

**REFUGEES' SETTLEMENT  
PATTERN, EDUCATIONAL  
BACKGROUND AND  
FORMAL DANISH LANGUAGE  
ACQUISITION**

ANNA PIIL DAMM

AHMAD HASSANI

TRINE SKRIVER HØHOLT JENSEN

MATHIAS SCHINDLER

MARIE LOUISE SCHULTZ-NIELSEN

Refugees' Settlement Pattern,  
Educational Background and  
Formal Danish Language Acquisition

---

Study Paper No. 168

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](mailto:kontakt@rff.dk)

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

January 2022

# Refugees' Settlement Pattern, Educational Background and Formal Danish Language Acquisition<sup>\*</sup>

Anna Pii Damm<sup>†</sup>, Ahmad Hassani<sup>‡</sup>, Trine Skriver Høholt Jensen<sup>§</sup>, Mathias Schindler<sup>\*\*</sup>,

Marie Louise Schultz-Nielsen<sup>††</sup>

## Abstract:

This study describes the geographic settlement patterns of the population of adult asylum seekers who obtained refugee status in Denmark between 2004 and 2015 and their educational background and formal Danish language acquisition during their first years in Denmark. Upon receipt of a residence permit, refugees were subject to spatial dispersal across municipalities to balance the non-Western immigrant share across municipalities. Our results are consistent with the spatial dispersal policy. Municipalities with a relatively low share of non-Western immigrants receive a disproportionate share of new refugees. Refugees are provided with housing in neighbourhoods with relatively high shares of public housing, co-nationals and individuals of low socioeconomic status. The refugees' educational attainment from abroad is generally modest; among refugees with self-reported education from abroad, 62% have completed at most 10 years of schooling. While most refugees enrol in a Danish language course offered by the municipality as part of the three-year introduction programme, within four years of asylum, only 54% have completed a language course. Refugees with a longer education tend to enrol in, and complete, a higher Danish course level. Completion rates are higher among men, refugees with a longer education, those settled outside population dense areas, and among certain nationalities (Eritreans and Iranians compared with Russians and Iraqis).

**Keywords:** Human Capital; Host Country Language; Humanitarian Migrants; Migration

**JEL Codes:** J24; J61; R23

---

<sup>\*</sup> 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 thank Francesco Fasani for helpful comments and suggestions for improvements and Marie Steiness at the language school A2B for explaining how language courses are administrated and students' results are reported. 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. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper. Workplace of authors: Damm, Hassani and Jensen: Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V; Schindler and Schultz-Nielsen: ROCKWOOL Foundation Research Unit, Ny Kongensgade 6, 1472 København K.

<sup>†</sup> Email: apd@econ.au.dk.

<sup>‡</sup> Email: ahass@econ.au.dk.

<sup>§</sup> Email: tshj@econ.au.dk.

<sup>\*\*</sup> Email: mathias.schindler@bse.eu.

<sup>††</sup> Email: mls@rff.dk. Corresponding author.

## 1. Introduction

This study provides an overview of refugees' geographical settlement in Denmark 2004-2015, their educational background and the Danish language acquisition they obtain during their first four years in Denmark. Furthermore, the study explores how host-country language acquisition vary with refugees' characteristics (e.g. education) and settlement. The paper provides descriptive analysis of the micro data used in our related study on the effects of residence-based co-ethnic networks on host-country language acquisition, see Damm, Hassani, Jensen and Schultz-Nielsen (2021).

Refugees' ability to speak the language of the host country can be vital not only for their labour market attachment and for supporting their children in school, but also for their participation in civic life in general. Accordingly, language training is a key part of the introduction programme for newly arrived refugees in many countries, including the Nordic countries (Hernes et al., 2019).

In Economics, language skill is considered 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).

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). Lazear provides empirical evidence. He shows how smaller minority groups in the US 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.

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. 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 favourably selected and, consequently, to be less efficient in learning a new language. 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 ethnic enclaves is expected to lower exposure to the host-country language.

The aim of our study is twofold: to provide an overview of newly arrived refugees' settlement in Denmark during 2004-2015, their educational background and formal Danish language acquisition; and to document how refugees' host-country language investments vary according to refugees' own characteristics as well as neighbourhood characteristics, in particular the share of language enclave

members. We argue that in studies of host-country language acquisition, the share of language enclave members can be a better measure of the size of the minority group than the share of co-nationals, because refugees can easily communicate with both their co-nationals and with other residents who share their language.

The current Danish language course system has been in place since 2004. To study formal Danish language acquisition during the first four years after asylum, we focus on refugee cohorts arriving between 2004 and 2015. Our sample consists of the 20,390 adult refugees who obtained a residence permit in Denmark during that period. The analysis includes information about their formal Danish language acquisition, the educational skills they bring from abroad, and the municipality and neighbourhood in which they are initially settled in Denmark. This was done by combining administrative registers from Statistics Denmark with information from ‘Danskundervisningsdatabasen’<sup>7</sup> and recently constructed data on the individual’s neighbourhood of residence by Damm et al. (2019a). Refugees’ Danish acquisition is measured by their enrolment in and completion of the formal Danish language courses.

Upon receipt of their residence permit, refugees are subject to spatial dispersal across municipalities to obtain a more even distribution of non-Western immigrants across municipalities. Our results are consistent with the spatial dispersal policy. Refugees tend to be settled initially in municipalities with a relatively low share of non-Western immigrants. 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.

The refugees’ educational attainment from abroad is generally modest; among refugees with self-reported education from abroad, 62% have completed at most 10 years of education, 12% have attained between 11 and 13 years of education, and the remaining 26% have attained more than 13 years of education.

Although most refugees (91%) are enrolled in a Danish language course, the share having completed a language course after four years is considerably lower (54%). The most ‘popular’ course is Danish II, which is attended by half of all refugees, followed by the less demanding Danish I, and the most advanced course, Danish III. Attained education from the source country is considered in the assignment of refugees to the entry-level Danish language course. Refugees with a longer education tend to enrol in and complete a higher Danish course level. Completion rates are higher among men, those with a longer education, among certain nationalities (e.g., Eritreans and Iranians rather than Russians and Iraqis) and among refugees settled outside population dense areas.

Our study contributes to two strands of literature. First, it adds to the literature on immigrant settlement patterns. Research regarding immigrants’ settlement (including refugees) in Denmark is found in some previous studies including Damm, Tranæs and Schultz-Nielsen (2006), Andersen (2015), Andersen (2017), Damm, Hassani and Schultz-Nielsen (2019b). Some studies describe immigrant settlement by country of origin (Statistics Denmark, 2016) or world region (Nielsen, 2014). Only few studies describe the settlement of refugees despite the use of dispersal policies for refugees in the US

---

<sup>7</sup> The Danish Language Education Database.

and several European countries (including Germany, Ireland, the Nordics, Switzerland and the UK). For an overview of the literature see Fasani, Frattini and Minale (2021). For Denmark these studies include; Damm (2005; 2009a), Nielsen and Jensen (2006). Studies that identify refugee status without measurement error are rare; Azlor, Damm and Schultz-Nielsen (2020) is a notable exception. Damm (2005; 2009a) describes the geographical settlement pattern of refugees who were subjected to the first Spatial Dispersal in Denmark, which was in place between 1986 and 1998. Our study describes the geographical settlement pattern of refugees who were subject to the second Spatial Dispersal Policy in Denmark in place between 1999 and June 2016, focusing on the cohorts arriving between 2004 and 2015.

Second, our study is related to the literature on immigrants' host-country language acquisition. For a general overview of the literature, we refer to Chiswick and Miller (2014) and a more recent study by Zorlu and Hartog (2018). A few Danish studies describe immigrants' language acquisition by country of origin (Larsen, 2002; Schultz-Nielsen, 2010; Arendt, Nielsen and Jacobsen, 2016). These studies have commonly used surveys on Danish language proficiency, since registers on formal Danish language course participation and completion have existed only since 2004. An exception is Arendt and Bolvig (2020); based on register data, they find that early on-the-job training reduces refugees' language test scores. However, refugees' labour market attachment may in the long run benefit from an extensive language training (Arendt, Bolvig, Foged, Hasager and Peri, 2021). The spatial dimension of immigrants' language proficiency is often limited to regional or municipal differences and is used for benchmarking analysis (Husted and Heinesen, 2004; Husted, 2008; Husted, Bonfils, Lauritzen and Nielsen, 2010; Jacobsen, Kiil and Kolodziejczyk, 2016); in such studies the main outcome of interest is the language schools' competence to teach the Danish language to students (including refugees, family-reunified individuals, labour migrants etc.). A recent RCT study explored the effect of bilingual instruction on immigrant children's Danish proficiency and finds no effect (Andersen, Humlum and Guul, 2017). Our study is the first to study individual investment in and acquisition of formal Danish language proficiency.

The rest of the study is structured as follows. In Section 2 we describe the institutional setting. Section 3 presents data and sample selection. Section 4 explains the geographical settlement patterns of refugees upon receipt of asylum, and Section 5 describes refugees' education from abroad. Section 6 tracks formal Danish language acquisition by levels and progression, and Section 7 investigates the interplay between refugees' Danish language acquisition, education and settlement. Section 8 concludes our analysis and proposes avenues for future research.

## **2. Institutional Setting**

The following section describes the main features of the legislation on the right to asylum and residency in Denmark, initial settlement of refugees, and the introduction programme with emphasis on the Danish language courses system in place since 2004. We focus on the legislation in place at the time of arrival of individuals in our refugee sample, 2004–2015.

## 2.1 Asylum Policy in Denmark 2004–2015

The rules regarding refugees' right to residency in Denmark are defined in the Aliens Act, which is founded on the United Nations' 1951 Refugee Convention. The convention states that people with a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group, or political opinion should be afforded protection (Christensen et al., 2000, p. 22). The refugees are sometimes referred to as having 'convention status'. People having substantive reasons for fleeing other than those listed in the UN Refugee Convention may also be granted residence. However, the rules were changed in July 2002. The 'de facto' refugee status was replaced by a new 'B status', while the criteria for 'convention status' remained unchanged. 'B status' provides less protection for war refugees and asylum seekers with subjectively perceived fear (ibid). Rules regarding 'first country of asylum' were tightened, which makes it more difficult to obtain a residence permit if the asylum seeker has a close connection to a third country.

Denmark received close to 500 'quota refugees' every year during 1989–2015. These refugees are settled in Denmark in agreement with the United Nations High Commissioner for Refugees (UNHCR). However, during 2015 and 2018, Denmark suspended the acceptance of quota refugees. In 2019 and 2020, Denmark agreed to receive 50 and 200 'quota refugees', respectively.<sup>8,9,10</sup>

Denmark joined the Dublin Convention in 1997. The Convention aims to ensure that asylum seekers' cases are processed in the first EU country of arrival. The agreement means that fewer asylum cases are processed in Denmark, as many refugees first arrive in other EU countries.

The Aliens Act was revised close to 80 times in 2004–2015, and during that time the refugees in our study obtained residence permits (Hvidtfeldt and Schultz-Nielsen, 2018). Commenting on all the revisions are beyond the scope of this article. Based on Hvidtfeldt and Schultz-Nielsen (2018), a summary of the main trends is given below.

Danish rules regarding refugees' right to residency and family reunification were generally tightened during 2002–2011 under the right-of-centre government. Apart from the change from 'de facto' to 'B status', other major changes in the Aliens Act include the introduction of the '24-year rule' and a tightening of 'the attachment rule', which make it more difficult to obtain family reunification. However, in some cases, refugees are exempt from the provisions, namely when the refusal of a residence permit is in conflict with Article 8 of the European Convention on Human Rights (the right to respect for private and family life).

In 2005, the criteria for selecting quota refugees underwent some changes. Previously, the selection was based solely on the need for protection, but since 2005, refugees' potential for integration in Denmark has also been considered. The regulations regarding 'tolerated stay' had become stricter by the end of 2008. Individuals with 'tolerated stay' shall now report regularly to a specific asylum

---

<sup>8</sup> <https://uim.dk/nyheder/danmark-tager-200-kvoteflygtninge-fra-rwanda> (visited 19/10/2020). <https://uim.dk/COVID-19:Konsekvenser-for-de-resterende-kvoteflygtninge-pa-2019-kvoten---Udlændinge-og-Integrationsministeriet> (visited 06/01/2021).

<sup>9</sup> A residence permit can also be granted due to significant humanitarian considerations. However, these cases are rare. Residence permits can also be granted on 'other grounds', e.g., in relation to unaccompanied minors and to accompanying children.

<sup>10</sup> Most refugees in the current study have 'convention status', followed by refugees with 'B status' and quota refugees.

centre. In 2010, a point system for obtaining permanent residence after four years or more was introduced and rules on expulsion were tightened. In July 2011, stricter rules for family reunification were introduced: a point system was implemented, the ‘attachment rule’ regarding spouses’ aggregated ties with Denmark was tightened further and the guarantee sum that must be provided was raised.

With the election of the new left-of-centre government in October 2011, most of the restrictions imposed from 2011 were abolished. From May 2013, asylum seekers have been allowed to live outside the asylum centre after six months of stay, and they are also allowed to work; however, few work in practice.<sup>11</sup> In May 2014, the criteria for selecting quota refugees were changed once more as resettlement in Denmark should lead to ‘a lasting improvement in life situation’ for the refugee.

In February 2015, due to the increased inflow of asylum seekers from Syria, the Danish government decided to introduce a new one-year temporary protection status for those refugees who would otherwise had been granted ‘B status’. If the refugee still needs protection after one year, the temporary stay can be prolonged.

In June 2015, when the new right-of-centre government took over office, the influx of asylum seekers in Denmark reached a level unparalleled since the civil war in Yugoslavia in the early 1990s. The police were granted extended powers to deal with the situation, and since December 2015, heavy fines can be levied on transportation operators if they carry passengers into Denmark without valid documentation. In reality, the increased inflow of asylum seekers cancelled the Schengen Agreement and, as in many other Schengen countries, temporary border control was introduced in Denmark. Later, in February 2016, the length of time rules regarding family reunification for refugees were deferred by two years (from one to three years after the asylum case has been processed). As we examine only settlement and skill attainment of refugees who arrived before 2016, we do not describe later changes to the Aliens Act.

#### a. Danish Spatial Dispersal Policy on Refugees 1999–2016

When the asylum seeker obtains a residence permit, that individual becomes a ‘refugee’ in our terminology. Shortly after receiving a residence permit, the Danish Immigration Service (DIS) assigns the refugee to a local government. The local government is responsible for housing provision.<sup>12</sup> We refer to this as the ‘first settlement’.

DIS disperses new refugees across municipalities according to a quota system. The calculations are explained in detail in Azlor et al. (2020). The main feature of the quota system is that DIS makes an annual forecast of the number of new refugees (expected to obtain a residence permit) in the following calendar year. Municipal quotas of refugees are then calculated based on the population statistics such that municipalities with a relatively low share of non-Western immigrants receive a higher share of newly arrived refugees.<sup>13</sup> Refugees are always settled together with close family members and are

---

<sup>11</sup> <https://www.dr.dk/ligetil/kun-31-asylansoegere-kom-i-job-sidste-aar> (accessed on 21 January 2021).

<sup>12</sup> For refugees arriving 2005–2010 it took an average of 40 days from receiving a residence permit to being registered in a municipality (Hvidtfeldt and Schultz-Nielsen, 2018).

<sup>13</sup> ‘Immigrants’ in the refugee quota calculation are defined as foreign citizens from non-EU/EAA countries and non-Nordic countries up until July 2016. Hereafter, immigrants are defined as non-Western immigrants excluding not only citizens from EU28 and EEA (as well as Andorra, Monaco, San Marino, Switzerland, Vatican City, Canada, the USA,

often settled in a municipality with others of the same nationality in order for them to be able to create a network of co-nationals, i.e., fellow countrypersons. Other settlement criteria taken into account are refugees' educational qualifications and need for medical treatment (Executive Order no. 630 of 25/08/1998). These rules remained unchanged until 2016, after which employment prospects have been considered in the settlement of refugees (Executive Order no. 980 of 28/06/2016).

#### b. The Danish Language Education as Part of the Danish Integration Programme

Denmark has offered a three-year introduction programme to newly arrived refugees and family reunified individuals since 1999. Participation in the programme is a condition for eligibility to means-tested social assistance (introduction allowance). It is the responsibility of the municipality in which refugees are settled to offer the introduction programme soon after settlement. The programme consists of an introduction to Danish society and culture, job-oriented guidance as well as trainee periods, and an extensive Danish language course corresponding to 1.2 years of full-time participation (Act of Parliament 474 of 10/07/1998). The introduction programme must be scheduled to enable immigrants to simultaneously participate in the labour market. The duration and the main components of the introduction programme were unchanged until October 2016. Refugees arriving from October 2016 have been subject to the 'job-first reform'. The reform reduces the introduction programme from three years to one. The programme can be prolonged up to five years if the refugee has not found employment or has begun ordinary education.<sup>14</sup>

The Danish language courses are mandatory and are at no cost to participants in the introduction programme (called 'I-participants').<sup>15</sup> Other immigrants, such as students and immigrants with a residence permit for work (called 'S-participants'), can also participate in the language courses, but the conditions for payment etc. vary for the S-participants over time. As our analysis focuses on refugees and their family members, we concentrate on the rules related to I-participants.

The language courses are divided into three distinct course levels: 'Danish I' for individuals with very short or no previous educational background<sup>16</sup>, 'Danish II' for individuals with a short educational background, and 'Danish III' for individuals with medium or long educational background.<sup>17</sup> Course providers expect participants in Danish II to have a relatively slow language acquisition, while participants in Danish III are expected to have a faster acquisition of Danish.<sup>18</sup> Before starting a language

---

Australia and New Zealand), but also immigrants with residence permits in connection with work, study, internships and au pair work in accordance with Executive Order no. 980 of 28/06/2016 (Hvidtfeldt and Schultz-Nielsen, 2018).

<sup>14</sup> For refugees arriving from July 2013 to October 2016, parts of the introduction programme may have changed after October 2016 (Act of Parliament 665 of 08/06/2016; Arendt, 2019).

<sup>15</sup> If I-participants are on welfare benefits, participation in language courses is a requirement for receiving financial aid.

<sup>16</sup> Poor in the sense that individuals cannot read or write in their native tongue or do not have knowledge of the Latin alphabet.

<sup>17</sup> Author's own translation. 'Danish I', 'Danish II' and 'Danish 3' are officially called "Danskuddannelse 1", "Danskuddannelse 2" and "Danskuddannelse 3", respectively. Source: <https://uim.dk/arbejdsomrader/danskundervisning-og-prover-for-udlaendinge/danskuddannelse/danskuddannelse-1-2-og-3/> (accessed on 21 January 2021).

