual Controls (3) in the tables.

As mentioned above, we also include a battery of fixed effects which control for unobserved aggregate factors that might affect the refugee's labor market outcomes. Commuting zone by year fixed effects, α_{kt}, capture unobserved time-varying heterogeneity at the CZ level including productivity shocks. Nationality by year fixed effects, α_{jt}, control for unobserved time-varying heterogeneity at the nationality level including changes in the quality of education in the origin country. Nationality by CZ fixed effects, α_{jk} control for the fact that resettlement agencies may place refugees from certain nationalities in specific locations where they have better labor-market outcomes on average over time. Finally, the error terms are clustered at the commuting zone-year level.[22]

6 Results

Table 1 presents the main results for the probability that a refugee is employed at 90 days after arrival, as a function of the size of her network. We introduce the fixed effects sequentially, starting with a simple set of country, year and commuting zone (CZ) fixed effects in column (1) of Table 2. In regression (2) of Table (2) we augment the specification with a more comprehensive set of fixed effects, at the community zone-year, nationality-year, and commuting zone-year levels. In columns (3) and (4) we add the individual characteristics observed by the resettlement agencies at the time of the placement decision (Individual Controls (1) and (2)). Finally, we control for other individual-level variables from follow-up surveys (Individual Controls (3)) in regression (5). Across these five specifications, the coefficient of the network variable is far from being statistically different from zero.[23]

---

[22]Our results are robust to using nationality-commuting zone level clustering.

[23]We find the same result when we use the network of *migrants* from the same country of origin and in the same CZ as the refugee (see Table A.2 in Appendix). Also note that we restrict the samples of Tables 1 and A.2 to be the same as the sample of Table 2, for comparability reasons. If we had not imposed this restriction, there would be 20,000 additional individuals in Table 1. The difference between the two samples is due to unavailability of data from the ACS for some CZs. We re-estimated the regressions in Table 1 using the larger sample and obtained similar results. The coefficient of the network variable is lower in magnitude and remains insignificant.

Table 1: The impact of the network

Dep. var.	(1)	(2)	(3)	(4)	(5)
			Employed		
Network (up to t-1)	-0.0009	0.0016	0.0026	0.0029	0.0042
	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)
Observations	75,832	75,645	75,645	75,645	73,578
R-squared	0.130	0.202	0.271	0.276	0.357
Year, nationality, and CZ FE	Y	N	N	N	N
CZ-Year FE	N	Y	Y	Y	Y
Nationality-Year FE	N	Y	Y	Y	Y
CZ-Nationality FE	N	Y	Y	Y	Y
Ind. Controls (1)	N	N	Y	Y	Y
Ind. Controls (2)	N	N	N	Y	Y
Ind. Controls (3)	N	N	N	N	Y

Notes: Robust standard errors clustered at the Commuting Zone-Year level in parentheses. *, **, ***: significant at 10%, 5%, and 1%, respectively. All network variables are divided by 100. The first set of individual control variables ($Ind.Controls(1)$) include age of the refugee, whether he or she receives a matching grant source of support, and the household size. The second set of individual control variables ($Ind.Controls(2)$) include education levels (primary, secondary, university, graduate, none, vocational or adult education). The third set of individual control variables ($Ind.Controls(3)$) include indications of whether the refugee has received all required core services, has source(s) of support from relatives or other non-government, has government cash assistance support, has a medical assistance source of support, has a source of support from Social Security, has other source(s) of support, the amount of R & P funds spent on behalf of this family, and the amount of R & P cash provided to this family.

In Appendix we show the estimates on the individual-level variables from the same regressions as in Table 1 (see Table A.3). The coefficients on age (and its square), gender, and the education levels have the expected signs. Age follows an inverse U-shaped pattern over the life cycle; male refugees are more likely to have a job (at 90 days) than female ones. Compared to the reference group with no formal education, the probability to have a job is higher the greater the level of education. Controlling for education, age, gender and marital status, we also find that the larger the number of household

members who accompany the refugee, the lower the probability that the refugee is employed after 90 days from arrival. From a theoretical point of view, the relationship between the two variables could be either positive or negative: so for example a greater number of kids might increase the incentive of the parents to look for a job right after arrival; at the same time, a greater number of kids might increase the adjustment costs right after arrival since the parents will be busy finding a school for them and might need to stay home with them. Also, to the extent that our individual-level controls are not able to fully capture unobserved characteristics, the sign might be driven by selection (i.e., refugees with more kids tend to come from a different socio-economic-cultural background than refugees with fewer kids). Interestingly, having a matching grant source of support increases the job prospects of the refugee. We also find that the probability that a refugee is employed after three months from arrival is negatively correlated with all "support" variables – denoting for example if the applicant has a government cash, medical, etc assistance source of support – with the exception of the indicator for having received all core services. We do not give strong interpretation to these coefficients since these variables are likely to be endogeneous, since programs are targeted to the most in need.

