

CO-ETHNIC NEIGHBORS AND INVESTMENT IN HOST- COUNTRY LANGUAGE SKILLS

ANNA PIIL DAMM

AHMAD HASSANI

TRINE SKRIVER HØHOLT JENSEN

MARIE LOUISE SCHULTZ-NIELSEN

Co-Ethnic Neighbors and Investment in Host-Country Language Skills

Study Paper No. 169

Published by:

© The ROCKWOOL Foundation Research Unit

Address:

The ROCKWOOL Foundation Research Unit

Ny Kongensgade 6

1472 Copenhagen, Denmark

Telephone +45 33 34 48 00

E-mail: kontakt@rff.dk

<https://www.rockwoolfonden.dk/en>

January 2022

Co-Ethnic Neighbors and Investment in Host-Country Language Skills*

Anna Piil Damm[†] Ahmad Hassani[‡] Trine Skriver Høholt Jensen[§]

Marie Louise Schultz-Nielsen^{**}

Abstract: This paper investigates the effects of having co-ethnic neighbors on investments in formal host-country language acquisition and labor market outcomes exploiting a natural experiment between 2004 and 2015 when refugee immigrants were assigned to neighborhoods quasi-randomly. In our context of low shares of co-ethnic neighbors, we find robust evidence that the share of co-ethnic neighbors reduces the probability of formal host-country language courses enrolment of both men and women. In addition, language course progression is not affected by the share of co-ethnic neighbors, but it is negatively affected by both the employment rate among co-ethnic neighbors and the local employment rate. Further results show that the quality of the co-ethnic network at arrival promotes immigrant employment one to four years since arrival, irrespective of gender. Our results lend support to the theory that living in an ethnic enclave reduces the individual's incentive to invest in host-country language acquisition because of increased chances of finding a job through the co-ethnic network.

Keywords: Immigrants, Refugees, Asylum Seekers, Neighborhoods, Second Language, Human Capital Investment, Natural Experiment.

JEL codes: J68, I26, I28, R23

* This paper acknowledges the support from and access to Statistics Denmark provided by the ROCKWOOL Foundation Research Unit. We acknowledge financial support from the ROCKWOOL Foundation Research Unit. We appreciate helpful comments from Francesco Fasani, Valérie Smeets, Nabanita Datta Gupta, Timo Hener, Astrid Würtz Rasmussen, Claes Bäckman, Anders Holm, Mikkel Mertz, Rasmus Landersø, Lars Skipper, Marianne Simonsen and Helena Skyt Nielsen. We also thank Bente Herbst Bendiksen and Janne Lindblad at the Danish Immigration Service for sharing their internal administrative statistics and knowledge about the Danish Spatial Dispersal Policy 1999-2016 with us, and Marie Steiness at the language school A2B for explaining how language courses are administrated and students' results are reported. Finally, we thank Mathias Schindler for research assistance. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

[†] Department of Economics and Business, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V. Email: apd@econ.au.dk.

[‡] Department of Economics and Business, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V. Email: ahass@econ.au.dk.

[§] Department of Economics and Business, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V. Email: tshj@econ.au.dk.

^{**} ROCKWOOL Foundation Research Unit, Ny Kongensgade 6, 1472 København K. Email: mls@rff.dk.

I. INTRODUCTION

Increasing rates of immigrants over the past decades have spurred debates about how to design integration policies for immigrant labor market assimilation and cultural integration (Bauer, Lofstrom and Zimmermann 2000; Dustmann, Vasiljeva and Damm 2019; Bisin 2000). In Europe, the 2004 and 2007 enlargements of the common European labor markets have triggered substantial migration flows from East European countries to the old EU countries, and more recently, there has been massive influx of asylum seekers, notably from Syria, peaking in the autumn 2015.

Although immigrant labor market integration improves with years spent in the host-country (Chiswick 1978; Borjas 1985, 1995; LaLonde and Topel 1992; Dustmann 1993; Lubotsky 2007; Algan, Dustmann, Glitz and Manning 2010; Sarvimäki 2011; Dustmann and Görlach 2015), the native-immigrant gap remains large, notably for non-labor immigrants (Edin, Fredriksson and Åslund 2003; Damm 2009; Schultz-Nielsen 2017; Bratsberg, Raaum and Røed 2017; Fasani, Frattini and Minale 2018).

It is well documented in the literature that immigrant labor market assimilation is promoted by human capital such as education and host-country language skills (e.g. Dustmann and van Soest 2002; Dustmann and Fabbri 2003; Berman, Lang and Siniver 2003; Cortes 2004; Bleakley and Chin 2004, 2010; Auer 2018; Lochmann, Rapoport and Speciale 2019) and favorable labor market conditions (e.g. Åslund and Rooth 2007; Azlor, Damm and Schultz-Nielsen 2019). Moreover, previous quasi-experimental studies find that living in an (established) ethnic enclave improves labour market outcomes of (low-skilled) immigrants (e.g. Edin, Fredriksson and Åslund 2003; Munshi 2003; Damm 2009, 2014; Beaman 2012; Martén, Hainmüller and Hangartner 2019), documenting that potential negative effects of living in an ethnic enclave on labor market outcomes such as slower acquisition of host-country language are more than outweighed by positive effects from job referral through ethnic networks. These results show that ethnic enclaves promote immigrant labor market assimilation by reducing the earnings and employment gap between immigrants and natives.

Integration policies may also aim at cultural integration which can be defined as two groups having shared values, beliefs, expectations, customs, jargon and rituals without sacrificing the characteristics of its own culture that makes it unique. Common language is vital for cultural integration. Language is the set of common sounds and symbols by which individuals communicate. Immigrants' ability to speak the official or majority language of the host country is thus important for cultural integration as well as their labor market attachment. It may also be important for effective school-parent collaboration and their participation in civic life in general. Accordingly, host-country language training is a key part of the introduction program for newly arrived refugees in many countries, including the Nordic countries (Hernes, Arendt, Joonas and Tronstad, 2019).

In economics, language skill is considered as a form of human capital. Immigrants are expected to invest in learning the host-country's language up to the point where the marginal return of learning the language equals the marginal cost (Becker and Chiswick 1966; Dustmann 1999).

The values of having a common culture and language within a society are discussed in the seminal article by Lazear (1999). He argues that a common culture and language decreases the cost of interaction between different groups within a society. Based on Becker's human capital theory, immigrants who experience lower marginal cost and higher marginal benefit of host-country language proficiency are likely to invest more in acquisition of the host-country language (Becker 1957, 1964). An empirical evidence by Lazear (1999) shows how smaller minority groups in the U.S. learn English faster than larger minority groups and that immigrants are more likely to be fluent in English if they live in a community with a low share of co-nationals. Lazear's theory also predict spatial 'sorting': immigrants who are not fluent in the host-country language tend to move to ethnic enclaves in order to maximize their revenues from trade.

The importance of educational attainment for immigrants' abilities to learn the host-country language is described by Chiswick and Miller (2007). They elaborate further on Becker's standard human capital model. Chiswick and Miller present a framework where language abilities are modelled as a function of three main sources: exposure to the host-country language, efficiency in learning the language, and economic incentives to learn it. Attained education from the source-country is mentioned as one of the main predictors of efficiency in learning a new language. The linguistic distance between languages in the source- and host-country is decisive for the efficiency. In this regard, host countries in which the official or majority language is a "small-area language" have a particular challenge.¹ In case of small-area-languages, the immigrant's investment in host-country language acquisition is a sunk cost when the person out-migrates. Other important factors related to efficiency are age at migration and refugee status. According to Chiswick and Miller (ibid), refugees and their tied movers (following spouses) are expected to be less favorably selected and, consequently, to be less efficient in learning a new language. In addition, exposure to a new language is affected by the duration of stay in the host-country and the intensity with which the new language is heard or spoken. Therefore, being acquainted with many co-nationals or living in an ethnic enclave is expected to lower exposure to the host-country language.

While living in an (established) ethnic enclave promotes labor market entry of recent (low-skilled) immigrants and increase their earnings, living in an ethnic enclave may lower the speed of acquisition of host-country language proficiency. The aim of our paper is to test empirically how exposure to co-ethnic neighbors affects immigrants' investment in host-country language acquisition in the first years after first-time immigration to the host-country.

To this end, our study draws on an ideal spatial allocation experiment with refugees in Denmark in place from 1999 until mid-2016 to estimate the effects of living in an ethnic enclave on

¹ E.g. the Danish language which is only spoken in Denmark and only understood in other Nordic countries.

investments in formal host-country language acquisition. Since we study the effects in case of a small-area language and refugee migrants, we can ignore the common concern of self-selection of migrants into destination countries, e.g. to minimize linguistic distance (Adséra and Pytlikova 2015) and pull factors like ethnic networks and favorable earnings prospects at destination (Pedersen, Pytlikova and Smith 2008; Hatton 2016). In our context, the main challenges for identification of neighborhood effects are location sorting of immigrants across locations in the host-country (Bauer et al. 2005; Edin et al. 2003; Damm 2009) and correlated effects. Our use of quasi-experimental data for refugee immigrants enables us to address both of these challenges. Below, we detail how.

The current Danish language course system has been in place since 2004. To study formal Danish language acquisition four years after asylum, we focus on refugee cohorts arriving between 2004 and 2015. Our estimation sample consists of refugee household heads and their jointly arrived spouses, who were quasi-randomly assigned across municipalities (94 out of 98) and neighborhoods (914 out of 1961) during that period. To exploit the spatial dispersal of refugees during that period as a quasi-experiment, we follow the novel identification strategy developed by Azlor, Damm and Schultz-Nielsen (2019). The analysis includes information about refugees' formal Danish language acquisition, the educational skills they bring from abroad, and the municipality and neighborhood in which they are initially settled in Denmark. This was done by combining administrative registers from Statistics Denmark with information from The Danish Language Education Database, as well as recently constructed data on the individual's neighborhood of residence by Damm, Hassani and Schultz-Nielsen (2021a). Refugees' Danish acquisition is measured by their enrolment in and completion of the formal Danish language courses.

It is crucial that the neighborhood data incorporate the actual spatial area wherein people interact with each other, e.g., taking account of physical barriers like main roads. Fortunately, we have such a data set to our disposal as described in Damm et al. (2021a). This data includes a persistent neighborhood measure covering all housing units in the years 1986 to 2019. It is also crucial to identify the assigned neighborhood without measurement error. We have a new address register with complete spells of addresses for all residents in Denmark at our disposal, which allows us to do that. Finally, it is key to identify refugee status without measurement error. We do that by exploiting a register on type of residence permit available for all immigrants to Denmark since 1997.

With few exceptions (e.g. Åslund et al. 2011; Damm 2014), studies on the effects of residence-based ethnic networks on immigrant outcomes fail to distinguish between the size and the quality of the ethnic network. In this study, the large number of neighborhoods of assignment creates sufficient variation across and within municipalities for disentanglement of the effects of the size of the ethnic enclave from the effects of the quality of the ethnic enclave. The sustained period of the policy creates cohort-variation in neighborhood characteristics within municipalities that we

exploit to estimate the effects of size and quality of the ethnic enclave relying only on within-municipality variation in neighborhood characteristics.

Since 1986, upon receipt of their residence permit, refugees in Denmark have been subjected to spatial dispersal across municipalities to obtain a more even distribution of non-Western immigrants across municipalities. We, thus, study the effects of co-ethnic neighbors on outcomes for a subgroup of immigrants who experience low levels of residential segregation by ethnicity. Moreover, we study the effects of co-ethnic neighbors on host-country language acquisition during the first years since arrival during which refugees could participate at no charge in a three-year introduction program which included host-country language training as a key element. During the first three years after arrival, participation in the introduction program was required for refugees and family-reunified spouses of refugees in order to receive means-tested social assistance. Participation in the introduction program was optional for refugees in households with sufficient earnings to provide for the family and courses should be scheduled to allow employed refugees to participate in a course. Failure to find any effects of the share of co-ethnic neighbors on host-country language acquisition on outcomes for our subgroup of immigrants who experience low levels of residential segregation does not preclude the existence of effects for subgroups of immigrants who live separate from natives. If, on the other hand, such effects exist for our subsample of immigrants, we would also expect such effects to exist for subgroups of immigrants who experience high residential segregation from natives.

Despite the fact that individuals in our sample experience are assigned to neighborhoods with a mixed population in terms of ethnicity, we find robust evidence that the share of co-ethnic neighbors decreases the probability to enroll in formal host-country language training in the first years after immigration. We find this effect irrespective of whether we define the co-ethnic group narrowly as fellow countrymen, or use a broader definition as language fellows or non-Western immigrants. Estimation of the effect of the share of co-ethnic neighbors on the employment probability allows us to rule out that the negative effect of the non-Western immigrant group share on the probability to ever enroll in formal host-country language training is due to job entry. Recalling that take-up of the integration program is a condition for receipt of social assistance, we investigate to which extent public transfers constitute personal income during the first four years after immigration and find that it constitutes 100% of personal income for 70% of never takers compared to 38% among ever takers. Since only 5% of never takers have a child during that period so that max 5% of never takers have been on paternity leave, our finding that 5% are never takers suggest that the requirement is not fully enforced.

At the intensive margin, we find no effect of the share of co-ethnic neighbors on the acquired level of host-country language, and for men, negative effects of the quality of the co-ethnic network. Similarly, we find negative effects of the overall employment rate in the assigned municipality on the acquired level of host-country language. Our investigation of potential mechanisms suggest that the quality of the residence-based co-ethnic network at arrival promotes employment, irrespective of gender, while the size of the residence-based co-ethnic network has little effect on

the employment probability. These findings lend support to the hypothesis that better local job prospects promotes early job entry but crowds out formal host-country language investments. Moreover, to investigate whether informal host-country language acquired on the job early on compensates for lack of formal host-country language training, we describe the host-country language skill speaking requirements across the most common occupations held by individuals in our sample during the first four years after immigration. We find that the common occupations among individuals in our sample employed early on have low requirements for host-country language speaking skills. For women, our regression analysis suggest that a higher ethnic group share is associated with lower host-country language speaking requirements, while the effect is insignificant for men. These two findings suggest that the negative effect of the ethnic group share on the ever take-up rate of formal host-country language training is not outweighed by a positive effect of the ethnic group share on informal host-country language training on the job.

Our study contributes to two strands of literature. First, the literature on the effects of residence-based ethnic networks on host-country language acquisition. Second, the literature on the effects of residence-based ethnic networks on labor market assimilation of immigrants.

The literature on the effects of residence-based ethnic networks on host-country language acquisition is limited to observational data studies (e.g. Beenstock 1993; Dustmann 1994; Chiswick and Miller 1996, 2002; Lazear 1999; Beckhusen et al. 2013; Danzer and Yaman 2016). Observational data studies for Israel, Germany, US, Australia and Canada report a negative relation between the size of the ethnic (language) enclave in the region and host-country language proficiency. Similarly, Danzer and Yaman (2016) finds a negative statistical association between the co-national share on self-reported German language proficiency exploiting variation in the co-national share in the individual's region of residence in 1985 for the guest-worker sample of the German Socio-Economic Panel in 1984. The negative relation is found around two decades after the household head was subject to spatial regional placement as part of the West-German guest-worker program during the 1960s and early 1970s, at which point the least successful guest-workers have left the country (Dustmann 1996) and the remaining guest-workers may have moved region. Since the data used lacks information about the region of placement, a falsification check is conducted that hypothetically relocates the most language proficient immigrants into ethnic enclaves to test the extent of cross-regional sorting necessary to render the results purely spurious. The test shows that 45% of individuals must be extremely self-selected to in order to explain the ethnic enclave effect. Our study offers the first causal estimates of the effects of the size and quality of residence-based ethnic networks on host-country language investments.

The literature on the effects of residence-based ethnic networks on labor market assimilation of immigrants builds on the stylized fact that there is widespread use of friends, relatives and acquaintances to search for jobs (see e.g. Rees 1966; Granovetter 1974; 1995; Montgomery 1991; Pellizari 2010). Job seekers may receive information about vacant jobs from contacts such as employed neighbors, notably employed co-ethnic neighbors since social networks are likely to be ethnically stratified due to friendship formation in accordance with the homophily principle.

Residence in an established ethnic enclave may promote labor market outcomes of recent immigrants through referrals to job which match the qualifications of the individual job seeker. Exploiting information about the share of the individual's origin-community who are in the U.S. and variation in immigrant influx from Mexico to the U.S. from rainfalls in the origin community, Munshi (2003) finds positive effects of the size of the origin community network in the U.S. on individual employment and wages. A number of studies have exploited quasi-experimental variation in the size and quality of ethnic networks in the destination from government policies that resettle refugees in the destination. The studies by Edin et al. (2003) and Damm (2009) measure the size and quality of the ethnic network in the municipality of resettlement and find that immigrants in Scandinavia sort into ethnic enclaves. Taking into account location sorting, the studies find positive effects of the size of the ethnic enclave on immigrant annual earnings (for low-skilled) which increase with the average skill level in the ethnic enclave. Similarly, measuring the size and quality of the ethnic network at the metropolitan level, Beaman (2012) provide evidence that living in an established enclave increases employment and earnings of resettled political refugees in the U.S. These studies define co-ethnics on the basis of shared nationality. Measuring the individual's ethnic network in the neighborhood of resettlement and using a narrow (co-nationals) and broad definition (non-Western immigrants) of co-ethnics, Damm (2014) disentangles the effects of the size and quality of the ethnic enclave and provides quasi-experimental evidence that the quality of the ethnic enclave per se improves labor market outcomes of recent immigrants. Leveraging exogenous placement of asylum seekers in Switzerland, Martén et al. (2019) defines the co-ethnic network in three ways, as the number of immigrants in the assigned canton and year who have a shared nationality, ethnicity, or language, and find refugees assigned to locations with many co-nationals are more likely to enter the labor market. Our study provides causal estimates of the effects of residence-based ethnic networks on the employment probability exploiting exogenous variation in ethnic concentration at the neighborhood level from a more recent spatial dispersal policy on refugees than Damm (2014), identifying refugees and the assigned neighborhood without measurement error. With better data, we can include jointly arrived spouses into our estimation sample, thus enabling us to provide causal estimates for female as well as male refugees. We also extend the study by Damm (2014) by providing causal estimates of the size and quality of residence-based co-ethnic networks, defining the co-ethnic group as language fellows.

The remainder of our study is structured as follows. Section 2 describes the institutional background. Section 3 explains our methodological considerations and empirical model. Section 4 describes our raw data and sampling strategy, followed by a presentation of summary statistics. Section 5 presents our baseline results, robustness checks and tests of potential mechanisms. Section 6 concludes and discusses the implications of our findings for integration policies.

II. INSTITUTIONAL BACKGROUND

II.A. Danish Spatial Dispersal Policy on Refugees 1999-2016

In order to distribute refugees evenly across municipalities with suitable facilities for integration such as institutions for obtaining qualifying education and relatively low shares of immigrants, Denmark adopted its first policy of geographical dispersal of refugees in 1986 (Damm 2005). The policy was active until the Danish Parliament's enactment of the 'Integration Act' from the 1st of January 1999, which included a major reform of the policy and was in place until mid-2016. Henceforth, we will refer to the reform as *Danish Spatial Dispersal Policy on Refugees 1999-2015*.

In this study, we exploit natural variation in neighborhood characteristics of newly arrived refugees created by the Danish Spatial Dispersal Policy on Refugees for the period of 2004-2015 because micro data on Danish language course enrollment and completion only exists for refugees who arrived in 2004 or later. Below, we describe the reform in 1999 in detail and its implications for the settlement pattern of the 2004-2015 refugee cohorts.

The aim of the reform was to obtain an even more equal distribution of refugees relative to the local population size across municipalities and decrease the subsequent migration out of the assigned municipality of residence. The former goal is obtained by allocating the annual influx of refugees across municipalities according to a quota system that settles refugees in municipalities with lower shares of immigrants. The formula for calculation of annual municipal quotas is given in the third chapter of the 'Integration Act'. In this regard, it is important to note that the method for calculation of quota did not change during 1999-2016.²

During the asylum process, a caseworker from Danish Immigration Service (DIS) meets the asylum seeker to confirm the identity of the asylum seeker and asks other questions related to the asylum case. During the meeting, the asylum seeker can express her wishes about the municipality of assignment, but only under very special conditions will a refugee be settled in a municipality with a full quota (DIS-interview³). At the beginning of each year, it is possible to assign refugees to all municipalities with positive quota. However, as months pass and more refugees are assigned to municipalities, the municipal quotas become filled. It means that refugees' wishes about the municipalities can be satisfied in the beginning of each year easily. But in the next months, it becomes harder and harder. Then, if a refugee would like to go to a municipality which has fulfilled the annual quota, she cannot go there and will be assigned to another municipality with vacant slots. The important point is that the date at which refugees are granted residence permit is not under control of the refugees. Therefore, despite their wishes about specific municipalities,

² For more detailed information about the calculation of annual quota, see Appendix A in Azlor et al. (2020).

³ Two of the authors (Damm and Schultz-Nielsen) conducted an interview that addressed the administration of the Danish Spatial Dispersal Policy in January 18th, 2017 with Bente Herbst Bendiksen and Janne Lindblad at the Danish Immigration Service (DIS).

refugees who receive residence permit in later months are less likely to be assigned to the municipalities they wish, which means more attractive municipalities for refugees are fulfilled earlier. As a result, they are assigned to other municipalities with vacant slots. Importantly, this aspect of the refugee settlement policy is a novel finding of our interview with DIS and has not been discussed in public and was not noted in the literature until the publication of Azlor et al. (2020). Besides, the date on which a refugee is assigned to a municipality can be considered out of the control of the refugee, given that the month in which asylum is granted can be regarded as non-strategic and municipal assignment that takes place shortly after receiving asylum. Furthermore, asylum seekers' waiting period to obtain a Danish residence permit can last months (or even years). The municipality of assignment is responsible for providing refugees assigned to the municipality with affordable rental housing of appropriate size, given the household size, shortly after receipt of residence permit. To solve acute housing problems among individuals who need social assistance, the municipality can use its right to assign individuals in this group to every fourth vacant unit in public housing⁴ as well as private rental that has received public subsidies for building renovation within the last five years; applicants are assigned according to a waiting list and in a way to secure a mix of residents in each public housing section, e.g. in terms of immigrant status (Ministry for Children, Equality, Integration and Social Affairs 2015). Moreover, to solve acute housing problems among individuals who need social assistance, the municipality can negotiate the right to assign such individuals to vacant private rental housing with compensation to the owner, among others to obtain a better mix of residents in certain apartment blocks, as well as to purchase apartments from cooperative housing associations for re-sale or rental (Ministry for Children, Equality, Integration and Social Affairs 2015). Finally, the municipality can purchase and re-build existing properties to house refugees. However, the municipality is not allowed to let housing units at a rental rate below the market rent (Ministry for Children, Equality, Integration and Social Affairs 2015). Refugees can obtain a means-tested loan from the municipality for residential deposits and a means-tested rent subsidy. Due to housing shortage of affordable rental housing, there were effectively queues of individuals with the same individual characteristics waiting for housing. This meant that whenever the municipality received a vacant housing unit in a block with relatively few immigrants, it was offered to the next refugee in line with household characteristics matching the housing. This prevented the municipality from placing refugees with the same household characteristics in a selective matter.⁵ As expected, refugees tend to be settled initially in municipalities with a relatively low share of non-Western immigrants (Damm et al. 2021b). Within the municipalities, refugees are provided with housing in neighbourhoods with relatively high shares of public housing, co-nationals and individuals of low socioeconomic status (Damm et al. 2021b).⁶ In this context, thus, investigating the effects of area characteristics for the

⁴ The municipality can negotiate a higher share than every fourth vacant unit in public housing (Ministry for Children, Equality, Integration and Social Affairs 2015).

⁵ For a similar argument related to refugees' settlement during the first Danish Spatial Dispersal Policy (1986-1998), see Hasager and Jørgensen (2021).

⁶ Since 2010, municipalities are not allowed to offer housing to newly recognized refugees in socially deprived neighborhoods (public housing areas listed as "Ghettos" on the annual list from the Ministry of Housing).

subsample of refugees assigned to a municipality in the later months of the year resembles a field experiment.

Using the formula to predict the annual municipal quota of refugees, Azlor, Damm, and Schultz-Nielsen (2020) find a correlation between the predicted and actual refugee quota across municipalities is as high as 0.96. Importantly, socioeconomic characteristics of the municipality (e.g., job and rental housing vacancies) do not enter the formula for calculation of municipal quota. The Danish Immigration Service also does not use information about educational qualifications obtained prior to asylum in their municipal assignment, because educational qualifications from the refugees' home countries are typically not easily transferred to the Danish labor market, and the municipalities' desire for special educational groups are, in any case, modest (Azlor et al. 2020). In addition, close family already living in Denmark is primarily considered, and spouses and children are always settled in the same municipality as the first arrived family member settled.