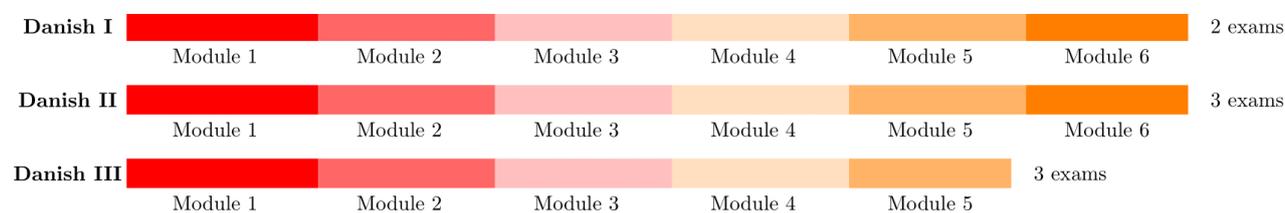
<sup>18</sup> According to the Curriculum of the language course, participants are expected to reach Danish language level 'A2' ('B1') after 4 (6) modules following Danish I, 3 (5) modules following Danish II and 2 (4) modules following Danish III (Ankestyrelsen, 2013). The classification follows the European Council (2001)

course, a screening test is conducted at a language school to determine the appropriate course level (Interview, Steiness).<sup>19</sup>

Each course level is subdivided into several single modules, as illustrated in Figure 1. The current structure dates back to 2004, when the current law ‘Danish Education for Adult Foreigners’ was first implemented (Act of Parliament 375 of 28/05/2003). Since 1999, there has been a comprehensive Danish language programme with three tracks. However, instead of modules, each language course consisted of 2–4 steps and the teacher decided when the participant should progress to the next step. The final step was completed through exams (Act of Parliament 487 of 01/07/1998; Husted, 2008).

As is evident from Figure 1, 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 to their skill level assessed by the screening test (Interview, Steiness). Each course level concludes with an exam. The exam is structured differently across course levels. Danish I concludes with a written and an oral test. Danish II and Danish III conclude with a written, an oral and a reading test.<sup>20,21</sup> After the Module 5 exam in Danish III, course participants can choose to take the 6<sup>th</sup> module, culminating in an exam called ‘Studieprøven’. ‘Studieprøven’ grants access to tertiary education in Denmark. Since Danish III is considered completed after passing Module 5’s exam and only 4% of refugees in our sample complete Studieprøven, Module 6 (Danish III) is not included in our analysis. 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.<sup>22</sup>

**Figure 1.** Structure of formal Danish language courses.



*Source:* Authors’ construction.

*Notes:* Duration of modules not to scale. The final exam for ‘Danish I’ consists of a written and an oral test. Final exams for ‘Danish II’ and ‘Danish III’ consist of a written, an oral and a reading test.

Figure 2 illustrates the typical timeline for refugees’ participation in a Danish course. After arriving in Denmark, refugees spend time in an asylum centre waiting to have their asylum application pro-

<sup>19</sup> The authors conducted an interview addressing the administration of language courses and reporting of students’ results with Marie Steiness at the A2B language school in Aarhus.

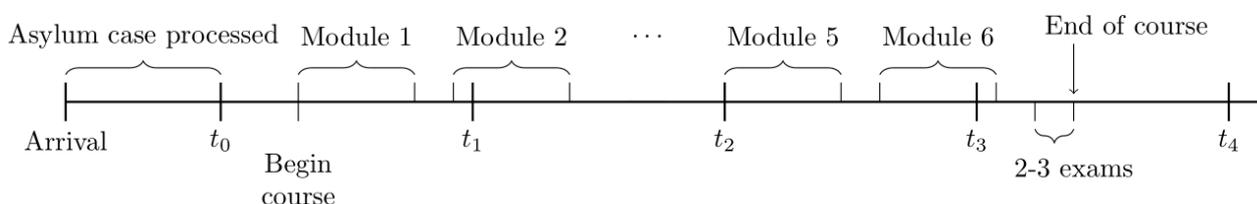
<sup>20</sup> The level of Danish skills required for passing the exams are the following: Danish I: written (A2), oral (B1), Danish II: written/reading (B1), oral (B1-B2), Danish III: (B2), ‘Studieprøven’: (C1). The classification follows the European Council (2001).

<sup>21</sup> The oral test has double weight compared with the written and reading test when calculating the GPA (Ankestyrelsen, 2012, 2013; Jacobsen et al., 2016).

<sup>22</sup> 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.

cessed. The waiting period varies between individuals and the year in which the application is processed (Hvidtfeldt and Schultz-Nielsen, 2018).<sup>23</sup> After obtaining asylum (denoted by  $t_0$  in Figure 2), refugees usually begin Danish courses within a year. The refugees sequentially complete modules that constitute a Danish course at a given level, as illustrated in Figure 1. The duration of the period also varies but usually takes 2–3 years (corresponding to  $t_2$  and  $t_3$  in Figure 2) for individuals in the considered sample. Each module level ends with a test. The test determines whether the individual can progress to the next module level. Module levels are usually completed sequentially, but in some situations the refugee may leave the course and return later, in which case the individual’s language proficiency is reassessed. A refugee completes a language course by taking the respective final exam for that course. The final exam is conducted after completion of the last module of the course level.

**Figure 2.** Timeline typical of an adult refugee’s first four years in Denmark.



*Source:* Authors’ construction.

*Notes:* Duration of modules not to scale.

### 3. Data and Sample Selection

Our access to high-quality administrative register data for the full population of refugees arriving in Denmark since 2004 allows us to distinguish individuals by date of their residence permit, country of origin, neighbourhood of residence, educational background, and language acquisition from Danish language courses. Consequently, we can explore how newly arrived refugees are settled across neighbourhoods in Denmark, as well as how they differ in terms of educational attainment from abroad and Danish language acquisition during the first four years of asylum. In particular, we study locations of settlement of refugees in Denmark over the last decade; examine their overrepresentation in particular types of municipality and neighbourhood; and examine the differences in refugee characteristics in terms of country of origin, gender, number of children, educational attainment on arrival and attainment of Danish language proficiency across settlement areas.

Our sample consists of individuals who were granted a residence permit in Denmark for reasons of asylum during 2004–2015. The term ‘arrival cohort’ is used when referring to refugees who receive their (first) residence permit for reasons of asylum in a given year. The performance of individuals (in a multitude of dimensions) in the formal Danish language courses is recorded in the register called ‘Danskundervisningsdatabasen’ (DUB). The DUB has recorded information since 2004. The most recent record in the DUB is from 30 September 2019, and we wish to track individuals for 4 years (allowing for potential extension of the introduction programme, e.g., due to maternity leave); 2015 is the latest arrival cohort considered in this analysis. We consider only adult refugees (18 years and

<sup>23</sup> Refugees arriving 2005–2010 waited on average 376 days from arrival until receipt of residence.

older) because individuals obtaining residence at younger ages are usually referred to the standard Danish education system.

In addition to Danish register data from Statistics Denmark, we used DUB to obtain detailed individual information regarding Danish language education for adults received by refugees. DUB contains information about individual participation and completion of module levels according to course level, as well as the start and end dates of each module. DUB provides individual information on completion of exams and the grades obtained. Course levels (Danish I, II and III) are defined as completed when all relevant exams (see Section 2.3) are registered as ‘completed’ in the data set.<sup>24</sup>

The initial address was identified using the residence change register (BEFBOP). BEFBOP provides information on the unique housing address where each person has lived since registration in Denmark’s population register (BEF). BEFBOP documents spells (durations) of residence in each housing unit. We can precisely identify the unique housing address of each refugee from the (exact) date of receiving a residence permit (i.e., initial housing address). By linking the unique housing addresses to the neighbourhood dataset constructed by Damm et al. (2019a), we can find the neighbourhood initially assigned (and those subsequent) for each refugee.

As explained in Appendix Table 1, the final sample was extracted using the following selection criteria. First, persons without a unique identifier (PNR) and persons with admission category imputed by Statistics Denmark in the residence permit register (OPHG) during 1997–2019 were dropped from our sample. Only the first residence permit of each person was considered. Next, we restricted our sample to individuals with refugee status who were granted residency during 2004–2015, observed in the population register (BEF) at least once during 2004–2019, and who were at least 18 years old on the date of arrival to Denmark. Third, we dropped from the sample individuals who were not observed in the population register in the same year as or the year after receiving their residence permit. Finally, we dropped persons without an identified neighbourhood or for whom the initial neighbourhood of residence was different from the first registered municipality of residence. This resulted in a sample of 20,390 adult refugees. We refer to this final sample as ‘the sample’.

As described in Table 1, the sample of adult refugees are settled across the five regions of Denmark: Northern Jutland, Central Jutland, Southern Denmark, Zealand (including adjacent islands), and the Capital Region. Panel A reports mean demographic and educational background characteristics of the individuals on receipt of asylum. Panel B reports ever enrolment in and completion rates four years after asylum of any Danish language course registered in DUB.

---

<sup>24</sup> The data show that it is possible for course participants to receive their diploma even though their average of grades obtained is less than 02.00, which is the general threshold for passing courses in Denmark.

**Table 1.** Individual characteristics of refugees, overall and according to region of initial settlement.

	Total	Northern Jutland	Central Jutland	Southern Denmark	Capital Region	Zealand
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Panel A: Individual characteristics at time of asylum</b>						
Age	32.04 (10.43)	31.17 (9.39)	32.30 (10.67)	32.68 (10.43)	32.43 (11.64)	31.13 (9.69)
Men	0.73 (0.44)	0.73 (0.45)	0.73 (0.45)	0.74 (0.44)	0.70 (0.46)	0.74 (0.44)
Youngest child aged 0–2 years	0.12 (0.33)	0.14 (0.35)	0.13 (0.34)	0.13 (0.34)	0.11 (0.31)	0.11 (0.31)
Youngest child aged 3–17 years	0.22 (0.41)	0.24 (0.43)	0.24 (0.42)	0.24 (0.43)	0.17 (0.38)	0.20 (0.40)
Top 5 countries of origin:						
Syria	0.48 (0.50)	0.39 (0.49)	0.51 (0.50)	0.54 (0.50)	0.41 (0.49)	0.45 (0.50)
Eritrea	0.13 (0.33)	0.19 (0.40)	0.12 (0.32)	0.09 (0.29)	0.11 (0.31)	0.16 (0.36)
Iran	0.07 (0.25)	0.06 (0.24)	0.06 (0.25)	0.04 (0.21)	0.10 (0.30)	0.08 (0.27)
Afghanistan	0.07 (0.25)	0.04 (0.20)	0.05 (0.23)	0.06 (0.24)	0.09 (0.28)	0.09 (0.28)
Somalia	0.05 (0.21)	0.05 (0.22)	0.06 (0.23)	0.04 (0.20)	0.04 (0.19)	0.05 (0.21)
Arrival cohorts:						
2004–2007	0.09 (0.29)	0.11 (0.31)	0.08 (0.28)	0.11 (0.31)	0.11 (0.32)	0.07 (0.26)
2008–2011	0.16 (0.37)	0.17 (0.37)	0.15 (0.36)	0.16 (0.36)	0.14 (0.34)	0.18 (0.39)
2012–2015	0.75 (0.43)	0.73 (0.44)	0.76 (0.43)	0.74 (0.44)	0.75 (0.43)	0.75 (0.44)
Education from abroad:						
≤10 years	0.36 (0.48)	0.41 (0.49)	0.36 (0.48)	0.36 (0.48)	0.31 (0.46)	0.35 (0.48)
10 < years ≤ 13	0.07 (0.25)	0.06 (0.24)	0.07 (0.25)	0.07 (0.25)	0.08 (0.27)	0.07 (0.25)
>13 years	0.15 (0.36)	0.13 (0.33)	0.13 (0.34)	0.15 (0.36)	0.18 (0.39)	0.16 (0.36)
Unknown	0.43 (0.49)	0.40 (0.49)	0.45 (0.50)	0.42 (0.49)	0.42 (0.49)	0.42 (0.49)
<b>Panel B: Formal Danish language acquisition zero to four years after asylum</b>						
Ever observed in DUB (0/1)	0.95 (0.21)	0.97 (0.17)	0.96 (0.20)	0.95 (0.22)	0.92 (0.26)	0.97 (0.17)
Ever enrolled (0/1) on 30 Sep.	0.91 (0.28)	0.93 (0.26)	0.91 (0.28)	0.92 (0.28)	0.86 (0.34)	0.94 (0.24)
Ever completed course (0/1)	0.54 (0.50)	0.58 (0.49)	0.50 (0.50)	0.56 (0.50)	0.47 (0.50)	0.60 (0.49)
No. of observations	20,390	2,848	5,252	5,342	3,096	3,852

Source: Authors' calculations based on administrative register data from Statistics Denmark, 2004–2020.

Notes: Sample: Adult refugees settled in Denmark 2004–2015 (N=20,390). Standard deviation in parentheses.

As Panel A shows, refugees are aged on average 32 years at the time of receipt of their residence permit, and 73% are men. The majority (66%) are not accompanied by children, while 12% are accompanied by a youngest child aged 0–2 years, and 22% have a youngest child aged 3–17 years on arrival. Of the adult refugees, most come from Syria (48%), with others coming from Eritrea (13%), Iran (7%), Afghanistan (7%), and Somalia (5%). Among refugees arriving during 2004–2015, the majority (75%) arrived during the later years (2012–2015) when the European countries experienced a sudden increase in the influx of refugees, notably from Syria, which culminated in the autumn of 2015. Refugees' education from abroad is unknown in the registers for 43% of the sample. Around 36% have at most 10 years' education from abroad, while another 7% have more than 10 to 13 years' education, and the remaining 15% have more than 13 years' education. The DUB holds data on 95% of the refugees, meaning they had contact with the Danish language schools in 2004 or later. At the end of September of any given year within the first four years of obtaining a residence permit, 91% were registered as having been enrolled in a Danish language course at least once, with 54% having completed a language course.

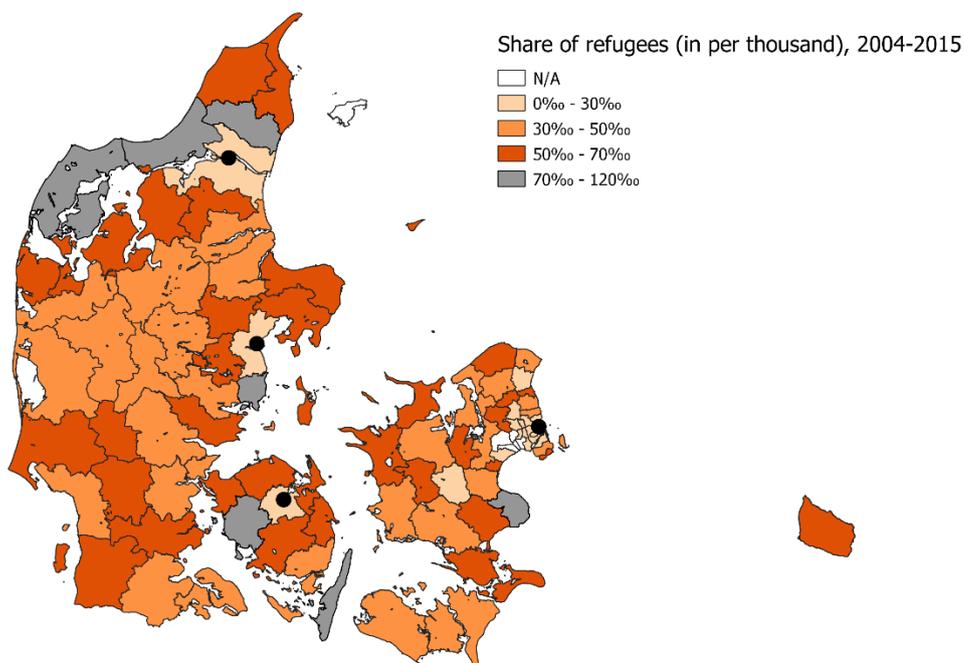
Although refugees settled in various regions tend to share many of the same characteristics, there are some differences. Iranians and refugees with many years of education are initially overrepresented in the Capital Region, Syrians have more often settled in Central and Southern Jutland, and Eritreans in Northern Jutland. The Danish language course enrolment and completion rates are relatively low among refugees in the Capital Region. However, in general, the regional differences are modest.

#### **4. Settlement of Refugees in Denmark 2004–2015**

##### **4.1 Settlement across Municipalities**

During 2004–2015, a total of 20,390 adult refugees (with a known residence permit status) settled in Denmark. Among these refugees, 92% were still living in Denmark in 2020, which emphasises the permanency of refugees' settlement. Figure 3 shows the refugees' distribution by the share of refugees per thousand inhabitants across the 98 municipalities, calculated as the total number of settled refugees during 2004–2015 in the municipality relative to the number of inhabitants in the municipality in 2015.

**Figure 3.** Number of settled refugees during 2004–2015 per thousand of the municipal population in 2015.



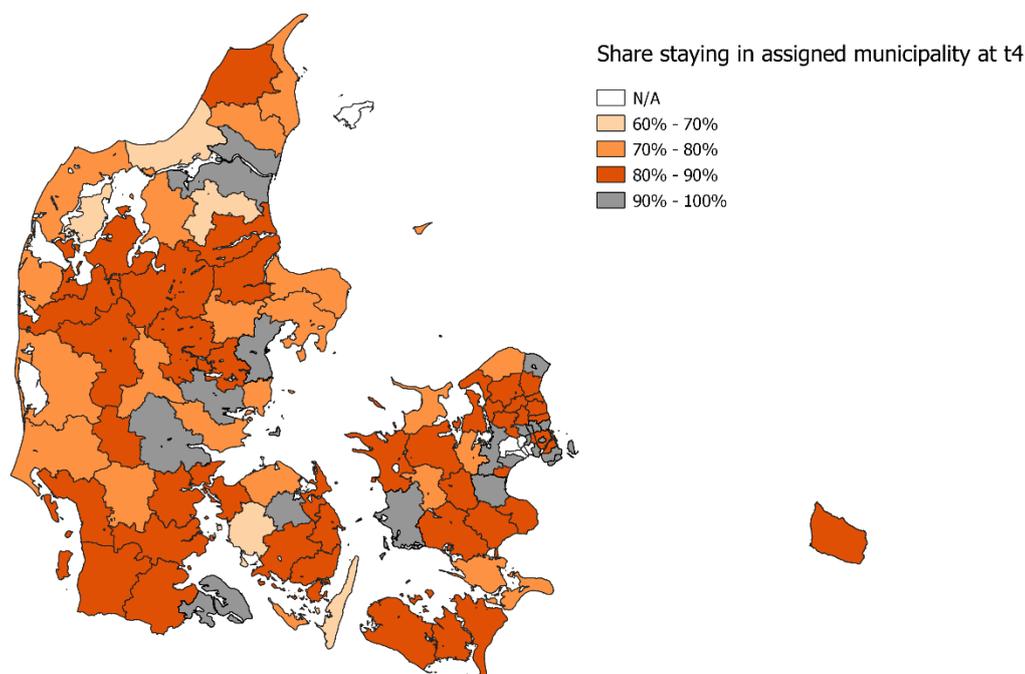
*Source:* Authors' calculations based on administrative register data from Statistics Denmark, 2004–2020.