Table 2: The impact of business owners and employees in the network

	(1)	(2)	(3)	(4)	(5)
Dep. var.			Employed		
Panel A			Impact of business owners		
Network (up to t-1)	-0.0018	-0.0029	-0.0035	-0.0033	-0.0015
	(0.001)	(0.004)	(0.003)	(0.003)	(0.003)
Share of self.-employed (2004)	0.0039**	0.0078**	0.0105***	0.0108***	0.0098***
× Network up to (t-1)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)
Observations	75,832	75,645	75,645	75,645	73,578
R-squared	0.130	0.202	0.271	0.276	0.357
Panel B			Impact of business owners and employees		
Network (up to t-1)	-0.0021	-0.0010	-0.0011	-0.0009	0.0008
	(0.002)	(0.004)	(0.004)	(0.004)	(0.003)
Share of self.-employed (2004)	0.0035	0.0125***	0.0162***	0.0166***	0.0153***
× Network up to (t-1)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
Share of employees (2004)	0.0007	-0.0065	-0.0079**	-0.0081**	-0.0077*
× Network up to (t-1)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	75,832	75,645	75,645	75,645	73,578
R-squared	0.130	0.202	0.271	0.276	0.357
Year, nationality, and CZ FE	Y	N	N	N	N
CZ-Year FE	N	Y	Y	Y	Y
Nationality-Year FE	N	Y	Y	Y	Y
CZ-Nationality FE	N	Y	Y	Y	Y
Ind. Controls (1)	N	N	Y	Y	Y
Ind. Controls (2)	N	N	N	Y	Y
Ind. Controls (3)	N	N	N	N	Y

Notes: Robust standard errors clustered at the Commuting Zone-Year level in parentheses. *, **, ***: significant at 10%, 5%, and 1%, respectively. Control variables are described below Table 1.

In Table 2 we estimate equation (1) and follow the same format as Table 1 by introducing progressively the set of fixed effects and the control variables. Table 2 shows evidence of heterogeneity

in the impact of networks consistent with our hypothesis. In Panel A we introduce the size of the network both linearly and in interaction with the 2004 share of entrepreneurs in the migrant network – note that the linear effect of the time-invariant entrepreneurship rate is absorbed by the nationality by CZ fixed effects. While the direct effect of the network is insignificant, the interaction term has a positive and significant impact. Hence the probability that the refugee is employed 90 days after arrival is positively affected by the number of business owners in the network. In columns (3), (4) and (5), where we add the individual controls such as the level of education or the participation in support programs, the results are very similar. Broadly speaking, the results in panel A, Table 2 are consistent with a framework in which newly-arrived refugees and migrant entrepreneurs complement each other. While our evidence is not direct, since we do not observe the hiring decisions of migrant entrepreneurs, one interpretation of our findings is that network members who are entrepreneurs hire refugees in their business. Note that it is very unlikely that refugees who have just arrived are business owners themselves, for the following reasons: first, we observe refugees only 90 days after arrival, which is too short a time to open a business; in addition, Beaman (2012) documents that the average wage of refugees at 90 days after arrival is consistent with minimum wage occupations. Hence our results are likely not capturing the fact that refugees of a certain nationality (both newcomers and other refugees) have an easier time to open a business in a given CZ at a certain point in time. In addition, given the very short time horizon, we can also rule out stories in which newcomers learn the skills to open a business in areas and times where co-nationals have higher entrepreneurship rates.

In Panel B, Table 2, we add the interaction of the network variable with the 2004 share of employees in the migrant network. We find that the impact is negative (and significant when we control for the individual-level variables). This means that the probability that the refugee is employed is negatively affected by the number of network members who are employees. This result is consistent with a framework in which newly-arrived refugees and network members working as employees compete with each other in the labor market. Indeed workers from the same country of origin are likely to be highly substitutable since they might have similar skills. To the extent they look for similar jobs, they penalize each other in the labor market. In particular, if tenured refugees fill some job positions, these are not available for newly arrived refugees.

Based on the coefficient estimates in regression (4), Panel B, Table 2, our findings indicate that at the mean, doubling the number of business owners in the refugee's network raises the probability that the refugee is employed by about 1.3 percentage points; similarly, at the mean, doubling the number of employees in the refugee's network decreases the probability that the refugee is employed by about 4.9 percentage points. To put these marginal effects into perspective, the greater number of business owners in the network implies that a refugee from Vietnam (who finds on average 244 entrepreneurs in the network in the CZ of first placement) has 4 additional percentage points probability to be employed compared to a refugee from Liberia (who finds on average only 1 entrepreneur). At the same time, the higher number of employees in the network implies that a refugee from Vietnam (who finds on average 1,465 employees in the network in the CZ of first placement) has 10 percentage points probability to be employed compared to a refugee from Liberia (who finds on average only 210 employees).

7 Extensions and robustness tests

In this section we extend the model and test the robustness of the results in columns (3), (4), and (5) of Table 2, panel B.

In Table 3, panel A, we extend the empirical model by accounting for the number of *unemployed* members in the network of the refugee. In particular, we introduce the interaction between the network variable and the share of unemployed people in the migrant network. Consistent with our interpretation of our main results based on labor-market complementarity and substitutability, we find that the predicted number of unemployed in the refugee' s network decreases the probability that the newly-arrived refugee is employed. Workers from the same country of origin are likely to have similar skills and, to the extent they are both actively looking for a job, they penalize each other in the labor market.

In Table 3, Panel B, we extend the empirical model by controlling for the direct effect of the overall *migrant* network. In Section 4 we argued that it is very unlikely that resettlement agencies are able to carry out strategic placement with respect to time-varying variables. But in the remote possibility they do, we would like to make sure our results are robust. Resettlement agencies may try

to place refugees in locations with a large community of migrants from the same country of origin – the latter in turn may be following economic opportunities. To the extent resettlement agencies can predict the size of the migrant community at the time of arrival of the refugee, our estimates might be biased. However, when we control for the overall *migrant* network at time *(t-1)*, our results do not change. Finally, resettlement agencies might take into account the number of tenured migrants, to the extent they are aware of their beneficial impact (Beaman 2012). This does not seem to be the case. Our results are unaltered when we control for the migrant network up to three years prior to the arrival of the refugee (Table 3, Panel C).