The latter goal of low relocation rates is reached through conditioning the eligibility for social assistance to residing in the assigned municipality during the introductory program, which is 36 months and can be extended in case of parental leave. A study covering the years 1997-2005 finds that the reform succeeded in distributing refugees more evenly across municipalities relative to the local population size (and away from larger cities) and reduced refugees' tendency to move out of the municipality of assignment (Nielsen and Jensen 2006). The geographical mobility pattern of refugees arrived from 2004 to 2015 is shown in Figure 1. The dotted line illustrates the percentage of refugees that out-migrate from the municipality of assignment bi-annually since asylum. The solid line shows the share of refugees that still live in the municipality of assignment bi-annually since asylum. During the first three years after asylum, the bi-annual out-migration rate from the municipality of assignment is very low (0.3-2.2%), and by the end of the third year, 92% of the refugees are still living in the municipality of assignment. As soon as the three-year introduction program period ends, the relocation rate increases to 3.6% in first half of the fourth year since asylum, but declines thereafter. Hence, the postponed desire of moving that some refugees seem to have is fulfilled rather quickly. By the end of the fourth year since asylum, 86% of the refugees are still living in the municipality of assignment. Figure 1 documents that the spatial dispersal policy in place since 1999 influences refugees' settlement for a long period.

The specific date in which the household head has been assigned to the municipality is key for our identification. The date of municipality assignment is the date of residence permit (asylum) recorded in the residence permit register, provided by DIS. We also refer to this date as the date of arrival.⁷ With the exception of UN quota refugees (of which Denmark invited 500 annually until 2016), applicants for asylum apply after arrival to Denmark and live in a refugee reception centre until their asylum application has been processed.

⁷ It takes on average 40 days from refugees are permitted residence until they are registered in the municipality population register in the period 2005-2010, for which the calculation has been made (Hvidtfeldt and Schultz-Nielsen, 2018).

[Figure 1: Geographical stability of settlement for refugees arriving 2004-2015]

[Figure 2: Cumulative distribution function for municipalities that have met their annual quota of refugees by a given calendar month.]

Figure 2 shows the accumulated number and share of municipalities that have reached their annual quota of refugees by a given calendar month for each year in our sample period 2004-2015, i.e. the cumulative distribution function of municipalities that have reached their full capacity.⁸ The figure shows the following: First, due to expectations of very large influx of refugees in 2012 and 2015, few municipalities reached their full capacity in those years. Second, with those two exceptions, at least 16 municipalities had reached their annual refugee quota by the fall of the year. Second, in around half of the calendar years, all municipalities had met their annual refugee quota by the end of the year (i.e., 2004-2009 and 2014). Therefore, our preferred estimation sample includes observations from all calendar years, except for 2012 and 2015.

II.B. Formal Danish-As-Additional Language Courses

During the period from 1999 until June 2016, the municipality of assignment has offered a 3-year long introduction program to adult refugees and jointly arrived spouses upon municipal assignment at no charge as well as cost of transportation paid by the municipality, in case of low household income. During the first three years after settlement in the assigned municipality, refugees and family reunified spouses faced two conditions for receipt of means-tested social assistance: continued residence in the assigned municipality and program participation (Ministry for Children, Equality, Integration and Social Affairs 2015). This group of course participants would lose entitlement to social assistance in case of program drop-out. By contrast, program participation was optional for refugees and family-reunified in households with sufficient earnings to provide for the family and for this subgroup there was no penalty for dropping out of the program. However, courses should be scheduled to allow refugees who had found work during the first three years after municipal assignment to participate in a course.

Host-country language courses constituted the key element of the introduction program; it included 1,800 hours of language instruction (Arendt et al. 2020). The courses are divided into three distinct *course levels* of ‘Danish I’ for individuals with very short or no previous educational background, ‘Danish II’ for individuals with a short educational background, and ‘Danish III’ for individuals with medium or long educational background. The assignment of refugees to course levels is based on screening tests, which are conducted at language schools.

⁸ Our approach to construct the date at which each municipality reaches its annual quota follows the approach by Azlor et al. (2020). To find the date, we use the register information on the date of arrival of refugees to each municipality and compare this number with the municipal refugee quota. However, in contrast to Azlor et al. (2020), we include municipalities with a zero annual quotas in the number of municipalities which have reached full capacity by a given date in each year.

Each course level, in turn, is subdivided into several single *modules*, which are graphically illustrated in Figure 3. As is evident from this figure, Danish I and Danish II consist of 6 modules, while Danish III consists of 5 modules. Participants do not necessarily start with Module 1 but start with a module appropriate for their skill level assessed by the screening test.⁹ Each module level ends with a *test*, which determines whether a participant can progress to the next module level. Module levels are usually completed sequentially, but in some situations the participant may leave the course and return later, in which case her language proficiency is reassessed. The participant completes the language course by taking the final *exam*, which is conducted by the Ministry of Immigration and Integration¹⁰ at national level and structured differently across course levels. That is, while Danish I concludes with a written and an oral exam, Danish II and Danish III end with a written, an oral, and a reading exams.¹¹ The language programme is designed to be completed within 3 years, but it can be prolonged due to interruptions such as maternity leave and periods with full-time work.¹²

[Figure 3: Structure of formal Danish language courses]

[Figure 4: Timeline of a typical adult refugee's first four years in the country]

Figure 4 illustrates the typical timeline for refugees' participation in Danish courses. After arriving in Denmark, refugees spend time in asylum centres waiting to have their asylum application processed. The waiting period varies between individuals and the year in which the application is processed (Hvidtfeldt and Schultz-Nielsen, 2018).¹³ After obtaining asylum (denoted by t_0 in Figure 4), refugees usually begin Danish courses within a year. The refugees sequentially complete modules that constitute their Danish course levels and finish their language education by the final exams. The duration of the language education period varies but usually takes 2–3 years (corresponding to t_2 and t_3 in Figure 4).¹⁴

⁹ Authors' interview with Marie Steiness at the language school A2B.

¹⁰ Ministry of Foreigners and Integration ('Udlændinge- og Integrationsministerium').

¹¹ The level of Danish skills required for passing the exams are the following: Danish I: written (A2), oral (B1); Danish II: written/reading (B1), oral (B1-B2); Danish III: (B2). The classification follows the Common European Framework of Reference for Languages (CEFR), which is launched in 2001. Exam grades are registered after final exams, but not after each completed module.

¹² From 2004, it was emphasized that the language course should be completed within three years, which was the case for all refugees up until 2016.

¹³ Refugees arriving in 2005–2010 waited on average 376 days from arrival until receipt of residence.

¹⁴ The study by Damm et al. (2021b) includes a detailed description of provision of formal Danish Language Courses for immigrants.

III. METHODOLOGICAL CONSIDERATIONS AND EMPIRICAL MODEL

III.A. Methodological considerations

The main challenges for identification of neighborhood effects are location sorting of immigrants and correlated effects. Our use of quasi-experimental data for refugee immigrants enables us to address both of these challenges. Our estimation sample consists of refugee household heads and their jointly arrived spouses who were quasi-randomly assigned across municipalities (94 out of 98) and neighborhoods (914 out of 1,961) during the period 2004 and 2015. Individuals in our sample originate from around 80 different countries. The large number of neighborhoods of assignment and countries of origin creates sufficient variation across and within municipalities for disentanglement of the effects of the size of the ethnic enclave from the effects of the quality of the ethnic enclave. The sustained period of the policy creates cohort-variation in neighborhood characteristics within municipalities that we exploit to estimate the effects of size and quality of the ethnic enclave relying only on within-municipality variation in neighborhood characteristics. Our identification strategy thus controls for correlated effects that are constant across ethnic groups living in the same municipality such as local labor demand.

The aim of our paper is to estimate the effects of the size and quality of the ethnic enclave on formal host-country language course participation and progression. Since formal host-country language participation and employment are to some extent substitutes due to time rivalry, we also investigate the effects on the probability of employment and the probability of being active. Employment is defined as being employed as the main activity in November of the year. Active is defined as being enrolled in a Danish course in September of the year or being employed as the main activity in November of the year. To explore possible mechanisms, we further investigate effects on fertility, defined as the number of children born after residence permit, and the required level of Danish speaking skills in the job, conditional on having found a job.

We measure language acquisition at both the extensive and intensive margins. Our extensive margin measure is an indicator for whether individuals are enrolled in a Danish course module in September of the year in any year $t + s, s = 0, 1, \dots, 4$, where t denotes the year of asylum. Henceforth, we refer to this indicator as the “ever take-up” status. Our main measure at the intensive margin is a discrete variable for highest *completed* Danish module level by (September of) each year. We prefer this intensive margin measure to a discrete variable of the highest *enrolled* Danish module level because it is more relevant, from both a policy and an employer perspective. Whereas for individuals enrolled in a module at course level 1 or 2 the variable takes discrete values from 1-6, for individuals enrolled in a module at course level 3, the variable takes discrete values from 1-5. While extensive margin regressions are run using the full sample (‘Estimation Sample’), our intensive margin regressions exclude observations for never takers (i.e. individuals for whom the “ever take-up” status is equal to 0). The intensive margin regressions control for the enrolled course level using indicators for the individual’s highest enrolled course level by (September of) that year. We conduct robustness checks using as an alternative intensive margin

measure, in particular a discrete variable for highest *enrolled* Danish module level by (September of) of each year.¹⁵

We estimate effects during the first four years since asylum because we wish to estimate effects during the period of individuals' entitlement to participate in the introduction program. Recall that the introduction program lasts three years, but can be extended in case of parental leave. Thus, our period of observation exceeds the standard introduction program period by one year to allow for potential extension due to parental leave.

We provide estimates of the effects of the size and quality of the ethnic enclave using three alternative definitions of ethnicity. According to our most narrow definition of ethnicity, individuals from the same country of origin share ethnicity, i.e. co-nationals. This is a commonly used measure of ethnicity in migration analyses using administrative registry data (see e.g. Borjas 1992, 1995, 1998; Bertrand et al. 2000; Edin et al. 2003; Damm 2009; 2014; Beaman 2012). According to our broadest definition of ethnicity, individuals from countries in the same income category (high versus low) share ethnicity. Since all refugee cohorts in Denmark over our observation period originate from a low-income country, by this broad definition individuals in our sample share ethnicity with immigrants from low-income countries, henceforth referred to as non-Western immigrants. According to our third definition of ethnicity, individuals from countries who share at least one official language share ethnicity due to their common language (Constant and Zimmermann 2008). We henceforth refer to such individuals as language fellows. See Section IV for variable definitions.

We define a neighborhood using the well-defined macro-neighborhoods by Damm et al. (2021a). To obtain neighborhoods that are homogenous in terms of population size and have unchanged and exogenously defined boundaries over the 1986-2016-period, Damm et al. (2021a) construct neighborhoods that divide 459,497 inhabited hectare cells of Denmark into 8,359 micro- and 1,961 macro-neighborhoods. Particularly, these neighborhoods are formed by clustering housing units that are in close distance from each other and not separated by physical barriers like rivers or main roads. These neighborhoods are also compact, homogeneous in terms of size, housing, and ownership types, and have constant boundaries over time. Macro-neighborhoods have on average 2,717 inhabitants (median 2,520) and are comparable to the U.S. census blocks groups in terms of the number of inhabitants. Additionally, the constructed neighborhood clusters are linked with the Danish population register identifying the neighborhoods for 99.53% of residents of Denmark during 1986-2016. For this study, we have updated the dataset on the individual's neighborhood of residence described in Damm et al. (2021a) to cover the period 1986-2019.

We define the size of the ethnic enclave as the number of enclave members relative to the number of inhabitants in the neighborhood, following e.g. Damm (2014). Since our neighborhoods have been constructed to be homogeneous in size, our baseline results are robust to using the alternative

¹⁵ Instead of standardizing the outcome variables to have mean 0 and standard deviation 1 across each entry cohort to capture potential time-variation in available slots at language schools, we captured the potential time-variation by inclusion of arrival year and arrival month fixed effects in our regressions.

definition of the number of enclave members, used in e.g. Edin et al. (2003) and Damm (2009). Turning to the definition of the quality of the ethnic enclave, we define it as the employment rate, following Damm (2014), in our baseline analysis. Our baseline results are robust to using more narrow definitions of the co-ethnic network, (i) co-ethnic neighbors who have the same skill-level as individual i and (ii) low-skilled co-ethnic neighbors.

III.B. Empirical model

Our baseline model estimates the reduced form effects of assigned ethnic enclave size and quality using Equation (1):

$$Y_{icjma(t+s)} = \beta_0 + \beta_1 ES_{cj(t)} + \beta_2 EQ_{cj(t)} + \mathbf{X}_{it}\boldsymbol{\beta} + \boldsymbol{\delta}_c + \boldsymbol{\delta}_m + \boldsymbol{\delta}_a + \boldsymbol{\delta}_t + \varepsilon_{icjma(t+s)}, \quad (1)$$

where $Y_{icjma(t+s)}$ denotes language acquisition for refugee i at time $t + s$, $s = 1, 2, 3, 4$, who comes from country c and was assigned to neighborhood j in municipality m upon receipt of asylum in calendar month a of calendar year t . Our two explanatory variables of interest are $ES_{cj(t)}$ which measures the size of the assigned ethnic enclave and $EQ_{cj(t)}$ which measures the quality of the assigned ethnic enclave. In the baseline analysis, we measure $ES_{cj(t)}$ by the co-ethnic share in neighborhood j in year t and $EQ_{cj(t)}$ by the employment rate of co-ethnic group in neighborhood j in year t . \mathbf{X}_{it} is a vector of personal characteristics for refugee i in year of asylum t . The vectors $\boldsymbol{\delta}_c, \boldsymbol{\delta}_m, \boldsymbol{\delta}_a, \boldsymbol{\delta}_t$ contains dummies for, respectively, country of origin, municipality of assignment, calendar month of asylum, and calendar year of asylum. ε is an error term.

We also report estimated effects from a specification with observable municipality characteristics (measured in the year of asylum) instead of municipality of assignment fixed effects. We refer to this specification as baseline specification 1 and the specification in Equation (1) as baseline specification 2. Specification 1 controls for the two municipality characteristics which enter the formula for calculation of the annual refugee quota of each municipality, the municipal share of the population in the country and the share of residents in the municipality who are non-Western immigrants. The coefficient estimates of these two variables should not be given a causal interpretation. Moreover, it controls for the co-national share for two reasons. First, because refugees are always settled together with close family members and are often settled in a municipality with others of the same nationality in order to enable them to create a network of countrymen. Second, because previous studies provide quasi-experimental evidence that (established) ethnic networks promote immigrants' labor market outcomes (Edin et al. 2003, Damm 2009, Beaman 2012). In addition, Specification 1 controls for three characteristics of economic opportunities in the assigned municipality at arrival: the employment rate, the crime conviction rate, and the share of households with high income. These variables have a causal interpretation. Previous quasi-experimental studies on labor market integration of immigrants show that favorable local labor market conditions at arrival promote immigrants' employment and earnings (e.g., Åslund and Rooth 2007; Azlor et al. 2020).

To further account for common shocks that may affect the individual’s host-country language investment, we cluster the standard errors at the level of municipality of assignment.

To identify the effects we make two assumptions. First, individuals in our estimation sample were randomly assigned to neighborhoods, conditional on a few demographic characteristics of the household head observed by DIS. For reasons of efficiency, we also control for the level of education at arrival. Second, there are no correlated neighborhood characteristics which vary by country of origin. If household heads’ initial assignment to neighborhoods were completely random, there would not be a correlation between personal characteristics and characteristics of the assigned neighborhoods. In other words, existence of such correlations raises the concern that more able individuals have realized more favorable settlement conditions. In this regard, we examine to see whether the individual’s educational attainment at arrival (i.e., less than 11 years of education, 11 to 13 years of education, and more than 13 years of education) is correlated with the observed demographic and socioeconomic characteristics of the assigned neighborhood after controlling for individual’s age, gender, marital status, having children in different age groups, country of origin, year and month of asylum. The detailed explanation and results of this balancing test are presented in Section IV.B. Based on the results of this balancing test, we find that the individual’s educational attainment at arrival is not correlated with the observed characteristics of the assigned neighborhood when the first 15 municipalities had filled with their annual quota. As a result, the preferred subsample for estimation is the subsample of the balanced panel of household heads who were granted asylum after the first 15 municipalities within a given year had filled their quotas. We, therefore, limit the estimation sample to refugee household heads who arrived after the first 15 municipalities within a given year had filled their quotas and their spouses arrived on the same date.

Our baseline analysis also includes the results of the model given by Equation (2) which allows for heterogeneous effects of the assigned size and quality of ethnic enclave across gender:

$$Y_{icjma(t+s)} = \beta_0 + \beta_1^M G_i ES_{cj(t)} + \beta_1^F (1 - G_i) ES_{cj(t)} + \beta_2^M G_i EQ_{cj(t)} + \beta_2^F (1 - G_i) EQ_{cj(t)} + \beta_3 G_i + \mathbf{X}_{1it} \boldsymbol{\beta}_4 + \delta_c + \delta_m + \delta_a + \delta_t + e_{icjma(t+s)}, \quad (2)$$

where G_i is a gender dummy ($G_i=1$ for male) and e is an error term.

IV. DATA

IV.A. Data sources, sample selection and variable definitions

For our analysis, we use micro data from three main sources: (i) longitudinal administrative register data from Statistics Denmark for the years between 1997 and 2020 (2019), (ii) the Danish Language Education Register (Danskundervisningsdatabasen, henceforth abbreviated DUB), and (iii) the full population panel dataset with annual information about the individual’s neighborhood

of residence, constructed by Damm, Hassani and Schultz-Nielsen (2021a). We link these three main data sources using a unique person identifier.

In this study, the identification of refugees is based on the residence permit register provided by Statistics Denmark, which includes detailed information on granted residence permits from 1997 and onwards, and hence, enables us to perfectly identify refugees for the period of interest. Since the residence permit register does not contain exact information on the initial municipality of placement, we use the unique person identifier to link the data from residence permit register with the Danish population register that contains demographic characteristics, such as nationality, gender, age, as well as information on the municipality and housing of residence. This allows us to trace people's municipality and housing of residence, determined at the end of each year over time. To find the neighborhood of residence, we use the unique identifier of housing addresses from the population register and link the dataset to the neighborhood clusters constructed by Damm et al. (2021a). Then we treat the first registered municipality and neighborhood as the initial municipality and neighborhood of assignment. In addition, we link the population register with employment and education registers to find, among others, level of attained education in source-country as well as labor market status of refugees.

We also extend our dataset by adding information about the language(s) a refugee can speak. To do so, we use a document from the United Nations (2017) stating the name of countries and the official languages (national official) of those countries.¹⁶ We consider the official languages of countries rather than any other languages spoken because we have information only about the refugees' source-countries instead of detailed information about the area or region of source-country from where a refugee came to Denmark. Therefore, we cannot identify the region and hence the specific spoken language in that region. Accordingly, we assume that refugees can speak the official language(s) of their source-countries. This assumption means that we use the official language of each country as a proxy for spoken language in that country. Identifying the official languages spoken in each country allows us to find the number of persons (in addition to co-nationals) in municipality and neighborhood of residence with whom a refugee can speak by a common language(s). Detailed information about identifying the official language(s) of each refugee-sending country is provided in Appendix A.

Moreover, to find the required level and importance of Danish language proficiency in several occupations, we firstly link the refugee's dataset to the work classification module register and find the International Standard Classification of Occupation code (ISCO) for the most dominant occupation of refugees in each year, and secondly, link the result to the task content from the O*NET database provided by the U.S. Bureau of Labor Statistics.¹⁷

¹⁶ "11th United Nations Conference on the Standardization of Geographical Names" on 17 July 2017. https://unstats.un.org/unsd/geoinfo/ungegn/docs/11th-uncsgn-docs/E_Conf.105_13_CRP.13_15_UNGEGN%20WG%20Country%20Names%20Document.pdf (accessed on 20 June 2021).

¹⁷ <https://www.onetcenter.org/database.html#all-files> (accessed on 20 June 2021).

The performance of individuals (in a multitude of dimensions) in the formal Danish language courses is recorded in the register called ‘Danskundervisningsdatabasen’ (DUB), which has the recorded information since 2004. DUB contains information about individuals’ participation and completion of module levels according to course level, as well as the start and end dates of each module and the obtained grades at the final course exams.¹⁸ Therefore, we link our dataset to the DUB dataset using the unique person identifier and track individuals for 4 years.

We extract three samples of refugees from the linked residence permit register and other Danish administrative registers and neighborhood cluster dataset, as described above. Each sample is in turn described below and explained in more detail in Appendix Table A1.

The Balanced Panel is extracted using the following selection criteria. First, persons without a unique identifier and persons with admission category imputed by Statistics Denmark in the residence permit register during 1997–2019 are dropped from our sample. Only the first residence permit of each person is considered. Second, we restrict our sample to individuals with refugee status who were granted residency during 2004–2015, observed in the population register at least once during 2004–2019, and who were at least 18 years old on the date of arrival to Denmark. Third, we drop from the sample individuals who were not observed in the population register in the same year or the year after receiving their residence permit. Fourth, we drop persons without an identified neighborhood or for whom the initial neighborhood of residence was different from the first registered municipality of residence. Fifth, we drop persons whose partners’ country of origin is Denmark and not observed in the population register during the first four years of arrival. The choice of four years allows us to consider the potential extension of the three-year introduction program, e.g., due to maternity leave. Finally, we limit the balanced panel to the first-arrived person of each family (i.e., household head), because the first-arrived person is the main subject of the dispersal policy, and the settlement of other members of the family depends on household head’s settlement. For the case of married couples who are granted asylum at the same time, we consider the husband as the household head. The balanced panel consists of 15,992 persons.

Recall that our balancing tests support the identifying assumption of random assignment to neighborhood conditional on a few observable individual characteristics, if we only use a subsample of the Balanced Panel, in particular refugee household heads who arrived after the first 15 municipalities within a given year had filled their quotas. This subsample includes 5,113 household heads, and we refer to it as the *Subsample of the Balanced Panel*. It is important to note that the subsample of the balanced panel includes household heads who are not only subjected to the Danish Spatial Dispersal Policy, but also arrived after the first municipality quotas had been filled. That is, the subsample of the balanced panel consists of refugees who are even less likely to influence the municipality, and hence, the neighborhood of assignment.

¹⁸ Information about exam grades (from one or more final course exams) exists for around 60% of our sample. We have chosen not to use the exam grades as an outcome in our study as they are only available for a selected part of our sample.

Finally, in order to increase the efficiency and external validity of the estimation results, we augment the subsample of the balanced panel by adding the spouses of head persons (mostly female) who get the asylum on the same date as the household head is granted asylum, because when the assigned neighborhood of the head person is random, the settled neighborhood of the jointly arrived spouse would be random as well. Inclusion of such spouses increases the number of observations in the subsample of the balanced panel by 715 individuals. Consequently, the *Estimation Sample* that we use for our estimations consist of 5,828 individuals.

We also use the Danish population, employment, and education registers and calculate the demographic and socioeconomic characteristics of the municipalities and neighborhoods over time.¹⁹ Then we link this dataset to each of the three samples and find the characteristics of the municipalities and neighborhoods of residence during the first four years since granting asylum. Appendix Table A2 provides detailed information about the definition of each variable and the data source used to construct the variables.

IV.B. Summary statistics

Table 1 reports summary statistics for refugees in the year of arrival across the four samples: (1) the balanced panel, (2) the subsample of the balanced panel of refugee household heads, (3) individuals in the balanced panel who are excluded from the subsample, and (4) the Estimation Sample.²⁰ Refugees tend to be young, unmarried men travelling alone. In cases where spouses have arrived at the same date, we have defined the man as the household head, causing male overrepresentation in the first three-mentioned samples (men constitute around 82% of observations). By adding spouses who have arrived at the same date of the household head to the subsample, we obtain an estimation sample without male overrepresentation. The level of education attained before arrival is missing for around 38% of individuals in our samples. Among individuals with known level of education, the majority has at most 10 years of education. The top-five countries of origin in the balanced panel are Syria, Afghanistan, Myanmar, Iran, and Iraq (listed in descending order). The most represented official languages in the country of origin of individuals in the balanced panel is by far Arabic, followed by Farsi, Burmese, Russian, and French. Nevertheless, around 2% come from a country without an official language.

[Table 1: Summary Statistics: Individual Characteristics at arrival. Overall and by Gender]

¹⁹ The calculated demographic and socioeconomic characteristics include the share of population in neighborhood (municipality) relative to the total population of the municipality (country), the shares of co-nationals, language fellows, and non-Western immigrants, as well as employment rates (overall, high-skilled, and low-skilled) among each of these three groups at neighborhood and municipality levels over time.

²⁰ See Table 1 in the companion paper by Damm, Hassani, Jensen, Schultz-Nielsen and Schindler (2021b) for summary statistics of refugees in the sample of newly arrived adult refugees in the period 2004 until 2015 (20,390 individuals) from which we have extracted our four samples. Refugees in the sample are on average 32 years old at arrival and 73% are men.