*Notes:* Sample: Adult refugees settled in Denmark during 2004–2015 (N=20,390). Number of municipalities: 98. Number of settled refugees 2004–2015 per thousand of the municipal population in 2015. Some municipalities are categorised as 'N/A' for compliance with data privacy rules due to few observations (Vallensbæk, Læsø, Brøndby, Albertslund, Ishøj and Høje-Taastrup). The four largest municipalities (Copenhagen, Aarhus, Odense and Aalborg) are marked with a dot.

Figure 3 shows the share of refugees settled (per thousand) is lowest (0–30‰) in the four largest municipalities (Copenhagen, Aarhus, Odense and Aalborg) and in other municipalities with high overrepresentation of non-Western immigrants (e.g. Albertslund, Brøndby, Høje Tåstrup and Ishøj). In many municipalities of Region Zealand, the shares are 30–70‰, which is also the case in Central and Southern Jutland. The larger shares of refugees (>70‰) are found in parts of Northern Jutland, such as Thisted, Morsø, Jammerbugt and Brønderslev, and also in other municipalities across the country, such as Odder, Fanø, Assens, Langeland and Stevns. The shares show that many of the existing immigrants in Denmark live in larger cities. According to the formula for calculation of the

annual municipal quota of refugees, municipalities with relatively lower shares of non-Western immigrants receive higher annual quotas relative to their population size (Azlor et al., 2020). Under the spatial dispersal programme, refugees are more often settled in less urban areas.

**Figure 4.** Percentage of settled refugees during 2004–2015 who remain living in the municipality of initial settlement four years after asylum.



*Source:* Authors' calculations based on administrative register data from Statistics Denmark, 2004–2020.

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). Number of municipalities: 98. Some municipalities are categorised as 'N/A' for compliance with data privacy rules due to few observations (Vallensbæk, Læsø, Brøndby, Albertslund, Ishøj and Høje-Taastrup).

Eligibility for means-tested welfare benefits during the three-year introductory programme is conditional on residing in the assigned municipality. The eligibility creates a financial incentive to remain in the municipality for the first three years. Figure 4 shows the percentage of settled refugees who remain living in the initial municipality of settlement four years after asylum. More than 90% of refugees who settled in municipalities with larger cities, such as Aarhus, Odense and Aalborg, remained living in the municipality four years after asylum. The rate of staying is slightly between 80 and 90% in the municipality of Copenhagen and many surrounding municipalities. The rates of remaining are the lowest in municipalities with more rural areas such as Samsø (<60%), Jammerbugt, Morsø, Assens and Langeland (60–70%).

While an overwhelming majority (79%) of refugees stay at least four years in the initially assigned municipality, fewer (35%) stay in the initially assigned neighbourhood after four years. To measure refugees' duration of stay in the initially assigned neighbourhood and characteristics of the neighbourhood, we used the dataset on the individual's neighbourhood of residence constructed by Damm, Hassani, and Schultz-Nielsen (2019a). The dataset covers the full population in Denmark during 1986–2019. Neighbourhoods are constructed based on Danish geo-referenced housing units. The

housing units are clustered according to adjacent housing units into 1,961 (macro) residential neighbourhoods with at least 600 households<sup>25</sup>. The neighbourhoods are delineated by physical barriers, homogenous in terms of housing type and house ownership, are compact and have had unchanged boundaries for 34 years (1986–2019).

#### 4.2 Characteristics of Municipalities and Neighbourhoods of Settled Refugees

In the following, we analyse whether refugees are settled in neighbourhoods and municipalities that resemble the average composition of inhabitants in Denmark based on a set of key socioeconomic characteristics.

With a random distribution of refugees across locations in Denmark, we would expect the average characteristics of the municipalities and neighbourhoods of assignment of refugees to mirror those of the overall Danish population. Columns 2 and 3 in Table 2 present mean characteristics of residents in the municipalities and neighbourhoods of initial settlement of refugees, respectively. For comparison, Column 1 reports mean characteristics of the Danish population from 2004 to 2015. Refugees are initially overrepresented in certain types of area. Refugees are settled in municipalities with an average share of non-Western immigrants (4.7%) close to the national average (4.5%)<sup>26</sup>; within municipalities they tend to be settled in neighbourhoods with a relatively high share of non-Western immigrants (7.7% versus a national average of 4.5%). Regarding employment, refugees are settled in municipalities with an average employment rate of 76.2%, which is almost identical to the national average (76.3%), but in neighbourhoods with a relatively low employment rate (71.8%). Similarly, refugees are settled in more ‘favourable’ municipalities (in terms of the shares of low-income individuals and criminals), but within those municipalities, they are referred to less ‘favourable’ neighbourhoods.

As noted by Lazear (1999), smaller minority groups (sharing a common culture and/or language) may be more likely to learn the host country language than would larger minorities. He provides evidence for US immigrants, showing they are more likely to be fluent in English if they live in a community with a low share of co-nationals. In the analysis of acquisition of host-country language, co-nationals or the ethnic enclave may not be the most relevant measure of the individual’s minority group. The individual may share a common language not only with fellow countrypeople, but also with immigrants from other source countries. To take the common language aspect into account, we developed a new measure called ‘language enclaves’. We identified language enclaves by counting the number

---

<sup>25</sup> On average 1,241 households live in each neighbourhood (ibid).

<sup>26</sup> Excluding the large 2015-arrival cohort from the refugee sample, the share of non-Western immigrants in the assigned municipalities is below the national average (4.3%). This is as expected since the settlement of refugees across municipalities is based on a quota system that assigns a lower share of refugees to municipalities with a relatively high share of non-Western immigrants.

of individuals in each municipality and neighbourhood who share the same official language(s) with a refugee’s origin country. A detailed description of the methodology is provided in Appendix 1.

**Table 2.** Characteristics of total population and the municipalities and neighbourhoods where refugees are settled, 2004–2015. Mean (standard deviation).

	Total Population	Refugees	
	Country (1)	Municipality (2)	Neighbourhood (3)
Non-Western immigrant share (%)	4.49 (0.35)	4.67 (2.41)	7.66 (9.07)
Employment rate (%)	76.27 (1.43)	76.23 (3.83)	71.81 (10.17)
Share of low-income individuals (%)	7.38 (0.82)	6.27 (2.17)	7.83 (5.10)
Share of criminals (%)	0.86 (0.05)	0.74 (0.18)	0.91 (0.55)
Public housing share (%)	19.18 (0.24)	16.01 (6.94)	23.93 (22.32)
Other rental housing share (%)	19.05 (0.82)	17.87 (4.76)	19.04 (12.99)
<b>Inhabitants</b>	<b>5,518,475</b>	<b>71,440</b>	<b>3,038</b>
Total # of municipalities	98	98	
Total # of neighbourhoods	1,961		1,358
Total # of individuals		20,390	20,390

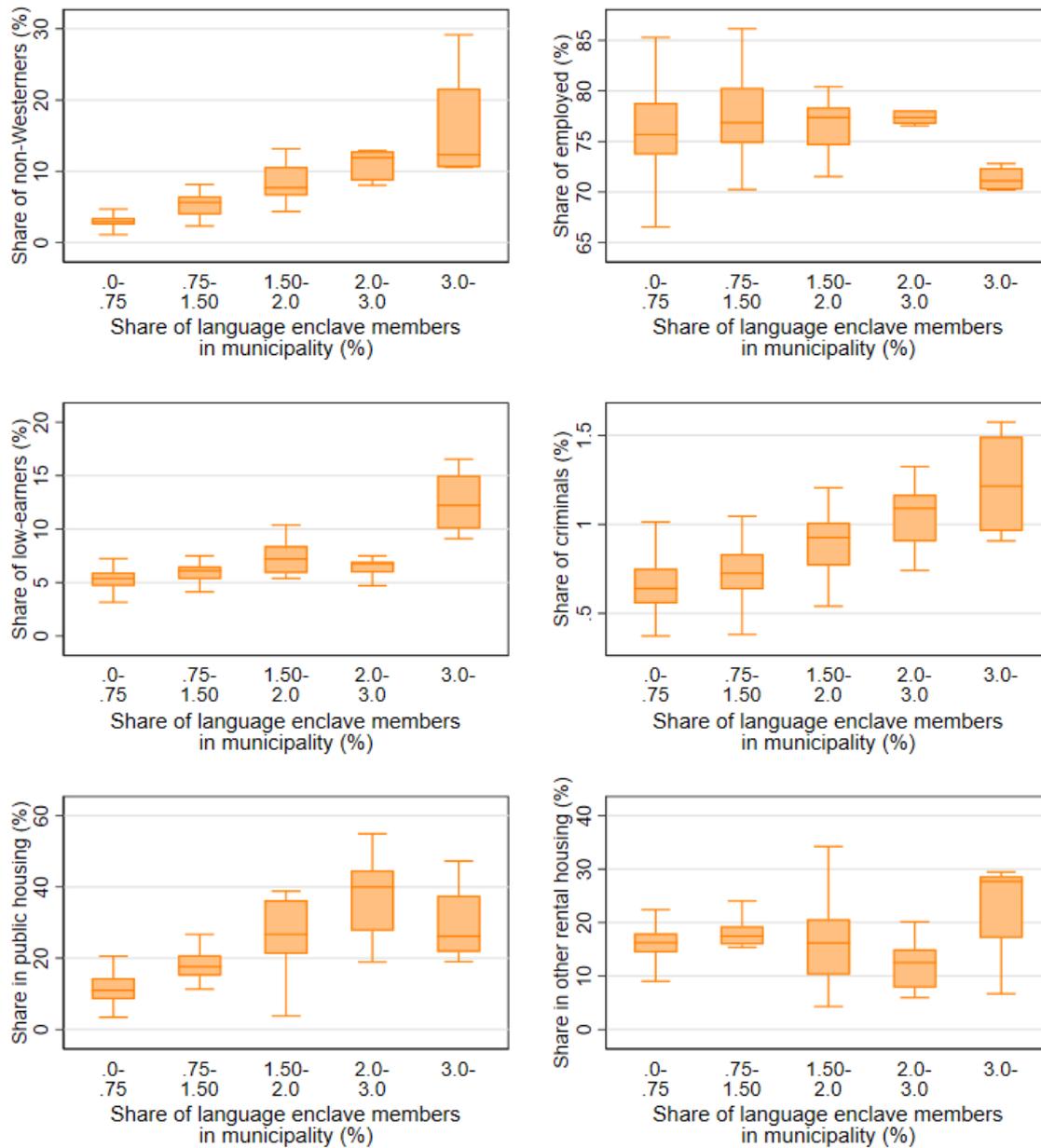
*Source:* Authors’ calculations based on data from Statistics Denmark, linked with dataset on neighbourhood of residence constructed by Damm et al. (2019a).

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). Neighbourhood corresponds to a macro-neighbourhood in Damm et al. (2019a). Non-Western immigrants include all immigrants except those from EU28 and EEA, Andorra, Monaco, San Marino, Switzerland, Vatican City, Canada, the USA, Australia and New Zealand. Low-income individuals are defined as residents with family-equivalent disposable income less than half of the median of such income in the region of residence. Standard deviation in parentheses.

In Figures 5 and 6 we analyse the correlation between the share of language enclave members<sup>27</sup> and the socioeconomic characteristics as well as housing ownership type listed in Table 2. The boxplot illustrates the distribution of each characteristic for the share of language enclave members in a given interval. At the municipal level (Figure 5), a positive correlation is seen between the share of language enclave members and other socioeconomic characteristics, such as share of low-income individuals, criminals, and non-Western residents. A positive correlation is also found for the share of public housing until language enclave members constitute 3% of the residents in the municipality. By contrast, we do not find a clear-cut relation between the share of language enclave members and the employment rate or the share of ‘other rental housing’.

<sup>27</sup> The number of language enclave members of individual  $i$  in location  $j$  is the number of residents who share the same official language(s) as individual  $i$  in location  $j$ . The share of language enclave members of individual  $i$  in location  $j$  is the number of residents who share the same official language(s) relative to the overall number of residents in location  $j$ .

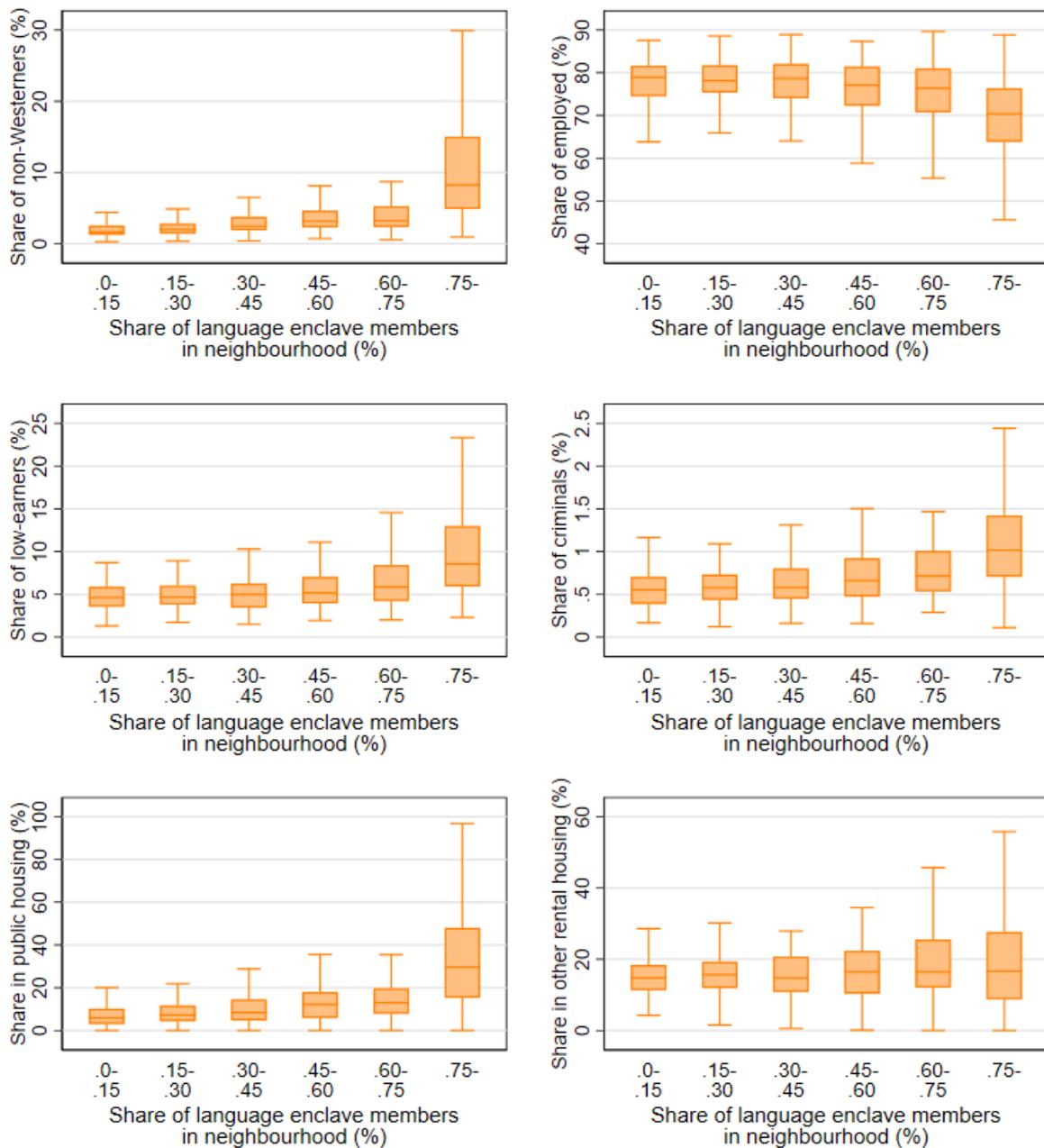
**Figure 5.** Correlation between share of language enclave members and other municipal characteristics.



*Source:* Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020.

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). Number of municipalities of settlement: 98. Non-Western immigrants include all immigrants except those from EU28 and EEA, Andorra, Monaco, San Marino, Switzerland, Vatican City, Canada, the USA, Australia and New Zealand. Low-income individuals are defined as residents with family-equivalent disposable income less than the half of the median of such income in the region of residence. The bottom and top of the box form the first and third quartile, while the line in the box shows the median and the whiskers illustrate the lower and upper adjacent values.

**Figure 6.** Correlation between share of language enclave members and other neighbourhood characteristics.



*Source:* Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020, linked with register on neighbourhood of residence constructed by Damm et al. (2019a).

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). 'Neighbourhood' corresponds to a macro-neighbourhood in Damm et al. (2019a). Non-Western immigrants include all immigrants except those from EU28 and EEA, Andorra, Monaco, San Marino, Switzerland, Vatican City, Canada, the USA, Australia and New Zealand. Low-income individuals are defined as residents with family-equivalent disposable income less than the half of the median of such income in the region of residence. The bottom and top of the box form the first and third quartile, while the line in the box shows the median and the whiskers illustrate the lower and upper adjacent values.

The relation between the share of language enclave members and socioeconomic characteristics as well as housing ownership type is more distinct at the neighbourhood level (Figure 6). The share of language enclave members in the neighbourhood is negatively correlated with the local employment rate and positively correlated with each of the following neighbourhood characteristics: share of non-Western residents, share of low-income individuals, share of criminals, and share of public housing. The correlations are as expected because refugees in Denmark tend to have low household incomes and cannot afford housing in more affluent neighbourhoods, apart from public housing.

The new measure of language enclave members is compared with other minority group measures in Table 3. The table reports means, standard deviations and the correlation between share of co-nationals, language enclave members and non-Western immigrants.<sup>28, 29</sup>

**Table 3.** Correlations of characteristics for municipalities and neighbourhoods of refugee settlement, 2004–2015.

	Share of co-nationals	Share of language enclave members	Share of non-Western immigrants	Mean (%)	SD
<b>Panel A: Municipality</b>					
Share of co-nationals	1.00	-	-	0.17	0.18
Share of language enclave members	0.14	1.00	-	1.03	1.11
Share of non-Western immigrants	0.22	0.67	1.00	4.67	2.41
Observations	20,390				
<b>Panel B: Neighbourhood</b>					
Share of co-nationals	1.00	-	-	0.41	1.09
Share of language enclave members	0.58	1.00	-	1.88	4.00
Share of non-Western immigrants	0.52	0.76	1.00	7.66	9.07
Observations	20,390				

*Source:* Authors' calculations based on data from Statistics Denmark, linked with dataset on neighbourhood of residence constructed by Damm et al. (2019a).

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). Neighbourhood corresponds to a macro-neighbourhood in Damm et al. (2019a). Non-Western immigrants include all immigrants except those from EU28 and EEA, Andorra, Monaco, San Marino, Switzerland, Vatican City, Canada, the USA, Australia and New Zealand.

The 20,390 refugees are settled in municipalities (Panel A) in which co-nationals constitute 0.2%, language enclave members 1.0%, and non-Western immigrants 4.7%. The language enclave is larger than the ethnic enclave and includes a lower number of immigrants than the total number of non-Western immigrants. As expected, the share of language enclave members is positively correlated with the share of non-Western immigrants and the share of co-nationals. For individuals in our sample, the share of language enclave members has a higher correlation with share of non-Western immigrants (0.67) than with the share of co-nationals (0.14).