As mentioned in Section 4, attrition of newly-arrived refugees is a minor issue in our data. But we still further explore this issue, in Table 3, Panel D, where we restrict the network to refugees arriving up to two years before the newcomer's arrival (instead of up to the previous year) and find results which are similar to our main findings. This suggests that refugees do not move much even after the first year of settlement.[24] We may also be concerned that displacement of natives might take place as a result of refugees' placement. However, refugees' networks do not seem to give rise to native displacement (Table A.4 in Appendix).

Another legitimate concern is the risk to confound the heterogeneous effect coming from more entrepreneurial nationality groups with the effect linked to other intrinsic group characteristics. We construct a measure of alternative dimensions of heterogeneity across nationalities – besides entrepreneurship. Specifically, we focus on nationality-specific in-marriage rates, which capture the extent of group cohesiveness within each network and have also been found to be correlated with the entrepreneurship rate of the group.[25] Kerr and Mandorff (2015) indeed show how social interactions lead to occupational stratification along ethnic lines. They conjecture that members of small social

---

[24]For example, if almost all refugees in the network left their initial placement after one year, our results would be radically different.

[25]We compute in-marriage rates - within group marriage rates - by refugee nationality following the methodology put forward by Kerr and Mandorff (2015). We use the US censuses from 1990 and 2000 restricted to married couples in which males migrated to the US between the age of 1 and 15. The latter is meant to avoid counting in someone who married before migrating and hence, confounding late migration with high in-marriage rates. We further restrict the sample to couples in which the husband is aged between 18 and 65. The in-marriage rate for nationality i is the ratio of the number of in-marriages for males of nationality i and the total number of marriages involving males of nationality i.

26

Table 3: Extensions

	(1)	(2)	(3)
Dep. var.		Employed	
Panel A		With the share of unemployed members	
Share of self-employed (2004)	0.0140***	0.0146***	0.0131***
× Network up to (t-1)	(0.005)	(0.005)	(0.005)
Share of employees (2004)	-0.0073*	-0.0076*	-0.0071*
× Network (up to t-1)	(0.004)	(0.004)	(0.004)
Share of unemployed (2004)	-0.0007***	-0.0007***	-0.0007***
× Network up to (t-1)	(0.000)	(0.000)	(0.000)
Panel B	Controlling for migration networks (up to t-1)		
Share of self-employed (2004)	0.0162***	0.0166***	0.0153***
× Network up to (t-1)	(0.005)	(0.005)	(0.005)
Share of employees (2004)	-0.0079**	-0.0081**	-0.0077**
× Network (up to t-1)	(0.004)	(0.004)	(0.004)
Migration (t-1)	-0.0000	-0.0000	-0.0001
	(0.000)	(0.000)	(0.000)
Observations	75,645	75,645	73,578
R-squared	0.271	0.276	0.357
CZ-Year, Nationality-Year & CZ-Nationality FE	Y	Y	Y
Ind. Controls (1)	Y	Y	Y
Ind. Controls (2)	N	Y	Y
Ind. Controls (3)	N	N	Y

Notes: Robust standard errors clustered at the Commuting Zone-Year level in parentheses. *, **, ***: significant at 10%, 5%, and 1%, respectively. All network variables are divided by 100. The *Network* (up to t-1) variable is included and remains insignificant in all regressions. The three sets of individual control variables are described under Table 1.

Table 3 (Continued): Extensions

	(1)	(2)	(3)
Dep. var.		Employed	
Panel C		Controlling for migration networks (up to t-3)	
Share of self-employed (2004)	0.0162***	0.0167***	0.0153***
× Network up to (t-1)	(0.005)	(0.005)	(0.005)
Share of employees (2004)	-0.0079**	-0.0081**	-0.0077*
× Network (up to t-1)	(0.004)	(0.004)	(0.004)
Migration (t-3)	0.0000	0.0001	0.0001
	(0.000)	(0.000)	(0.000)
Observations	75,645	75,645	73,578
R-squared	0.271	0.276	0.357
Panel D		Considering the network of tenured refugee	
Share of self-employed (2004)	0.0003**	0.0003***	0.0002**
× Network up to (t-3)	(0.000)	(0.000)	(0.000)
Share of employees (2004)	-0.0001	-0.0001	-0.0001
× Network up to (t-3)	(0.000)	(0.000)	(0.000)
Observations	75,645	75,645	73,578
R-squared	0.271	0.276	0.357
CZ-Year, Nationality-Year & CZ-Nationality FE	Y	Y	Y
Ind. Controls (1)	Y	Y	Y
Ind. Controls (2)	N	Y	Y
Ind. Controls (3)	N	N	Y

Notes: Robust standard errors clustered at the Commuting Zone-Year level in parentheses. *, **, ***: significant at 10%, 5%, and 1%, respectively. All network variables are divided by 100. The *Network* (up to t-1) and the *Network* (up to t-3) variables are included in Panels C and D, respectively and remain insignificant in both panels. The three sets of individual control variables are described under Table 1.

networks develop business specific skills through informal exchanges of information on their business activities and posit that these social interactions are complementary to production. Panels A and B of Table 4 control for the interaction term between the network and the in-marriage rate, defined in 1990 and in 2000, respectively. The number of self-employed in the network still has a positive and significant impact on the probability to be employed.

Finally, the results may be sensitive to the construction of the main variables. Our definition of networks based on common nationality may be too restrictive. We construct alternative measures of the network of refugees by including in the same network all refugees with similar language or religion as the newly-arrived refugee. Specifically, we construct *Linguistic Network* and *Religious Network* which are based, respectively, on the bilateral linguistic or religious proximity between the refugee's country of origin and all the other nationalities present at the CZ level. We further describe the construction of these alternative networks in Appendix C. Panels A and B of Table 5 show that our results are largely unaltered when we use these alternative measures of networks. Second, one may be concerned that our results are driven by nationalities associated with tiny networks. In Panel C, we show that our estimates are robust to dropping observations corresponding to small networks (network below 1000 individuals).