Table 2 allows for a comparison of mean characteristics of the assigned municipality across our four samples. Column (5) reports t-test of significant differences in mean values of assigned municipality characteristics between individuals in the subsample and individuals in balanced panel excluded from the subsample. Individuals who arrived early during the calendar year are more likely to be assigned to the most populated and immigrant dense municipalities and be assigned to a municipality with a higher co-national share, albeit low due to spatial dispersal of newly arrived refugees since 1986. Previous research shows that refugees whose location choice was no longer restricted by a spatial dispersal policy were attracted to such locations (Åslund 2005; Damm 2009b). This provides empirical evidence that refugees who arrived early during the calendar year were more likely to realize their preferred location of assignment.

Table 3 allows for a comparison of characteristics of assigned neighborhoods. On average, individuals who arrived early during the calendar year were more likely to be assigned to an immigrant dense neighborhood and to a neighborhood with a higher representation of language fellows, co-nationals and non-Western immigrants and lower employment rate among co-nationals and non-Western immigrants, but similar employment rate among language fellows.

Note that the average co-national share in the assigned neighborhood is low across samples. For individuals in the Subsample, it is 0.3% with a standard deviation of 0.8. Since each neighborhood has on average 2,717 residents, a standard deviation increase in the co-national share around the mean corresponds to an increase from around 8 to 30 co-language fellows. By construction, the share of co-language fellows in the area exceeds the share of co-nationals in the area. Individuals in the Subsample are assigned to a neighborhood which has on average 1.5% co-language fellows, with a large standard deviation of 3.5, which means a standard deviation increase in the co-language share around the mean corresponds to an increase from around 41 to 136 co-language fellows. Similarly, the share of non-Western immigrants in the neighborhood exceeds the share of co-language fellows in the neighborhood. Individuals in the subsample are assigned to a neighborhood which has on average 5.1% non-Western immigrants, with a large standard deviation of 5.5. A standard deviation increase in the share of non-Western immigrants around the mean corresponds to an increase from around 139 to 288 non-Western neighbors. Note also that the share of co-nationals, co-language fellows and non-Western immigrants in the assigned neighborhood exceeds the corresponding share in the assigned municipality, providing evidence of residential segregation by ethnicity across neighborhoods within the municipality. Such residential segregation by ethnicity is due to concentration of affordable rental housing in certain neighborhoods (Damm, Hassani and Schultz-Nielsen, 2019). In Appendix Table A3, we report the correlation between each neighborhood characteristic for individuals in the Subsample. The correlation between the share of co-ethnic neighbors using our three alternative definitions is, as expected, positive and high, ranging between 0.51 and 0.7, confirming that for individuals in our sample, one can regard the share of co-nationals, share of language fellows and share of non-Western immigrants as alternative measures of the size of the co-ethnic network in the neighborhood. The correlation between the share of co-ethnic neighbors and the employment rate among co-ethnic neighbors is close to zero, irrespective of whether we define co-ethnics as co-

nationals, language fellows or non-Western immigrants. Such neighborhood variation enables us to distinguish between the size and quality of the co-ethnic network in the neighborhood. In Appendix Table A4, we report the correlation between employment rate of the ethnic group in the neighborhood and the employment rate in the municipality for each of the three definitions of co-ethnics. We find, as expected, a positive correlation between the employment rate of co-ethnics living in the assigned neighborhood and the employment rate in the assigned municipality. It ranges from 0.13 defining co-ethnics as co-nationals to 0.36 defining co-ethnics as non-Western immigrants.

While Denmark has received non-Western immigrants for several decades as guest workers and family reunified, a substantial influx of refugees started more recently during the mid-1980s. The majority of refugees in Denmark originate from non-Western countries and come from a vast number of different countries. As in other Scandinavian countries, the native-immigrant employment gap is large, but it is particularly large for refugees (Schultz-Nielsen 2017). Due to spatial dispersal of refugees away from the larger cities with established multi-ethnic enclaves and slow pace of labor market integration of refugees, in particular among women, the average individual in the Subsample is assigned to a neighborhood with an employment rate among co-nationals of only 18%, while the employment rate of co-language neighbors is on average 33% and the employment of non-Western immigrants in the neighborhood is 50% and only slightly below the employment rate of non-Western immigrants in the country.

[Table 2: Summary Statistics: Characteristics of the Assigned Municipality]

[Table 3: Summary Statistics: Characteristics of the Assigned Neighborhood]

If household heads were initially completely randomly distributed across the country, we would expect personal characteristics and characteristics of the neighborhood of assignment to be uncorrelated. If the individual educational attainment is correlated with neighborhood of assignment characteristics, it raises our concern that more able individuals have realized more favorable settlement conditions. Therefore, for each neighborhood of assignment characteristic that we observe, we investigate whether individual educational attainment is correlated with the neighborhood of assignment characteristics, controlling for individual demographic characteristics observed by DIS: age, indicators for being male, marital status, having children in different age groups, country of origin, year and month of asylum, and the municipality characteristics which enter the formula for determination of the municipal annual refugee quota, the population share and non-Western immigrant share, as well as the co-national share which may also have affected the municipal assignment due to DIS' intention of assignment to municipalities with existing concentration of co-nationals. We observe the following assigned neighborhood characteristics: the population share, co-national share, share of language fellows, non-Western immigrants share, employment rate among co-ethnics, employment rate among high-skilled co-ethnics, and employment rate among low-skilled co-ethnics, where co-ethnics are defined alternatively as co-nationals, language fellows and non-Western immigrants. We report the coefficient estimates of the level of education (11-13 years, more than 13 years, using below 11 years as the reference

category) in Appendix Table A5. The results for the balanced sample are shown in Panel A. We find that high-skilled individuals were significantly less likely to be assigned to a neighborhood with a higher co-ethnic share and more likely to be assigned to a neighborhood with a high employment rate among non-Western immigrants. In Panels B-E, we report the coefficient estimates of the level of education in each regression for various subsamples of the balanced panel excluding household heads who receive asylum before respectively the first 6, 9, 12, and 15 municipalities have filled their annual refugee quota.²¹ For the subsamples drawn by using a lower cut-off than 15 full municipalities, we find that individuals with high education were significantly less likely to be assigned to a neighborhood with a high share of non-Western immigrants and more likely to be assigned to a neighborhood with a high employment rate among (low-skilled) non-Western immigrants (significant at the 5 or 10% significance level). By contrast, for the subsample of household heads who get asylum after the first 15 municipalities within a given year have their quotas filled (Panel E), we find that educational groups have been randomly distributed across municipalities, conditional on personal demographic characteristics observed by DIS. Therefore, the preferred subsample for estimation of the effects of co-ethnic networks would be the largest subsample that provides quasi-random allocation, i.e., passes the balancing test. Panels B-E show that the preferred subsample for estimation is therefore the subsample of the balanced panel of household heads who got asylum after the first 15 municipalities within a given year have their quotas filled. This selection criterion reduces our subsample to 5,113 household heads.

Table 4 reports the balancing test for the Balanced Panel in Panel A, and for the preferred subsample of household heads for estimation in Panel B from Table A5. Controlling for municipality of assignment by fixed effects instead of demographic municipality of assignment characteristics at arrival also shows that the Balanced Panel is imbalanced for three out of 13 neighborhood characteristics (Table A6, Panel A), while the preferred subsample of house heads for estimation is balanced for all, but one neighborhood characteristic (Table A6, Panel B).²²

[Table 4: Balancing test: Assigned Neighborhood Characteristics and Individual Characteristics of Assignee (Households Heads)]

Table 5 summarizes the outcome variables for individuals in the Estimation Sample for the full sample (col. 1), men (col. 2) and women (col. 3). Recall that participation in the introduction program primarily consists of participation in Danish language courses and that it gives refugees entitlement to social assistance. Therefore, in Panel A, we find, as expected, that the overwhelming majority of individuals in the Estimation Sample have been enrolled in a Danish language course

²¹ We ran the balancing tests for all subsamples of the balanced sample including the individuals that are granted refugee status once j municipalities had been filled, where $j=1,2,\dots,30$.

²² According to Pei, Pischke and Schwandt (2017) a generally more powerful way of testing the relationship is to use the proxy for the candidate confounder (in our case educational level at the time of asylum) on the left-hand side of the regression instead of the right-hand side. Therefore, we have conducted this balancing test as well and shown it in Appendix Table A7. This test confirms that there is no correlation between individuals' educational attainment (as measured by a dummy for having at least 11 years of education) and any of the 13 neighborhood characteristics.

in September of each year in at least one year during the first four years since arrival (94.9%). Men are more likely to have been enrolled than women (96.1% vs. 91.8%). The enrollment rate decreases steadily from year one to four, from 80% in year one to 14.9% in year four. However, the enrollment rate of women exceeds that of men in years three and four, consistent with longer paternity leave periods of women, postponing their course participation.

During years one to four since arrival (Panel B), 58.6% of individuals in the Estimation Sample have been employed (in Nov.) in at least one year. The share of ever employed is substantially higher among men than women (69.5% vs. 29.1%). The employment rate increases steadily by years since arrival, irrespective of gender, reaching 53.8% for men and 20.9% for women.

Despite the fact that the language courses should be designed such that individuals can combine full-time employment with course participation, as reported in Panel C, the share of individuals in the Estimation Sample who are both enrolled in a language course in September of the year and work is low. It peaks in year 2 since arrival at 23.8% for men and 6.1% for women. As reported in Panel D, the share of individuals in the sample who are active (enrolled in language course in September and/employed in November of the year) increases steadily from 87.1% to 94.2% from year 1 to four, increasing from a lower level for women than men (77% vs. 90.8%) and reaching a lower level for women than men (89.6% vs. 95.9%).

[Table 5: Summary Statistics of Individual Outcomes by Years since Migration]

V. RESULTS

In this section, we present our baseline estimates of the effects of the size and quality of the co-ethnic network in the assigned neighborhood of residence, sensitivity checks and results from investigation of possible mechanisms. We end the section by presentation of the effects of the size and quality of the co-ethnic network in the current neighborhood of residence, estimated using an instrumental variable approach.

V.A. Baseline results

We first report our baseline estimates of the effect of the size of the co-ethnic network in the assigned neighborhood of residence on host-country language investments at the extensive margin, measured as the probability of ever enrolling in a host-country's language course. Table 6 presents the results for our three alternative definitions of the co-ethnic group: co-nationals (cols. 1-4), language fellows (cols. 5-8), and non-Western immigrants (cols. 9-12). For each definition, we present results from four specifications, two specifications without/with control for the quality (measured as the employment rate among co-ethnic neighbors). The idea of presenting results for four specifications for each definition of the co-ethnic group is to investigate how robust the

estimated effects of size of the co-ethnic network are to control for the quality of the co-ethnic network and municipality characteristics. Results reported in the odd numbered columns controls for the observed demographic and socioeconomic characteristics of the assigned municipality at arrival (specification 1), while results reported in the even numbered columns control for time-constant municipality of assignment characteristics using municipality fixed effects (specification 2). Panel A reports effects assuming homogeneous effects by gender, while Panel B reports effects allowing for heterogeneous effects by gender. The last specification reported for each definition of the co-ethnic group in Panel A is given by Eq. 1, whereas the last specification reported for each definition of the co-ethnic group in Panel B is given by Eq. 2.²³ Note that neighborhood and municipal characteristics enters the regressions as rates and not in percent. Interestingly, we find that the ethnic group share decreases the probability of ever enrolling in the host-country language course. The magnitude of the estimated effect for each definition of co-ethnic group varies little across specifications, and the effect is precisely estimated defining co-ethnics as non-Western immigrants. A doubling of the share of non-Western immigrants among neighbors at arrival (corresponding to a standard deviation increase around the mean of 5.1%) decreases the probability of ever enrolling in a host-country language course by around 1 percentage point, corresponding to 1%. The magnitude of the coefficient estimate is higher for women, but not significantly different. Turning to the estimated effects of the quality of the co-ethnic network, if anything, it promotes the ever enrollment rate. However, the coefficient estimate of the employment rate of the ethnic group is only significant using the narrow definition of the co-ethnic group and only significant at a 10% level. Besides, the reported results in Panel B of Table 6 suggest that, everything else equal, men are 1.5 percentage points more likely to ever enroll than women, although the effect is only significant at a 10% significance level.

[Table 6: Effects of assigned ethnic group share and employment rate of ethnic group on ever enrollment in host-country language course]

Turning to our baseline estimates of the effects of the size and quality of the co-ethnic network in the assigned neighborhood of residence on host-country language investments at the intensive margin, measured as the level of completed modules in September of each year, we report the results from linear regressions in Table 7. In Panel A, we report estimated effects using specifications which assume homogeneous effects by gender, and report results from estimation of Eq. (1) in the fourth specification for each definition of the co-ethnic group. In Panel B, we report estimated effects using specifications which allow for heterogeneous effects by gender. In particular, we report results from estimation of Eq. (2) in the fourth specification for each definition of the co-ethnic group. For each of the three alternative definitions of the co-ethnic group, we report estimated effects of the quality of the co-ethnic network using four different specifications.

²³ Since the dependent variable in Table 6 is a dummy variable, in fact we estimate linear probability models. Our model specification uses dummy variables as control variables in order to satisfy the requirements of a saturated model. The saturated model with a discrete outcome will identify identical coefficient estimates and standard errors to a logit/probit model (Angrist 2001).

In contrast to the two last specifications for each definition of the co-ethnic group, the first two specifications for each definition of the co-ethnic group provides the estimated effects of the quality of the ethnic group without controlling for the size of the ethnic group. As in Table 6, results reported in the even numbered columns controls for municipality of assignment characteristics using municipality of assignment fixed effects, while results reported in the odd numbered columns are from specifications which control for municipality of assignment characteristics through inclusion of demographic and socioeconomic characteristics. For the two narrower definitions of the co-ethnic group, we find evidence that the employment rate of the co-ethnic group in the assigned neighborhood at arrival decreases the module level completed, conditional on course enrollment. Using the broad definition of the co-ethnic group, we find little evidence that the quality of the co-ethnic network influences the module level completed. As shown in the Appendix Table A8, the results in Table 7 are robust to using the ‘Estimation sample’, i.e., including never takers of host-country language course after replacing the missing value of the dependent variable with 0 for never takers. The results in Table 7 are also robust to measuring effects on module level enrolled instead of completed. We find no effects of the ethnic group share on the module level completed (Table 7 and Table A8).

In Table 7, we also report the coefficient estimate of the employment rate in the assigned municipality at arrival because it is significant at a 1 percent significance level. The sign is negative and large in absolute terms. Interestingly, the coefficient estimate of the employment rate in the municipality of assignment is lower in absolute terms in tables A8 and A9 than in Table 7, but still significant at a 1% significance level, whereas it does not influence the ever enrollment rate (Table 6). Viewed together, these results suggest that local labor market condition primarily influences the decision to complete a language module, and influences the decision through its effect on the opportunity cost: Worse local employment prospects decrease the opportunity cost of host-country language progression. According to the estimates in Table A8, 4 percentage points decrease in the employment rate in the assigned municipality at arrival (corresponding to a standard deviations decrease around the mean of 76.9%) increases the module level completed by 4% around the mean of the sample distribution.

Across specifications and samples, we find that, everything else equal, men tend to complete modules at a higher level than women. Since men also tend to have higher level of education at arrival (Table 1), a likely explanation is that the return to host-country language is higher for men than women. The coefficient estimate of being male of around 0.17 corresponds to 6 percent around the mean of the sample distribution. Consistent with this result, we find some evidence of heterogeneous effects of the quality of the ethnic group by gender (Table 7 and Appendix Table A8, Panel B). The employment rate of the ethnic group has a negative and significant effect on completed module level for men, whereas the magnitude is lower in absolute terms for women and estimated with imprecision. This further supports our interpretation that language progression depends on the opportunity costs, which are higher for men. That is, for men, 30 percentage points decrease in the employment rate of language fellows (corresponding to a standard deviation)

increases the module level completed by around 0.05, corresponding to 2% increase around the mean of the sample distribution.

[Table 7: Effects of assigned ethnic group share and employment rate of ethnic group on the level of host-country language course completion]

V.B. Robustness analyses

In this subsection, we investigate whether our baseline results in Table 6 and 7 are robust to using alternative samples and ethnic group definitions. Appendix Table A10 reports effects at the extensive margin (Panel A) and intensive margin (Panel B) using alternative samples, allowing for heterogeneous effects by gender and defining the ethnic group as language fellows. For comparison, we repeat our baseline results (Table 6 and Table 7, cols. 7-8) in cols. 1 and 2. In cols. 3 and 4, we report the effects restricting the Estimation Sample to the 2004-2012 cohorts. The reason for estimating the effects after restricting our baseline sample to cohorts that arrived before 2013 is that individuals in the cohorts who arrived between 2013 and 2015 were still enrolled in the three-year integration program transferred to a new integration program called “work-first” after the reform of the Integration Act in mid-2016, which speeded up entry into regular jobs, crowding out formal host-country language investments (for men only) (Arendt 2020). This restriction reduces the sample size from 5,828 to 3,383. The point estimates of the ethnic group share on ever enrollment in host-country (Panel A) are negative across gender and higher in absolute terms than the baseline estimates but estimated with imprecision. The point estimates of the employment rate of the ethnic group on the module level completed (Panel B) are negative as the baseline estimates, but estimated with imprecision. By contrast, the negative effect of the employment rate in the assigned municipality at the intensive margin remains significant. In cols. 5 and 6, we report the effects using the Balanced Panel of Refugee Household Heads and Jointly Arrived Spouses. Recall that estimates using this sample may be biased since early arrivals during the calendar year may have self-selected into municipalities. As for the baseline sample, use of this sample yields negative point estimates of the ethnic group share on ever enrollment in host-country (Panel A) across gender, however, the point estimates are biased towards zero, and so are the point estimates of the employment rate of the ethnic group on module level completed for this sample (Panel B). Leaving out early arrivals during the calendar year from the estimation sample is thus important for consistency of the estimated peer effects. Similarly, the negative estimate of the employment rate in the assigned municipality at the intensive margin is biased towards zero, but it remains significant.

Appendix Table A11 reports effects at the extensive margin (Panel A) and intensive margin (Panel B) allowing for heterogeneous effects by gender and defining the ethnic peer group as language fellows (for comparison, baseline estimates in cols. 7-8 in Tables 6 and 7 are repeated in cols. 1-2) or subgroups of language fellows. In cols. 3-4, we report the effects of the ethnic group share and employment rate of the ethnic group in the assigned neighborhood at arrival, defining the

individual's ethnic peer group as language fellows who have the same skill level as the individual, distinguishing between two skill levels: Max 10 years of education at arrival versus more than 10 years of education. The idea is to test empirically, whether ethnic networks are stratified by skill level. The results in Panel A for this narrow peer definition shows little effect of the assigned share of peers at the extensive margin, while the negative effect of the employment rate of peers remains significant for men at the intensive margin (Panel B). In cols. 5-6, we estimate whether ethnic network effects are stronger if we restrict the ethnic group to low-skilled language fellows defined as max 10 years of education. This could be the case, because low-skilled members in established enclaves are more likely to have found jobs through informal channels, in particular through ethnic networks (Damm 2014). They may therefore disseminate more information about job vacancies in their ethnic group than high-skilled members. The results in Panel A for this narrow peer definition shows little effect of the assigned share of peers at the extensive margin, while the negative effect of the employment rate of peers remains significant for men at the intensive margin. Moreover, the assigned share of peers has a negative and significant effect for women at the intensive margin. Our take-away from the comparison of estimates across peer definitions is that all language fellows - rather than those with the same skill level and low-skilled language fellows - living in the assigned neighborhood is the better measure of residence-based ethnic networks.

V.C. Mechanisms

Viewed together, the negative effects of the ethnic group share in the assigned neighborhood at the extensive margin and the negative effects of the employment rate of the ethnic group in the assigned neighborhood and the overall employment rate in the assigned municipality at the intensive margin suggest that early entry into employment crowds out formal host-country language investments. To explore this mechanism further, we first estimate the effects of ethnic group characteristics on individual employment (in November) using pooled cross-sectional analysis for year one to four since arrival. For each of the three alternative specifications of the co-ethnic group, we report results using four different specifications in Table 8. The two first specifications estimate the effects of the employment rate of the ethnic group in the assigned neighborhood without controlling for the size (share) of ethnic group, while the two last specifications control for that. As in Tables 6 and 7, the odd numbered columns control for demographic and socioeconomic characteristics of the assigned municipality at arrival, whereas the even numbered columns instead control for municipality of assignment through municipality of assignment dummies. The coefficient estimates of the employment rate of the ethnic group is positive in most of the specifications, though estimated with imprecision and only significant at a 10 percent significance level using a narrow definition of the co-ethnic group. In other words, ten percentage points increase in the employment rate among co-nationals (corresponding to one third of a standard deviation) increases the individual's employment probability by around 0.2 percentage points, corresponding to 1.3%. The coefficient estimate of the employment rate in the assigned municipality also has a positive sign, but it is insignificant across specifications. The

positive signs of the employment rate of the ethnic group and the municipal employment rate at arrival are consistent with employment crowding out formal host-country language progression. However, the ethnic group size has a negative sign across specifications in Table 8, and it is significant using the broader definition of co-ethnic group. In other words, the non-Western immigrant share in the neighborhood at arrival decreases the employment probability one to four years after arrival. Thus, we can rule out that the negative effect of the non-Western immigrant group share on the probability to ever enroll in formal host-country language training is due to the job entry. Consistent with this result, descriptive statistics on employment rates and personal income sources by year four since migration reported in Appendix Table A12 shows that never takers are substantially less likely to have worked by year four than ever takers (22% versus 46%). We can thus rule out for the majority of never takers of formal host-country language training that they learn the host-country on the job which is likely to be an alternative powerful method to learn the host country language. Viewed together, these findings suggest that never takers on average have lower fluency in the host-country language by year four since migration than ever takers. Recalling that take-up of the integration program is a condition for receipt of social assistance, we also calculate the share for whom public transfers constitute their sole source of personal income during the first four years since migration. We find that the share is much higher among never takers and ever takers (70% versus 38%). Since only 5% of never takers have had a child by year four, at most 5% of never takers have been on paternity leave during that period. Since language training constitutes the main element of the integration program, these descriptive statistics suggest that the integration program participation requirement for social assistance receipt is not fully enforced.²⁴

[Table 8: Effects of assigned ethnic group share and employment rate of ethnic group on employment status]

To explore further whether early entry into employment crowds out formal host-country language progression, we estimate the effects of the employment rate of the ethnic group in the assigned neighborhood at arrival on an indicator for being active, using pooled cross-sections for years one to four after arrival. A null effect of the employment rate of the ethnic group in the assigned neighborhood and the municipal employment rate at arrival would be consistent with such crowding out. Recall that active is defined as being enrolled in a Danish language course (in September) and/or having a job (in November). In Table 9, we report the coefficient estimates for each of the three alternative definitions of co-ethnic group, using four different specifications for each definition; the four specifications are identical to those in Table 8. We do in fact find a null effect of both the employment rate of the ethnic group in the assigned neighborhood and the municipal employment rate at arrival, consistent with employment crowding out formal host-

²⁴ Future research should investigate whether incomplete enforcement is due to receipt of social assistance during sick leave.

country language progression. The ethnic group share in the assigned neighborhood at arrival has a negative sign, but it is insignificant across ethnic group definitions and specifications.

[Table 9: Effects of assigned ethnic group share and employment rate of ethnic group on dummy for being active]

Next we investigate another channel through which the ethnic group share in the assigned neighborhood of arrival may decrease the probability of take up of formal host-country language training by year four since arrival, recalling from Table 6 that women are less likely to take up formal host-country language training by four years since arrival. The channel investigated is the fertility channel, i.e., that women who are assigned to a neighborhood with more co-ethnic neighbors are more likely to follow traditional gender roles of having more children instead of having regular work. To explore this potential channel, we estimate the effects of the ethnic group share in the assigned neighborhood on a variable for the number of child births since arrival until year four since arrival, using cross-sectional data for year four after arrival. On average, women in the Estimation Sample have had 0.51 child births by year four, with a standard deviation of 0.73. In other words, every other women in our sample have had a child during the first four years since arrival. A positive and significant effect for women would be consistent with that channel. In Table 10, we report the coefficient estimates of the effects of the ethnic group share in the assigned neighborhood for each of the three alternative definitions of co-ethnic group, using four different specifications for each definition. The specifications are identical to the specifications in Table 6. The results reported in Panel B allows for heterogeneous effects by gender. While the coefficient estimate of the effects of the ethnic group share in the assigned neighborhood on the probability of child birth is negative and insignificant for men across definitions and specifications, it is positive for women, albeit only significant at a 10 percent significance level using the broader definition of the co-ethnic group and only using the last specification. According to the significant estimate, a doubling of the share of non-Western immigrant neighbors at arrival (corresponding to a standard deviation increase) increases the number of child births by around 0.71 (corresponding to a standard deviation increase) around the mean number of child birth by year four for women in the Estimation Sample (i.e., 0.51). Notice further from Table 10 that for women we also find a consistently positive effect of the employment rate of the co-ethnic group in the neighborhood of arrival. The magnitude of the coefficient estimate is smaller for men and only significant using the two narrow definitions of the co-ethnic group. For women, ten percentage points increase in the employment rate of non-Western immigrants (corresponding to three quarters of a standard deviation) increases the annual probability of child birth by around 46 percentage points. The positive effects of the employment rate of co-ethnic neighbors at arrival found for women on both the annual probability of employment (Table 8) and the annual probability of child birth (Table 10) provides a further explanation for why the employment probability of co-ethnic neighbors at arrival has little effect on the annual probability of being active (Table 9).