Consistent with residential segregation of immigrants and ethnic minority groups, minority group shares and standard deviations are larger at the neighbourhood level (Panel B) than at the municipality level (Panel A) and more correlated. At the neighbourhood level, 0.4% of refugees' neighbours are co-nationals, 1.9% are the same language enclave, and 7.7% are non-Western immigrants. At the neighbourhood level, the share of language enclave members has an even higher correlation with both the share of non-Western immigrants (0.76%) and the share of co-nationals (0.58%).

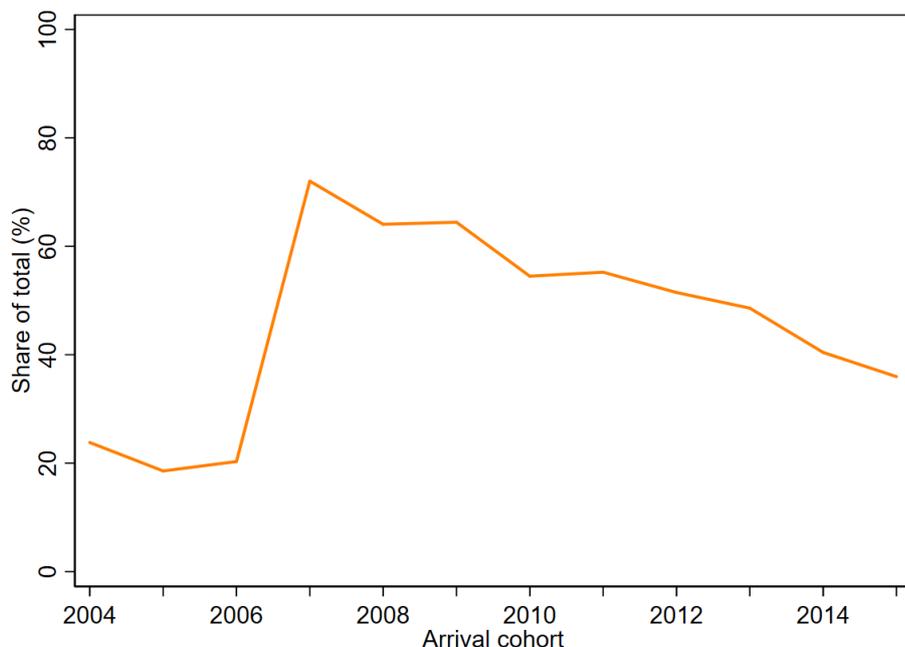
## 5. Refugees' Education upon Settlement in Denmark

Education upon settlement can be important for immigrants' (incl. refugees') efficiency in learning the host-country language (Chiswick and Miller, 2007). A previous Danish study finds the immigrants more likely to be good or fluent in Danish if they have a university degree from the home country (Schultz-Nielsen, 2010). The three language courses available for newly arrived immigrants in Denmark are intended for individuals with different educational backgrounds. An initial screening test is used to find the right course level for each participant before starting the course (see Section 2.3).

### 5.1 Differences in Education Attainment according to Arrival Cohort and Country of Origin

Information regarding refugees' education from the home country is typically gathered by Statistics Denmark through surveys. However, Statistics Denmark did not collect such annual surveys during 2007–2015. The survey was reintroduced in 2016. All adult immigrants in Denmark without information regarding education from abroad were contacted and asked to participate in the survey (Skaksen and Schultz-Nielsen, 2017). The result is shown in Figure 7.

**Figure 7.** Share of refugees with missing information on education from abroad according to arrival cohort



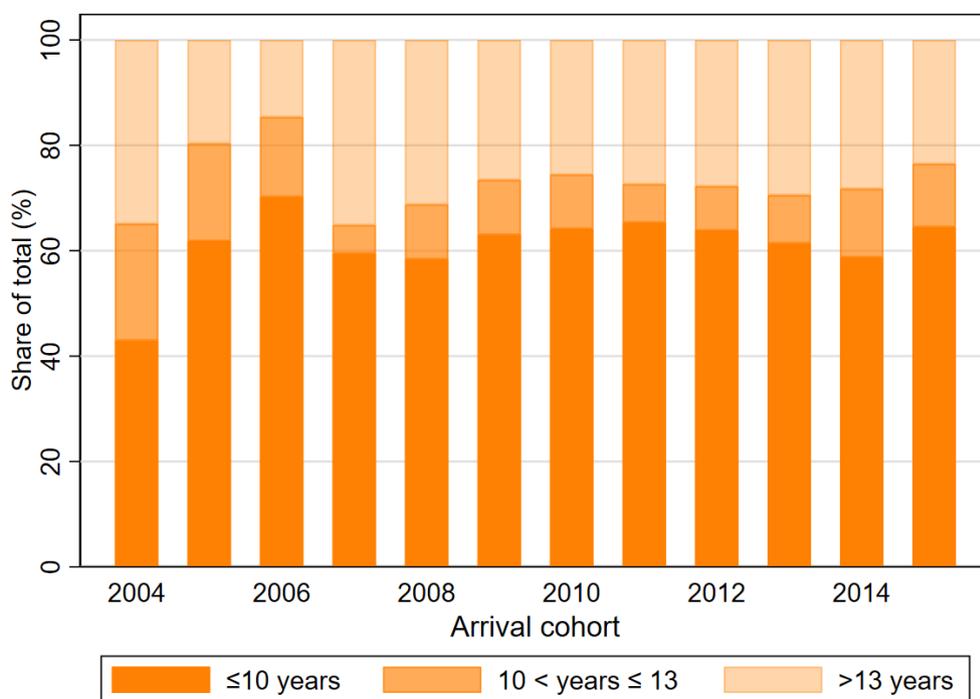
*Source:* Authors' calculations based on administrative register data from Statistics Denmark, 2004–2020.

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390).

Close to one fifth of the refugees who obtained asylum before 2007 have missing educational information from abroad. The share jumps to 72% for refugees arriving in 2007 due to the lapse in survey collection. As a result of participation in the 2016 survey, the share with missing educational information from abroad is lower for arrivals after 2007; the figure shows the share declines with arrival year and reaches 36% for the 2015 cohort.

Focusing on individuals in our sample for whom educational information from abroad exists, 62% have attained at most 10 years' education, 12% have attained between 11 and 13 years', and the remaining 26% have attained more than 13 years'. Figure 8 shows the shares vary according to arrival cohort. We divide the sample into three groups: 10 or less years of education, more than 10 and up to 13 years of education, and more than 13 years of education. The share of refugees with 10 or less years of education rises from close to 43% for refugees arriving in 2004 to around 70% for the 2006-asylum cohort and declines to a level close to 60% for the following cohorts. The share with 10–13 years of education declines from around 22% for the 2004-asylum cohort to 12% for the last (2015) cohort. The share with more than 13 years' education declines from 2004 to 2006 but rises again in 2007; the rise could be related to the change in data collection, and the share remains fairly stable thereafter. Overall, the share of refugees with short education ( $\leq 10$  years) has increased over the period 2004 to 2015, with most of the change occurring in the early years (2004–2006) *before* the rise in the share of refugees with missing education (Figure 7). The rising share with missing education information highlights the difficulty in assigning the time trend with certainty.

**Figure 8.** Level of education among refugees with available educational information according to arrival cohort.

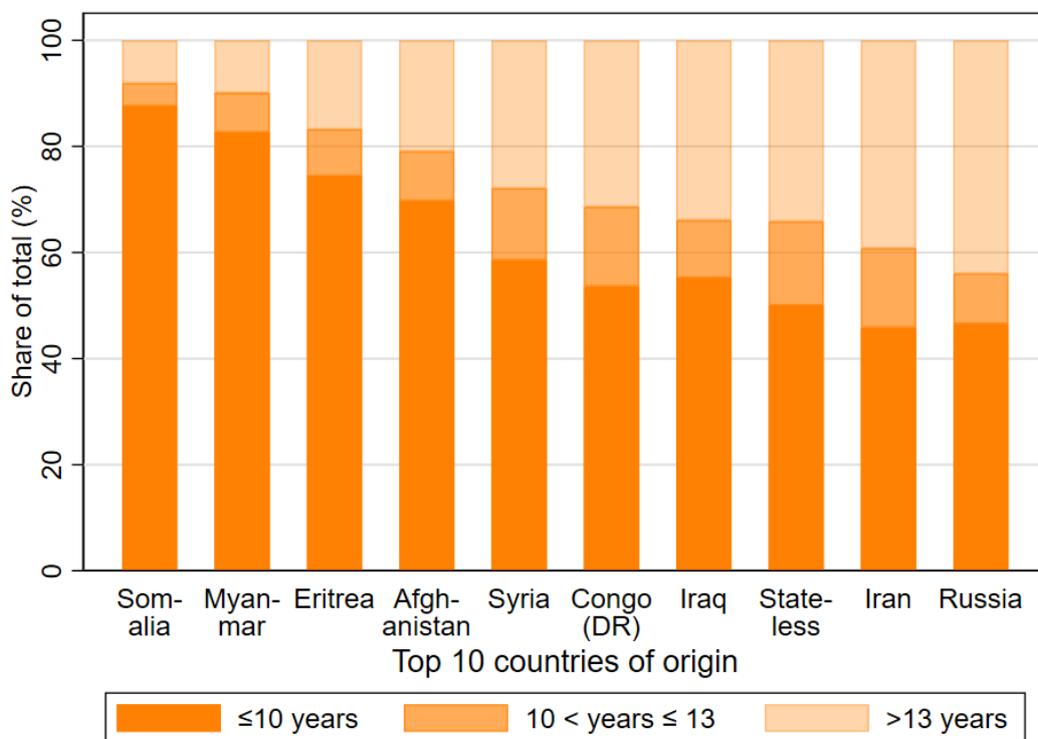


Source: Authors' calculations based on administrative register data from Statistics Denmark, 2004–2020.  
 Notes: Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390).

The level of education from abroad tends to be similar among refugees from the same country of origin. Firstly, because refugees' educational level reflects the supply of education provided by the education system of their home country; secondly because there are similarities in the circumstances under which co-nationals need to flee, e.g., if specific ethnic groups or professions are persecuted.

Figure 9 illustrates years of education from abroad by country of origin for the ten largest refugee groups arriving in Denmark from 2004 to 2015. The share with longest education (more than 13 years) is the highest among refugees from Russia (44%), followed by Iran (39%), Iraq (34%), DR Congo (31%), Syria (28%), Afghanistan (21%), Eritrea (17%), Myanmar (10%) and Somalia (8%). By contrast, the share with up to 10 years of education is the highest among refugees from Somalia (86%), Myanmar (83%) and Eritrea (74%).

**Figure 9.** Years of education among refugees with self-reported information about educational attainment on arrival according to country of origin.



Source: Authors' calculations based on administrative register data from Statistics Denmark, 2004–2020.

Notes: Sample: Adult settled refugees in Denmark 2004–2015 with self-reported information about highest attained education on arrival (N=11,627).

The majority of Iraqis in the sample obtained their residence permit in the period 2007–2009, during which Iraqi interpreters (for the Danish Military Forces in Iraq) and their families were invited to seek asylum in Denmark due to possible persecuted in Iraq for their collaboration with the Danish Military Forces. Russian refugees (including ethnic Chechens) arrived throughout the period 2004–2015, especially during 2012–2014. Iranians and Afghans mainly obtained refugee status in the very early years or after 2008, while refugees from Somalia arrived during 2012–2015. By contrast, the

refugees from Myanmar mainly obtained a residence permit during 2005–2008 and 2010–2011. Refugees from Syria and Eritrea, the two countries from which Denmark received the greatest number of refugees, obtained their residence permits in the later years. Refugees from Eritrea primarily obtained residency in 2015, while the Syrian refugees have obtained residency in increasing numbers since 2011.

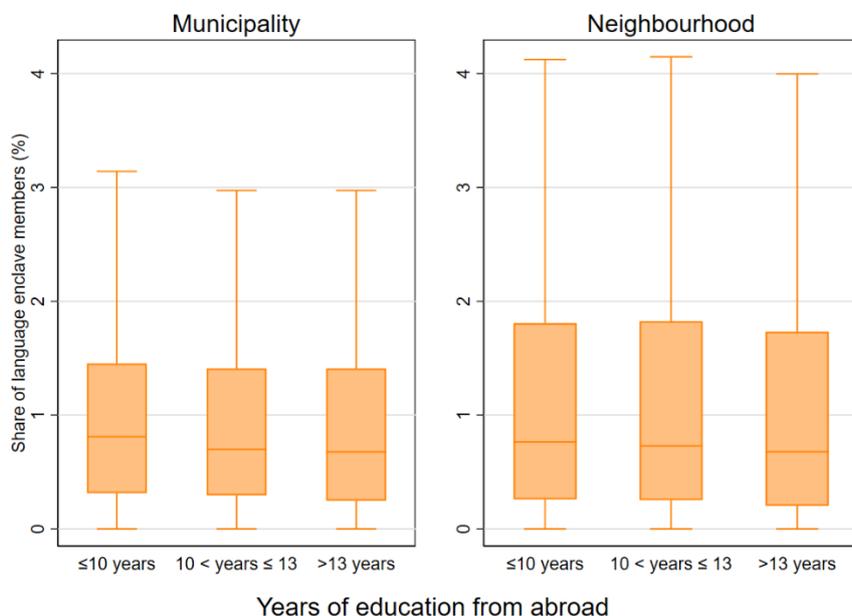
## 5.2 Do Refugees Differ in Educational Attainment from Abroad across Settlement Areas?

Within the municipal capacity to receive new refugees, determined by the annual quota, the authorities settle each refugee on the background of their country of origin; the authorities are more likely to settle a refugee in a municipality where their co-nationals already live. The presence of other minority group members can facilitate the settlement process due to (i) the existence of an established ethnic network, facilitating the introduction of newly arrived individuals to the local facilities; and (ii) the local authorities' prior experience with integration of refugees from the particular country of origin and with the established network of interpreters. Quasi-experimental studies find positive effects of living in established ethnic enclaves on labour market outcomes of immigrants (Edin, Fredriksson and Åslund, 2003; Damm, 2009b, 2014; Beaman, 2012). If the cost of learning the Danish language is lower among refugees with longer educations, one hypothesis could be that individuals with more education are settled in municipalities with fewer language enclave members, even though this is not a criterion in the official settlement policy.

Figure 10 presents the correlation between educational level and share of language enclave members in the location of initial settlement of refugees. The first part of the figure shows the distribution of the share of language enclave members in the municipality in a boxplot across educational levels. The second part shows a similar distribution using the share of language enclave members in the neighbourhood. The variance in the share of language enclave members is larger at the neighbourhood level than the municipality level. However, it is not evident that refugees with longer education have been settled in local areas with fewer enclave members, at either the municipality or the neighbourhood level. In Appendix Figure A3 the co-national share is used as a measure of the size of the individual's minority group and a similar conclusion is reached.

The low correlation between refugees' educational level from abroad and the share of language enclave members upon settlement is illustrated in Figure 10. For tests of whether the share of language enclave members in the local area upon settlement is indeed exogenous, with respect to refugees' educational attainment, conditional on individual demographic background, we refer to Damm, Hasani, Jensen and Schultz-Nielsen (2021).

**Figure 10.** Share of language enclave members in the local area upon settlement according to educational level.



*Source:* Authors’ calculations based on administrative register data from Statistics Denmark, 2004–2020, linked with register on neighbourhood of residence constructed by Damm et al. (2019a).

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). The bottom and top of the box form the first and third quartile, while the line in the box shows the median and the whiskers illustrate the lower and upper adjacent values.

## 6 Formal Danish Language Acquisition: Levels and Progression

For analysis of the Danish Language acquisition by refugees in Denmark, we utilised data from the DUB database as described in Section 3. The refugees’ enrolment in Danish education is recorded on 30 September in each of the four years after obtaining a residence permit. The completion of a course level is also recorded on 30 September (as defined by whether all relevant exams for the course level have been completed, see Figure 1). For refugees enrolled in multiple course levels during their education, we use information about the highest completed course level if one or more course level is completed. If the refugees never complete a course, we use the information about the course level with the last attended module as the best assessment by the language school regarding the refugees’ Danish language education level.

In our sample, 95% of individuals had been in contact with the approved Danish language schools within four years of asylum (see Table 1). According to our definition, 91% of the sample had enrolled in a language school, i.e., enrolled on 30 September in a year within the first four years of asylum. The remaining 9%, corresponding to 1,835 individuals, who had never been in contact with the Danish education system are recorded as never enrolled because they had had only minimal contact with the language school and were not enrolled on 30 September in any of the first four years, which is the time dimension of our recordings.

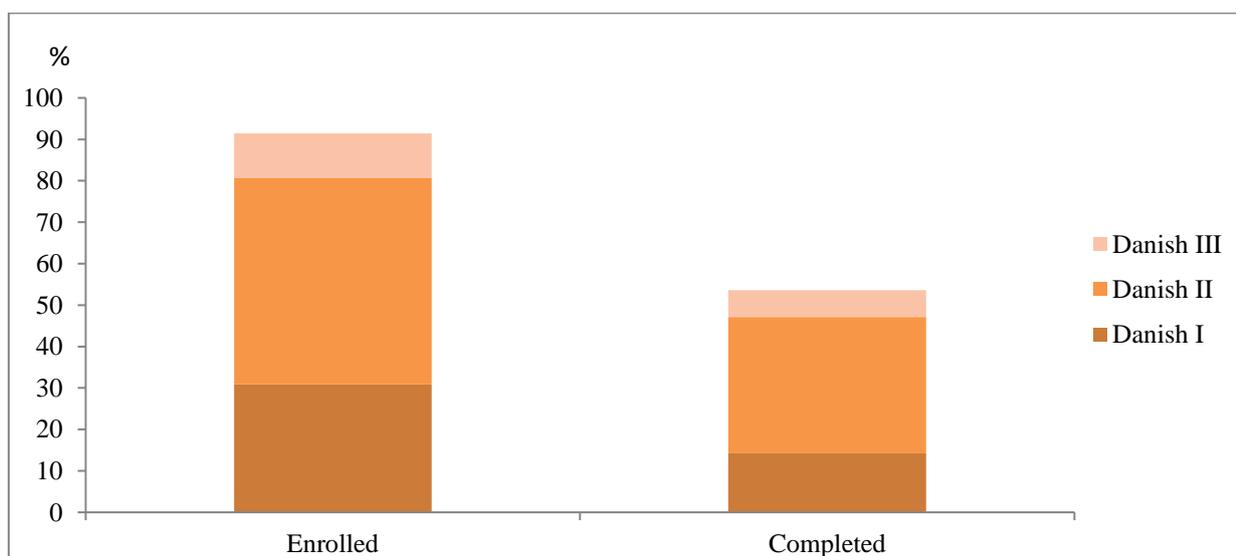
## 6.1 How Many Refugees Enrolled in a Danish Language Course and Complete it?