Table 4: Robustness

	(1)	(2)	(3)
Dep. var.		Employed	
Panel A		Controlling for social isolation (1990)	
Share of self-employed (2004)	0.0162***	0.0165***	0.0154***
× Network up to (t-1)	(0.006)	(0.006)	(0.005)
Share of employees (2004)	-0.0102**	-0.0103**	-0.0099**
× Network up to (t-1)	(0.005)	(0.004)	(0.004)
In-Marriage rate (1990)	-0.0004	-0.0004	-0.0005
× Network up to (t-1)	(0.000)	(0.000)	(0.000)
Observations	62,929	62,929	61,026
R-squared	0.276	0.280	0.359
Panel B		Controlling for social isolation (2000)	
Share of self-employed (2004)	0.0165***	0.0168***	0.0161***
× Network up to (t-1)	(0.005)	(0.005)	(0.005)
Share of employees (2004)	-0.0101**	-0.0100**	-0.0093**
× Network up to (t-1)	(0.004)	(0.004)	(0.004)
In-Marriage rate (2000)	-0.0001	-0.0001	-0.0001
× Network up to (t-1)	(0.000)	(0.000)	(0.000)
Observations	63,157	63,157	61,145
R-squared	0.270	0.274	0.357
CZ-Year, Nationality-Year & CZ-Nationality FE	Y	Y	Y
Ind. Controls (1)	Y	Y	Y
Ind. Controls (2)	N	Y	Y
Ind. Controls (3)	N	N	Y

Notes: Robust standard errors clustered at the Commuting Zone-Year level in parentheses. *, **, ***: significant at 10%, 5%, and 1%, respectively. All network variables are divided by 100. The *Network* variable is included and remains insignificant in all regressions. The three sets of individual control variables are described below Table 1.

Table 5: Robustness (cont.)

	(1)	(2)	(3)
Dep. var.	Employed		
Panel A	Alternative Networks (linguistic)		
Share of self. employ. among Linguistic network	0.0178***	0.0182***	0.0169***
× Linguistic Network up t-1	(0.005)	(0.005)	(0.005)
Share of employ. among Linguistic network	-0.0082**	-0.0084**	-0.0080**
× Linguistic Network up to (t-1)	(0.004)	(0.004)	(0.004)
Observations	75,645	75,645	73,578
R-squared	0.271	0.276	0.357
Panel B	Alternative Networks (religious)		
Share of self-employed among Religious network	0.0167***	0.0171***	0.0143***
× Religious Network up to t-1	(0.005)	(0.005)	(0.005)
Share of employees among Religious network	-0.0083**	-0.0083**	-0.0067*
× Religious Network up to (t-1)	(0.004)	(0.004)	(0.004)
Observations	75,645	75,645	73,578
R-squared	0.271	0.276	0.357
Panel C	Dropping tiny networks		
Share of self-employed (2004)	0.0162***	0.0166***	0.0152***
× Network up to (t-1)	(0.005)	(0.005)	(0.005)
Share of employees (2004)	-0.0079**	-0.0080**	-0.0076*
× Network up to (t-1)	(0.004)	(0.004)	(0.004)
Observations	74,293	74,293	72,230
R-squared	0.270	0.274	0.356
CZ-Year, Nationality-Year & CZ-Nationality FE	Y	Y	Y
Ind. Controls (1)	Y	Y	Y
Ind. Controls (2)	N	Y	Y
Ind. Controls (3)	N	N	Y

Notes: Robust standard errors clustered at the Commuting Zone-Year level in parentheses. *, **, ***: significant at 10%, 5%, and 1%, respectively. All network variables are divided by 100. The *Network* (up to t-1) variable is included and remains insignificant in all regressions. The three sets of individual control variables are described below Table 1.

8 Conclusions

The successful labor-market integration of refugees and immigrants matters well beyond the labor market. Recently Verwimp (2015) provides anecdotal evidence that bad labor-market outcomes of refugees and migrants may contribute to their political radicalization.[26] Similarly the NPR (2016) piece mentioned in the Introduction suggests that one reason why Belgian Turks are less likely to be radicalized than Belgian Moroccans is that the former end up faring better in the labor market.[27] The literature on migration and crime also underlines the importance of labor market integration. The link between migration and crime has been found to be weak or in many cases non-existent (Bell and Machin, 2013). However, when asylum seekers are found to increase crimes, such impact seems to be mostly explained by barriers to labor market integration (Butcher and Piehl, 1998; Bell et al., 2013; Spenkuch, 2014). Couttenier et al. (2016) even show that offering labor market access to asylum seekers eliminates the impact of asylum seekers on violent crimes. Similarly, Mastrobuoni and Pinotti (2015) and Freedman et al. (2018) find that, when labor market opportunities for migrants improve once they are granted legal status, the risk of crime recidivism decreases.

In this paper we posit that network members who are entrepreneurs help refugees from the same country of origin, by hiring them in their businesses. On the other hand, network members who are employees are likely to compete with newly-arrived refugees in the labor market. Our estimates are consistent with these hypotheses. They show that doubling the number of business owners in refugee' network increases their probability of being employed by one percentage point. At the same time, doubling the number of employees in refugees' network decreases their probability of being

---

[26]Verwimp (2015) shows that the number of migrants – both first- and second-generation ones – who leave European countries to become foreign fighters in Syria and Iraq (per million inhabitants) is positively associated with the gap in employment between natives and migrants in each country.