[Table 10: Effects of assigned ethnic group share and employment rate of ethnic group on fertility in the host-country]

[Table 11: Effects of assigned ethnic group share and employment rate of ethnic group on the Danish language skill requirements in the job]

Summing up, we find negative effects of the ethnic group share in the assigned neighborhood at arrival on enrollment into a Danish language course by year four since arrival and negative effects of the employment rate of the ethnic group in the assigned neighborhood at arrival on the module level completed during year one to four since arrival. By contrast, we find some evidence of positive effects of the employment rate of the ethnic group in the assigned neighborhood at arrival on the annual probability of employment and no effects on being active. Viewed together, the effects of the employment rate of the ethnic group are consistent with employment crowding out formal host-country language investments. This should not be of concern if workers learn the host-country language on the job. To explore whether this is the case, we analyze the Danish skill speaking requirements in the individual's job. The language skill speaking requirements in the occupation is rated on a scale from 0 to 100. On average, workers in the Estimation Sample work in an occupation with Danish language speaking requirements of 20. The most common occupation of individuals in the Estimation Sample is cleaning and it requires a modest level of Danish speaking skills. In this regard, Table 11 investigate the effects of the ethnic group share and the employment rate in the assigned neighborhood at arrival on the level of Danish speaking skills required in the occupation of workers in the Estimation Sample, using pooled cross-sections for years one to four since arrival. On average, men work in occupations which require a lower level of Danish speaking skills than women. For men, the coefficient estimate of the employment rate of the ethnic group is close to zero, while the coefficient estimate of the ethnic group share is positive for the two narrow definitions of the co-ethnic group and large, but still insignificant for the narrow definition of the co-ethnic group. By contrast, for women, we find some evidence that a higher ethnic group share is associated with a lower level of Danish speaking language requirements, although the effect is only significant at a ten percent level. According to the significant estimates, one percentage points increase in the share of language fellows (corresponding to 27% of a standard deviation) decreases the Danish language requirement level by 24.8. For women, the employment rate of co-ethnic neighbors at arrival is associated with a higher level of Danish speaking skill requirements, although the effect is only significant using one of the two narrow definitions of the co-ethnic group. According to the significant estimates, ten percentage points increase in the employment rate of language fellows in the assigned neighborhood at arrival (corresponding to one third of a standard deviation) increases the Danish language speaking skills in the job by around 20, corresponding to a doubling. Overall, the results in Table 11 lend little support to the hypothesis that host-country language training on the job is a substitute for formal host-country language training among individuals who are assigned to neighborhoods with a higher share of their co-ethnic group.

V.D. Effects of current exposure to co-ethnic neighbors on formal host-country investment

As a final analysis, we estimate the effects of current exposure to co-ethnic neighbors on formal host-country investment using an instrumental variable (IV) approach. In case of heterogeneous treatment effects, but monotonic response to treatment, the IV-estimates of the ethnic group share in the current neighborhood of residence can be interpreted as local average treatment effects. If one is willing to assume homogeneous treatment effects, the IV-approach yields estimates of average treatment effects. Due to the combination of low relocation rates from the municipality of assignment and tendency to move to a similar neighborhood within the assigned municipality in case of neighborhood relocation, the ethnic group share in the assigned neighborhood at arrival is highly correlated with the ethnic group share in the neighborhood of residence four years after arrival, and thus a relevant instrument for it. Assuming that the ethnic group share in the assigned neighborhood at arrival only affects the decision to take up of formal host-country language by year four through its effect on the individual's exposure to co-ethnic neighbors, the ethnic group share in the assigned neighborhood of arrival is also a valid instrument for the ethnic group share in the neighborhood of residence in year four. According to Panel B of Table 12, which presents the first-stage regression estimate of the excluded variable in the second stage, the excluded variable is a relevant instrument. The F-test statistic is far above 10, across the three alternative definitions of the co-ethnic group and across specifications. Across definitions and specifications, the 2SLS estimate of the ethnic group share in year four on the probability of having enrolled in a Danish language course by year four is negative, and it is significant at least at a ten percent significance level except for one specification (Panel A of Table 12). Defining the co-ethnic group as non-Western immigrants, the 2SLS estimates show that a doubling of the share of non-Western immigrants in the current neighborhood of residence (corresponding to a standard deviation increase) decreases the probability of ever enrolling in a Danish language course by 4.6 percentage points, or by 4.8%.

[Table 12: IV-estimates of the effects of assigned ethnic group share and employment rate of ethnic group on ever enrollment in formal Danish language course]

VI. DISCUSSION AND CONCLUSIONS

To answer the question of how exposure to more co-ethnic neighbors in the first years after immigration affects host-country language acquisition, this paper draws on an almost ideal spatial allocation experiment with refugees in Denmark in place from 1999 until mid-2016. The quasi-random assignment of newly arrived refugees to Denmark who arrived later during the calendar year after the first municipalities had reached their annual capacity offers a unique solution to the fundamental problem of neighborhood sorting. Despite a very high take-up rate of the formal host-country language training offered at no charge as part of the three-year integration program for refugees during which participants are entitled to means-tested social assistance, we find robust

evidence that the share of co-ethnic neighbors decreases the probability to enroll in formal host-country language training in the first years after immigration. We find this effect irrespective of whether we define the co-ethnic group narrowly as fellow countrymen, or use a broader definition as language fellows or non-Western immigrants. A doubling of the share of non-Western immigrants among neighbors at arrival (corresponding to a standard deviation increase around the mean of 5.1%) decreases the probability of ever enrolling in a host-country language course by around 1 percentage point, corresponding to around 1%. The magnitude of the coefficient estimate is higher for women, but not significantly different. Everything else equal, compared to women, men are 1.5 percentage points more likely to have enrolled by year four since arrival. Everything else equal, men also acquire a higher level of formal host-country language than women. Since men also tend to have higher level of education at arrival, a likely explanation is that the return to host-country language is higher for men than women. Moreover, we do not find positive effects of the share of co-ethnic neighbors on the employment probability. Therefore, we can rule out that the negative effect of the non-Western immigrant group share on the probability to ever enroll in formal host-country language training is due to the job entry. Consistent with this, we find that only 22% of never takers have had regular employment by year four since migration compared to 46% of ever takers. We can thus rule out for the majority of never takers of formal host-country language training that they learn the host-country language on the job. Viewed together, these findings suggest that never takers have lower fluency in the host-country language by year four since migration than ever takers. Recalling that take-up of the integration program is a condition for receipt of social assistance, we investigate to which extent public transfers constitute personal income during the first four years after immigration and find that it constitutes 100% of personal income for 70% of never takers compared to 38% among ever takers. Since only 5% of never takers have a child during that period so that max 5% of never takers have been on paternity leave, our finding that 5% are never takers suggest that the requirement is not fully enforced.

Turning to effects of exposure to co-ethnic neighbors at the intensive margin of formal host-country language acquisition, we find no effect of the share of co-ethnic neighbors on the acquired level of formal host-country language in year one to four since arrival. However, for men, we find robust evidence of negative effects of the quality (measured by employment rate) of the co-ethnic network at the acquired level of formal host-country language; the magnitude of the effect for women is smaller and imprecisely estimated. For men, 30 percentage points decrease in the employment rate of language fellows (corresponding to a standard deviation) increases the module level completed by around 2% around the mean of the sample distribution. Depending on the course level enrolled, participants can complete at most five or six modules. Similarly, we find negative effects of the overall employment rate in the assigned municipality on the acquired level of formal host-country language. That is, 4 percentage points decrease in the employment rate in the assigned municipality at arrival (corresponding to a standard deviation decrease around the mean of 76.9%) increases the completed module level by 4% around the mean of the sample distribution.

Our investigation of potential mechanisms suggests that the quality of the residence-based co-ethnic network at arrival promotes immigrant employment one to four years since arrival, irrespective of gender. Exploiting exogenous variation in neighborhood characteristics for (only) male refugees in Denmark who were quasi-randomly assigned to neighborhoods during the first Danish Spatial Dispersal on refugees 1986-1998, Damm (2014) finds that an increase in the quality of co-ethnic neighbors promotes both immigrant employment and annual earnings two to six years after arrival. Our study, thus, adds to the study by Damm (2014) that the quality of the residence-based co-ethnic network at arrival promotes not only employment of male refugees, but also of female refugees. Our investigation of potential mechanisms lends support to the hypothesis that better local job prospects promotes early job entry but crowds out formal host-country language investments.

Overall, our results lend support to the theory proposed by Lazear (1999) that living in an ethnic enclave reduces the individual's incentive to invest in host-country language acquisition because of increased chances of finding a job through the co-ethnic network. In the Danish context of a low or negative correlation between the co-ethnic group share and the employment rate of the co-ethnic group, housing policies which assign newly arrived immigrants to housing outside immigrant-dense neighborhoods (e.g. spatial dispersal policies on refugees) promote participation in formal host-country language training, but lower host-country language progression because of increased chances of finding a job through the co-ethnic network which crowds out formal host-country language investments. By offering host-country language training for free as part of the three-year introduction program for newly recognized refugees and making eligibility to means-tested social assistance conditional on participation in the three-year introduction program and by spatial dispersal of refugees across municipalities with relatively low shares of non-Western immigrants, the Danish Integration Act achieved very high take-up of host-country language training, but our results suggest that there may be a need for improvement of language courses to facilitate individual participation in host-country language training after having found regular employment.

Having a job may be a powerful method to learn the language, if the working language is the host-country language. However, due to lack of data on host-country language acquired informally on the job, our study cannot fully answer the question of whether host-country language training on the job substitutes for formal host-country language training. Without such data we cannot investigate whether formal host-country language training delivers better host-country language skills than informal host-country language training on the job. Having that said, our study does provide descriptive evidence that the most common occupations of refugees during the first four years after asylum have low requirements for host-country language speaking skills. Moreover, for women, our regression analysis suggest that a higher ethnic group share is associated with lower host-country language speaking requirements, while the effect is insignificant for men. These two findings suggest that the negative effect of the ethnic group share on the ever take-up rate of formal host-country language training is not outweighed by a positive effect of the ethnic group share on informal host-country language training on the job. Recent findings show that the

Danish 1999-introduction program did improve refugees' long-term employment (Arendt et al. 2021). However, since it is still the case in Denmark and Norway that the employment gap between natives and refugees widens 6-10 years after asylum (Schultz-Nielsen 2017; Bratsberg et al. 2017), future research should investigate how the introduction program should be designed to further promote long-term employment of refugees.

Having a job may also be a powerful way for cultural integration as long as immigrants are exposed to natives (as co-workers or customers or both). A related question of relevance for design of integration policies for new immigrants is whether refugees acquire faster cultural integration in the jobs held early on than through participation in an introduction program which consists of class-room training in the host-country culture and host-country language. This is another question for future research.

References

- Adserà, Alícia and Mariola Pytlikova (2015), "The Role of Language in Shaping International Migration", *Economic Journal*, 125 (Aug.), F49-F81.
- Angrist, Joshua (2001), "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors", *Journal of Business and Economic Statistics*, 19(1), 2-28.
- Arendt, Jacob N., Iben Bolvig, Mette Foged, Linea Hasager, and Giovanni Peri, (2020) "Integrating Refugees: The Role of Language Training and Work Incentives", UC Davis WP.
- Arendt, Jacob N., Iben Bolvig, Mette Foged, Linea Hasager, and Giovanni Peri, (2021) "Language Training and Refugees' Integration", NBER WP 26834.
- Auer, D. (2018), "Language roulette – the effect of random placement on refugees' labour market integration", *Journal of Ethnic and Migration Studies*, 44(3), 341–362.
- Azlor, Luz, Anna P. Damm and Marie Louise Schultz-Nielsen (2020), "Local Labour Demand and Immigrant Employment", forthcoming in *Labour Economics*.
- Bauer, Thomas, Gil S. Epstein and Ira N. Gang, I. (2005). "Enclaves, language, and the location choice of migrants", *Journal of Population Economics*, 18(4), 649–62.
- Beaman, Lori A. (2012), "Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S.", *Review of Economic Studies*, 79, 128-161.
- Becker, Gary S., and Chiswick, Barry R. (1966), "Education and the Distribution of Earnings." *American Economic Review Papers and Proceedings*, 56 (May), 358-69.

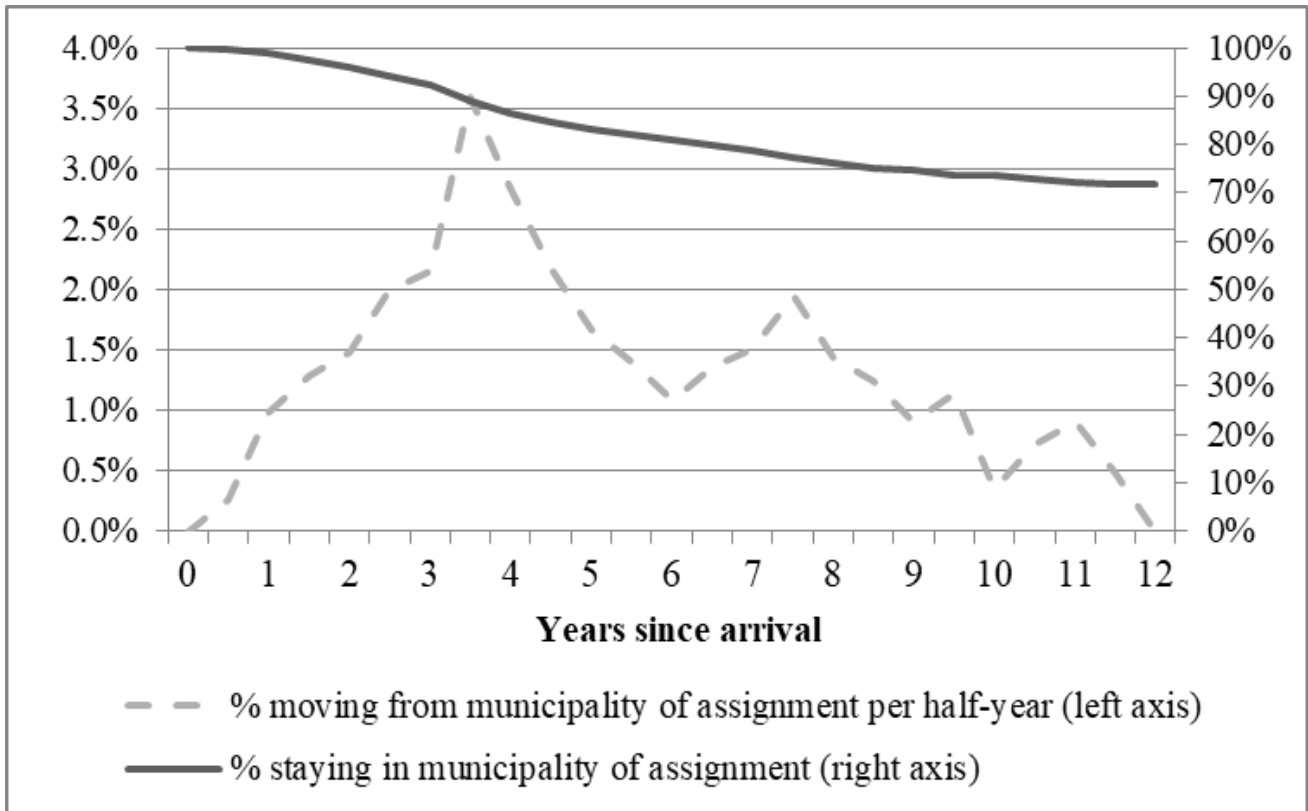
- Beckhusen, Julia, Raymond J.G.M. Florax, Thomas de Graaf, Jacques Poot and Brigitte Waldorf (2013), "Living and working in ethnic enclaves: English Language proficiency of immigrants in US metropolitan areas", *Papers in Regional Science*, 92(2), 305-328.
- Beenstock, Michael (1993), "Learning Hebrew and Finding a Job: An Econometric Analysis of Immigrant Absorption in Israel", Maurice Falk Institute for Economic Research in Israel, discussion paper 93.05, Israel.
- Berman, Eli, Kevin Lang and Erez Siniver (2003) "Language-skill complementarity: Returns to immigrant language acquisition". *Labour Economics*, (10), 265–290.
- Bertrand, Marianne, Erzo F.P. Luttmer and Sendhil Mullainathan. 2000. "Network Effects and Welfare Cultures". *Quarterly Journal of Economics* 115: 1019-1055.
- Bisin, Alberto and Thierry Verdier (2000). "Beyond the melting pot: cultural transmission, marriage, and the evolution of ethnic and religious traits". *Quarterly Journal of Economics*, 115, 955–988.
- Bleakley, Hoyt and Aimee Chin (2004), "Language skills and earnings: evidence from childhood immigrants", *Review of Economics and Statistics*, 84(2), 481–96.
- Bleakley, Hoyt and Aimee Chin (2010), "Age at arrival, English proficiency, and social assimilation among US immigrants", *American Economic Journal: Applied Economics*, 2(1), 165–92.
- Borjas, George J. (1985), "Assimilation, Changes in Cohort Quality and the earnings of immigrants", *Journal of Labor Economics*, 3(4), 463–489.
- Borjas, George J. (1992), "Ethnic Capital and Intergenerational Mobility", *Quarterly Journal of Economics* 107(1), 123-150.
- Borjas, George J. (1995), "Ethnicity, Neighborhoods, and Human-Capital Externalities", *American Economic Review* 85(3), 365-390.
- Bratsberg, Bernt, Oddbjørn Raaum and Knut Røed (2017), "Immigrant labor market integration across admission classes", *Nordic Economic Policy Review*, 17-54.
- Chiswick, Barry R. (1978), "The Effect of Americanization on the Earnings of Foreign-born Men", *Journal of Political Economy*, 86(5), 897-921.
- Chiswick, Barry R. and Paul W. Miller (1992), "Language in the labor market. The immigrant experience in Canada and the United States". In: Chiswick, BR (ed) *Immigration, language and ethnic issues: Canada and the United States*. American Enterprise Institute, Washington DC, pp 229-296
- Chiswick, Barry R. and Paul W. Miller (1996), "Ethnic networks and language proficiency among immigrants", *Journal of Population Economics*, 9, 19–35.

- Chiswick, Barry R. and Paul W. Miller (2002) “Immigrant earnings: Language skills, linguistic concentrations and the business cycle”, *Journal of Population Economics* 15: 31–57.
- Constant, A.F. and K.F. Zimmermann (2008), “Measuring ethnic identity and its impact on economic behavior”, *Journal of European Economic Association*, 6, 424–433.
- 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.
- Damm, A.P. (2005), “The Danish Dispersal Policy on Refugee Immigrants 1986-1998: A Natural Experiment?”, Aarhus School of Business, Department of Economics WP 05-3.
- Damm, A.P. (2009), “Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence”, *Journal of Labor Economics*, 27(2), 281-314.
- Damm, A.P. (2014), “Neighborhood Quality and Labor Market Outcomes: Evidence from Quasi-Random Neighborhood Assignment of Immigrants”, *Journal of Urban Economics*, 79, 139-166.
- Damm, Anna P., Hassani, Ahmad and Marie L. Schultz-Nielsen, “Residential Segregation in Denmark during 1986-2016”, in: Hassani, A. “Residential Segregation and Public Housing”, PhD dissertation, Aarhus BSS, Aarhus University, Department of Economics and Business Economics, 2019, Ch. 2.
- Damm, Anna P., Hassani, Ahmad and Marie L. Schultz-Nielsen, (2021a), “Data construction of Danish Residential Neighbourhoods for 1986-2016”, forthcoming in *Danish Journal of Economics (Nationaløkonomisk Tidsskrift)*.
- Damm, A.P., A. Hassani, T. S. H. Jensen, M. Schindler, M.L. Schultz-Nielsen, (2021b), “Refugees’ Settlement Pattern, Educational Background and Formal Danish Language Acquisition“, forthcoming as RFF WP.
- Danzer, Alexander M. and Firat Yaman (2016) “Ethnic concentration and language fluency of immigrants: Evidence from the guest-worker placement in Germany”, *Journal of Economic Behavior & Organization*, 131, 151-165.
- Dustmann, Christian (1993), “Earnings Adjustment of Temporary Migrants”, *Journal of Population Economics*, 6, 153-168.
- Dustmann, Christian (1994), “Speaking fluency, writing fluency and earnings of migrants in Germany”. *Journal of Population Econ* 7, 133-156.
- Dustmann, Christian (1996), “Return migration—the European experience”. *Economic Policy*, 11(22), 214–250.
- Dustmann, Christian (1999), “Temporary Migration, Human Capital, and Language Fluency of Migrants”, *Scandinavian Journal of Economics*, 101(2), 297-314.

- Dustmann, Christian and Francesca Fabbri, (2003), “Language proficiency and labour market performance of immigrants in the UK”, *Economic Journal*, 113(489), 695–717.
- Dustmann, C. and A. van Soest (2002), “Language and the Earnings of Immigrants”, *ILR Review*, 55(3), 473-492.
- Dustmann, C., K. Vasiljeva and A.P. Damm (2019), “Refugee Migration and Electoral Outcomes”, *Review of Economic Studies*, 86, 2035-2091.
- Edin, P-A, P. Fredriksson and O. Åslund (2003), “Ethnic Enclaves and the Economic Success of Immigrants - Evidence from a Natural Experiment”, *Quarterly Journal of Economics*, 118, 329-357.
- Fasani, Francesco, Tommaso Frattini and Luigi Minale (2018), “(The Struggle for) Refugee Integration into the Labour Market: Evidence from Europe”, IZA Discussion Paper 11333.
- Granovetter, Mark. 1973. The Strength of Weak Ties, *American Journal of Sociology*, vol. 78(6) (May):1360-80.
- Granovetter, Mark. 1974, 1995. *Getting a Job: A Study of Contacts and Careers*, Chicago: University of Chicago Press.
- Hasager, Linea and Mia Jørgensen. 2021. “Sick of Your Poor Neighborhood?”. CEBI WP 02/21.
- Hatton, Timothy (2016), “Refugees, Asylum Seekers, and Policy in the OECD Countries”, *American Economic Review: Papers & Proceedings*, 106(5), 441-445.
- Hvidtfeldt, Camilla and Marie L. Schultz-Nielsen (2018), “Refugees and Asylum seekers in Denmark 1992-2016”, The ROCKWOOL Foundation Research Unit, Study Paper 133.
- Lalonde, R.J. and R.H. Topel, (1992), “The Assimilation of Immigrants in the U.S. Labor Market,” In: Borjas, George J. and Richard B. Freeman (eds.) “*Immigration and the Workforce: The Economic Consequences for the United States and Source Areas*”, Chicago and London: University of Chicago Press, 67-92.
- Lazear, Edward P. (1999), “Culture and Language”, *Journal of Political Economy*, 107(56), S95-S126.
- Lubotsky, D. (2007), “Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings”, *Journal of Political Economy*, 115(5), 820-867.
- Ministry for Children, Equality, Integration and Social Affairs [Ministeriet for Børn, Ligestilling, Integration og Sociale Forhold]. 2015. Letter of information on housing assignment of new refugees [Informationsbrev om boligplacering af nye flygtninge]. 6th of Feb. 2015. Case number 2015-418.
- Montgomery, James D. 1991. Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. *American Economic Review*, vol. 81(5) (Dec.): 1408-1418.

- Munshi, K. (2003), “Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market”, *Quarterly Journal of Economics*, 118(2), 549-599.
- Nielsen, Chantal P. and Kræn B. Jensen (2006), “The Danish Integration Act’s Significance for the Settlement Patterns of Refugees”, Copenhagen: Institute for Local Government Studies.
- Pedersen, Peder J., Mariola Pytlikova and Nina Smith (2008), “Selection and network effects – migration flows into OECD countries, 1990-2000”, *European Economic Review*, 52(7), 1160–86.
- Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt (2017), “Poorly Measured Confounders Are More Useful On the Left than on the Right”, NBER WP 23232.
- Pellizari, Michele. 2010. Do Friends and Relatives really Help in Getting a Good Job? *Industrial and Labor Relations Review*, vol. 63(3): 494-510.
- Rees. Albert. 1966. Labor Economics: Effects of More Knowledge. Information Networks in Labor Market. *American Economic Review*, (Papers and Proceedings), vol. 56 (½) (Mar. 1.): 559-566.
- Sarvimäki, Matti (2011), “Assimilation to a Welfare State: Labor Market Performance and Use of Social Benefits by Immigrants to Finland”, *Scandinavian Journal of Economics* 113(3), 665–688.
- Schultz-Nielsen, Marie L. (2017), “Labour market integration of refugees in Denmark”, *Nordic Economic Policy Review*, 55-90.
- Åslund, Olof (2005), “Now and forever? Initial and subsequent location choices of immigrants”, *Regional Science and Urban Economics*, 35,141–165.
- Åslund, Olof, and Dan-Olof Rooth (2007), “Do Where and When Matter? Initial Labour Market Conditions and Immigrant Earnings”, *Economic Journal*, 117(518), 422-448.
- Åslund, Olof, Per-Anders Edin, Peter Fredriksson and Hans Grönqvist (2011), “Peers, Neighborhoods, and Immigrant Student Achievement”, *American Economic Journal: Applied Economics*, 3(2), 67-95.