Figure 11 shows the shares of refugees' enrolment in and completion of a Danish language course, respectively. The course level the refugee enrolled in corresponds to the level on which the individual was initially placed. Because individuals can move between course levels during their education, the completion of a given course level, as illustrated in Figure 11, does not uniquely correspond to the individual's enrolment on a given course. Figure 11 shows that 91% of refugees in the sample were enrolled in a Danish language course on 30 September in a year during the first four years ( $t_0-t_4$ ) after obtaining a residence permit, while 54% had completed a course during that period.

Figure 11 illustrates how the highest share of both enrolment and completion is in Danish II, which is the middle category (50% of the sample were enrolled initially and 33% completed Danish II). Of the refugees in our sample, 31% had been enrolled in Danish I, and 11% in Danish III. After four years, 14% had completed Danish I, 33% had completed Danish II, and 6% had completed Danish III.

Among the individuals enrolled, 34% had enrolled in Danish I, but only 27% had completed it, conditional on having completed a course. By contrast, those who completed a course are overrepresented among individuals enrolled on Danish II: 54% enrolled versus 61% completed. These course differences suggest that refugees enrolled on Danish I (which is intended for participants with the lowest level of education from abroad) may be less qualified or less motivated compared with refugees enrolled on Danish II. Of the refugees enrolled, 12% were enrolled in Danish III, while 12% of refugees who completed a course completed Danish III. A large discrepancy between enrolment and completion is apparent across all course levels.

**Figure 11.** Share enrolled in a Danish language course and completion rate within four years of asylum.



*Source:* Authors' calculations based on administrative register data from Statistics Denmark, 2004–2020.

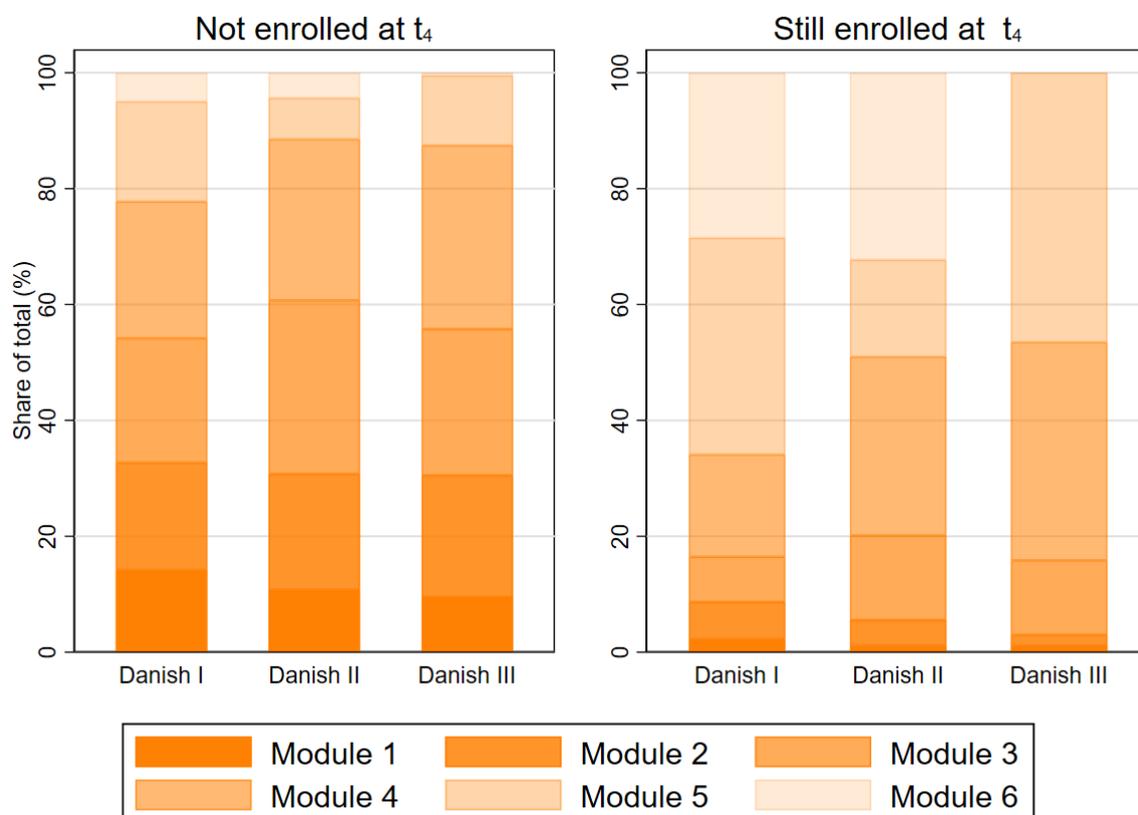
*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390).

## 6.2 What is the Level of Danish Language Proficiency of Refugees Who Have Not Completed a Language Course?

Refugees may not have completed a language course within four years of asylum for several reasons. Some are still enrolled after four years (9%), but the majority of refugees have either dropped out or have completed a course (91%).

Figure 12 presents the latest module level four years after arrival among those refugees not completing a course. Refugees who are no longer enrolled (38% of the sample) are distributed relatively evenly across modules, apart from the last module; seemingly, refugees drop out of the language course at all module levels except the last module (Module 6 for Danish I and Danish II, Module 5 for Danish III, see Section 2.3), i.e., close to completion.<sup>30</sup> A high share of refugees who are still enrolled are currently participating in one of the final modules and are close to course completion.

**Figure 12.** Most recent module among refugees not completing a Danish language course within four years of asylum.



*Source:* Authors' calculations based on administrative register data from Statistics Denmark, 2004–2020.

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). Subsample in left graph: 7,710; Subsample in right graph: 1,752.  $t_4$  denotes four years after asylum.

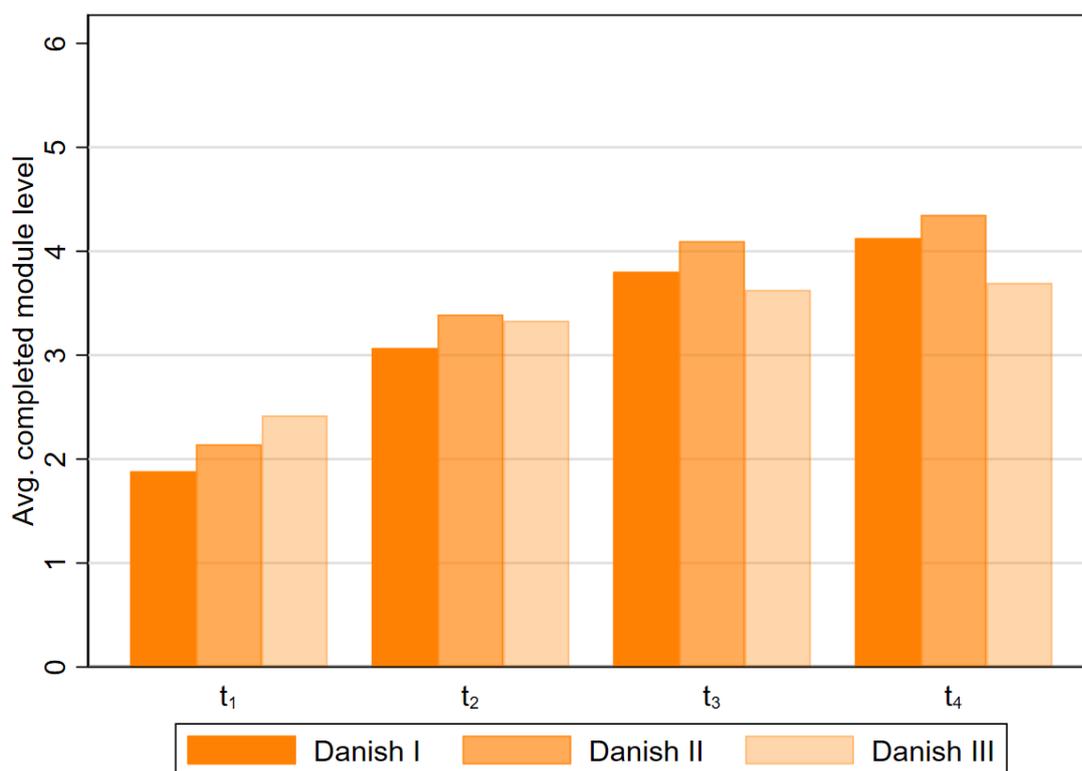
<sup>30</sup> We cannot rule out that some refugees who are not enrolled on a course on 30 September four years after asylum, re-enroll the following year. However, this seems unlikely given that the legislation emphasizes that the formal language courses should be completed within 3 years (see Section 2.3).

The majority of refugees either complete or terminate their formal Danish language education within four years of asylum. A small fraction of refugees require more than four years to complete their Danish language education, which might be due to maternity leave or periods with full-time work. It is not possible to make a general conclusion about the acquired Danish level among dropouts, as dropout happens across all modules except the last one.

### 6.3 How Fast is Danish Learned (Language Progression)?

Another important factor in acquiring a new language is the intensity of exposure, and hence, the speed of acquisition. Figure 13 presents the speed of progression of refugees in the formal Danish language education system. The figure shows the average level of the completed module on 30 September conditional on being enrolled at any point within four years of asylum (i.e., among the 91% of the sample). Figure 13 is divided according to course level. Since Danish III has one less module than Danish I and Danish II, caution is recommended when comparing progression patterns between course levels.

**Figure 13.** Progression in average module completed among refugees enrolled in a Danish language course within four years of asylum according to course level.



*Source:* Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020.

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390).  $t_1$  denotes one year after asylum,  $t_2$  two years after asylum and so forth.

Focusing on completion of modules in the Danish I and II courses one year after asylum ( $t_1$ ), refugees

have on average completed Module 2.<sup>31</sup> Two years after asylum, refugees have on average completed Module 3. Three years after, refugees have on average completed Module 4. Between  $t_1$  and  $t_2$  the change in average completed module level increases by 1.1 for Danish I and Danish II. Between  $t_2$  and  $t_3$  the change is 0.7, between  $t_3$  and  $t_4$  the change is 0.3. For completion of modules in the Danish III course, the pattern of module progression is similar. An upward trend in average module completion level across time periods is apparent, although progression is relatively faster in the earlier periods.

Progression between the final year ( $t_4$ ) and the year before ( $t_3$ ) is virtually stagnant for participants in Danish III, while some progression remains for participants in Danish I and II. Participants complete Danish III within three years, while some participants in Danish I and II may use extra time, but as mentioned earlier, these courses have an extra module.

## 7 Interplay Between Language Acquisition and other Factors

In Sections 3–6 we analysed refugees' demographics, settlement pattern, education from abroad on arrival, and formal Danish language acquisition, respectively. In this section we turn to the interplay between these factors and formal Danish language acquisition.

### 7.1 Educational Differences in Enrolment and Completion Rates

Refugees' education upon settlement is expected to be important for the speed at which they learn the Danish language (Chiswick and Miller, 2007). The three language courses are intended for individuals with different educational backgrounds (Ankestyrelsen, 2012, 2013). Our results show that the level of education on asylum may have been decisive for which language course refugees have been enrolled on, but there is no one-to-one relation between educational length and language course level.

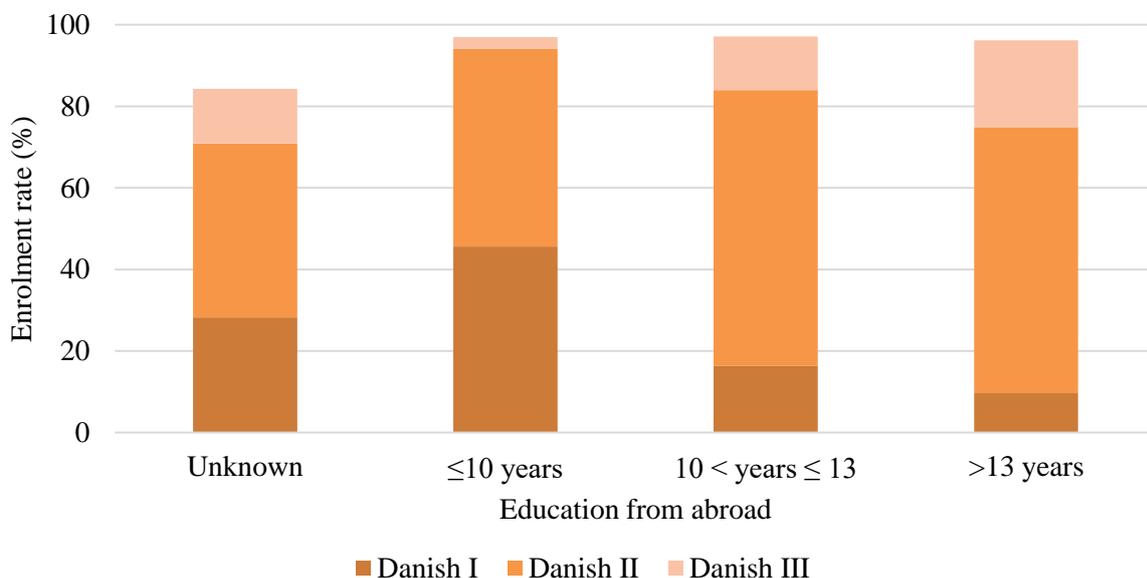
Figure 14 presents the correlation between years of education from abroad and the enrolment rate in formal Danish language courses within four years of asylum. The enrolment rate is high both for refugees with a short education (97%) and for those with a long education (96–97%).

Refugees without information on education from abroad have a significantly lower enrolment rate (84%) than those with a low level of education from abroad. A reason for this could be that refugees who do not have the possibility (or interest) to participate in a language course are also less likely to participate in an education survey. Furthermore, as shown in Figure 7, the coverage of the educational information is larger in the later years (close to the 2016 survey) and correlates positively with a rising share of refugees having enrolled in later years, see Figure 16. As reported in Figure 14, the share of refugees enrolled in Danish I is the highest among those with at most ten years' education (46%), followed by those with 10 to 13 years' education (16%), and those with more than 13 years' education (10%). At the opposite end of the scale, the share of refugees enrolled in Danish III is the lowest

---

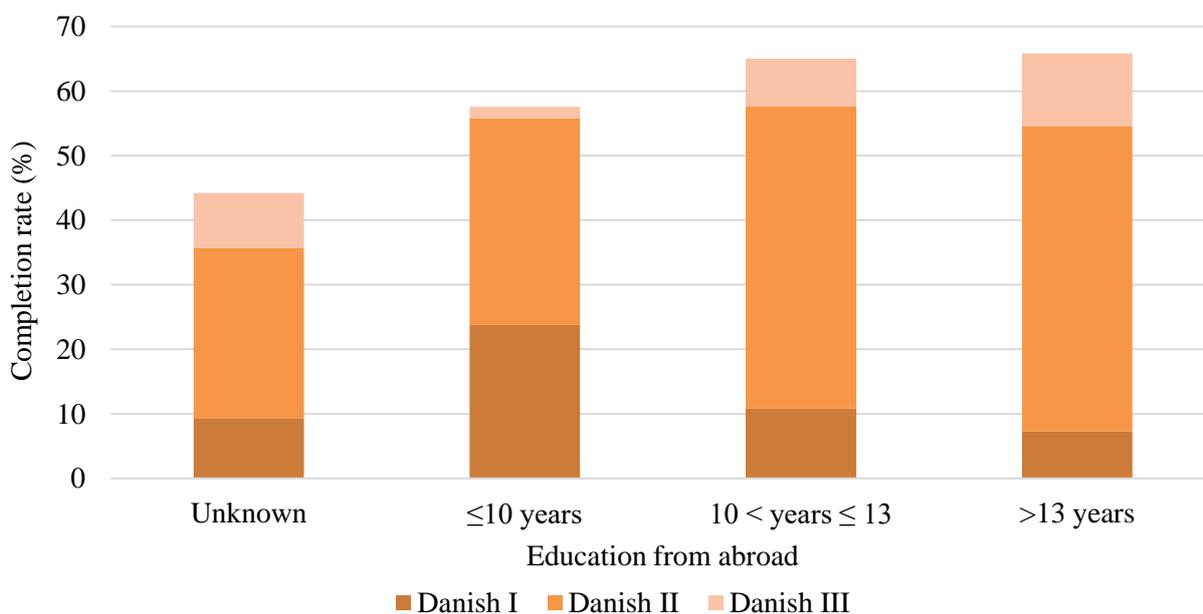
<sup>31</sup> Since individuals can begin their formal Danish language education on the module level deemed appropriate by the language school (as a result of the screening test, see Section 2.3), it is not possible to conclude that refugees have actually completed two entire module levels by  $t_1$ .

**Figure 14.** Share of refugees enrolled in Danish language courses within four years of asylum according to years of education.



Source: Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020.  
 Notes: Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390).

**Figure 15.** Share of refugees completing a Danish language course within four years of asylum according to years of education.



Source: Authors' calculations based on administrative register data from Statistics Denmark, 2004–2020.  
 Notes: Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390).

among those with at most ten years' education (3%), followed by those with 10 to 13 years' education

(13%), those with unknown education (13%), and those with more than 13 years' education (21%). It is an interesting finding that the share of individuals enrolled on course level III is significantly higher among individuals with an unknown level of education compared with the group with at most 10 years' education. It suggests that the the former group includes a non-ignorable share of refugees with a long education.

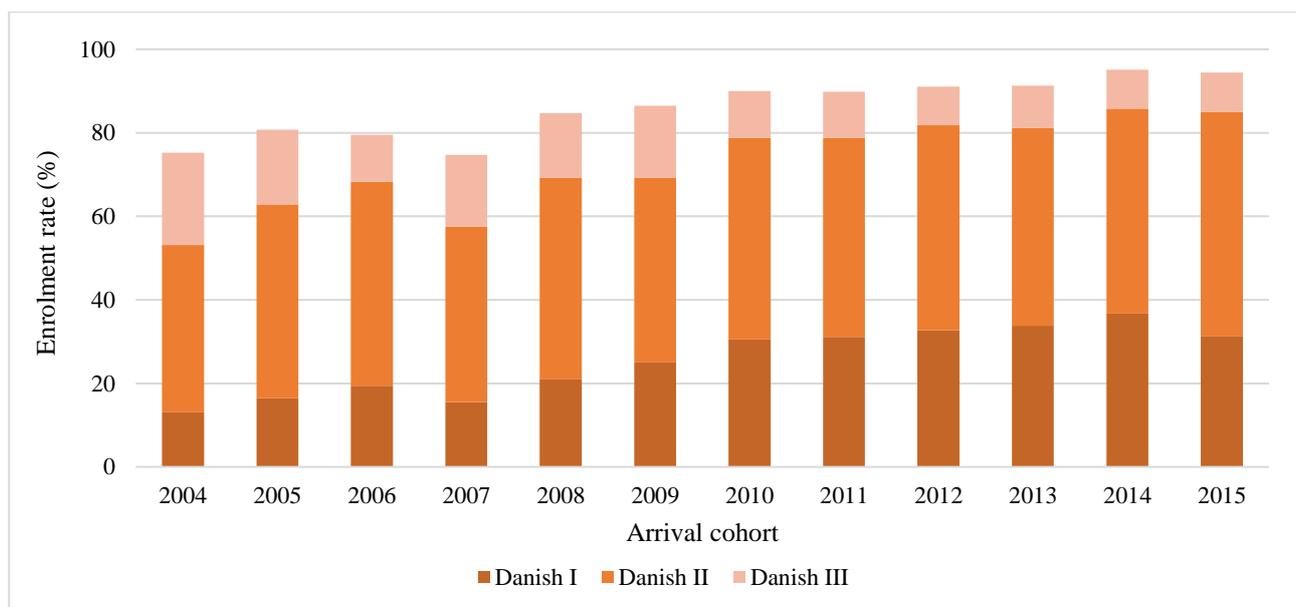
For the share of refugees completing a language course within four years of asylum (Figure 15), we find that even though refugees with a short education are more likely to enrol in a language course, they are significantly less likely to complete it. The language course completion rate is significantly lower among refugees without educational information (44%) and refugees with at most 10 years' education (58%) compared with refugees with 10–13 years' education (65%) and refugees with more than 13 years' education (66%).