[27]"Turks and Moroccans immigrated to Belgium around the same time in the 1970s. And yet, when it comes to radicalization, the two groups couldn't be more different. Scores of Moroccans have left for Syria [to join ISIS], and there is not one recorded Turk who has followed the same path. ... Both Belgian Turks and Moroccans face labor-market discrimination in the Belgian labor market. Yet, Belgian Turks end up faring better in the labor market. ... Belgian Turks are somewhat more insulated because when they don't get a job they think they're qualified for, they turn to entrepreneurs in their own communities for help." (NPR story, April 4, 2016: "When It Comes To Radicalization In Belgium, Turks and Moroccans Are Different.")

employed by about two percentage points. Our results are not driven by refugees self-selecting into specific labor markets. Moreover, given the battery of controls and fixed effects we include, and the allocation process of refugees without U.S. ties, our estimates are not driven by strategic placement by resettlement agencies as a function of location characteristics. Based on these findings, our paper points to a new policy option which calls for providing business incentives and opportunities to tenured refugees and migrants to both facilitate their self-employment and to ease the labor market integration of newcomers.

References

Andersson, F., M. García-Pérez, J. Haltiwanger, K. McCue, and S. Sanders (2014). Workplace concentration of immigrants. *Demography 51*(6), 2281–2306.

Autor, D. and D. Dorn (2013). The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market. *American Economic Review 103*(5), 1553–1597.

Bandiera, O., I. Barankay, and I. Rasul (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica 77*(4), 1047–1094.

Beaman, L. A. (2012). Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S. *Review of Economic Studies 79*(1), 128–161.

Bell, B., F. Fasani, and S. Machin (2013). Crime and immigration: Evidence from large immigrant waves. *Review of Economics and Statistics 95*(4), 1278–1290.

Bell, B. and S. Machin (2013). Immigration and Crime. In A. Constant and K. Zimmerman (Eds.), *International Handbook on the Economics of Migration*. Northampton, MA: Edward Elgar.

Borjas, G. (1994). Economics of immigration. *Journal of Economic literature 32*, 1667–1717.

Butcher, K. and A. Piehl (1998). Cross-City Evidence on the Relationship between Immigration and Crime. *Journal of Policy Analysis and Management 17*, 457–493.

Calvó-Armengol, A. and M. O. Jackson (2004, June). The effects of social networks on employment and inequality. *American Economic Review 94*(3), 426–454.

Chiswick, B. (1991). Speaking, Reading, and Earnings among Low-Skilled Immigrants. *Journal of Labor Economics 9*(2), 149–170.

Chiswick, B. and P. Miller (1995). The endogeneity between language and earnings: International analyses. *Journal of Labor Economics 13*(2), 246–288.

Clemens, M. (2017). Violence, developmet and migration waves: vidence from central american chil migrant apprehensions. *CGD Working Paper 459*.

Cortes, K. (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., V. Preotu, D. Rohner, and M. Thoenig (2016). The Violent Legacy of Victimization: Post-conflict Evidence on Asylum Seekers, Crimes and Public Policy in Switzerland. *CEPR Discussion Paper 11079*.

Dolphin, S. and G. Genicot (2010). What do networks do? the role of networks on migration and coyote use. *Review of Development Economics 14*(2), 343–359.

Dustmann, C. and F. Fabbri (2003). Language proficiency and labor market performances of immigrants in the UK. *Economic Journal 113*, 695–717.

Dustmann, C. and A. van Soet (2012). Language and the Earnings of Immigrants. *Industrial and Labor Relations Review 55*(3), 473–492.

Fairlie, R. and M. Lofstrom (2013). Immigration and entrepreneurship. In B. Chiswick and P. Miller (Eds.), *The Handbook on the Economics of International Migration*. North-Holland Publishing.

Freedman, M., E. Owens, and S. Bohn (2018). Immigration, Employment Opportunities, and Criminal Behavior. *American Economic Journal: Economic Policy*, forthcoming.

Giuliano, L., D. I. Levine, and J. Leonard (2009). Manager race and the race of new hires. *Journal of Labor Economics 27*(4), 589 – 631.

Giuliano, L., D. I. Levine, and J. Leonard (2011). Racial bias in the manager-employee relationship: An analysis of quits, dismissals, and promotions at a large retail firm. *Journal of Human Resources 46*(1), 26 – 52.

Kerr, S. and W. R. Kerr (2017). Immigrant Entrepreneurship. *harvard Business School Working paper 17-011*.

Kerr, W. R. and M. Mandorff (2015, September). Social Networks, Ethnicity, and Entrepreneurship. NBER Working Papers 21597, National Bureau of Economic Research, Inc.

Mahajan, P. and D. Yang (2017). Taken by storm: Hurricanes, migrant networks, and u.s. immigration. *mimeo*.

Mastrobuoni, G. and P. Pinotti (2015). Legal Status and the Criminal Activity of Immigrants. *American Economic Journal: Applied Economics 7*(2), 175–206.

Melitz, J. and F. Toubal (2014). Native language, spoken language, translation and trade. *Journal of International Economics 93*, 351–363.

Montgomery, J. D. (1991). Social networks and labor-market outcomes: Toward an economic analysis. *American Economic Review 81*(5), 1407–18.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the u. s. labor market. *The Quarterly Journal of Economics 118*(2), 549–599.

Munshi, K. (2014a). Community Networks and Migration. In Y. Bramoulle, A. Galeotti, and B. Rogers (Eds.), *The Oxford Handbook of the Economics of Networks*, pp. 630–648. Oxford University Press.