Figure 1 Geographical stability of settlement for refugees arriving 2004-2015

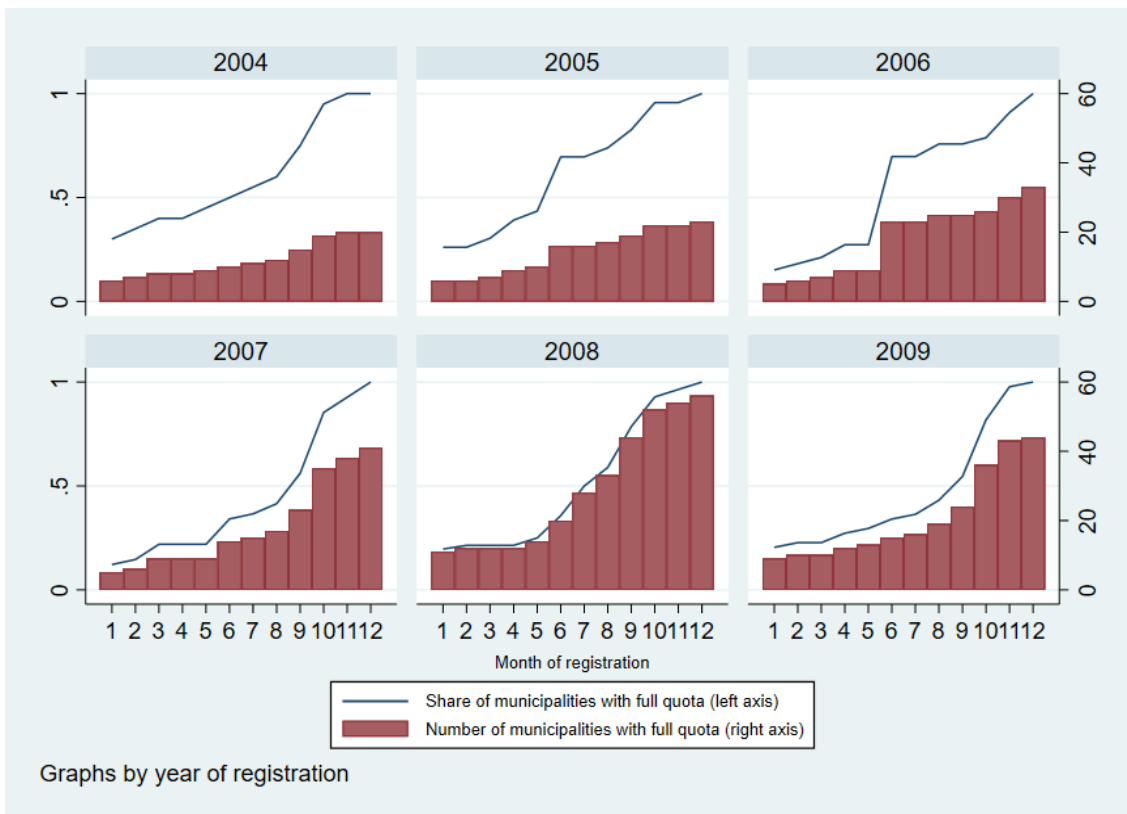


Source : Administrative register information from Statistics Denmark 1997-2020.

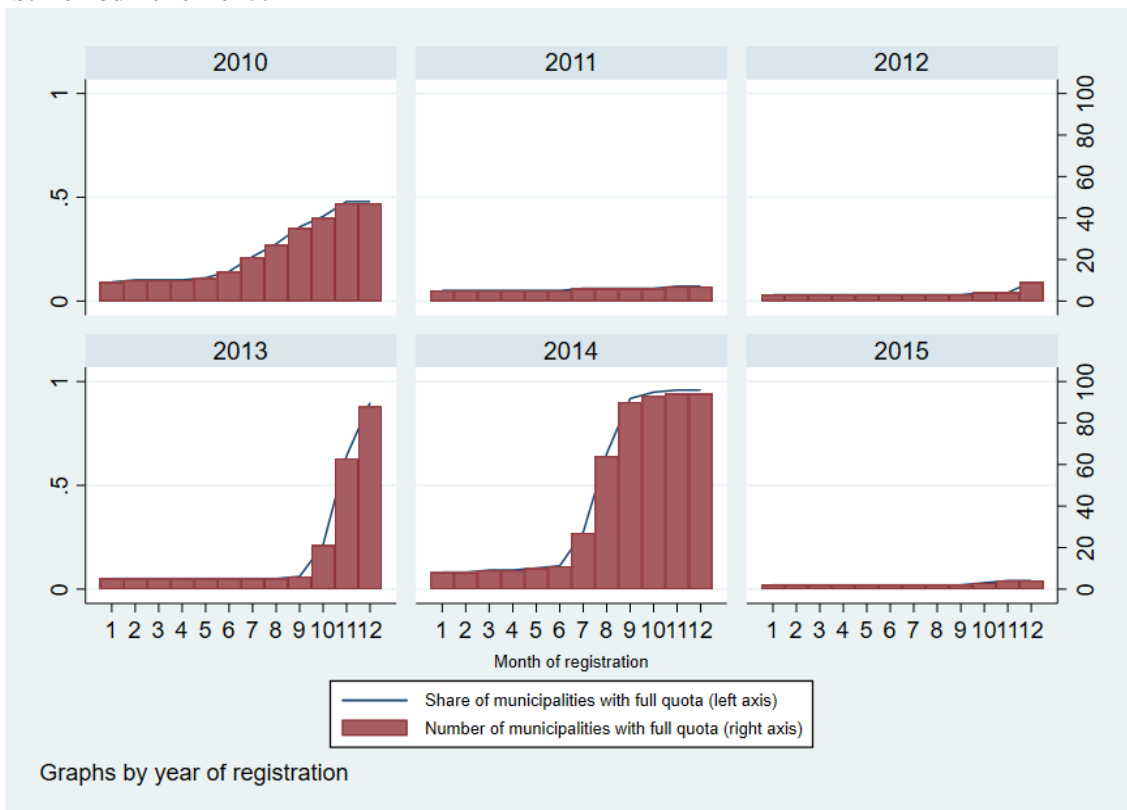
Note : Calculation is based on a gross sample of refugees including all adult refugees arriving to Denmark from 2004-2015. The out-migration rate from the municipality of assignment is calculated as Kaplan-Meier empirical hazard rates. The rate of staying in the municipality of assignment is calculated as Kaplan-Meier empirical survivor rates.

Figure 2 Cumulative distribution function for municipalities that have met their annual quota of refugees by a given calendar month. Separate figures for each calendar year, 2004-2015.

a. Period 2004-2009.



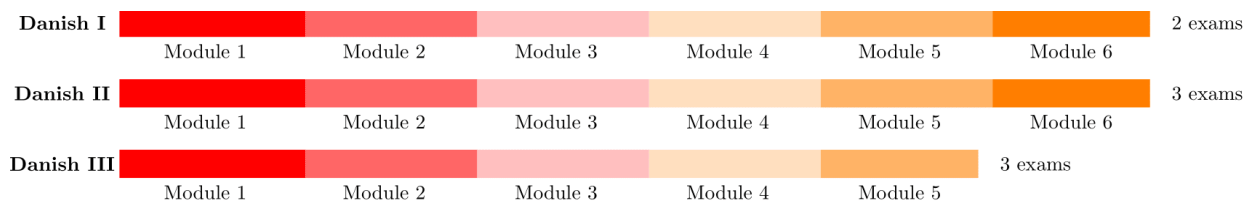
b. Period 2010-2015.



Source : Administrative register information from Statistics Denmark from 1997-2020.

Note : Sample: The gross sample of refugee household heads who got asylum during 2004-2015.

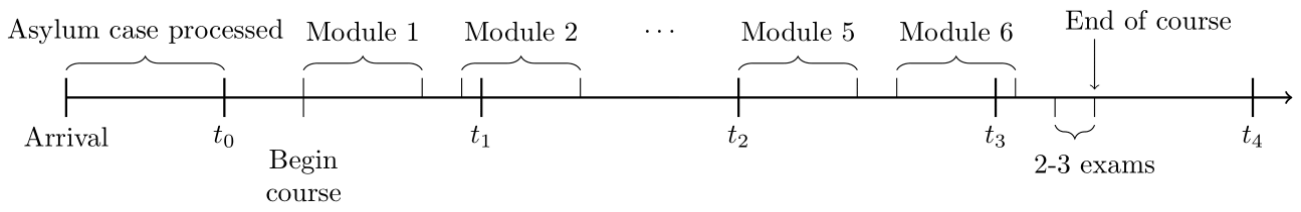
Figure 3 Structure of formal Danish language courses



Source : Authors' construction.

Notes : Duration of modules not to scale.

Figure 4 Timeline of a typical adult refugee's first four years in the country



Source : Authors' construction.

Notes : Duration of modules not to scale.

Table 1 Summary Statistics: Individual Characteristics at arrival. Overall and by Gender. Mean (standard deviation).

	Balanced Panel			Subsample of Balanced panel			Individuals in balanced panel who are excluded from the Subsample			Estimation Sample		
	Total (1)	Female (2)	Male (3)	Total (4)	Female (5)	Male (6)	Total (7)	Female (8)	Male (9)	Total (10)	Female (11)	Male (12)
Individual demographic characteristics:												
Men (0/1)	0.818 (0.386)	0.000 (0.000)	1.000 (0.000)	0.834 (0.372)	0.000 (0.000)	1.000 (0.000)	0.811 (0.392)	0.000 (0.000)	1.000 (0.000)	0.731 (0.443)	0.000 (0.000)	1.000 (0.000)
Age at Arrival	30.897 (8.99)	31.077 (9.24)	30.858 (8.89)	31.268 (9.17)	31.708 (10.46)	31.181 (8.89)	30.723 (8.788)	30.816 (9.672)	30.701 (8.569)	31.421 (8.72)	32.077 (9.16)	31.181 (8.58)
Marital status at Arrival	0.401 (0.490)	0.302 (0.459)	0.423 (0.494)	0.382 (0.486)	0.218 (0.413)	0.415 (0.493)	0.410 (0.492)	0.337 (0.473)	0.427 (0.495)	0.455 (0.498)	0.565 (0.496)	0.415 (0.493)
Children Aged 0-2 at Arrival (0/1)	0.089 (0.284)	0.138 (0.345)	0.078 (0.268)	0.086 (0.281)	0.138 (0.345)	0.076 (0.265)	0.090 (0.286)	0.139 (0.346)	0.079 (0.269)	0.123 (0.328)	0.250 (0.433)	0.076 (0.265)
Children Aged 3-17 at Arrival (0/1)	0.174 (0.379)	0.295 (0.456)	0.147 (0.354)	0.181 (0.385)	0.340 (0.474)	0.149 (0.357)	0.171 (0.376)	0.276 (0.447)	0.146 (0.353)	0.246 (0.431)	0.509 (0.500)	0.149 (0.357)
Level of education:												
Missing education (0/1)	0.384 (0.486)	0.453 (0.498)	0.368 (0.482)	0.406 (0.491)	0.485 (0.500)	0.390 (0.488)	0.373 (0.484)	0.440 (0.497)	0.358 (0.479)	0.417 (0.493)	0.489 (0.500)	0.390 (0.488)
<= 10 years of education	0.360 (0.480)	0.337 (0.473)	0.366 (0.482)	0.332 (0.471)	0.304 (0.460)	0.337 (0.473)	0.374 (0.484)	0.351 (0.477)	0.379 (0.485)	0.328 (0.469)	0.301 (0.459)	0.337 (0.473)
11-13 years of education	0.117 (0.322)	0.094 (0.292)	0.123 (0.328)	0.119 (0.324)	0.093 (0.291)	0.125 (0.330)	0.117 (0.321)	0.095 (0.293)	0.122 (0.327)	0.118 (0.322)	0.099 (0.299)	0.125 (0.330)
13 < years of education	0.139 (0.345)	0.116 (0.320)	0.144 (0.351)	0.143 (0.350)	0.119 (0.324)	0.148 (0.355)	0.136 (0.343)	0.114 (0.318)	0.142 (0.349)	0.138 (0.345)	0.111 (0.314)	0.148 (0.355)
No. of obs.	15,992	2,906	13,086	5,113	850	4,263	10,879	2,056	8,823	5,828	1,565	4,263

Table 1 (continued) Summary Statistics: Individual Characteristics at arrival. Overall and by Gender. Mean (standard deviation).

Official language of source country:												
Arabic	0.501 (0.500)	0.312 (0.463)	0.542 (0.498)	0.445 (0.497)	0.219 (0.414)	0.490 (0.500)	0.527 (0.499)	0.351 (0.477)	0.567 (0.495)	0.422 (0.494)	0.235 (0.424)	0.490 (0.500)
Burmese	0.028 (0.164)	0.021 (0.142)	0.029 (0.169)	0.071 (0.256)	0.054 (0.226)	0.074 (0.262)	0.008 (0.086)	0.007 (0.082)	0.008 (0.087)	0.080 (0.272)	0.098 (0.298)	0.074 (0.262)
Farsi	0.070 (0.254)	0.072 (0.259)	0.069 (0.253)	0.083 (0.276)	0.080 (0.271)	0.083 (0.276)	0.063 (0.244)	0.069 (0.254)	0.062 (0.241)	0.080 (0.271)	0.071 (0.257)	0.083 (0.276)
French	0.020 (0.142)	0.044 (0.204)	0.015 (0.123)	0.037 (0.190)	0.082 (0.275)	0.028 (0.166)	0.013 (0.111)	0.028 (0.164)	0.009 (0.094)	0.040 (0.196)	0.071 (0.257)	0.028 (0.166)
Russian	0.021 (0.143)	0.038 (0.192)	0.017 (0.129)	0.031 (0.175)	0.066 (0.248)	0.025 (0.155)	0.016 (0.124)	0.027 (0.161)	0.013 (0.114)	0.034 (0.182)	0.060 (0.238)	0.025 (0.155)
Other language	0.344 (0.475)	0.492 (0.500)	0.311 (0.463)	0.305 (0.460)	0.465 (0.499)	0.273 (0.446)	0.363 (0.481)	0.503 (0.500)	0.330 (0.470)	0.310 (0.463)	0.411 (0.492)	0.273 (0.446)
No official language	0.017 (0.129)	0.021 (0.145)	0.016 (0.125)	0.028 (0.164)	0.034 (0.182)	0.026 (0.160)	0.012 (0.108)	0.016 (0.126)	0.011 (0.103)	0.034 (0.181)	0.054 (0.227)	0.026 (0.160)
Official Danish language acquisition:												
Ever enrolled in DUB	0.971 (0.167)	0.957 (0.204)	0.974 (0.158)	0.955 (0.207)	0.927 (0.260)	0.961 (0.194)	1.596 (1.949)	1.573 (1.932)	1.602 (1.953)	0.949 (0.220)	0.918 (0.275)	0.961 (0.194)
Highest module level completed	2.37 (2.004)	2.235 (1.978)	2.4 (2.008)	2.174 (2.001)	2.012 (1.985)	2.207 (2.002)	2.462 (1.998)	2.327 (1.968)	2.493 (2.004)	2.132 (2.003)	1.929 (1.989)	2.207 (2.002)
Employed (0/1)	0.277 (0.447)	0.109 (0.311)	0.314 (0.464)	0.285 (0.451)	0.124 (0.329)	0.317 (0.465)	0.273 (0.445)	0.102 (0.302)	0.313 (0.464)	0.262 (0.440)	0.112 (0.315)	0.317 (0.465)
Either employed or in language course	0.868 (0.339)	0.836 (0.371)	0.875 (0.331)	0.831 (0.375)	0.801 (0.400)	0.838 (0.369)	0.895 (0.307)	0.860 (0.347)	0.903 (0.296)	0.821 (0.383)	0.784 (0.412)	0.838 (0.369)
No. of obs.	15,992	2,906	13,086	5,113	850	4,263	10,879	2,056	8,823	5,828	1,565	4,263

Source: Danish administrative registers for the period 1997-2020.

Table 2 Summary Statistics: Characteristics of the Assigned Municipality (reported in shares). Mean (standard deviation).

	Balanced Panel	Subsample of Balanced Panel	Individuals in Balanced Panel who are excluded from the Subsample	Estimation Sample	t-test of the difference of (2) and (3)	
	(1)	(2)	(3)	(4)	(5)	
Share of population	0.013 (0.011)	0.011 (0.009)	0.013 (0.011)	0.011 (0.009)	-8.903	***
Share of immigrant and descendents	0.077 (0.031)	0.068 (0.029)	0.081 (0.031)	0.067 (0.029)	-25.496	***
Employment rate	0.762 (0.038)	0.768 (0.040)	0.759 (0.037)	0.769 (0.040)	13.552	***
Co-national enclave characteristics:						
Share of co-nationals	0.002 (0.002)	0.001 (0.002)	0.002 (0.002)	0.001 (0.002)	-8.827	***
Employment rate of co-nationals	0.221 (0.218)	0.281 (0.236)	0.193 (0.204)	0.283 (0.238)	24.290	***
Employment rate of high-skilled co-nationals	0.232 (0.242)	0.291 (0.263)	0.204 (0.227)	0.292 (0.264)	21.410	***
Employment rate of low-skilled co-nationals	0.157 (0.187)	0.207 (0.217)	0.133 (0.166)	0.208 (0.219)	23.697	***
Co-language enclave characteristics:						
Share of co-language individuals	0.011 (0.011)	0.007 (0.009)	0.012 (0.012)	0.007 (0.009)	-27.431	***
Employment rate of co-language individuals	0.367 (0.148)	0.387 (0.176)	0.357 (0.131)	0.388 (0.181)	12.122	***
Employment rate of high-skilled co-language individuals	0.421 (0.157)	0.428 (0.189)	0.418 (0.138)	0.426 (0.195)	3.813	***
Employment rate of low-skilled co-language individuals	0.248 (0.141)	0.275 (0.175)	0.235 (0.119)	0.275 (0.179)	17.082	***
Non-western enclave characteristics:						
Share of non-Western immigrants	0.034 (0.014)	0.030 (0.013)	0.035 (0.014)	0.030 (0.013)	-21.687	***
Employment rate of non-Western immigrants	0.501 (0.054)	0.515 (0.064)	0.494 (0.047)	0.516 (0.065)	23.482	***
Employment rate of high-skilled non-Western individuals	0.549 (0.050)	0.556 (0.061)	0.545 (0.043)	0.557 (0.062)	13.300	***
Employment rate of low-skilled non-Western individuals	0.394 (0.068)	0.422 (0.078)	0.381 (0.059)	0.423 (0.078)	36.267	***
No. of obs.	15,992	5,113	10,879	5,828		

Source : Administrative registers from Statistics Denmark 1997-2020.

Table 3 Summary Statistics: Characteristics of the Assigned Neighborhood (reported in shares). Mean (standard deviation).

	Balanced Panel	Subsample of Balanced Panel	Individuals in Balanced Panel who are excluded from the Subsample	Estimation Sample	t-test of the difference of (2) and (3)	
	(1)	(2)	(3)	(4)	(5)	
Share of population	0.065 (0.059)	0.069 (0.073)	0.062 (0.052)	0.069 (0.072)	6.85	***
Share of immigrants and descendents	0.107 (0.094)	0.099 (0.093)	0.111 (0.094)	0.099 (0.094)	-7.480	***
Employment rate	0.719 (0.101)	0.726 (0.100)	0.716 (0.101)	0.727 (0.100)	6.241	***
Co-national enclave characteristics:						
Share of co-nationals	0.004 (0.009)	0.003 (0.008)	0.004 (0.009)	0.003 (0.008)	-4.172	***
Employment rate of co-nationals	0.132 (0.254)	0.180 (0.297)	0.109 (0.228)	0.179 (0.296)	16.757	***
Employment rate of high-skilled co-nationals	0.115 (0.258)	0.155 (0.299)	0.097 (0.234)	0.154 (0.298)	13.218	***
Employment rate of low-skilled co-nationals	0.085 (0.212)	0.113 (0.251)	0.072 (0.190)	0.112 (0.250)	11.452	***
Co-language enclave characteristics:						
Share of co-language individuals	0.019 (0.037)	0.015 (0.035)	0.021 (0.038)	0.014 (0.036)	-9.290	***
Employment rate of co-language individuals	0.327 (0.259)	0.328 (0.290)	0.326 (0.244)	0.323 (0.292)	0.39	
Employment rate of high-skilled co-language individuals	0.349 (0.287)	0.337 (0.317)	0.355 (0.271)	0.330 (0.319)	-3.630	***
Employment rate of low-skilled co-language individuals	0.209 (0.251)	0.207 (0.273)	0.210 (0.240)	0.204 (0.273)	-0.692	
Non-Western enclave characteristics:						
Share of non-Western immigrants	0.053 (0.055)	0.051 (0.055)	0.054 (0.055)	0.051 (0.056)	-3.986	***
Employment rate of non-Western immigrants	0.485 (0.124)	0.499 (0.128)	0.478 (0.122)	0.500 (0.129)	10.340	***
Employment rate of high-skilled non-Western individuals	0.525 (0.128)	0.536 (0.135)	0.519 (0.124)	0.535 (0.136)	7.516	***
Employment rate of low-skilled non-Western individuals	0.382 (0.145)	0.406 (0.156)	0.371 (0.138)	0.408 (0.158)	14.182	***
No. of obs.	15,992	5,113	10,879	5,828		

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Table 4 Balancing test: Assigned neighborhood characteristics and individual characteristics of assignee (households heads)

	Dependent variables (assigned neighbourhood characteristics at arrival)												
	Population share	Co-national share	Share of language fellows	Non-Western immigrant share	Employment rate among co-nationals	Employment rate among language fellows	Employment rate among non-Western immigrants	Employment rate among high-skilled co-nationals	Employment rate among high-skilled language fellows	Employment rate among high-skilled non-Western immigrants	Employment rate among low-skilled co-nationals	Employment rate among low-skilled language fellows	Employment rate among low-skilled non-Western immigrants
Panel A: Balanced panel													
(Ref. cat.: 0-10 yrs. educ.)													
11-13 yrs. of educ.	0.000403 (0.00131)	-0.0000695 (0.000161)	0.000806 (0.00105)	0.000362 (0.00137)	0.00181 (0.00623)	-0.0000156 (0.00626)	-0.000232 (0.00321)	-0.00125 (0.00540)	-0.000704 (0.000664)	-0.000722 (0.00358)	0.00605 (0.00520)	0.00701 (0.00697)	0.000742 (0.00360)
More than 13 yrs. of educ.	0.00334* (0.00179)	-0.000505** (0.000201)	-0.00159 (0.000981)	-0.00296** (0.00141)	0.00177 (0.00657)	0.00648 (0.00645)	0.00765** (0.00327)	-0.00493 (0.00650)	0.00587 (0.00705)	0.00600* (0.00347)	0.00123 (0.00592)	-0.00246 (0.00773)	0.00793** (0.00369)
R ²	0.204	0.208	0.158	0.164	0.218	0.169	0.072	0.227	0.180	0.061	0.127	0.116	0.068
No. of obs.							15,992						
Panel B: Subsample of refugee household heads arrived after the first 15 municipalities had filled their annual quota													
(Ref. cat.: 0-10 yrs. Educ.)													
11-13 yrs. of educ.	-0.00254 (0.00380)	-0.000110 (0.000248)	0.00107 (0.00161)	-0.000780 (0.00209)	0.00538 (0.0142)	-0.00142 (0.0146)	0.00172 (0.00652)	-0.00299 (0.0131)	-0.00400 (0.0148)	-0.000420 (0.00704)	0.0182 (0.0116)	0.0194 (0.0133)	0.00751 (0.00818)
More than 13 yrs. of educ.	0.00246 (0.00418)	-0.000228 (0.000323)	-0.000124 (0.00158)	-0.00419* (0.00227)	-0.0000938 (0.0141)	-0.00430 (0.0124)	0.00608 (0.00602)	-0.00843 (0.0137)	-0.00922 (0.0145)	-0.000279 (0.00637)	0.00474 (0.0112)	0.00474 (0.0112)	0.0163** (0.00738)
R ²	0.172	0.227	0.184	0.201	0.214	0.174	0.102	0.227	0.193	0.094	0.144	0.134	0.193
No. of obs.							5,113						
Controls:													
Country of origin F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year of arrival F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Month of arrival F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Demographic characteristics of assigned municipality at arrival	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: Significance level: *, p<0.1, **, p<0.05, ***, p<0.01. Reported coefficients are based on linear regression of neighbourhood characteristics in the year of arrival on individual characteristics in the year of arrival. Standard errors clustered by municipality of assignment in parentheses. Individuals in the Subsample used in Panel B have been assigned to 94 different municipalities and 914 different neighborhoods. See Table A1 for a detailed description of sample selection criteria. Demographic characteristics of assigned municipality are population share, non-Western immigrant share and co-nationals. Additional control: Dummy for missing information on educational attainment at arrival.