## 7.2 Enrolment and Completion Rates by Year of Registration and Country of Origin

Figures 16 and 17 illustrate the enrolment and completion rates of formal Danish languages courses within four years of asylum by arrival cohort, respectively.

According to Figure 16, the overall share of refugees enrolled in formal Danish education within four years of asylum increases according to arrival cohort, from 75% for arrival cohort 2004 to 91% for arrival cohort 2015. The composition also changes somewhat over the years as a higher share participate in Danish I and a lower share in Danish III (this also holds when restricting the sample to enrolled refugees). This finding is consistent with our expectation, given the importance of education from abroad in determining course placement and since refugees with at most 10 years' education from abroad constitute an increasing share over the observed period.

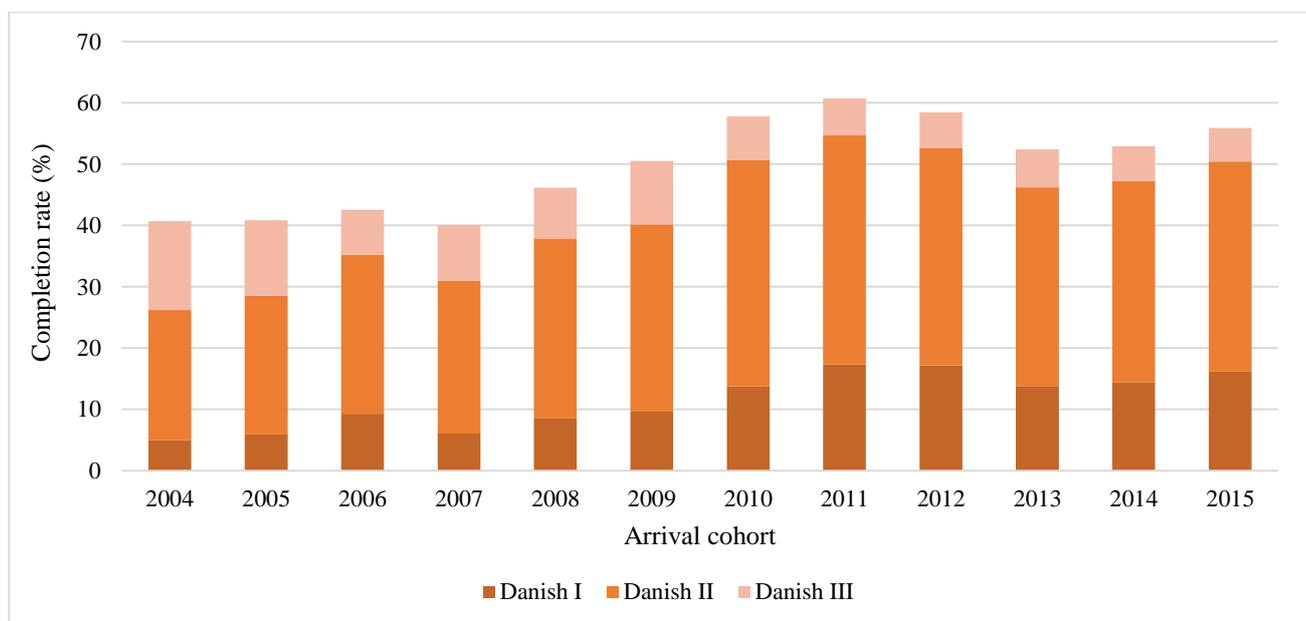
**Figure 16.** Share of refugees enrolled in Danish language courses within four years of asylum according to arrival cohort.



Source: Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020.  
Notes: Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390).

Completion rates are higher for cohorts that arrived later, as illustrated in Figure 17. The increase is pronounced for cohorts obtaining asylum during 2008–2011. This is consistent with the general rise in enrolment. One possible explanation for this surge in completion rates is that individuals opt for education (such as formal Danish language education) when outside options are limited, as during the financial crisis. The increase in completion rates from 2008 to 2011 coincides with a general sharp drop in employment participation rates in Denmark for non-Western immigrants in the period 2009 to 2015 due to the financial crisis.<sup>32</sup> Completion rates peak for the cohort obtaining asylum in 2011; the cohort are expected to finish their Danish language education in 2014 (the formal Danish language courses are designed to take 3 years, see Section 2.3). After 2015, employment rates rise for non-Western immigrants. The rise might explain the downturn in completion rates for refugees obtaining asylum after 2012. While enrolment and completion rates seem to be correlated with employment prospects, the relationships are not perfect.

**Figure 17.** Share of refugees who completed a Danish language course within four years of asylum according to arrival cohort.



Source: Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020.

Notes: Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390).

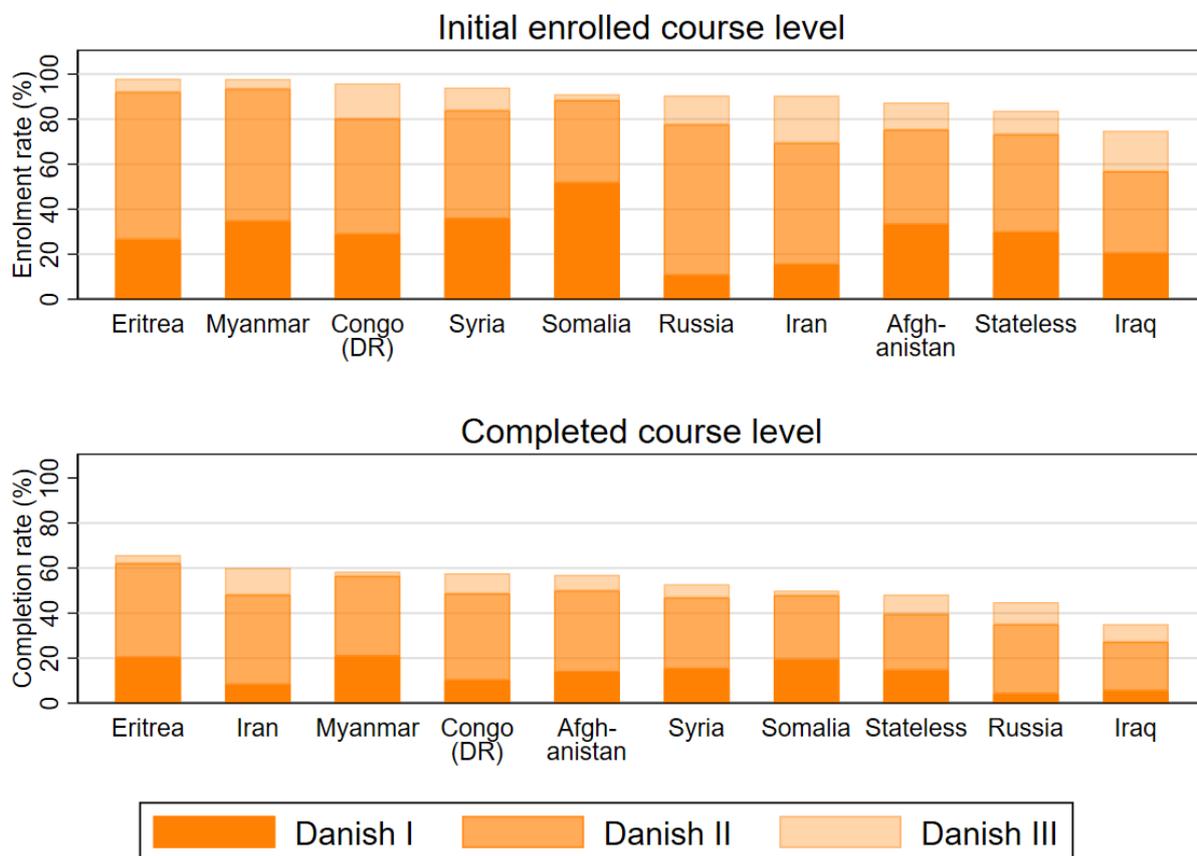
Figure 18 reports the share of refugees who, respectively, enrol in and complete a language course within four years of asylum separately for each of the ten largest refugee countries. The share of enrolment is the highest for the groups most recently arrived from Eritrea and Syria and for refugees from Myanmar and DR Congo. The lowest share is found for the Iraqis.

The course level on which refugees are enrolled differs between the countries of origin. The difference may be because some countries have alphabets more closely related to that in Denmark (Latin), but it may also reflect differences in the educational country profiles, as described in Section 5.1.

<sup>32</sup> Statistics Denmark, table 'RAS200' and 'RAS1F1'.

The highest completion rates are found for Eritreans followed by refugees from Iran, Myanmar, and DR Congo.

**Figure 18.** Share of refugees enrolled in and completing a Danish language course within four years of asylum according to country of origin.



*Source:* Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020.  
*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390).

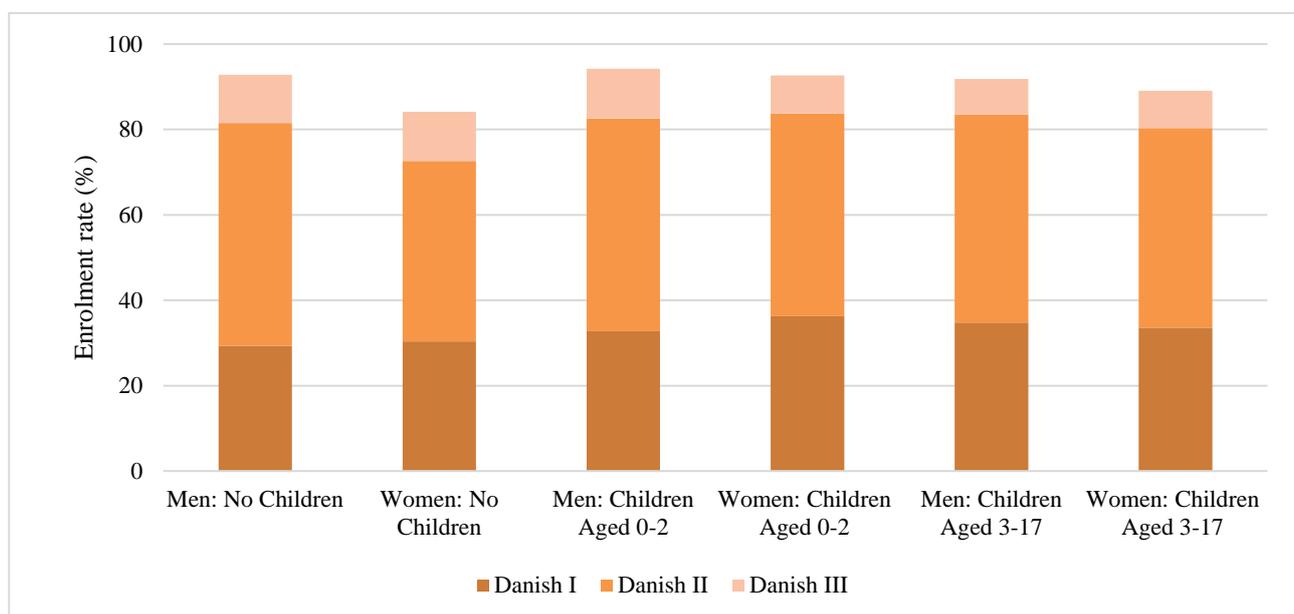
### 7.3 Are Men More Likely than Women to be Enrolled in and Complete Language Courses, and Do Children Have an Influence?

Earlier studies have shown that among immigrants, men tend to have higher host-country language proficiency than do women (Dustmann, 1994; Chiswick and Miller, 1996; Liebig and Tronstad, 2018). This is supported by the findings of previous Danish studies (Schultz-Nielsen, 2010; Arendt et al., 2016; Jacobsen et al., 2016). As Danish is spoken only in Denmark, no refugee speaks Danish upon asylum. Therefore, it is relevant to analyse gender differences in refugees' rates of enrolment and completion of the language courses. Male refugees may have a higher return from host-country language proficiency due to more work experience from the source-country and shorter parental leave periods (Djuve, Kavli, Sterri and Bråten, 2017). Empirically, there is a substantial gender gap in the employment rate of non-Western immigrants in Denmark (Schultz-Nielsen, 2019). A counter argu-

ment is that participation in the introduction programme (including language courses) is a requirement for obtaining social assistance, regardless of gender. Furthermore, the universal nursery system in Denmark provides public childcare during daytime and the service is free for families with low incomes. Consequently, both men and women are expected to be enrolled in language courses.

The share of refugees enrolled in the Danish language courses at any point within the first four years of asylum is shown in Figure 19. The figure shows the enrolment rate is generally high, above 80%, irrespective of gender and regardless of whether the individuals had children and the age of the youngest child (at the time of asylum): 0–2 years or 3–17 years.

**Figure 19.** Share of refugees enrolled in a Danish language course within four years of asylum according to gender, and to age of the youngest child on arrival.



Source: Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020.

Notes: Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390).

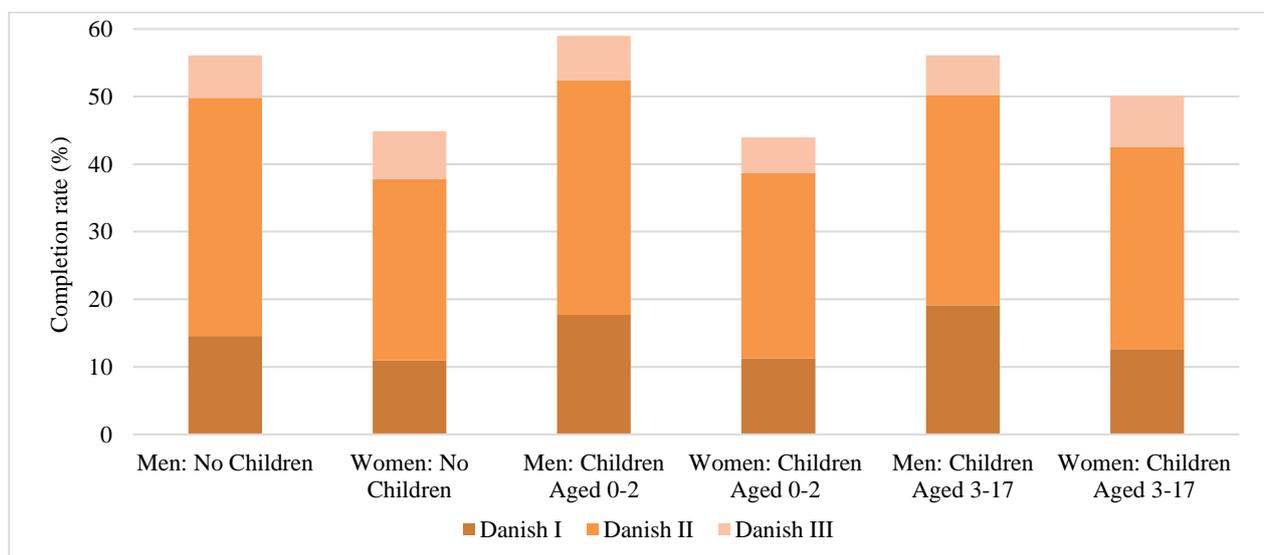
The lowest enrolment rate is found for women without children (84%) followed by women (89%) and men (92%) with children aged 3–17 years. Of the men without children, 93% are enrolled, followed by women with children aged 0–2 (93%). The highest share of enrolment (94%) is found for men with children aged 0–2 years. Mean differences within gender groups are small, albeit statistically significant. Having young children at the time of settlement does not seem to hinder language course enrolment. While the gender difference is modest, women in general have slightly lower enrolment rates. Across gender and child categories, close to 30% are enrolled in Danish I.

A similar calculation based on the completion rate of a Danish language course is shown in Figure 20. As shown in Figure 11, completion rates are much lower than enrolment rates. Completion rates after four years of residence ( $t_4$ ) are the highest for men with children aged 0–2 years at the time of residence (59%), followed by men with no children (56%) and men with children aged 3–17 years (56%). For women, the highest completion rates are found among those with children aged 3–17

years (50%), followed by women without children at the time of residence (45%), and women with children aged 0–2 years (44%).

By examining the share of refugees who completed their Danish courses within four years (conditional on enrolment), we find a similar ‘ranking’. Conditional on enrolment, completion rates are 63% for men with children aged 0–2 years, 61% for men with children aged 3–17 years, 61% for men with no children, 56% for women with children aged 3–17 years, 53% for women without children and 47% for women with children aged 0–2 years. Refugee women have significantly lower completion rates compared with refugee men, regardless of family status. As refugees are generally young upon asylum, many female refugees who did not have children at the time of residence may have later become mothers, while those with older children are less likely to give birth. A potential explanation for the lower completion rate for refugee women includes maternity leave periods, which prolong the language course period.

**Figure 20.** Share of refugees completing a Danish Language Course within four years of asylum according to gender and to age of the youngest child on arrival.



Source: Authors’ calculations based on administrative register data from Statistics Denmark, 2004-2020.

Notes: Sample: Adult settled refugees in Denmark 2004-2015 (N=20,390).