Munshi, K. (2014b). Community networks and the process of development. *Journal of Economic Perspectives 28*(4), 49–76.

Olof Åslund, Lena Hensvik, O. N. S. (2014). Seeking similarity: How immigrants and natives manage in the labor market. *Journal of Labor Economics 32*(3), 405–441.

Patel, K. and F. Vella (2012, 2016/06/01). Immigrant networks and their implications for occupational choice and wages. *Review of Economics and Statistics 95*(4), 1249–1277.

Spenkuch, J. L. (2014). Understanding the Impact of Immigration on Crime. *American Law and Economics Review 16*(1), 177–219.

Separate Appendixes with Supplemental Material for:

Labor market integration of refugees to the United States: Do entrepreneurs in the network help?

February 14, 2018

Abstract

This document contains a set of appendixes with supplemental material.

Appendix A   Tables

Table A.1: Descriptive Statistics

	Obs.	Mean	St. Dev.	Min.	Max.
Employed	73,578	0.3071842	0.4613296	0	1
Network (up to t-1)	73,578	4.922013	6.802575	0	70.61
Share of self-employed (2004) × Network up to (t-1)	73,578	0.7873741	2.137421	0	14.94566
Share of employees (2004) × Network up to (t-1)	73,578	2.97996	5.394778	0	53.02923
Age	73,578	32.55903	10.92816	18	65
Age^2	73,578	1179.513	813.5362	324	4225
Housheold size	73,578	3.309386	2.204692	1	15
Male	73,578	0.5785425	0.4937959	0	1
Married	73,578	0.5956536	0.4907684	0	1
No Education	73,578	0.3967083	0.4892177	0	1
Education: Kindergarten	73,578	0.0025279	0.0502153	0	1
Education: Primary	73,578	0.1738427	0.3789768	0	1
Education: Secondary	73,578	0.2919895	0.4546806	0	1
Education: Intermediate	73,578	0.0379325	0.1910344	0	1
Education: Pre-University	73,578	0.0104515	0.1016976	0	1
Education: Technical School	73,578	0.0218951	0.1463422	0	1
Professional	73,578	0.0058985	0.0765754	0	1
University/College	73,578	0.0657805	0.2478998	0	1
Graduate School	73,578	0.0051102	0.0713034	0	1
Matching grant source of support	73,578	0.4007448	0.4900527	0	1
All required core services have been provided	73,578	0.990663	0.0961768	0	1
Source(s) of support from relatives or other non-government	73,578	0.2078475	0.4057698	0	1
Government cash assistance support	73,578	0.47767	0.4995045	0	1
Medical assistance source of support	73,578	0.8246351	0.3802815	0	1
Source of support from Social Security	73,578	0.01635	0.1268183	0	1
Other source(s) of support	73,578	0.464541	0.4987445	0	1
Amount of R & P funds spent on behalf of this family	73,578	1397.566	1276.623	0	23687.28
Amount of R & P cash provided to this family	73,578	564.7767	889.4204	0	9000

Notes: Descriptive statistics have been computed for the most comprehensive analytical sample of table 3, i.e. including the full set of control variables for interpretation purposes. All network variables are divided by 100.

Table A.1 (continued): Descriptive Statistics

	Obs.	Mean	St. Dev.	Min.	Max.
Share of unemployed (2004) × Network up to (t-1)	63,772	19.99498	43.67504	0	609.5593
Linguistic network	73,578	8.707178	8.338295	0	71.75341
Share of self-employed (2004) in Ling. network × Linguistic Network up to (t-1)	73,578	1.013144	2.239999	0	15.83355
Religious network	73,578	17.25106	14.19942	0	72.61958
Share of self-employed (2004) in rel. network × Religious Network up to (t-1)	73,578	1.467041	2.595386	0	16.38391

Notes: Descriptive statistics have been computed for the most comprehensive analytical sample of table 3, i.e. including the full set of control variables for interpretation purposes. All network variables are divided by 100.

Table A.2: The impact of the network of migrants

	(1)	(2)	(3)	(4)	(5)
Dep. var.			Employed		
Migration (up to t-1)	-0.0001	0.0000	0.0000	-0.0000	-0.0000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	75,832	75,645	75,645	75,645	73,578
R-squared	0.130	0.202	0.271	0.276	0.357
Year, nationality, and CZ FE	Y	N	N	N	N
CZ-Year FE	N	Y	Y	Y	Y
Nationality-Year FE	N	Y	Y	Y	Y
CZ-Nationality FE	N	Y	Y	Y	Y
Ind. Controls (1)	N	N	Y	Y	Y
Ind. Controls (2)	N	N	N	Y	Y
Ind. Controls (3)	N	N	N	N	Y

Notes: Robust standard errors clustered at the Commuting Zone-Year level in parentheses. *, **, ***: significant at 10%, 5%, and 1%, respectively. All network variables are divided by 100. Controls are described below Table 1.