Table 5 Summary Statistics on individual outcomes by years since migration

	Estimation sample		
	All (1)	Men (2)	Women (3)
Panel A: Enrolled in Danish language course in Sept.			
Years since migration:			
1	0.800 (0.400)	0.836 (0.371)	0.702 (0.458)
2	0.680 (0.466)	0.698 (0.459)	0.633 (0.482)
3	0.366 (0.482)	0.350 (0.477)	0.409 (0.492)
4	0.149 (0.356)	0.133 (0.339)	0.194 (0.396)
Ever enrolled by year 4	0.949 (0.220)	0.961 (0.194)	0.918 (0.275)
Panel B: Employed			
Years since migration:			
1	0.146 (0.353)	0.178 (0.383)	0.059 (0.236)
2	0.293 (0.455)	0.360 (0.480)	0.112 (0.315)
3	0.392 (0.488)	0.476 (0.500)	0.162 (0.369)
4	0.449 (0.497)	0.538 (0.499)	0.209 (0.407)
Ever employed by year 4	0.586 (0.493)	0.695 (0.461)	0.291 (0.454)
Panel C: Enrolled in language course in Sept. and employed in Nov.			
Years since migration:			
1	0.108 (0.311)	0.133 (0.340)	0.039 (0.194)
2	0.191 (0.393)	0.238 (0.426)	0.061 (0.240)
3	0.130 (0.336)	0.159 (0.366)	0.050 (0.218)
4	0.060 (0.238)	0.074 (0.262)	0.023 (0.150)
Panel D: Active (enrolled in language course in Sept. or employed in Nov.)			
Years since migration:			
1	0.871 (0.335)	0.908 (0.289)	0.770 (0.421)
2	0.915 (0.279)	0.936 (0.244)	0.858 (0.350)
3	0.928 (0.259)	0.946 (0.226)	0.878 (0.327)
4	0.932 (0.251)	0.948 (0.221)	0.889 (0.314)
Ever active by year 4	0.942 (0.234)	0.959 (0.199)	0.896 (0.306)
Panel E: Child birth			
Years since migration:			
1	0.073 (0.259)	0.051 (0.220)	0.132 (0.338)
2	0.089 (0.284)	0.086 (0.280)	0.096 (0.295)
3	0.096 (0.295)	0.094 (0.292)	0.102 (0.303)
4	0.078 (0.269)	0.080 (0.271)	0.075 (0.263)
Number of childbirths by year 4	0.385 (0.639)	0.340 (0.595)	0.509 (0.732)
Number of observations in Panels A-E	5,828	4,263	1,565

Table 5 (continued) Summary Statistics on individual outcomes by years since migration**Panel F: Enrolled module level in Danish language course, contional on having enrolled**

Years since migration:

1	2.529 (1.094)	2.580 (1.060)	2.365 (1.179)
2	3.059 (1.990)	3.134 (1.994)	2.839 (1.963)
3	1.849 (2.374)	1.774 (2.375)	2.067 (2.359)
4	0.760 (1.786)	0.676 (1.711)	1.003 (1.968)

Panel G: Highest completed module in Danish language course, contional on having enrolled

Years since migration:

1	1.360 (1.116)	1.394 (1.103)	1.251 (1.148)
2	2.723 (1.577)	2.792 (1.561)	2.521 (1.606)
3	3.607 (1.661)	3.671 (1.639)	3.420 (1.710)
4	3.950 (1.609)	3.969 (1.595)	3.896 (1.648)

Source : Administrative register information from Statistics Denmark 1997-2020.*Notes* : Sample: Estimation Sample described in notes to Table 4. The ever enrolled in Danish language course in Sept. by year 4 (Panel A) is calculated across years 0-4 since arrival. The ever employed by year 4 (Panel B), the ever active by year 4 (Panel C) and the number of childbirths by year 4 (Panel D) are calculated across the years 1-4 since arrival.

Table 6 Effects of assigned ethnic group share and employment rate of ethnic group on ever enrollment in host-country language course.

	Dependent variable: Ever enrolled in Danish language course by year four (Mean 0.949)											
	Definition of ethnic group											
	Co-nationals				Language fellows				Non-Western immigrants			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Baseline specification												
Ethnic group share (neighborhood)	-1.047 (0.727)	-1.226* (0.722)	-1.064 (0.729)	-1.260* (0.725)	-0.191* (0.113)	-0.210* (0.122)	-0.18 (0.112)	-0.201 (0.123)	-0.154* (0.0671)	-0.143* (0.0843)	-0.153** (0.0711)	-0.141* (0.0845)
Employment rate of ethnic group (neighborhood)			0.0148 (0.00921)	0.0186* (0.0111)			0.0150 (0.0102)	0.0118 (0.0109)			0.000428 (0.0246)	0.00148 (0.0217)
Male	0.0153* (0.00900)	0.0138 (0.00919)	0.0154* (0.00900)	0.0140 (0.00919)	0.0153* (0.00903)	0.0140 (0.00924)	0.0154* (0.00899)	0.0141 (0.00921)	0.0153* (0.00901)	0.0140 (0.00922)	0.0153* (0.00901)	0.0140 (0.00922)
Panel B: Heterogeneous effects by gender												
Ethnic group share (neighborhood)*Male	-1.072 (0.825)	-1.350 (0.819)	-1.106 (0.822)	-1.400* (0.814)	-0.204* (0.108)	-0.238* (0.121)	-0.192* (0.108)	-0.228* (0.121)	-0.136* (0.0755)	-0.136 (0.0842)	-0.138 (0.0844)	-0.133 (0.0884)
Ethnic group share (neighborhood)*Female	-1.011 (1.013)	-1.047 (0.992)	-0.991 (1.002)	-1.039 (0.978)	-0.156 (0.163)	-0.135 (0.164)	-0.146 (0.163)	-0.129 (0.165)	-0.200** (0.0953)	-0.159 (0.114)	-0.192* (0.104)	-0.159 (0.119)
Employment rate of ethnic group (neighborhood)*Male			0.0183* (0.0101)	0.0227* (0.0121)			0.0165 (0.0125)	0.0136 (0.0138)			-0.00211 (0.0257)	0.00244 (0.0225)
Employment rate of ethnic group (neighborhood)*Female			0.00517 (0.0216)	0.00717 (0.0221)			0.0110 (0.0202)	0.00706 (0.0205)			0.00906 (0.0573)	-4.92e-05 (0.0578)
Male	0.0155* (0.00838)	0.0149* (0.00863)	0.0132 (0.00907)	0.0122 (0.00926)	0.0159* (0.00889)	0.0154* (0.00915)	0.0142 (0.0110)	0.0132 (0.0111)	0.0120 (0.0103)	0.0128 (0.0103)	0.0181 (0.0361)	0.0114 (0.0365)
Controls:												
Assigned municipality characteristics	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Assigned municipality FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Number of observations	5,828											

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$. Coefficient estimates from linear probability models. Standard errors clustered by municipality of assignment in parentheses. Sample: Estimation Sample described in notes to Table 4. Additional controls: personal characteristics at arrival, country of origin, year of arrival, month of arrival, dummy for missing information on educational attainment at arrival. Assigned municipality characteristics are population share, non-Western share, co-national share, employment rate, crime conviction rate, share of households with high-income. Individuals in the sample have been assigned to 94 different municipalities and 914 different neighborhoods. Mean (std. dev.) co-ethnic neighbor share based on co-nationals, language fellows, and non-Westerns are 0.003 (0.008), 0.014 (0.036), and 0.051 (0.056), respectively. Mean (std. dev.) employment rate of ethnic enclave based on co-nationals, language fellows, and non-Westerns are 0.179 (0.296), 0.323 (0.292), and 0.500 (0.129), respectively. Mean (std. dev.) employment rate in the assigned municipality: 0.769 (0.040).

Table 7 Effects of assigned ethnic group share and employment rate of ethnic group on level of host-country language course module completion

Dependent variable: Highest module completed in Danish language course in year t+s (mean: 2.94)												
	Co-nationals				Definition of ethnic group				Non-Western immigrants			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Baseline specification												
Ethnic group share (neighborhood)			-0.949 (3.590)	0.575 (3.664)			-0.336 (0.831)	-0.193 (0.811)			-0.0849 (0.485)	-0.354 (0.492)
Employment rate of ethnic group (neighborhood)	-0.172** (0.0706)	-0.116* (0.0657)	-0.171* (0.0705)	-0.117* (0.685)	-0.142* (0.0773)	-0.141* (0.0724)	-0.146* (0.0783)	-0.143** (0.0729)	0.0497 (0.158)	-0.132 (0.146)	0.0329 (0.177)	-0.203 (0.16)
Male	0.169*** (0.0512)	0.174*** (0.0520)	0.169*** (0.0512)	0.174*** (0.0519)	0.170*** (0.0514)	0.174*** (0.0521)	0.170*** (0.0514)	0.174*** (0.0521)	0.170*** (0.0514)	0.175*** (0.0523)	0.170*** (0.0514)	0.175*** (0.0523)
Employment rate (municipality)	-3.081*** (0.868)		-3.082*** (0.869)		-2.995*** (0.873)		-2.989*** (0.87)		-3.169*** (0.921)		-3.152*** (0.908)	
Panel B: Heterogeneous effects by gender												
Ethnic group share (neighborhood)*Male			-2.824 (4.742)	-0.674 (5.079)			0.107 (0.800)	0.374 (0.784)			0.107 (0.185)	-0.351 (0.244)
Ethnic group share (neighborhood)*Female			1.324 (5.151)	1.912 (5.181)			-1.596 (1.471)	-1.701 (1.405)			-0.171 (0.358)	0.0953 (0.419)
Employment rate of ethnic group (neighborhood)*Male	-0.212*** (0.0795)	-0.156** (0.0743)	-0.207** (0.0807)	-0.155** (0.0758)	-0.160* (0.0825)	-0.156* (0.0806)	-0.162** (0.0824)	-0.155* (0.0801)	0.103 (0.168)	-0.0868 (0.167)	0.0323 (0.479)	0.148 (0.125)
Employment rate of ethnic group (neighborhood)*Female	-0.0544 (0.131)	-0.00375 (0.135)	-0.0630 (0.132)	-0.0106 (0.136)	-0.0928 (0.130)	-0.101 (0.124)	-0.0977 (0.132)	-0.105 (0.126)	-0.105 (0.316)	-0.264 (0.306)	-0.395 (0.851)	0.336* (0.191)
Male	0.200*** (0.0577)	0.203*** (0.0598)	0.211*** (0.0582)	0.210*** (0.0607)	0.192*** (0.0645)	0.192*** (0.0676)	0.168** (0.0678)	0.163** (0.0712)	0.0657 (0.180)	0.0860 (0.187)	0.00876 (0.219)	-0.00273 (0.229)
Employment rate (municipality)	-3.106*** (0.869)		-3.107*** (0.871)		-2.992*** (0.873)		-2.993*** (0.870)		-3.164*** (0.921)		-3.142*** (0.909)	
Controls:												
Assigned municipality characteristics	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Assigned municipality FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Number of individuals	5,279											
Number of observations	20,663											

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$. Coefficient estimates from linear regressions. Standard errors clustered by municipality of assignment in parentheses. Sample: Unbalanced panel of individuals in Estimation Sample described in notes to Table 4 who are enrolled in a Danish language course in Sept. in the year. Additional controls: personal characteristics at arrival, country of origin, year of arrival, month of arrival, dummy for missing information on educational attainment at arrival. Assigned municipality characteristics are population share, non-Western share, co-national share, crime conviction rate, share of households with high-income. Individuals in the sample have been assigned to 94 different municipalities and 914 different neighborhoods. Mean (std. dev.) co-ethnic neighbor share based on co-nationals, language fellows, and non-Westerns are 0.003 (0.008), 0.014 (0.036), and 0.051 (0.056), respectively. Mean (std. dev.) employment rate of ethnic enclave based on co-nationals, language fellows, and non-Westerns are 0.179 (0.296), 0.323 (0.292), and 0.500 (0.129), respectively. Mean (std. dev.) employment rate in the assigned municipality: 0.769 (0.040).

Table 8 Effects of assigned ethnic group share and employment rate of ethnic group on employment status

	Dependent variable: Employed in November of year t+s (Mean: 0.26)											
	Co-nationals				Definition of ethnic group				Non-Western immigrants			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Baseline specification												
Ethnic group share (neighborhood)			-0.252 (0.555)	-0.616 (0.557)			-0.0857 (0.161)	-0.212 (0.132)			-0.207** (0.0845)	-0.226** (0.104)
Employment rate of ethnic group (neighborhood)	0.0261* (0.0154)	0.0161 (0.0166)	0.0263* (0.0154)	0.0170 (0.0166)	0.0151 (0.0135)	0.00496 (0.0143)	0.0142 (0.0133)	0.00264 (0.0145)	0.0392 (0.0437)	0.0348 (0.0399)	-0.00192 (0.0513)	-0.0106 (0.0490)
Male	0.223*** (0.00909)	0.223*** (0.00902)	0.223*** (0.00909)	0.223*** (0.00900)	0.223*** (0.00910)	0.223*** (0.00903)	0.223*** (0.00911)	0.223*** (0.00905)	0.223*** (0.00912)	0.223*** (0.00905)	0.223*** (0.00913)	0.223*** (0.00904)
Employment rate (municipality)	0.0720 (0.211)		0.103 (0.210)		0.0648 (0.212)		0.0659 (0.213)		0.0363 (0.209)		0.0773 (0.213)	
Panel B: Heterogeneous effects by gender												
Ethnic group share (neighborhood)*Male			-0.291 (0.788)	-0.595 (0.847)			-0.0560 (0.204)	-0.208 (0.158)			-0.300*** (0.104)	-0.316** (0.134)
Ethnic group share (neighborhood)*Female			-0.207 (0.854)	-0.670 (0.844)			-0.172 (0.227)	-0.230 (0.232)			0.0162 (0.138)	-0.0182 (0.132)
Employment rate of ethnic group (neighborhood)*Male	0.0223 (0.0192)	0.0109 (0.0203)	0.0225 (0.0193)	0.0114 (0.0204)	0.00936 (0.0175)	-0.000999 (0.0182)	0.00840 (0.0173)	-0.00359 (0.0182)	0.0189 (0.0515)	0.0208 (0.0463)	-0.0395 (0.0593)	-0.0410 (0.0575)
Employment rate of ethnic group (neighborhood)*Female	0.0365* (0.0205)	0.0300 (0.0216)	0.0366* (0.0210)	0.0318 (0.0220)	0.0298 (0.0217)	0.0197 (0.0225)	0.0291 (0.0216)	0.0180 (0.0227)	0.0956* (0.0500)	0.0735 (0.0501)	0.0906 (0.0570)	0.0602 (0.0541)
Male	0.226*** (0.00996)	0.227*** (0.0100)	0.226*** (0.0101)	0.227*** (0.0102)	0.230*** (0.0126)	0.230*** (0.0126)	0.228*** (0.0127)	0.230*** (0.0127)	0.262*** (0.0321)	0.250*** (0.0302)	0.305*** (0.0367)	0.289*** (0.0363)
Employment rate (municipality)	0.0704 (0.211)		0.0697 (0.211)		0.0662 (0.212)		0.0668 (0.213)		0.0346 (0.209)		0.0740 (0.213)	
Controls:												
Assigned municipality characteristics	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Assigned municipality FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Number of individuals	5,828											
Number of observations	23,312											

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$. Coefficient estimates from linear probability models. Standard errors clustered by municipality of assignment in parentheses. Sample: Estimation Sample described in notes to Table 4. Additional controls: personal characteristics at arrival, country of origin, year of arrival, month of arrival, dummy for missing information on educational attainment at arrival. Assigned municipality characteristics are population share, non-Western share, co-national share, crime conviction rate, share of households with high-income. Individuals in the sample have been assigned to 94 different municipalities and 914 different neighborhoods. Mean (std. dev.) co-ethnic neighbor share based on co-nationals, language fellows, and non-Westerns are 0.003 (0.008), 0.014 (0.036), and 0.051 (0.056), respectively. Mean (std. dev.) employment rate of ethnic enclave based on co-nationals, language fellows, and non-Westerns are 0.179 (0.296), 0.323 (0.292), and 0.500 (0.129), respectively. Mean (std. dev.) employment rate in the assigned municipality: 0.769 (0.040).

Table 9 Effects of assigned ethnic group share and employment rate of ethnic group on dummy for being active

Dependent variable: Enrolled in Danish language course in Sept. or employed in Nov. in year t+s (Mean: 0.87)												
	Definition of ethnic group											
	Co-nationals				Language fellows				Non-Western immigrants			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Baseline specification												
Ethnic group share (neighborhood)			-0.814 (0.799)	-0.975 (0.747)			-0.106 (0.155)	-0.117 (0.156)			-0.108 (0.0964)	-0.0885 (0.0991)
Employment rate of ethnic group (neighborhood)	0.0148 (0.0128)	0.0179 (0.0140)	0.0154 (0.0126)	0.0193 (0.0140)	0.0115 (0.0111)	0.00690 (0.0117)	0.0103 (0.0114)	0.00562 (0.0118)	0.0255 (0.0272)	0.0168 (0.0309)	0.00407 (0.0316)	-0.000978 (0.0338)
Male	0.0535*** (0.00957)	0.0527*** (0.00969)	0.0536*** (0.00955)	0.0526*** (0.00965)	0.0535*** (0.00958)	0.0526*** (0.00970)	0.0535*** (0.00960)	0.0526*** (0.00971)	0.0535*** (0.00961)	0.0525*** (0.00972)	0.0534*** (0.00962)	0.0525*** (0.00972)
Employment rate (municipality)	-0.0931 (0.118)		-0.0953 (0.118)		-0.0997 (0.118)		-0.0983 (0.118)		-0.117 (0.115)		-0.0955 (0.111)	
Panel B: Heterogeneous effects by gender												
Ethnic group share (neighborhood)*Male			-1.117 (1.055)	-1.361 (1.040)			-0.0677 (0.151)	-0.0913 (0.150)			-0.0533 (0.107)	-0.0404 (0.106)
Ethnic group share (neighborhood)*Female			-0.378 (1.108)	-0.428 (1.033)			-0.221 (0.203)	-0.196 (0.205)			-0.245* (0.141)	-0.201 (0.148)
Employment rate of ethnic group (neighborhood)*Male	0.0109 (0.0143)	0.0146 (0.0161)	0.0120 (0.0142)	0.0165 (0.0159)	-0.000520 (0.0137)	-0.00511 (0.0151)	-0.00171 (0.0140)	-0.00644 (0.0152)	0.0184 (0.0265)	0.0131 (0.0294)	0.00656 (0.0333)	0.00408 (0.0344)
Employment rate of ethnic group (neighborhood)*Female	0.0257 (0.0246)	0.0266 (0.0242)	0.0246 (0.0235)	0.0263 (0.0231)	0.0420* (0.0231)	0.0366 (0.0226)	0.0411* (0.0232)	0.0356 (0.0227)	0.0452 (0.0619)	0.0268 (0.0664)	0.00330 (0.0689)	-0.00785 (0.0726)
Male	0.0564*** (0.0101)	0.0550*** (0.0102)	0.0586*** (0.0104)	0.0577*** (0.0105)	0.0673*** (0.0126)	0.0660*** (0.0127)	0.0654*** (0.0131)	0.0648*** (0.0131)	0.0669** (0.0331)	0.0594* (0.0340)	0.0420 (0.0436)	0.0383 (0.0436)
Employment rate (municipality)	-0.0948 (0.117)		-0.0968 (0.117)		-0.0968 (0.118)		-0.0962 (0.118)		-0.117 (0.115)		-0.0949 (0.111)	
Controls:												
Assigned municipality characteristics	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Assigned municipality FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Number of individuals	5,828											
Number of observations	23,312											

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$. Coefficient estimates from linear probability models. Standard errors clustered by municipality of assignment in parentheses. Sample: Estimation Sample described in notes to Table 4. Additional controls: personal characteristics at arrival, country of origin, year of arrival, month of arrival, dummy for missing information on educational attainment at arrival. Assigned municipality characteristics are population share, non-Western share, co-national share, crime conviction rate, share of households with high-income. Individuals in the sample have been assigned to 94 different municipalities and 914 different neighborhoods. Mean (std. dev.) co-ethnic neighbor share based on co-nationals, language fellows, and non-Westerns are 0.003 (0.008), 0.014 (0.036), and 0.051 (0.056), respectively. Mean (std. dev.) employment rate of ethnic enclave based on co-nationals, language fellows, and non-Westerns are 0.179 (0.296), 0.323 (0.292), and 0.500 (0.129), respectively. Mean (std. dev.) employment rate in the assigned municipality: 0.769 (0.040).

Table 10 Effects of assigned ethnic group share and employment rate of ethnic group on the number of child births by year four

Dependent variable: Number of child births by year four (Mean: 0.385)												
	Co-nationals				Language fellows				Non-Western immigrants			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Baseline specification												
Ethnic group share (neighborhood)	0.281 (1.459)	0.996 (1.416)	0.124 (1.417)	0.740 (1.372)	-0.268 (0.246)	-0.231 (0.271)	-0.185 (0.241)	-0.157 (0.266)	-0.0246 (0.187)	0.00389 (0.199)	0.0909 (0.199)	0.200 (0.196)
Employment rate of ethnic group (neighborhood)			0.138*** (0.0355)	0.140*** (0.0369)			0.111*** (0.0397)	0.102** (0.0412)			0.107 (0.0743)	0.173** (0.0766)
Male	-0.108*** (0.0158)	-0.109*** (0.0158)	-0.107*** (0.0158)	-0.108*** (0.0157)	-0.108*** (0.0158)	-0.109*** (0.0158)	-0.108*** (0.0161)	-0.108*** (0.0159)	-0.108*** (0.0158)	-0.109*** (0.0158)	-0.108*** (0.0158)	-0.109*** (0.0158)
Panel B: Heterogeneous effects by gender												
Ethnic group share (neighborhood)*Male	-0.724 (1.340)	-0.0489 (1.414)	-0.690 (1.328)	-0.0916 (1.391)	-0.431* (0.237)	-0.410 (0.269)	-0.350 (0.233)	-0.340 (0.267)	-0.123 (0.155)	-0.114 (0.175)	-0.100 (0.172)	-0.0191 (0.183)
Ethnic group share (neighborhood)*Female	1.760 (2.214)	2.506 (2.151)	1.235 (2.186)	1.794 (2.121)	0.180 (0.467)	0.231 (0.490)	0.248 (0.472)	0.285 (0.493)	0.233 (0.356)	0.298 (0.370)	0.547 (0.367)	0.706* (0.373)
Employment rate of ethnic group (neighborhood)*Male			0.111*** (0.0337)	0.106*** (0.0345)			0.0848** (0.0390)	0.0700* (0.0394)			0.00558 (0.0763)	0.0593 (0.0761)
Employment rate of ethnic group (neighborhood)*Female			0.213*** (0.0585)	0.230*** (0.0591)			0.178*** (0.0594)	0.180*** (0.0613)			0.365*** (0.130)	0.456*** (0.135)
Male	-0.0997*** (0.0182)	-0.0997*** (0.0183)	-0.0812*** (0.0214)	-0.0771*** (0.0211)	-0.100*** (0.0178)	-0.100*** (0.0179)	-0.0697*** (0.0247)	-0.0646*** (0.0240)	-0.0901*** (0.0234)	-0.0877*** (0.0238)	0.106 (0.0784)	0.128 (0.0780)
Controls:												
Assigned municipality characteristics	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Assigned municipality FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Number of observations	5,828											

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$. Coefficient estimates from linear regressions. Standard errors clustered by municipality of assignment in parentheses. Sample: Estimation Sample described in notes to Table 4. Additional controls: personal characteristics at arrival, country of origin, year of arrival, month of arrival, dummy for missing information on educational attainment at arrival. Assigned municipality characteristics are population share, non-Western share, co-national share, employment rate, crime conviction rate, share of households with high-income. Individuals in the sample have been assigned to 94 different municipalities and 914 different neighborhoods. Mean (std. dev.) co-ethnic neighbor share based on co-nationals, language fellows, and non-Westerns are 0.003 (0.008), 0.014 (0.036), and 0.051 (0.056), respectively. Mean (std. dev.) employment rate of ethnic enclave based on co-nationals, language fellows, and non-Westerns are 0.179 (0.296), 0.323 (0.292), and 0.500 (0.129), respectively. Mean (std. dev.) employment rate in the assigned municipality: 0.769 (0.040).