#### 7.4 Are there Geographical Differences in Enrolment and Completion Rates?

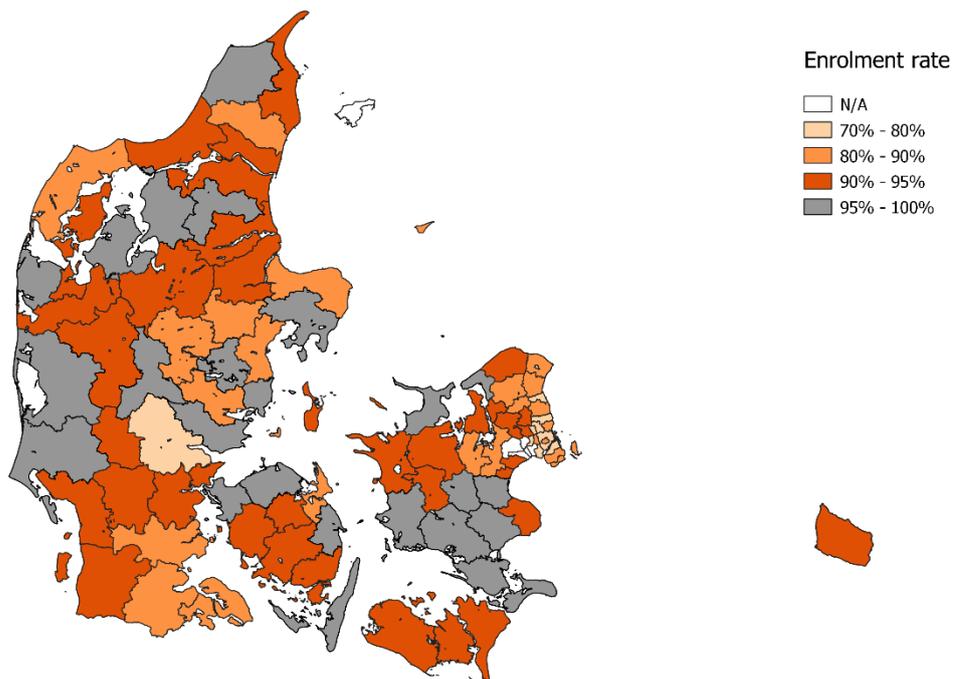
To examine geographical differences in the share of refugees participating in Danish language courses across municipalities, the enrolment and participation rates among refugees are calculated for all of the municipalities that host a sizeable number of refugees.<sup>33</sup> It is important to note refugees may differ in terms of skills across settlement areas. Accordingly, observed differences cannot be used to benchmark municipalities with respect to the effectiveness of their introduction programme.

According to Figure 21, the share of refugees enrolled in a Danish language course at any time during the first four years of asylum is generally high but varies between municipalities. Among refugees

<sup>33</sup> Due to Statistics Denmark’s data privacy rules, six municipalities with small numbers of settled refugees are excluded.

obtaining asylum during 2004–2015, the enrolment rate in a Danish language course is the highest (above 90%) in municipalities in Western Jutland, and also in Skanderborg, Syddjurs, Nyborg, Langeland, Holbæk, Halsnæs, Køge, Stevns and Faxe. The lowest rates of 70–80% are found in municipalities in the Capital Region (Copenhagen, Hvidovre and Lyngby-Taarbæk), while, as mentioned earlier, other municipalities received no or only very few refugees.

**Figure 21.** Share of refugees enrolled in a Danish Language Course within four years of asylum according to municipality.

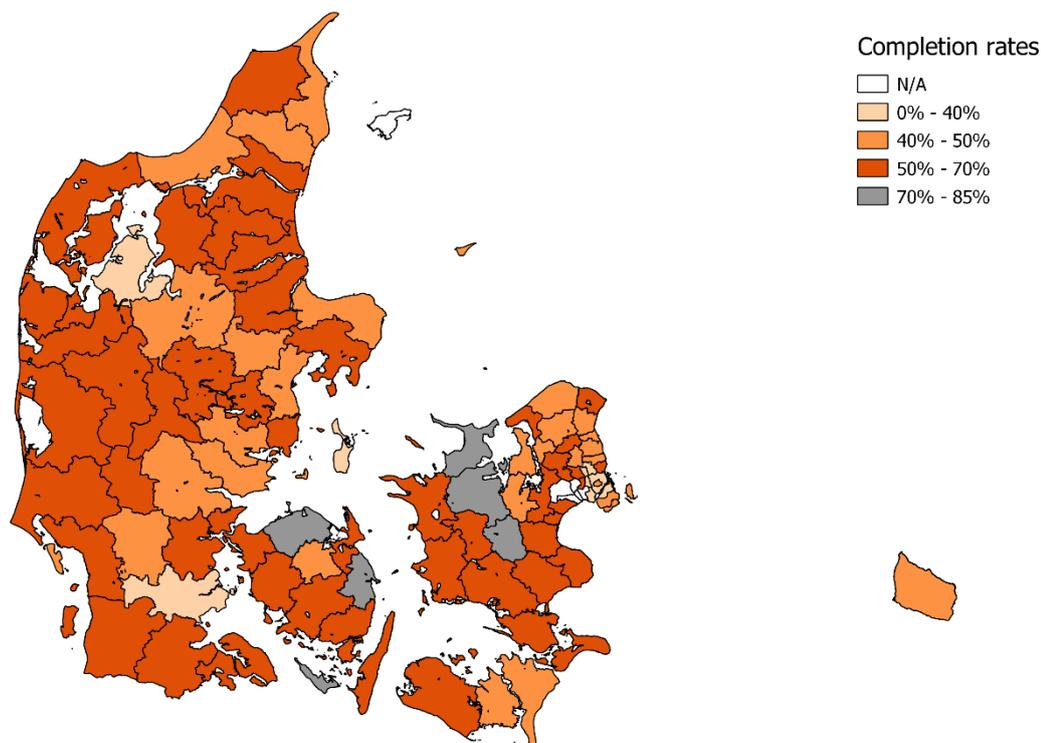


*Source:* Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020.

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). Some municipalities are categorised as 'N/A' for compliance with data privacy rules due to few observations (Vallensbæk, Læsø, Brøndby, Albertslund, Ishøj and Høje-Taastrup).

Figure 22 illustrates municipal variation in the share of refugees completing a Danish language course. The lower completion rates (below 40% and 40–50%) are found in the Capital Region and in several municipalities in Jutland, including Samsø, Skive, and Haderslev. The highest shares (above 70%) are found in more rural parts, such as Ærø.

**Figure 22.** Share of refugees completing a Danish language course within four years of asylum according to municipality.



*Source:* Authors' calculations based on administrative register data from Statistics Denmark, 2004–2020.

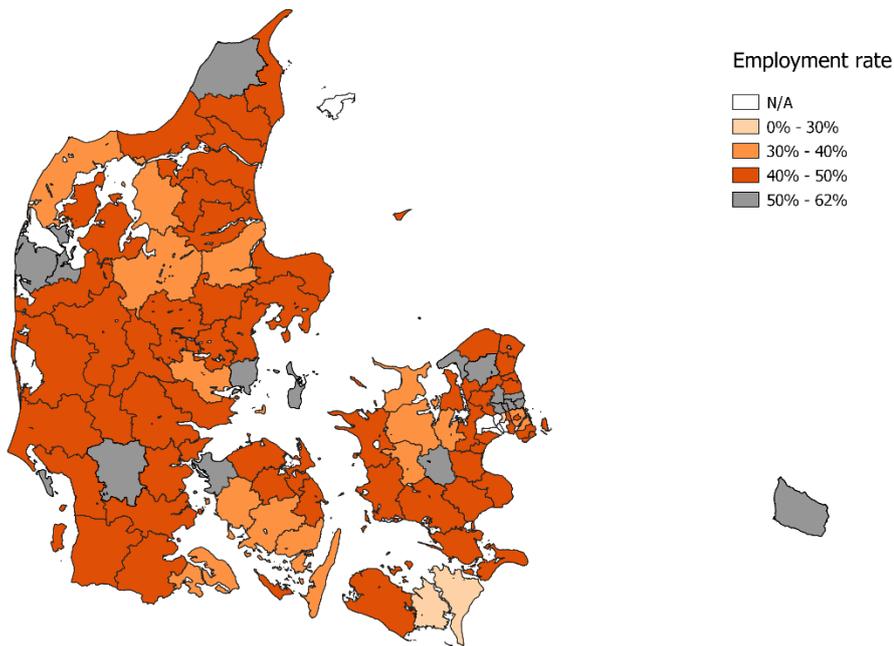
*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). Some municipalities are categorised as 'N/A' for compliance with data privacy rules due to few observations (Vallensbæk, Læsø, Brøndby, Albertslund, Ishøj and Høje-Taastrup).

### 7.5 What Explains the Municipal Variation in Language Course Completion Rates?

One hypothesis is that refugees settled in municipalities with fewer employment prospects are more likely to complete the language course in which case employment and language course completion are substitutes. Another hypothesis is that a completed language course gives access to employment (in the longer run) in which case employment and course completion are complements. Figure 23 shows the average employment rates of initially settled refugees in the municipality across municipalities four years after settlement. The highest employment rates of refugees are found in parts of the Capital Region, and also in Ringsted (with a high completion rate), Samsø (with a low completion rate), Bornholm and Lemvig (with an average completion rate).

The municipal correlation between the share of refugees completing a language course and the share of refugees being employed at time  $t_4$  is -0.1781. The correlation points in the direction of language courses and employment being substitutes in refugees' first years of residence. We cannot claim a causal relationship based on these figures.

**Figure 23.** Share of refugees in employment within four years of asylum according to municipality.



*Source:* Authors' calculations based on administrative register data from Statistics Denmark, 2004–2020.

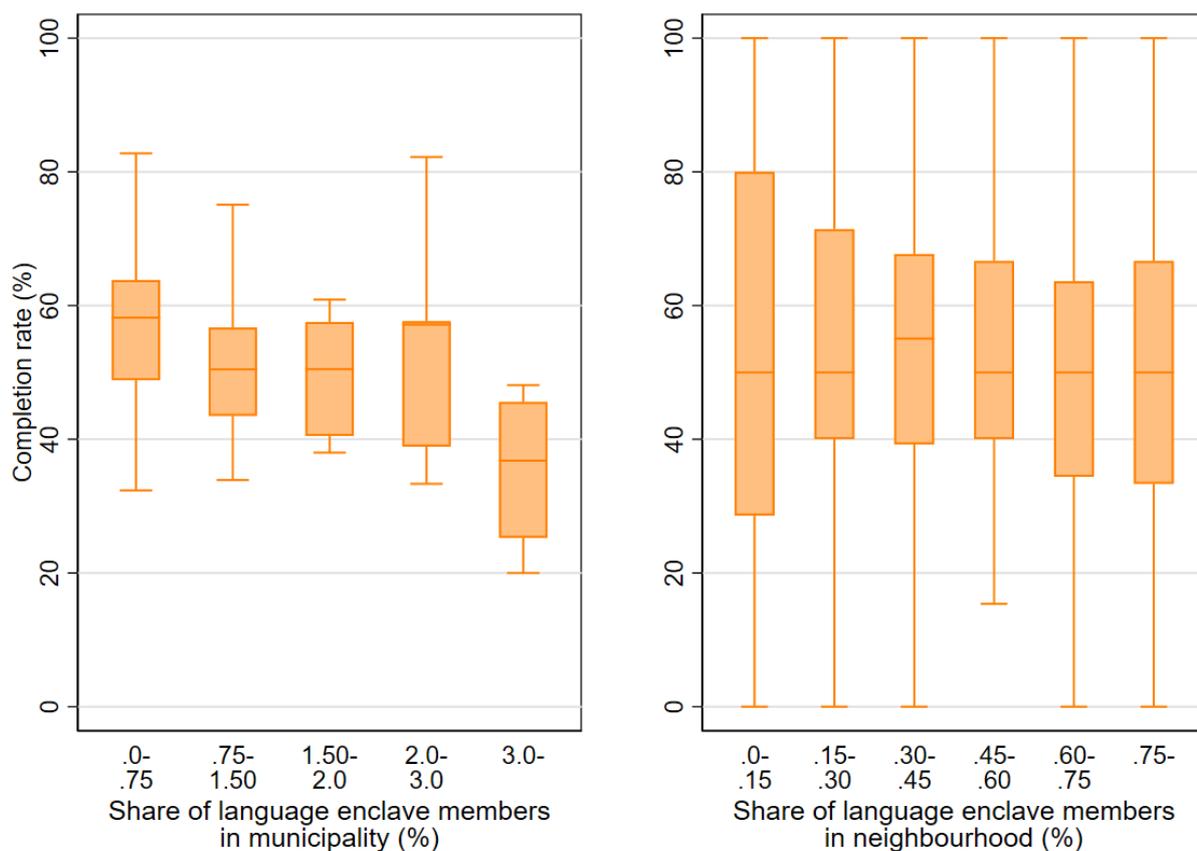
*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). For refugees arrived in 2015, employment status four years after asylum is imputed by their employment status three years after asylum. Some municipalities are categorised as 'N/A' for compliance with data privacy rules due to few observations (Vallensbæk, Læsø, Brøndby, Albertslund, Ishøj and Høje-Taastrup).

Another dimension of geographic settlement is the share of co-nationals living in the municipality at the time of settlement. Refugees are often settled in a municipality with others of the same nationality to allow them to create a network of co-nationals. A higher concentration of minority group members may decrease the need to learn the host-country language (Lazear, 1999). To analyse the causal relationship between the share of language enclave members and investments in Danish language acquisition, we refer to Damm et al. (2021). Our study merely describes the correlation between the two measures.

Figure 24 shows the correlation between having completed a language course and the share of language enclave members in the municipality and neighbourhood of settlement at the local area.<sup>34</sup> The boxplot illustrates the distribution of the completion rate for the share of language enclave members in a given interval. Based on this figure, there is a negative correlation between the share of language enclave members and the share completing a language course. At the neighbourhood level, the correlation is not clear and, within each interval of the share of language enclave member, there is a large variation in the share having completed a Danish language course; all intervals include neighbourhoods with completion rates of 0% or 100%. We can conclude that the share of language enclave members in a municipality is negatively correlated with completion rates, while the relationship does not seem to hold at the neighbourhood level.

<sup>34</sup> For similar results when the minority group is measured by share of co-nationals, see Appendix Figure A4.

**Figure 24.** Share of refugees completing a language course within four years of asylum according to share of language enclave members in the municipality and neighbourhood.



*Source:* Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020, linked with register on neighbourhood of residence constructed by Damm et al. (2019a).

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). The bottom and top of the box form first and third quartile, while the line in the box shows the median and the whiskers illustrate the lower and upper adjacent values.

## 8 Conclusion

The paper examines the settlement patterns, educational background, and Danish language acquisition among all 20,390 adult refugees obtaining residence in Denmark from 2004 to 2015. During the period, adult refugees were offered a three-year introduction programme, which included a substantial amount of language training, corresponding to 1.2 years' full-time participation. Using administrative registers from Statistics Denmark, we followed the refugees' integration process during the first four years after receipt of their residence permit. We describe the characteristics of the municipalities and neighbourhoods in which the refugees are settled, the educational skills the refugees bring from abroad, and the formal Danish language acquisition the refugees obtain.

The analysis is based on information from administrative registers including ‘Danskundervisningsdatabasen’ (DUB). The DUB includes detailed information about enrolment in and completion of all formal language courses during the first four years of asylum in Denmark.

The analysis shows that on settlement refugees are distributed to municipalities with a below average share of non-Western immigrants due to a ‘refugee quota system’. Within the municipality, refugees are often settled in neighbourhoods with higher shares of public housing and in less advantageous neighbourhoods in terms of socioeconomic composition. The most likely reason is that such neighbourhoods have vacant affordable housing. Refugees also tend to settle in neighbourhoods with a high concentration of their language enclave members.

The refugees’ educational attainment from abroad is modest; among refugees with self-reported education attained prior to arrival, 62% have 10 years’ education at most. The share varies considerably according to country of origin. Educational attainment from abroad influences the level of Danish course enrolment. Among refugees with 10 years’ education at most, 46% are enrolled on Danish I (the less demanding course), whereas the share enrolled is much lower among those with more education.

Most refugees enrolled in a Danish language course within the first four years of asylum (91%), and increasingly so during our observation period. The most ‘popular’ course is Danish II, which is attended by half of all refugees, followed by Danish I and the more advanced course Danish III.

The share of refugees completing a language course within the first four years of asylum (54%) falls considerably below the enrolment rate. We find that completion rates in general are higher among men compared with women, among those with a longer compared with a shorter education, and among Eritreans and Iranians compared with Russians and Iraqis. Completion rates are higher for refugees settled in some more rural areas compared with refugees settled in the Capital Region.

Our study raises a number of questions to be addressed in future research. Most importantly, which factors explain the high dropout rates from formal Danish language courses? In this relation, researchers could investigate demand-side factors and supply-side factors. Demand-side factors include the individual’s labour market status, e.g., parental leave periods, and environment in the neighbourhood of residence, e.g., whether refugees settled in a neighbourhood with a relatively large share of language enclave members are less likely to complete formal host-country language courses. Supply-side factors include the quality of the language courses provided as such, the match quality between participants and course levels, and the logistics of the course provision, e.g., whether the hours of teaching are compatible with the working hours of potential course participants.

In view of the finding that the immigrant-native employment gap remains large even 7–10 years after immigration, in particular for non-labour migrants (Bratsberg et al. 2017; Schultz-Nielsen 2017), another important area for future research is to investigate whether formal host-country language training is required for strong labour market attachment of immigrants in the long-run. It may be the case that skills obtained abroad are not fully transferable to the host country’s labour market and that host-

country language proficiency is a pre-requisite for enrolment in and completion of formal education in the host country.

## References

- Andersen, S. C., M. K. Humlum and T. S. Guul (2017), "Modersmålsbaseret Undervisning - Modersmålsundervisning på 1. klassetrin", TrykFondens Børneforskningscenter.
- Andersen, H. S. (2015), "Indvandring, integration og etnisk segregation: udviklingen i indvandrernes bosætning siden 1985", SBI forlag, vol. 2015:01.
- Andersen, H. S. (2017), "Selective moving behavior in ethnic neighbourhoods: white flight, white avoidance, ethnic attraction or ethnic retention?", *Housing Studies*, 32 (3), 296-318.
- Andersson, P. J. (2019) "Labour Market Policies: What Works for Newly Arrived Immigrants?" Chapter 4 in L. Calmfors and N. S. Gassen (Eds.), *Integrating Immigrants into the Nordic Labour Markets*, Nord 2019:024. Nordic Council of Ministers, 2019. <http://norden.diva-portal.org/smash/get/diva2:1317928/FULLTEXT01.pdf>
- Ankestyrelsen (2012), "Aktiviteten hos udbydere af danskuddannelse for voksne udlændinge m.fl. i 2012" [English: Activity of providers of formal Danish language education for adult aliens in 2012], Ankestyrelsens statistikker.
- Ankestyrelsen (2013), "Aktiviteten hos udbydere af danskuddannelse for voksne udlændinge m.fl. i 2013" [English: Activity of providers of formal Danish language education for adult aliens in 2013], Ankestyrelsens statistikker.
- Arendt, J. N., C. P. Nielsen and V. Jacobsen (2016), "Importance of origin and destination country skills for labour market attachment of immigrants from Pakistan, Iran and Turkey", *Nordic Journal of Migration Research*, 6(2), 72-79.
- Arendt, J. N. (2019), "Employment Effects of a Job-First Policy for Refugees", Rockwool Foundation Research Unit WP no. 139.
- Arendt, J. N. and I. Bolvig (2020), "Early labor market entry, language acquisition and labor market success of refugees", VIVE WP.
- Arendt, J. N., I. Bolvig, M. Foged, L. Hasager and G. Peri. (2021), "Language Training and Refugees' integration", NBER WP 26834.
- Azlor, L., A. P. Damm and M. L. Schultz-Nielsen (2020), "Local Labour Demand and Immigrant Employment", *Labour Economics*, 63.
- Beaman, L. A., (2012), "Social networks and the dynamics of labour market outcomes: evidence from refugees resettled in the US", *Review of Economic Studies*, 79, 128–161.
- Becker, G. S. (1957), *The Economics of Discrimination*, Chicago: Univ. Chicago Press.
- Becker, G. S. (1964), *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. NBER.
- Becker, G. S. and B. R. Chiswick (1966), "Education and the Distribution of Earnings", *American Economic Review*, 56, 358-69.
- Bratsberg, B., O. Raaum and K. Røed (2017), "Immigrant labor market integration across admission classes", *Nordic Economic Policy Review*, 17-54.
- Chiswick, B. R. and P. W. Miller (1996), "Ethnic networks and language proficiency among immigrants", *Journal of Population Economics*, 9, 19-35.