Table A.3: The impact of the network: Estimates of coefficients of individual-level covariates (Part 1)

	(1)	(2)	(3)	(4)	(5)
Dep. var.			Employed		
Network (up to t-1)	-0.0009	0.0016	0.0026	0.0029	0.0042
	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)
Age			0.0171***	0.0170***	0.0128***
			(0.001)	(0.001)	(0.001)
Age^2			-0.0003***	-0.0003***	-0.0002***
			(0.000)	(0.000)	(0.000)
Household size			-0.0183***	-0.0161***	-0.0138***
			(0.001)	(0.001)	(0.002)
Male			0.1954***	0.1907***	0.1809***
			(0.008)	(0.008)	(0.008)
Married			-0.0129**	-0.0129**	-0.0349***
			(0.006)	(0.006)	(0.005)
Education: Kindergarten				0.0030	-0.0004
				(0.039)	(0.039)
Education: Primary				-0.0016	-0.0083
				(0.006)	(0.005)
Education: Secondary				0.0584***	0.0424***
				(0.006)	(0.006)
Education: Intermediate				0.0326***	0.0239**
				(0.011)	(0.010)
Education: Pre-University				0.1449***	0.1098***
				(0.027)	(0.024)
Education: Technical School				0.0773***	0.0597***
				(0.016)	(0.016)

Table A.3 (continued) The impact of the network: Estimates of coefficients of individual-level covariates (Part 2)

	(1)	(2)	(3)	(4)	(5)
Professional				0.0815***	0.0707***
				(0.024)	(0.023)
University/College				0.1100***	0.0841***
				(0.011)	(0.010)
Graduate School				0.0938***	0.0539**
				(0.026)	(0.026)
Matching grant					-0.1853***
source of support					(0.012)
All core services					0.0875***
					(0.024)
Source(s) of support					-0.0659***
from others					(0.014)
Government cash					-0.3596***
assistance support					(0.013)
Medical assistance					-0.0552***
source of support					(0.013)
Source of support					-0.0410**
from Social Security					(0.017)
Other source(s) of support					-0.0708***
					(0.010)
Amount of R & P funds					0.0000**
					(0.000)
Amount of R & P cash					0.0000
					(0.000)
Observations	75,832	75,645	75,645	75,645	73,578
R-squared	0.130	0.202	0.271	0.276	0.357

Table A.4: The impact of refugees on native population

	(1)
Panel A	Native displacement
Dep. var.	Population
Network (up to t-1)	-6.8217
	(16.112)
Observations	661
R-squared	0.999
Panel B	Native displacement of employees
Network (up to t-1)	-2.7539
	(3.653)
Observations	661
R-squared	0.998
Panel C	Native displacement of entrepreneurs
Network (up to t-1)	-8.7836
	(25.107)
Observations	661
R-squared	0.999
CZ FE	Y
Year FE	Y

Notes: Robust standard errors clustered at the Commuting Zone level in parentheses. *, **, ***: significant at 10%, 5%, and 1%, respectively. All network variables are divided by 100.
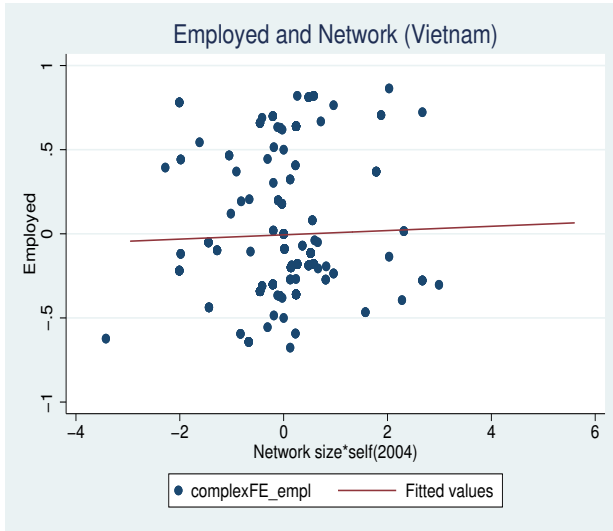
Appendix B   Figures

Figure B.1. Size of networks of refugees (without U.S. ties) across commuting zones within the United States, in 2010
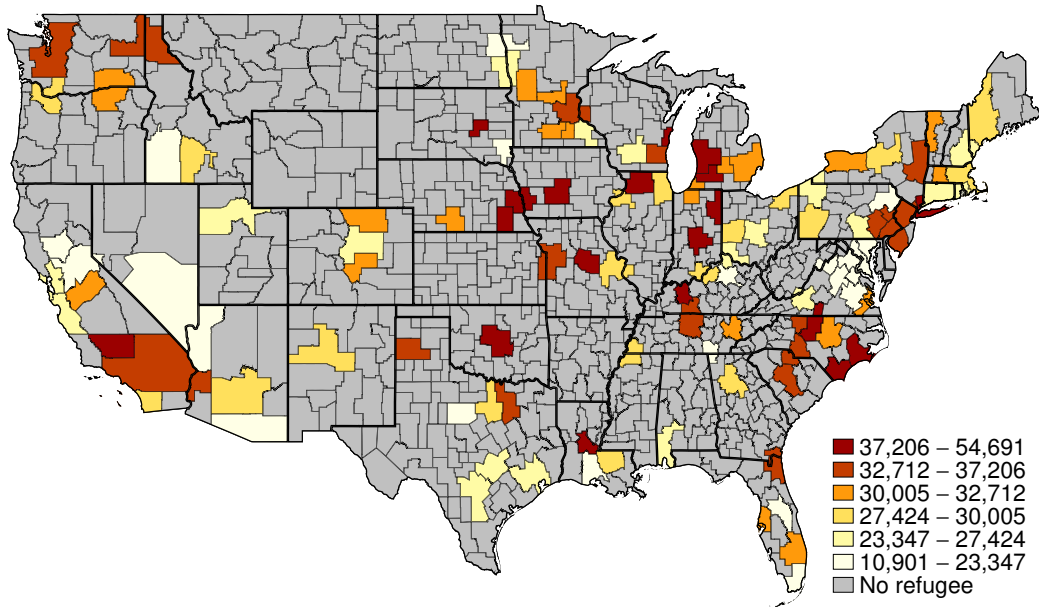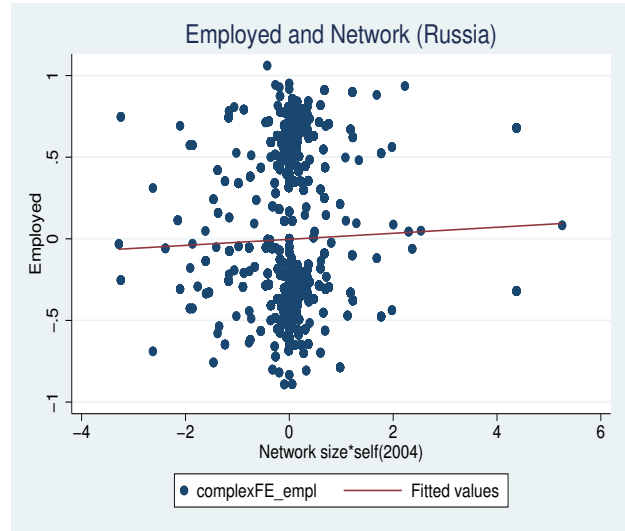