Table 11 Effects of ethnic enclave size and ethnic enclave quality on level of Danish language speaking skills required in the job

	Dependent variable: Danish language speaking skill requirement in the job (Mean: 20.183)											
	Definition of ethnic group											
	Co-nationals				Language fellows				Non-Western immigrants			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Baseline specification												
Ethnic group share (neighborhood)	11.01 (26.30)	14.40 (26.76)	10.43 (26.26)	12.88 (26.77)	1.123 (3.408)	1.562 (3.772)	1.727 (3.427)	1.805 (3.845)	-2.052 (2.412)	-1.292 (3.139)	-2.794 (2.975)	-2.877 (3.890)
Employment rate of ethnic group (neighborhood)			0.648 (0.554)	0.659 (0.557)			0.744 (0.491)	0.292 (0.497)			-0.657 (1.329)	-1.326 (1.394)
Male	-1.959*** (0.395)	-1.858*** (0.378)	-1.954*** (0.397)	-1.858*** (0.380)	-1.959*** (0.394)	-1.862*** (0.377)	-1.965*** (0.396)	-1.865*** (0.377)	-1.958*** (0.394)	-1.864*** (0.377)	-1.961*** (0.394)	-1.865*** (0.375)
Panel B: Heterogeneous effects by gender												
Ethnic group share (neighborhood)*Male	20.58 (30.19)	22.79 (31.81)	23.45 (30.62)	23.82 (32.16)	2.679 (3.288)	2.805 (3.739)	3.213 (3.319)	2.877 (3.805)	-1.943 (2.571)	-1.638 (3.526)	-2.328 (3.036)	-2.988 (4.104)
Ethnic group share (neighborhood)*Female	-21.13 (34.20)	-12.23 (35.93)	-35.45 (32.89)	-24.90 (34.76)	-24.67* (14.04)	-18.76 (14.42)	-24.84* (14.20)	-19.52 (14.73)	-2.685 (8.296)	0.563 (8.312)	-4.791 (8.611)	-2.024 (8.904)
Employment rate of ethnic group (neighborhood)*Male			0.311 (0.543)	0.364 (0.557)			0.471 (0.502)	-0.0231 (0.526)			-0.259 (1.357)	-1.017 (1.358)
Employment rate of ethnic group (neighborhood)*Female			2.453** (1.233)	2.110* (1.237)			2.013* (1.191)	1.674 (1.176)			-2.477 (3.246)	-2.861 (3.415)
Male	-2.075*** (0.397)	-1.956*** (0.382)	-1.690*** (0.458)	-1.646*** (0.453)	-2.219*** (0.424)	-2.068*** (0.410)	-1.731*** (0.633)	-1.532** (0.611)	-1.991*** (0.535)	-1.766*** (0.534)	-3.201* (1.908)	-2.759 (1.944)
Controls:												
Assigned municipality characteristics	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Assigned municipality FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Number of individuals							2,684					
Number of observations							5,825					

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$. Coefficient estimates from linear regressions. Standard errors clustered by municipality of assignment in parentheses. Sample: Employed individuals in Estimation Sample described in notes to Table 4. Additional controls: personal characteristics at arrival, country of origin, year of arrival, month of arrival, dummy for missing information on educational attainment at arrival. Assigned municipality characteristics are population share, non-Western share, co-national share, employment rate, crime conviction rate, share of households with high-income. Individuals in the sample have been assigned to 94 different municipalities and 914 different neighborhoods. Mean (std. dev.) co-ethnic neighbor share based on co-nationals, language fellows, and non-Westerns are 0.003 (0.008), 0.014 (0.036), and 0.051 (0.056), respectively. Mean (std. dev.) employment rate of ethnic enclave based on co-nationals, language fellows, and non-Westerns are 0.179 (0.296), 0.323 (0.292), and 0.500 (0.129), respectively. Mean (std. dev.) employment rate in the assigned municipality: 0.769 (0.040).

Table 12 IV-estimates of Co-ethnic group share on Ever Enrollment in Danish Language Course by Four Years Since Arrival

	Co-nationals				Definition of ethnic group Language fellows				Non-Western immigrants			
	Dependent variable (value measured by year four since arrival)											
	Ethnic group share	Ever enrolled in Danish Language course	Ethnic group share	Ever enrolled in Danish Language course	Ethnic group share	Ever enrolled in Danish Language course	Ethnic group share	Ever enrolled in Danish Language course	Ethnic group share	Ever enrolled in Danish Language course	Ethnic group share	Ever enrolled in Danish Language course
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	
Panel A: 2SLS estimates												
Ethnic group share (current neighborhood)		-1.590 (1.072)		-2.192* (1.231)		-0.381* (0.224)		-0.431* (0.250)		-0.366** (0.150)		-0.423* (0.245)
Male		0.0161* (0.00900)		0.0144 (0.00913)		0.0157* (0.00885)		0.0145 (0.00900)		0.0154* (0.00887)		0.0140 (0.00901)
Panel B: First-stage regression												
Ethnic group share (assigned neighborhood)	0.659*** (0.0866)		0.559*** (0.0749)		0.502*** (0.0825)		0.486*** (0.0840)		0.420*** (0.0243)		0.338*** (0.0307)	
Controls:												
Assigned municipality characteristics	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No
Assigned municipality FE	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Number of observations	5,828											

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$. Standard errors clustered by municipality of assignment in parentheses. Sample: Estimation Sample described in notes to Table 4. Additional controls: personal characteristics at arrival, country of origin, year of arrival, month of arrival, dummy for missing information on educational attainment at arrival. Assigned municipality characteristics are population share, non-Western share, co-national share, employment rate, crime conviction rate, share of households with high-income. Individuals in the sample have been assigned to 94 different municipalities and 914 different neighborhoods. Mean (std. dev.) co-ethnic neighbor share at assignment based on co-nationals, language fellows, and non-Westerns are 0.003 (0.008), 0.014 (0.036), and 0.051 (0.056), respectively. Mean (std. dev.) co-ethnic neighbor share in current neighborhood based on co-nationals and non-Westerns are 0.015 (0.023), and 0.124 (0.125), respectively. Mean (std. dev.) of Ever enrolled in Danish language course by year four: 0.949 (0.220).

Table A1 Sample selection procedure

#	Explanations	Reduction	Sample size
1	All the population in the residence permit register (OPHG) for the period 1997-2019		1,518,674
2	Drop all the observations without pnr (if pnr==.)	297,188	1,221,486
3	Drop all individuals with imputed residence permit type in any given year	240,770	980,716
4	Keep only 1st residence permit of each individual	265,125	715,591
5	Keep only the refugees	656,264	59,327
6	Residence permit between 2004 and 2015	30,705	28,622
7	Country of origin of the individual is not Denmark	6	28,616
8	The individual is found in the population register (BEF) at least once between 2004 to 2019	164	28,452
9	Age at arrival is 18 or over	6,578	21,874
10	First appearance in BEF is the same or one year after receiving residence permit; out of which:	1,361	20,513
11	The initial neighbourhood of residence is identified	87	20,426
12	Initially identified neighbourhood is located in the initially identified municipality	36	20,390
13	Country of origin of partner of the individual is not Denmark	34	20,356
14	Observed in the population register for the period of 2002-2020 in years 1-4	1,272	19,084
15	Household head (Balanced Panel)	3,092	15,992
16	Subsample of balanced panel when the first 15 municipalities had filled their annual quotas		5,113
17	Subsample of balanced panel and jointly arrived partners (Estimation Sample)		5,828

Table A2.A Variable definitions and primary data sources: Household head characteristics.

Variable	Definition	Primary data source
Refugee	Dummy for having the residence permit type of refugee.	Residence Permit Register (OPHG), Statistics Denmark (DST).
Date of initial immigration	Dates for first time of immigration.	Historical Migration Register (VNDS), DST.
Date of residence permit	Dates for residence permits imputed by the Immigration Service.	Residence Permit Register (OPHG), DST.
Country of origin	Dummy for source country.	Population register (BEF), DST.
Age at time of immigration	Age calculated as the observation year minus the year of birth observed in the population register.	Population register (BEF), DST.
Household head	Dummy for first-arrived adult in the household; if the spouses have arrived on the same date, the husband is defined as the household head.	Residence Permit Register (OPHG) and Population Register (BEF), DST.
Municipality of assignment	Municipality registered in the population registers in the year of receiving residence permit or the following year.	Population register (BEF), DST.
Neighborhood of assignment	neighborhood of the registered housing address in the population registers in the year of receiving residence permit or the following year.	Population register (BEF), DST, and constructed neighborhood clusters by Damm et al. (2021).
Education level	Education level before immigration, constructed based on an education code of the highest degree attained before immigration.	
Male	Dummy for male.	Population register (BEF), DST.
Married	Dummy for married at arrival.	Population register (BEF), DST.
Child aged 0-2	Dummy for having a child aged 0-2 years.	Population register (BEF), DST.
Child aged 3-17	Dummy for having a child aged 3-17 years.	Population register (BEF), DST.

Table A2.B Variable definitions and primary data sources: Outcome characteristics.

Variable	Definition	Primary data source
Ever enrolled in DUB	Dummy for ever enrolled in any Danish course during the first four years since registration as asylum.	Danskundervisningsdatabasen (DUB)
Highest module level completed	The highest completed module level in Danish course on 30 th September of each year.	DUB
Employed	Dummy for whether the refugee is employed in November of each year.	Population register (BEF), DST.
Either employed of language course	Dummy for whether the refugee is either employed in November or registered in Danish course on 30 th September of each year.	Employment register (RAS), DST.
Danish language speaking skill required in job	The weighted average (i.e., importance*level) of the Danish language speaking skill in the dominant occupation of refugees in each year.	Work classification module register (AKM), DST, and task content from (O*NET), the US Bureau of Labor Statistics.
Child birth	Dummy for whether the individual has a new born child (children) in each year.	Population register (BEF), DST.

Table A2.C Variable definitions and primary data sources: Area Characteristics.

Variable	Definition	Primary data source
Municipality quota	Annual maximum quota of refugees to be allocated in the municipality.	Danish Immigration Service (DIS)
Population share	Number of inhabitants in the neighborhood divided by the total population of municipality.	Population register, DST, and constructed neighborhood clusters by Damm et al. (2021). Authors' calculations based on full population data,
Non-Western immigrants share	Number of non-western immigrants living in the neighborhood divided by the number of inhabitants in the neighborhood.	Population register, DST, and constructed neighborhood clusters by Damm et al. (2021). Authors' calculations based on full population data.
Share of language fellows	Number of persons speaking the same language as a refugee speaks living in the neighborhood divided by the number of inhabitants in the neighborhood.	Population register, DST, constructed neighborhood clusters by Damm et al. (2021), and the list of official languages spoken in refugee-sending countries from UN (2017). Authors' calculations based on full population data.
Share of co-nationals	Number of conationals living in the municipality divided by the number of inhabitants in the municipality.	Population register, DST, and constructed neighborhood clusters by Damm et al. (2021). Authors' calculations based on full population data.
Employment rate among non-Western immigrants	Number of employed non-Western immigrants aged 25-64 divided by the number of non-Western immigrants aged 25-64 in the labor force of the neighborhood.	Population and employment registers, DST, and constructed neighborhood clusters by Damm et al. (2021). Authors' calculations based on full population data.
Employment rate among language fellows	Number of employed same language persons aged 25-64 divided by the number of same language persons aged 25-64 in the labor force of the neighborhood.	Population and employment registers, DST, constructed neighborhood clusters by Damm et al. (2021), and the list of official languages spoken in refugee-sending countries from UN (2017). Authors' calculations based on full population data.
Employment rate among co-nationals	Number of employed co-nationals aged 25-64 divided by the number of co-nationals aged 25-64 in the labor force of the neighborhood.	Population and employment registers, DST, and constructed neighborhood clusters by Damm et al. (2021). Authors' calculations based on full population data.
Employment rate among high-skilled non-Western immigrants	Number of employed non-Western immigrants aged 25-64 with more than 10 years of education divided by the number of non-Western immigrants aged 25-64 with more than 10 years of education in the neighborhood.	Population, education, and employment registers, DST, and constructed neighborhood clusters by Damm et al. (2021). Authors' calculations based on full population data.
Employment rate among high-skilled language fellows	Number of employed same language persons aged 25-64 with more than 10 years of education divided by the number of same language persons aged 25-64 with more than 10 years of education in the neighborhood.	Population, education, and employment registers, DST, constructed neighborhood clusters by Damm et al. (2021), and the list of official languages spoken in refugee-sending countries from UN (2017). Authors' calculations based on full population data.
Employment rate among high-skilled co-nationals	Number of employed co-nationals aged 25-64 with more than 10 years of education divided by the number of co-nationals aged 25-64 with more than 10 years of education in the neighborhood.	Population, education, and employment registers, DST, and constructed neighborhood clusters by Damm et al. (2021). Authors' calculations based on full population data.
Employment rate among low-skilled non-Western immigrants	Number of employed non-Western immigrants aged 25-64 with at most 10 years of education divided by the number of non-Western immigrants aged 25-64 with at most 10 years of education in the neighborhood.	Population, education, and employment registers, DST, and constructed neighborhood clusters by Damm et al. (2021). Authors' calculations based on full population data.
Employment rate among low-skilled language fellows	Number of employed same language persons aged 25-64 with at most 10 years of education divided by the number of same language persons aged 25-64 with at most 10 years of education in the neighborhood.	Population, education, and employment registers, DST, constructed neighborhood clusters by Damm et al. (2021), and the list of official languages spoken in refugee-sending countries from UN (2017). Authors' calculations based on full population data.
Employment rate among low-skilled co-nationals	Number of employed co-nationals aged 25-64 with at most 10 years of education divided by the number of co-nationals aged 25-64 with at most 10 years of education in the neighborhood.	Population, education, and employment registers, DST, and constructed neighborhood clusters by Damm et al. (2021). Authors' calculations based on full population data.

Table A3 Correlations between the size and quality of co-ethnic network in the initially assigned neighborhood

	Share of co-nationals	Share of language fellows	Share of non-Western immigrants	Employment rate of co-nationals	Employment rate of language fellows	Employment rate of non-Western immigrants
Share of co-nationals	1	0.5669	0.5136	0.122	-0.0144	-0.2438
Share of language fellows	0.5669	1	0.6966	-0.0144	-0.0229	-0.3014
Share of non-Western immigrants	0.5136	0.6966	1	0.047	-0.0575	-0.3968
Employment rate of co-nationals	0.122	-0.0144	0.047	1	0.5738	0.1343
Employment rate of language fellows	-0.0144	-0.0229	-0.0575	0.5738	1	0.2578
Employment rate of non-Western immigrants	-0.2438	-0.3014	-0.3968	0.1343	0.2578	1
Number of municipalities				94		
Number of neighborhoods				914		

Source : Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: Sample: Estimation Sample described in notes to Table 4.

Table A4 Correlations between quality of co-ethnic network in the initially assigned municipality and neighborhood

	Employment rate of co-nationals (neighborhood)	Employment rate of language fellows (neighborhood)	Employment rate of non-Western immigrants (neighborhood)	Employment rate (neighborhood)
Employment rate of co-nationals (municipality)	0.4035	0.1926	0.1436	0.0767
Employment rate of language fellows (municipality)	0.252	0.3584	0.1946	0.1437
Employment rate of non-Western immigrants (municipality)	0.1603	0.1496	0.4687	0.3268
Employment rate (municipality)	0.1292	0.1571	0.3645	0.5026
Number of municipalities			94	
Number of neighborhoods			914	

Source : Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes : Sample: Estimation Sample described in notes to Table 4.

Table A5 Balancing test: Assigned neighborhood characteristics and individual characteristics of assignee (households heads)

	Dependent Variables (assigned neighbourhood characteristics at arrival)												
	Population share	Co-national share	Share of language fellows	Non-Western immigrant share	Employment rate among co-nationals	Employment rate among language fellows	Employment rate among non-Western immigrants	Employment rate among high-skilled co-nationals	Employment rate among high-skilled language fellows	Employment rate among high-skilled non-Western immigrants	Employment rate among low-skilled co-nationals	Employment rate among low-skilled language fellows	Employment rate among low-skilled non-Western immigrants
Panel A: Balanced Panel													
(Ref. cat.: 0-10 yrs. educ.)													
11-13 yrs. of educ.	0.000403 (0.00131)	-0.0000695 (0.000161)	0.000806 (0.00105)	0.000362 (0.00137)	0.00181 (0.00623)	-0.0000156 (0.00626)	-0.000232 (0.00321)	-0.00125 (0.00540)	-0.000704 (0.000664)	-0.000722 (0.00358)	0.00605 (0.00520)	0.00701 (0.00697)	0.000742 (0.00360)
More than 13 yrs. of educ.	0.00334* (0.00179)	-0.000505** (0.000201)	-0.00159 (0.000981)	-0.00296** (0.00141)	0.00177 (0.00657)	0.00648 (0.00645)	0.00765** (0.00327)	-0.00493 (0.00650)	0.00587 (0.00705)	0.00600* (0.00347)	0.00123 (0.00592)	-0.00246 (0.00773)	0.00793** (0.00369)
R ²	0.204	0.208	0.158	0.164	0.218	0.169	0.072	0.227	0.180	0.061	0.127	0.116	0.068
No. of obs.	15,992												
Panel B: Subsample of refugee household heads arrived after the first 6 municipalities had filled their annual quota													
(Ref. cat.: 0-10 yrs. educ.)													
11-13 yrs. of educ.	0.000182 (0.00171)	-0.000124 (0.000184)	0.00144 (0.00136)	-0.000639 (0.00200)	0.00541 (0.00862)	0.00675 (0.00894)	0.00278 (0.00458)	0.00289 (0.00853)	0.00434 (0.00922)	0.00205 (0.00495)	0.00573 (0.00549)	0.0104 (0.00756)	0.0211** (0.00909)
More than 13 yrs. of educ.	0.00389 (0.00275)	-0.000375 (0.000246)	-0.00126 (0.00125)	-0.00331* (0.00197)	0.00945 (0.00895)	0.00606 (0.00811)	0.00726* (0.00390)	-0.00357 (0.00876)	0.00253 (0.00922)	0.00375 (0.00410)	0.0104** (0.00466)	0.0104 (0.00823)	-0.00189 (0.00903)
R ²	0.194	0.189	0.174	0.175	0.203	0.154	0.08	0.216	0.163	0.07	0.125	0.119	0.7
No. of obs.	10,621												
Panel C: Subsample of refugee household heads arrived after the first 9 municipalities had filled their annual quota													
(Ref. cat.: 0-10 yrs. educ.)													
11-13 yrs. of educ.	-0.000810 (0.00237)	-9.23e-05 (0.000207)	0.00173 (0.00126)	-0.000181 (0.00188)	0.00340 (0.0103)	0.00578 (0.0103)	0.00180 (0.00500)	-0.00188 (0.00963)	0.00235 (0.0109)	0.00130 (0.00546)	0.0122 (0.00917)	0.0206** (0.0102)	0.00558 (0.00595)
More than 13 yrs. of educ.	0.00392 (0.00325)	-0.000268 (0.000264)	-0.000875 (0.00147)	-0.00367 (0.00224)	0.00646 (0.00939)	0.00796 (0.00872)	0.00685 (0.00457)	-0.00910 (0.00924)	0.00197 (0.0102)	0.00324 (0.00480)	0.0105 (0.00910)	0.000198 (0.00956)	0.0110** (0.00539)
R ²	0.192	0.188	0.181	0.179	0.201	0.153	0.088	0.215	0.161	0.077	0.124	0.116	0.075
No. of obs.	8,931												
Panel D: Subsample of refugee household heads arrived after the first 12 municipalities had filled their annual quota													
(Ref. cat.: 0-10 yrs. educ.)													
11-13 yrs. of educ.	-0.00219 (0.00332)	-0.000280 (0.000253)	0.000839 (0.00167)	-0.000499 (0.00198)	0.00286 (0.0130)	-0.00219 (0.0130)	-0.00124 (0.00599)	-0.00295 (0.0118)	-0.00492 (0.0130)	-0.00264 (0.00647)	0.0209* (0.0112)	0.0206 (0.0126)	0.00429 (0.00763)
More than 13 yrs. of educ.	0.00413 (0.00453)	-0.000484 (0.000328)	-0.000912 (0.00160)	-0.00492** (0.00235)	-0.000415 (0.0121)	-0.000662 (0.0110)	0.00685 (0.00559)	-0.0112 (0.0114)	-0.00671 (0.0137)	0.00162 (0.00589)	0.00759 (0.0107)	-0.00611 (0.0109)	0.0141* (0.00714)
R ²	0.179	0.203	0.185	0.196	0.096	0.216	0.166	0.226	0.182	0.087	0.143	0.128	0.82
No. of obs.	6,250												
Panel E: Subsample of refugee household heads arrived after the first 15 municipalities had filled their annual quota													
(Ref. cat.: 0-10 yrs. educ.)													
11-13 yrs. of educ.	-0.00254 (0.00380)	-0.000110 (0.000248)	0.00107 (0.00161)	-0.000780 (0.00209)	0.00538 (0.0142)	-0.00142 (0.0146)	0.00172 (0.00652)	-0.00299 (0.0131)	-0.00400 (0.0148)	-0.000420 (0.00704)	0.0182 (0.0116)	0.0194 (0.0133)	0.00751 (0.00818)
More than 13 yrs. of educ.	0.00246 (0.00418)	-0.000228 (0.000323)	-0.000124 (0.00158)	-0.00419* (0.00227)	-0.0000938 (0.0141)	-0.00430 (0.0124)	0.00608 (0.00602)	-0.00843 (0.0137)	-0.00922 (0.0145)	-0.000279 (0.00637)	0.00474 (0.0112)	0.00474 (0.0112)	0.0163** (0.00738)
R ²	0.172	0.227	0.184	0.201	0.214	0.174	0.102	0.227	0.193	0.094	0.144	0.134	0.193
No. of obs.	5,113												
Controls:													
Country of origin F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year of arrival F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Month of arrival F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Demographic characteristics of assigned municipality at arrival	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: ***, p<0.01, **: p<0.05, *: p<0.1. Reported coefficients are based on linear regression of neighbourhood characteristics in the year of arrival on individual characteristics in the year of arrival. Standard errors clustered by municipality of assignment in parentheses. Sample: Estimation Sample described in notes to Table 4. Reported coefficients are based on linear regression of neighborhood characteristics in the year of arrival on individual characteristics in the year of arrival. Demographic characteristics of assigned municipality at arrival are population share, non-Western immigrant share and co-nationals. Additional control: Dummy for missing information on educational attainment at arrival. Individuals in the Subsample used in Panel E have been assigned to 94 different municipalities and 914 different neighborhoods.

Table A6 Balancing test: Assigned neighborhood characteristics and individual characteristics of assignee (households heads)

	Dependent Variables (assigned neighborhood characteristics at arrival)												
	Population share	Co-national share	Share of language fellows	Non-Western immigrant share	Employment rate among co-nationals	Employment rate among language fellows	Employment rate among non-Western immigrants	Employment rate among high-skilled co-nationals	Employment rate among high-skilled language fellows	Employment rate among high-skilled non-Western immigrants	Employment rate among low-skilled co-nationals	Employment rate among low-skilled language fellows	Employment rate among low-skilled non-Western immigrants
Panel A: Balanced Panel													
(Ref. cat.: 0-10 yrs. educ.)													
11-13 yrs. of educ.	0.000324 (0.000507)	-7.78e-05 (0.000164)	0.000889 (0.000916)	0.000473 (0.00122)	0.000343 (0.00576)	-0.00102 (0.00590)	-0.00137 (0.00288)	-0.00128 (0.00561)	-0.00130 (0.00659)	-0.000782 (0.00301)	-0.00128 (0.00561)	-0.00130 (0.00659)	-0.00109 (0.00349)
More than 13 yrs. of educ.	0.00114** (0.000501)	-0.000248 (0.000164)	-0.000644 (0.000834)	-0.000982 (0.00117)	0.000500 (0.00581)	0.00557 (0.00597)	0.00483* (0.00282)	-0.00508 (0.00567)	0.00310 (0.00650)	0.00286 (0.00298)	-0.00508 (0.00567)	0.00310 (0.00650)	0.00804** (0.00335)
R ²	0.9	0.168	0.235	0.309	0.244	0.229	0.241	0.251	0.239	0.212	0.152	0.179	0.203
No. of obs.	15,992												
Panel B: Subsample of refugee household heads arrived after the first 15 municipalities had filled their annual quota													
(Ref. cat.: 0-10 yrs. educ.)													
11-13 yrs. of educ.	-0.000121 (0.000895)	-0.000149 (0.000244)	0.000244 (0.00150)	-0.00138 (0.00207)	0.00464 (0.0122)	-0.00449 (0.0121)	0.000189 (0.00360)	-0.00726 (0.0118)	-0.00809 (0.0133)	0.000147 (0.00574)	0.0183 (0.0113)	0.0149 (0.0122)	0.00554 (0.00693)
More than 13 yrs. of educ.	-0.000218 (0.000827)	-0.000140 (0.000280)	0.000111 (0.00143)	-0.00287 (0.00204)	-0.00468 (0.0119)	-0.0129 (0.0117)	-0.00139 (0.00346)	-0.0119 (0.0116)	-0.0187 (0.0126)	-0.00521 (0.00570)	0.000567 (0.0107)	-0.0152 (0.0113)	0.0150** (0.00636)
R ²	0.942	0.355	0.193	0.4	0.275	0.25	0.304	0.289	0.268	0.264	0.185	0.209	0.261
No. of obs.	5,113												
Controls:													
Country of origin F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year of arrival F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Month of arrival F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Assigned municipality F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Source : Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes : ***: p<0.01, **: p<0.05, *: p<0.1. Reported coefficients are based on linear regression of neighbourhood characteristics in the year of arrival on individual characteristics in the year of arrival. Standard errors clustered by municipality of assignment in parentheses. Sample: Estimation Sample described in notes to Table 4. Reported coefficients are based on linear regression of neighborhood characteristics in the year of arrival on individual characteristics in the year of arrival. Additional control: Dummy for missing information on educational attainment at arrival. Individuals in the Subsample used in Panel B have been assigned to 94 different municipalities and 914 different neighborhoods.