Note: Only communities with individual observations are considered.

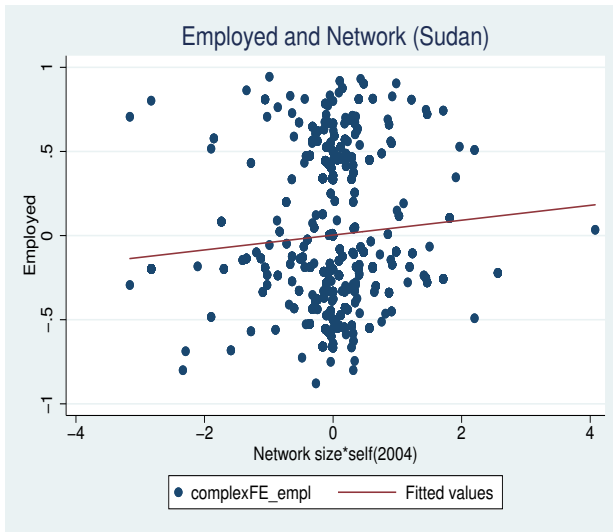Source: Worldwide Refugee Admissions Processing System (WRAPS).

Figure B.2. The impact of network size on the probability of employment (Vietnam, Russia, Sudan, Iran
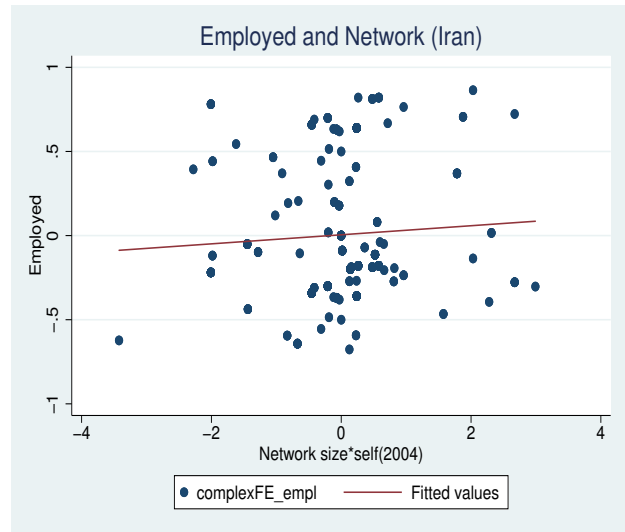

(a) Correlation between employment and network size (Vietnam)

(b) Correlation between employment and network size (Russia)

(c) Correlation between employment and network size (Sudan)

(d) Correlation between employment and network size (Iran)

Note: All variables are partialled-out by commuting zone-year, nationality-year, and commuting-nationality fixed effects.

Appendix C Description of alternative networks

We produced alternative networks for refugees of nationality i based on the bilateral linguistic / religious proximity between country i and all the other j nationalities present at the commuting zone (CZ) level. We used Melitz and Toubal's dataset (2014) on linguistic proximity to compute, for each year, an alternative network for each refugee nationality present at the commuting level.

$$linguistic\,network_{it} = \sum_{j \neq i} refugee\,number_{jt} * linguistic\,proximity_{ij}$$

with i, the nationality of the refugees of interest, j, the other refugee nationalities present in the same CZ and t, the year.[28] These alternative networks come in addition to the networks of refugees from nationality i.

For each nationality i, we compute the predicted number of entrepreneurs of different nationalities based on linguistic or religious proximity, as follows:

$$entrepreneurs_{it} = \sum_{j \neq i} refugee\,number_{jt} * linguistic\,proximity_{ij} * self - employment\,rate_j$$

In the case of the predicted number of employees of different nationalities for refugees of nationality i, we use the employment rate of the j nationalities present in the commuting zone.

Networks based on religion are computed similarly using the bilateral religious proximity instead of linguistic proximity.

The bilateral dataset on religious proximity was built thanks to data on religious composition by country in 2010 taken from "The Global Religious Landscape, a Pew Research Center publication. Bilateral religious proximity is the probability that one random citizen from country i is of the same religion as one random citizen from country j. It is computed as follows

$$religious\,proximity_{ij} = \sum_{r=1}^{6} population\,percentage_{ri} * population\,percentage_{rj}$$

with $population\,percentage_{ri}$ being the percentage of citizens of religion r from country i (same for country j). The six religions considered are Christianism, Islam, Hinduism, Buddhism, Judaism and Folk or Traditional religions.

---

[28]Note that since we lacked information for the bilateral linguistic proximities between Burma and Djibouti and Korea, we took the average of the linguistic proximities of their neighbouring countries. For Djibouti: Eritrea, Ethiopia, Somalia and Yemen and for Korea, Russia, China and Japan.