Table A7 Alternative balancing test: Level of education of assignee (household heads) and assigned neighborhood characteristics at arrival after the first 15 municipalities had filled their annual quota

	Dependent Variable: Household head has at least 11 years of education												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Assigned neighborhood characteristic at arrival:													
Population share	0.00452 (0.0791)												
Co-national share		-0.546 (0.659)											
Share of language fellows			0.0574 (0.165)										
Non-Western immigrant share				-0.146 (0.103)									
Employment rate among co-nationals					0.000483 (0.0198)								
Employment rate among language fellows						-0.00587 (0.0198)							
Employment rate among non-Western immigrants							0.424 (0.0434)						
Employment rate among high-skilled co-nationals								-0.0117 (0.0191)					
Employment rate among high-skilled language fellows									-0.0115 (0.0182)				
Employment rate among high-skilled non-Western immigrants										-0.00286 (0.0408)			
Employment rate among low-skilled co-nationals											0.0278 (0.0224)		
Employment rate among low-skilled language fellows												0.00593 (0.0207)	
Employment rate among low-skilled non-Western immigrants													0.0755** (0.0350)
R ²	0.292	0.293	0.292	0.293	0.292	0.292	0.293	0.293	0.293	0.292	0.293	0.292	0.293
No. of obs.								5,113					
Country of origin F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year of arrival F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Month of arrival F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Demographic characteristics of assigned municipality at arrival	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: ***: p<0.01, **: p<0.05, *: p<0.1. Coefficient estimates from linear probability models. Standard errors clustered by municipality of assignment in parentheses. Sample: Estimation Sample described in notes to Table 4. Reported coefficients are based on linear regression of binary variable of household head has at least 11 years of education on neighborhood characteristics in the year of arrival. Additional controls: personal characteristics at arrival, country of origin, year of arrival, month of arrival, dummy for missing information on educational attainment at arrival. Demographic characteristics of assigned municipality are population share, non-Western immigrant share and co-national share. Individuals in the sample have been assigned to 94 different municipalities and 914

Table A8 Effects of assigned ethnic group share and employment rate of ethnic group on level of host-country language course module completion

Dependent variable: Highest module level completed in Danish language course in year t+s (mean: 2.65)												
	Co-nationals				Definition of ethnic group				Non-Western immigrants			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Baseline specification												
Ethnic group share (neighborhood)			-2.746 (2.854)	-1.369 (2.762)			-0.451 (0.658)	-0.265 (0.587)			-0.327 (0.405)	-0.491 (0.414)
Employment rate of ethnic group (neighborhood)	-0.145** (0.0661)	-0.0929 (0.0631)	-0.143** (0.0661)	-0.0909 (0.0635)	-0.115* (0.0699)	-0.112* (0.0672)	-0.120* (0.0705)	-0.115* (0.0673)	0.0643 (0.141)	-0.0954 (0.138)	-0.000600 (0.151)	-0.194 (0.150)
Male	0.134*** (0.0450)	0.141*** (0.0458)	0.134*** (0.0451)	0.141*** (0.0458)	0.134*** (0.0452)	0.142*** (0.0459)	0.134*** (0.0453)	0.142*** (0.0459)	0.135*** (0.0452)	0.142*** (0.0461)	0.135*** (0.0453)	0.142*** (0.0461)
Employment rate (municipality)	-2.546*** (0.811)		-2.553*** (0.815)		-2.479*** (0.815)		-2.473*** (0.813)		-2.648*** (0.863)		-2.584*** (0.856)	
Panel B: Heterogeneous effects by gender												
Ethnic group share (neighborhood)*Male			-4.128 (3.282)	-2.143 (3.277)			-0.0956 (0.653)	0.185 (0.584)			-0.230 (0.411)	-0.356 (0.419)
Ethnic group share (neighborhood)*Female			-0.799 (4.352)	-0.356 (4.322)			-1.443 (1.093)	-1.440 (1.031)			-0.560 (0.628)	-0.803 (0.640)
Employment rate of ethnic group (neighborhood)*Male	-0.174** (0.0770)	-0.118 (0.0736)	-0.170** (0.0780)	-0.115 (0.0747)	-0.136* (0.0782)	-0.123 (0.0779)	-0.140* (0.0777)	-0.124 (0.0770)	0.0939 (0.154)	-0.0618 (0.157)	0.0457 (0.167)	-0.136 (0.172)
Employment rate of ethnic group (neighborhood)*Female	-0.0629 (0.114)	-0.0251 (0.115)	-0.0690 (0.114)	-0.0270 (0.117)	-0.0621 (0.109)	-0.0828 (0.103)	-0.0660 (0.110)	-0.0850 (0.104)	-0.0178 (0.271)	-0.189 (0.268)	-0.116 (0.300)	-0.334 (0.300)
Male	0.156*** (0.0513)	0.159*** (0.0527)	0.165*** (0.0516)	0.164*** (0.0531)	0.158*** (0.0599)	0.155** (0.0622)	0.141** (0.0620)	0.133** (0.0646)	0.0789 (0.160)	0.0783 (0.165)	0.0366 (0.195)	0.0200 (0.203)
Employment rate (municipality)	-2.559*** (0.809)		-2.566*** (0.813)		-2.474*** (0.815)		-2.475*** (0.814)		-2.646*** (0.863)		-2.580*** (0.857)	
Controls:												
Assigned municipality characteristics	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Assigned municipality FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
No. of individuals	5,828											
No. of obs.	23,312											

Source : Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$. Coefficient estimates from linear regressions. Standard errors clustered by municipality of assignment in parentheses. Sample: Estimation Sample described in notes to Table 4. Module level completed takes the value 0 for individuals who have never enrolled in a Danish language course by four years since arrival. Additional controls: personal characteristics at arrival, country of origin, year of arrival, month of arrival, dummy for missing information on educational attainment at arrival. Assigned municipality characteristics are population share, non-Western share, co-national share, crime conviction rate, share of households with high-income. Individuals in the sample have been assigned to 94 different municipalities and 914 different neighborhoods. Mean (std. dev.) co-ethnic neighbor share based on co-nationals, language fellows, and non-Westerns are 0.003 (0.008), 0.014 (0.036), and 0.051 (0.056), respectively. Mean (std. dev.) employment rate of ethnic enclave based on co-nationals, language fellows, and non-Westerns are 0.179 (0.296), 0.323 (0.292), and 0.500 (0.129), respectively. Mean (std. dev.) employment rate in the assigned municipality: 0.769 (0.040).

Table A9 Effects of assigned ethnic group share and employment rate of ethnic group on level of host-country language course module enrollment

Dependent variable: Highest module enrolled in Danish language course in year t+s (mean: 2.037)												
	Co-nationals				Definition of ethnic group				Non-Western immigrants			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Baseline specification												
Ethnic group share (neighborhood)			0.608 (1.437)	-0.226 (1.585)			-0.344 (0.369)	-0.433 (0.377)			0.256 (0.230)	-0.113 (0.233)
Employment rate of ethnic group (neighborhood)	-0.0453 (0.0302)	-0.0233 (0.0279)	-0.0457 (0.0302)	-0.0230 (0.0277)	-0.0518 (0.0318)	-0.0299 (0.0292)	-0.0558* (0.0320)	-0.0349 (0.0295)	-0.0223 (0.0858)	-0.0729 (0.0653)	0.0283 (0.0963)	-0.0954 (0.0659)
Male	0.123*** (0.0213)	0.127*** (0.0218)	0.123*** (0.0212)	0.127*** (0.0218)	0.123*** (0.0212)	0.127*** (0.0217)	0.123*** (0.0213)	0.127*** (0.0218)	0.123*** (0.0213)	0.127*** (0.0218)	0.123*** (0.0213)	0.127*** (0.0218)
Employment rate (municipality)	-1.837*** (0.545)		-1.836*** (0.545)		-1.802*** (0.547)		-1.796*** (0.547)		-1.823*** (0.574)		-1.874*** (0.575)	
Panel B: Heterogeneous effects by gender												
Ethnic group share (neighborhood)*Male			0.636 (1.398)	0.183 (1.735)			-0.0885 (0.315)	-0.112 (0.352)			0.331* (0.191)	0.0234 (0.186)
Ethnic group share (neighborhood)*Female			0.523 (2.423)	-0.809 (2.272)			-1.095 (0.827)	-1.314* (0.770)			0.0502 (0.505)	-0.456 (0.479)
Employment rate of ethnic group (neighborhood)*Male	-0.0628* (0.0350)	-0.0411 (0.0339)	-0.0632* (0.0353)	-0.0417 (0.0340)	-0.0570* (0.0343)	-0.0327 (0.0324)	-0.0602* (0.0349)	-0.0363 (0.0330)	-0.0189 (0.0775)	-0.0771 (0.144)	0.0454 (0.0901)	-0.0691 (0.0728)
Employment rate of ethnic group (neighborhood)*Female	0.00707 (0.0561)	0.0272 (0.0491)	0.00664 (0.0578)	0.0301 (0.0509)	-0.0375 (0.0575)	-0.0224 (0.0538)	-0.0416 (0.0568)	-0.0269 (0.0530)	-0.0323 (0.169)	-0.0715 (0.0662)	-0.0152 (0.180)	-0.153 (0.143)
Male	0.136*** (0.0239)	0.140*** (0.0237)	0.136*** (0.0242)	0.137*** (0.0241)	0.129*** (0.0293)	0.116 (0.0823)	0.116*** (0.0298)	0.114*** (0.0301)	0.130*** (0.0294)	0.124 (0.0801)	0.0784 (0.0983)	0.0608 (0.0963)
Employment rate (municipality)	-1.848*** (0.546)		-1.848*** (0.545)		-1.801*** (0.547)		-1.799*** (0.547)		-1.823*** (0.575)		-1.869*** (0.576)	
Controls:												
Assigned municipality characteristics	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Assigned municipality FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
No. of individuals	5,279											
No. of observations	20,663											

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$. Coefficient estimates from linear regressions. Standard errors clustered by municipality of assignment in parentheses. Sample: Unbalanced panel of individuals in Estimation Sample described in notes to Table 4 who are enrolled in a Danish language course in Sept. in the year. Additional controls: personal characteristics at arrival, country of origin, year of arrival, month of arrival, dummy for missing information on educational attainment at arrival. Assigned municipality characteristics are population share, non-Western share, co-national share, crime conviction rate, share of households with high-income. Individuals in the sample have been assigned to 94 different municipalities and 914 different neighborhoods. Mean (std. dev.) co-ethnic neighborhood share based on co-nationals, language fellows, and non-Westerns are 0.003 (0.008), 0.014 (0.036), and 0.051 (0.056), respectively. Mean (std. dev.) employment rate of ethnic enclave based on co-nationals, language fellows, and non-Westerns are 0.179 (0.296), 0.323 (0.292), and 0.500 (0.129), respectively. Mean (std. dev.) employment rate in the assigned municipality: 0.769 (0.040).

Table A10 Robustness Checks of Effects of Assigned Ethnic Group Share and Employment Rate of Ethnic Group on Host-Country Language Acquisition. Ethnic Group Defined as Language Fellows.

	Sample					
	Baseline sample		After filling of the first 15 munic. but restricted to cohorts 2004-2012		Balanced panel and jointly arrived spouses	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effects on Ever Enrollment in Host-Country Language Course						
Ethnic group share	-0.192*	-0.228*	-0.505	-0.441	-0.0616	-0.0503
(neighborhood)*Male	(0.108)	(0.121)	(0.316)	(0.294)	(0.0481)	(0.0500)
Ethnic group share	-0.146	-0.129	-0.322	-0.245	-0.0586	-0.0196
(neighborhood)*Female	(0.163)	(0.165)	(0.311)	(0.298)	(0.114)	(0.112)
Employment rate of ethnic group (neighborhood)*Male	0.0165	0.0136	0.0214	0.0147	0.00651	0.00767
	(0.0125)	(0.0138)	(0.0188)	(0.0193)	(0.00683)	(0.00735)
Employment rate of ethnic group (neighborhood)*Female	0.0110	0.00706	0.0175	0.0146	0.0133	0.00966
	(0.0202)	(0.0205)	(0.0267)	(0.0279)	(0.0111)	(0.0112)
Male	0.0142	0.0132	0.0220	0.0222	0.00977*	0.00831
	(0.0110)	(0.0111)	(0.0153)	(0.0151)	(0.00562)	(0.00564)
No. of obs.	5,828		3,383		17,602	
Panel B: Effects on Level of Host-Country Language Course Completion						
Ethnic group share	0.107	0.374	-0.0911	-0.337	0.00952	0.0984
(neighborhood)*Male	(0.800)	(0.784)	(1.493)	(1.377)	(0.449)	(0.399)
Ethnic group share	-1.596	-1.701	-1.427	-1.760	-0.581	-0.797
(neighborhood)*Female	(1.471)	(1.405)	(1.662)	(1.647)	(0.719)	(0.654)
Employment rate of ethnic group (neighborhood)*Male	-0.162**	-0.155*	-0.133	-0.0803	-0.0250	3.06e-05
	(0.0824)	(0.0801)	(0.0969)	(0.0964)	(0.0616)	(0.0524)
Employment rate of ethnic group (neighborhood)*Female	-0.0977	-0.105	-0.0979	-0.0950	4.61e-05	0.0421
	(0.132)	(0.126)	(0.161)	(0.157)	(0.0844)	(0.0834)
Male	0.168**	0.163**	0.124	0.101	0.159***	0.162***
	(0.0678)	(0.0712)	(0.0819)	(0.0873)	(0.0443)	(0.0447)
Employment rate (municipality)	-2.993***		-3.567***		-2.381***	
	(0.870)		(1.222)		(0.895)	
No. of individuals	5,279		2,915		16,506	
No. of obs.	20,663		11,388		64,928	
Controls:						
Assigned municipality characteristics	Yes	No	Yes	No	Yes	No
Assigned municipality FE	No	Yes	No	Yes	No	Yes

Source: Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$. Panel A: Coefficient estimates from linear probability models. Panel B: Coefficient estimates from linear regressions. Standard errors clustered by municipality of assignment in parentheses. Sample: Cols. (1)-(2): Estimation Sample described in notes to Table 4, cols. (3)-(4): Individuals in estimation sample who arrived before 2013, and cols. (5)-(6): Balanced Panel of refugee household heads and jointly arrived spouses. Additional controls: personal characteristics at arrival, country of origin, year of arrival, month of arrival, dummy for missing information on educational attainment at arrival. Assigned municipality characteristics are population share, non-Western share, co-national share, crime conviction rate, share of households with high-income. Individuals in the sample have been assigned to 94 different municipalities and 914 different neighborhoods. Mean (std. dev.) co-ethnic neighbor share based on co-nationals, language fellows, and non-Westerns are 0.003 (0.008), 0.014 (0.036), and 0.051 (0.056), respectively. Mean (std. dev.) employment rate of ethnic enclave based on co-nationals, language fellows, and non-Westerns are 0.179 (0.296), 0.323 (0.292), and 0.500 (0.129), respectively. Mean (std. dev.) employment rate in the assigned municipality: 0.769 (0.040). Mean (std. dev.) of Ever enrolled in Danish language course by year four: Cols. (1)-(2): 0.949 (0.220), cols. (3)-(4): 0.917 (0.276), and cols. (5)-(6): 0.969 (0.173). Mean (std. dev.) of Highest Completed Module Level in year $t+s$: Cols. (1)-(2): 2.936 (1.808), cols. (3)-(4): 2.527 (1.973), cols. (5)-(6): 2.880 (1.870).

Table A11 Effects of Assigned Ethnic Group Share and Employment Rate among Ethnic Peers on Host-Country Language Acquisition. Ethnic Group Defined as language fellows.

	Peer definition					
	Baseline specification: Peers defined as co-ethnic neighbors		Peers defined as neighbors who have the ethnic origin and skill level		Peers defined as low-skilled among co-ethnic neighbors	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effects on Ever Enrollment in Host-Country Language Course						
Ethnic group share (neighborhood)*Male	-0.192*	-0.228*	0.00866	0.00360	-0.000159	-0.00576
	(0.108)	(0.121)	(0.00933)	(0.00982)	(0.0103)	(0.0111)
Ethnic group share (neighborhood)*Female	-0.146	-0.129	-0.00169	-0.00356	0.00182	0.00180
	(0.163)	(0.165)	(0.0189)	(0.0189)	(0.0188)	(0.0188)
Employment rate of ethnic group (neighborhood)*Male	0.0165	0.0136	0.0176	0.0157	0.0187	0.0162
	(0.0125)	(0.0138)	(0.0125)	(0.0137)	(0.0126)	(0.0138)
Employment rate of ethnic group (neighborhood)*Female	0.0110	0.00706	0.0133	0.00934	0.0125	0.00831
	(0.0202)	(0.0205)	(0.0203)	(0.0205)	(0.0201)	(0.0203)
Male	0.0142	0.0132	0.00956	0.00899	0.0141	0.0143
	(0.0110)	(0.0111)	(0.0163)	(0.0164)	(0.0164)	(0.0165)
No. of obs.	5,828					
Panel B: Effects on Level of Host-Country Language Course Completion						
Ethnic group share (neighborhood)*Male	0.107	0.374	0.0410	0.0789	0.00137	0.0581
	(0.800)	(0.784)	(0.0875)	(0.0772)	(0.0939)	(0.0853)
Ethnic group share (neighborhood)*Female	-1.596	-1.701	-0.139	-0.103	-0.268**	-0.205*
	(1.471)	(1.405)	(0.0993)	(0.0977)	(0.119)	(0.118)
Employment rate of ethnic group (neighborhood)*Male	-0.162**	-0.155*	-0.166**	-0.167**	-0.160*	-0.155*
	(0.0824)	(0.0801)	(0.0820)	(0.0810)	(0.0827)	(0.0811)
Employment rate of ethnic group (neighborhood)*Female	-0.0977	-0.105	-0.0728	-0.0848	-0.0854	-0.0928
	(0.132)	(0.126)	(0.131)	(0.124)	(0.131)	(0.126)
Male	0.168**	0.163**	0.122	0.122	0.0887	0.0919
	(0.0678)	(0.0712)	(0.0964)	(0.0950)	(0.0908)	(0.0902)
Employment rate (municipality)	-2.993***		-2.983***		-3.028***	
	(0.870)		(0.884)		(0.879)	
No. of individuals	5,279					
No. of obs.	20,663					
Controls:						
Assigned municipality characteristics	Yes	No	Yes	No	Yes	No
Assigned municipality FE	No	Yes	No	Yes	No	Yes

Source : Administrative registers from Statistics Denmark 1997-2020 linked with dataset on the individual's neighborhood of residence constructed by Damm, Hassani and Schultz-Nielsen (2021a).

Notes : ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$. Panel A: Coefficient estimates from linear probability models. Panel B: Coefficient estimates from linear regressions. Standard errors clustered by municipality of assignment in parentheses. Sample: Estimation Sample described in notes to Table 4. Additional controls: personal characteristics at arrival, country of origin, year of arrival, month of arrival, dummy for missing information on educational attainment at arrival. Assigned municipality characteristics are population share, non-Western share, co-national share, crime conviction rate, share of households with high-income. Individuals in the sample have been assigned to 94 different municipalities and 914 different neighborhoods. Mean (std. dev.) co-ethnic neighbor share based on co-nationals, language fellows, and non-Westerns are 0.003 (0.008), 0.014 (0.036), and 0.051 (0.056), respectively. Mean (std. dev.) employment rate of ethnic enclave based on co-nationals, language fellows, and non-Westerns are 0.179 (0.296), 0.323 (0.292), and 0.500 (0.129), respectively. Mean (std. dev.) employment rate in the assigned municipality: 0.769 (0.040). Mean (std. dev.) of Ever enrolled in Danish language course by year four: 0.949 (0.220). Mean (std. dev.) of Highest Completed Module Level in year t+: 2.936 (1.808)

Table A12 Descriptive Statistics on Personal Income Source by Year Four Since Migration

	Estimation sample					Enrolled in a Danish language course by year four					Never enrolled in a Danish language course by year four				
	All	Single		Married		All	Single		Married		All	Single		Married	
		Men	Women	Men	Women		Men	Women	Men	Women		Men	Women	Men	Women
Employed (in Nov. of year)	0.450 (0.498)	0.558 (0.497)	0.201 (0.401)	0.519 (0.500)	0.215 (0.411)	0.462 (0.499)	0.568 (0.495)	0.197 (0.398)	0.529 (0.499)	0.234 (0.424)	0.220 (0.415)	0.378 (0.487)	0.246 (0.434)	0.161 (0.371)	0.000 (0.000)
Any child births	0.078 (0.269)	0.047 (0.211)	0.076 (0.265)	0.112 (0.315)	0.074 (0.262)	0.080 (0.271)	0.047 (0.211)	0.073 (0.260)	0.114 (0.318)	0.078 (0.268)	0.047 (0.213)	0.045 (0.208)	0.105 (0.310)	N.A.	N.A.
Number of child births	0.358 (0.639)	0.158 (0.447)	0.416 (0.709)	0.518 (0.665)	0.578 (0.743)	0.396 (0.644)	0.158 (0.446)	0.429 (0.711)	0.524 (0.666)	0.621 (0.755)	0.189 (0.499)	0.162 (0.458)	0.281 (0.675)	0.268 (0.556)	0.097 (0.298)
No public transfers	0.024 (0.154)	0.034 (0.182)	0.012 (0.108)	0.026 (0.159)	0.007 (0.082)	0.025 (0.155)	0.035 (0.184)	0.008 (0.090)	0.026 (0.160)	0.007 (0.085)	0.020 (0.141)	N.A.	N.A.	N.A.	0.000 (0.000)
Amount of public transfers	87,825 (66,859)	60,459 (53,185)	138,338 (77,969)	85,251 (59,820)	120,577 (67,234)	87,350 (67,066)	59,588 (52,864)	141,490 (78,329)	84,421 (59,463)	121,756 (67,889)	96,704 (62,314)	76,116 (56,666)	104,269 (65,418)	116,379 (65,225)	107,152 (57,996)
Public transfers constitute 100% of personal income	0.398 (0.489)	0.254 (0.436)	0.673 (0.469)	0.333 (0.471)	0.686 (0.464)	0.382 (0.486)	0.243 (0.429)	0.675 (0.469)	0.320 (0.466)	0.659 (0.474)	0.696 (0.461)	0.459 (0.501)	0.649 (0.481)	0.821 (0.386)	1.000 (0.000)
Earnings as share of personal income	0.371 (0.496)	0.483 (0.516)	0.150 (0.292)	0.413 (0.540)	0.171 (0.313)	0.380 (0.501)	0.492 (0.520)	0.144 (0.288)	0.420 (0.542)	0.186 (0.322)	0.187 (0.345)	0.322 (0.408)	0.214 (0.336)	0.134 (0.325)	0.000 (0.000)
Earnings constitute 100% of personal income	0.017 (0.129)	0.023 (0.151)	0.007 (0.086)	N.A.	N.A.	0.018 (0.131)	N.A.	N.A.	N.A.	N.A.	N.A.	0.000 (0.000)	N.A.	N.A.	0.000 (0.000)
N	5,828	2,107	673	2,156	892	5,532	1,996	616	2,100	820	296	111	57	56	72

Source : Administrative registers from Statistics Denmark 1997-2020.

Notes : N.A. means "not available" for reasons of discretion.