- Chiswick, B. R. and P. W. Miller (2007), *The Economics of Language*, Routledge Studies in the Modern World Economy.
- Chiswick, B. R. and P. W. Miller (2014). Chapter 5 ‘International migration and the economics of language’, *Handbook of the economics of International Migration*, Vol. 1a., 211-269. North Holland.
- Christensen, L. B., N.-E. Hansen, G. Homann, E. B. Jørgensen, K. U. Kjær, M. Kjærum, I. E. Koch, N. Lassen, J. Vedsted-Hansen and L. Wendland (2000), *Udlændingeret* [English: Immigration law], Jurist- og Økonomforbundets Forlag.
- Clark, A. E., N. Westergård-Nielsen and N. Kristensen (2009), “Economic Satisfaction and Income Rank in Small Neighbourhoods”. *Journal of the European Economic Association*, 7(2/3), 519–527.
- Damm, Anna P. “The Danish Dispersal Policy on Refugee Immigrants 1986-1998: A Natural Experiment?” 2005. Aarhus School of Business, Department of Economics WP 05-3.
- Damm, A. P. (2009a), “Determinants of Recent Immigrants’ Location Choices: Quasi-Experimental Evidence”, *Journal of Population Economics* 2009, 22(1): 145-174.
- Damm, A. P. (2009b), “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, A. P., and C. Dustmann (2014), “Does Growing Up in a High Crime Neighborhood Affect Youth Criminal Behavior?” *The American Economic Review*, 104(6), 1806–1832.
- Damm, A. P., A. Hassani and M. L. Schultz-Nielsen (2019a), “Construction of Danish Residential Neighbourhoods for 1986-2016” in Hassani, A. *Residential Segregation and Public housing*, PhD dissertation, Aarhus University, Department of Economics and Business Economics, 2019(10), 3-48, forthcoming in *Danish Journal of Economics (Nationaløkonomisk Tidsskrift)*.
- Damm, A. P., A. Hassani and M. L. Schultz-Nielsen (2019b), “Residential segregation in Denmark during 1986-2016” in Hassani, A. *Residential Segregation and Public housing*, PhD dissertation, Aarhus University, Department of Economics and Business Economics, 2019(10), 51-129.
- Damm, A. P., A. Hassani, T. H. S. Jensen and M. L. Schultz-Nielsen (2021), “Co-Ethnic Neighbors and Investment in Host-Country Language Acquisition”, forthcoming as RFF WP.
- Damm, A. P. and M. L. Schultz-Nielsen (2008), “Danish Neighbourhoods: Construction and Relevance for Measurement of Residential Segregation”, *Danish Journal of Economics*, 146 (2008), 241-262.
- Damm, A. P., M. L. Schultz-Nielsen and T. Tranæs (2006), *En befolkning deler sig op?* [English: Further Segregation of the Population?], Gyldendal.
- Djuve, A. B., H. C. Kavli, E. B. Sterri and B. Bråten (2017), *Introduksjonsprogram og norskopplæring – Hva virker – for hvem?* [English: Introduction programme and Norwegian acquisition – What works – and for whom?], Fafo-rapport 2017:31.
- Dustmann, C. (1994), “Speaking fluency, writing fluency and earnings of migrants”, *Journal of Population Economics*, 7, 133-156.

- 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, F., T. Frattini and L. Minale (2021), "(The struggle for) Refugee Integration into the Labour Market: Evidence from Europe", forthcoming in *Journal of Economic Geography*.
- Hernes, V., Arendt, J., Andersson Joonas, P. and Tronstad, K. R. (2019). Nordic comparative analysis of integration of refugees after completed integration programme. Copenhagen: Nordic Council of Ministers.
- Husted, L. (2008), *Danskuddannelse for voksne udlændinge 2004-2006: progression og benchmarking*, AKF Forlag.
- Husted, L., I. S. Bonfils, H. B. Lauritzen and S. B. Nielsen (2010), *Indvandrerkvinder i danskuddannelse* [English: Female immigrants' Participation in Danish Language Courses], AKF Forlag.
- Husted, L. and E. Heinesen (2004), *Benchmark-analyse af kommunernes integrationsindsats i forhold til udlændinge omfattet af integrationsloven*, AKF Forlag.
- Hvidtfeldt, C. and M. L. Schultz-Nielsen (2018), "Refugees and asylum seekers in Denmark 1992–2016", *Rockwool Foundation Research Unit WP*, no. 133.
- Jacobsen, R. H., A. Kiil and C. Kolodziejczyk (2016), *Benchmarking og effektivitet af danskuddannelse for udlændinge*, Kora.
- Knudsen, L. B. (2014), *Mapping og registerdata*. Presented at the Syddansk Universitetsforlag. Retrieved from [http://vbn.aau.dk/da/publications/mapping-og-registerdata\(96ae69f4-1c89-4059-99ec-de82f11bd80a\)/export.html](http://vbn.aau.dk/da/publications/mapping-og-registerdata(96ae69f4-1c89-4059-99ec-de82f11bd80a)/export.html) (visited 28/08/2020, 12:18).
- Larsen, C. (2002), "Uddannelse og dansk kundskaber", Chapter 4 in *Indvandrerne og arbejdsmarkedet* by Mogensen, G.V. and P.C. Matthiessen, Spektrum.
- Lazear, E. P. (1999), "Culture and Language", *Journal of Political Economy*, 107(56), 95-126.
- Liebig, T. and K. R. Tronstad (2018), "Triple Disadvantage?: A first overview of the integration of refugee women", *OECD Social, Employment and Migration Working Papers*, no. 216.
- Ministeriet for by, bolig og landdistrikter (2014), *Analyse – segregering i de fire største danske Byområder* [English: Analysis – segregation in the four larger Danish metropolitan areas].
- Nielsen, C.P. and K.B. Jensen. (2006). "The Danish Integration Act's Significance for the Settlement of Patterns of Refugees", Copenhagen: Institute for Local Government Studies.
- Nielsen, R. S. (2014), "A change for the better? Life course analysis of the housing careers of Somalis and Turks in the Danish housing market", Danish Building Research Institute, University of Aalborg, *PhD thesis*.
- Schultz-Nielsen, M. L. (2010), "The employment effect of immigrants' language proficiency". Chapter 4 in "Essays in Migration", Aarhus University, *PhD thesis 2010:1*.
- Schultz-Nielsen, M.L. (2017), "Labour market integration of refugees in Denmark", *Nordic Economic Policy Review*, 55-90.
- Schultz-Nielsen, M. L. and J. R. Skaksen (2017), "Indvandreres uddannelse", *Rockwool Foundation Research Unit WP*, no. 48.

Statistics Denmark (2016), *Indvandrere i Danmark 2016*, [English: Immigrants in Denmark 2016], Copenhagen: Statistics Denmark.

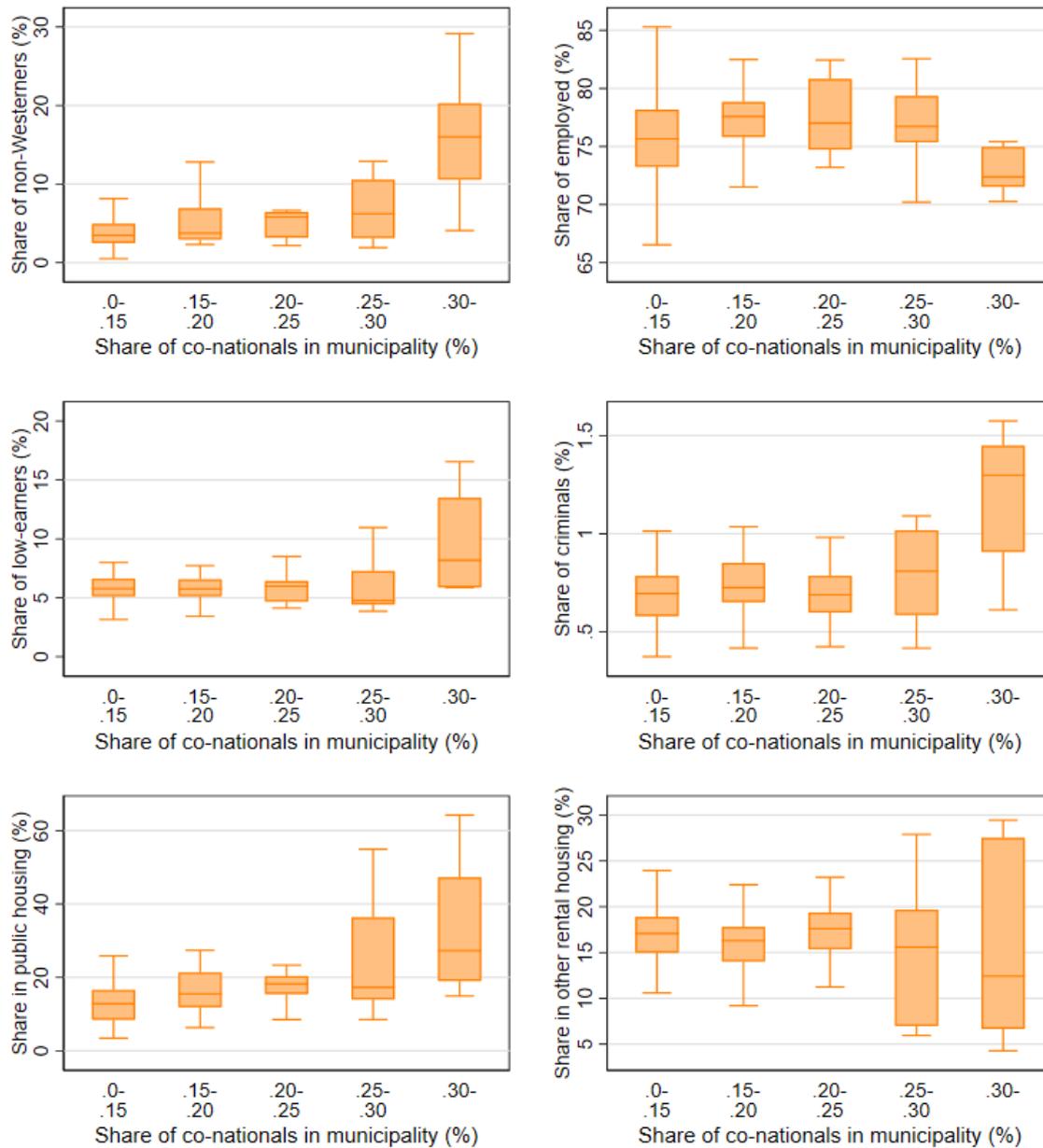
Zorlu, A. and J. Hartog (2018), 'The Impact of Language on Socioeconomic Integration of Immigrants', *IZA DP*. no. 11,485.

**Appendix Table 1.** Sample selection criteria.

#	Explanation	Reduction	Sample size
1	Entire 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 the individuals with imputed residence permit type in any given year If an individual's observation is imputed at any time	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 and 2019	164	28,452
9	Age on arrival is 18 years or over, (calculated as the year of receiving a residence permit minus the date of birth recorded in BEF)	6,578	21,874
10	First appearance in BEF is the same or one year after receiving a residence permit; out of which: <i>Observed in year 0</i> <i>Observed in year 1</i>	1,361	20,513 <i>10</i> <i>20,503</i>
11	The initial neighbourhood of residence is observed (identified by linking residence change register (BEFBOP) to neighbourhood dataset)	87	20,426
12	The municipality of initially identified neighbourhood should be the same as the first municipality of residence according to step 11 ( <b>Final Sample</b> )	36	<b>20,390</b>

Source: Administrative register information from Statistics Denmark from 1999–2019, and the constructed neighbourhood dataset by Damm et al. (2019a).

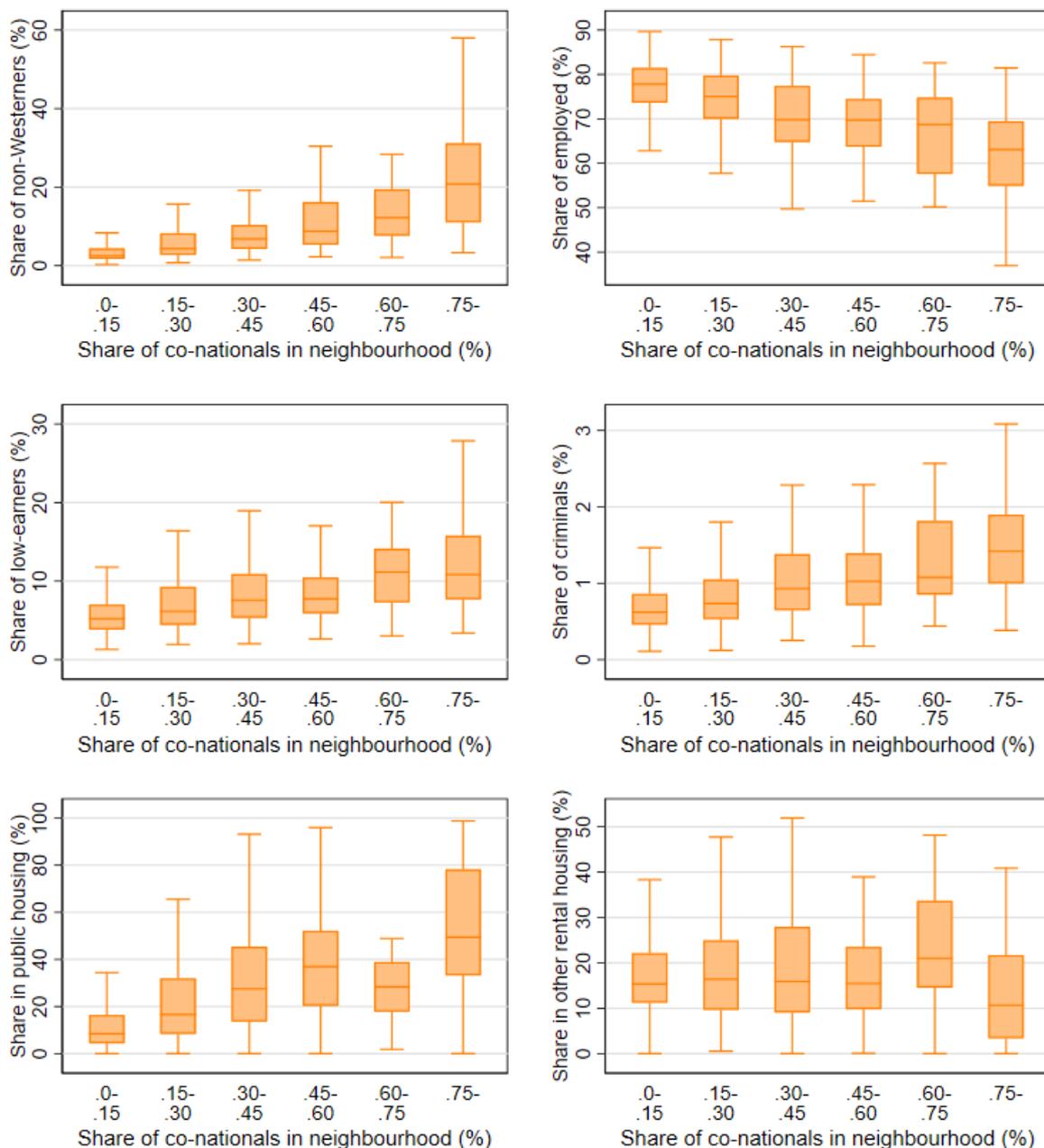
**Appendix Figure A1.** Correlation between share of co-nationals and other municipal characteristics.



*Source:* Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020.

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). Number of municipalities of settlement: 98. Non-Western immigrants include all immigrants except those from EU28 and EEA, Andorra, Monaco, San Marino, Switzerland, Vatican City, Canada, the USA, Australia and New Zealand. Low-income individuals are defined as residents with family-equivalent disposable income less than half the median of such income in the region of residence. The bottom and top of the box form the first and third quartile, while the line in the box shows the median and the whiskers illustrate the lower and upper adjacent values.

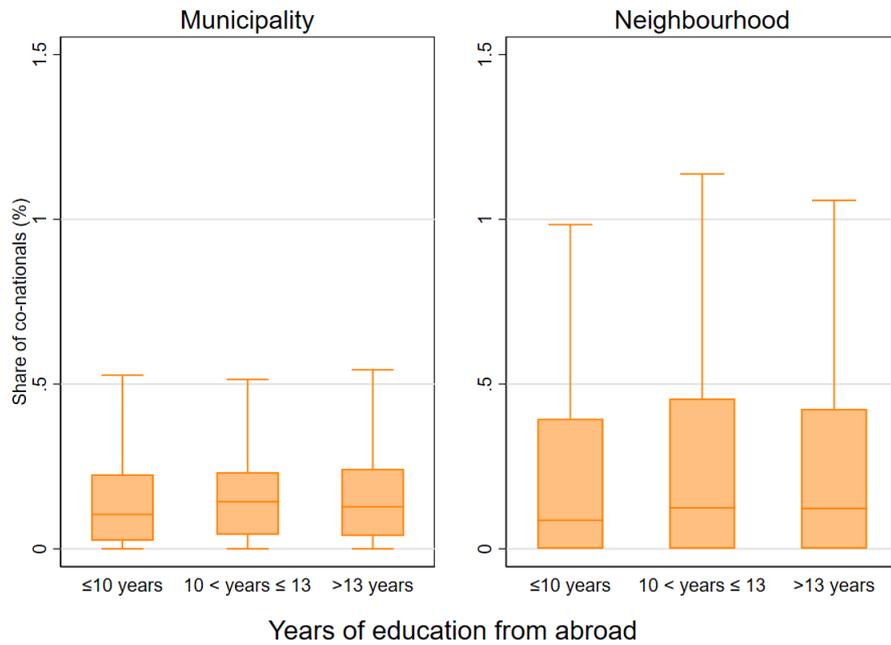
**Appendix Figure A2.** Correlation between share of co-nationals and other neighbourhood characteristics.



*Source:* Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020, linked with register on neighbourhood of residence constructed by Damm et al. (2019a).

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). 'Neighbourhood' corresponds to a macro-neighbourhood in Damm et al. (2019a). Non-Western immigrants include all immigrants except those from EU28 and EEA, Andorra, Monaco, San Marino, Switzerland, Vatican City, Canada, the USA, Australia and New Zealand. Low-income individuals are defined as residents with family-equivalent disposable income less than half the median of such income in the region of residence. The bottom and top of the box form the first and third quartile, while the line in the box shows the median and the whiskers illustrate the lower and upper adjacent values.

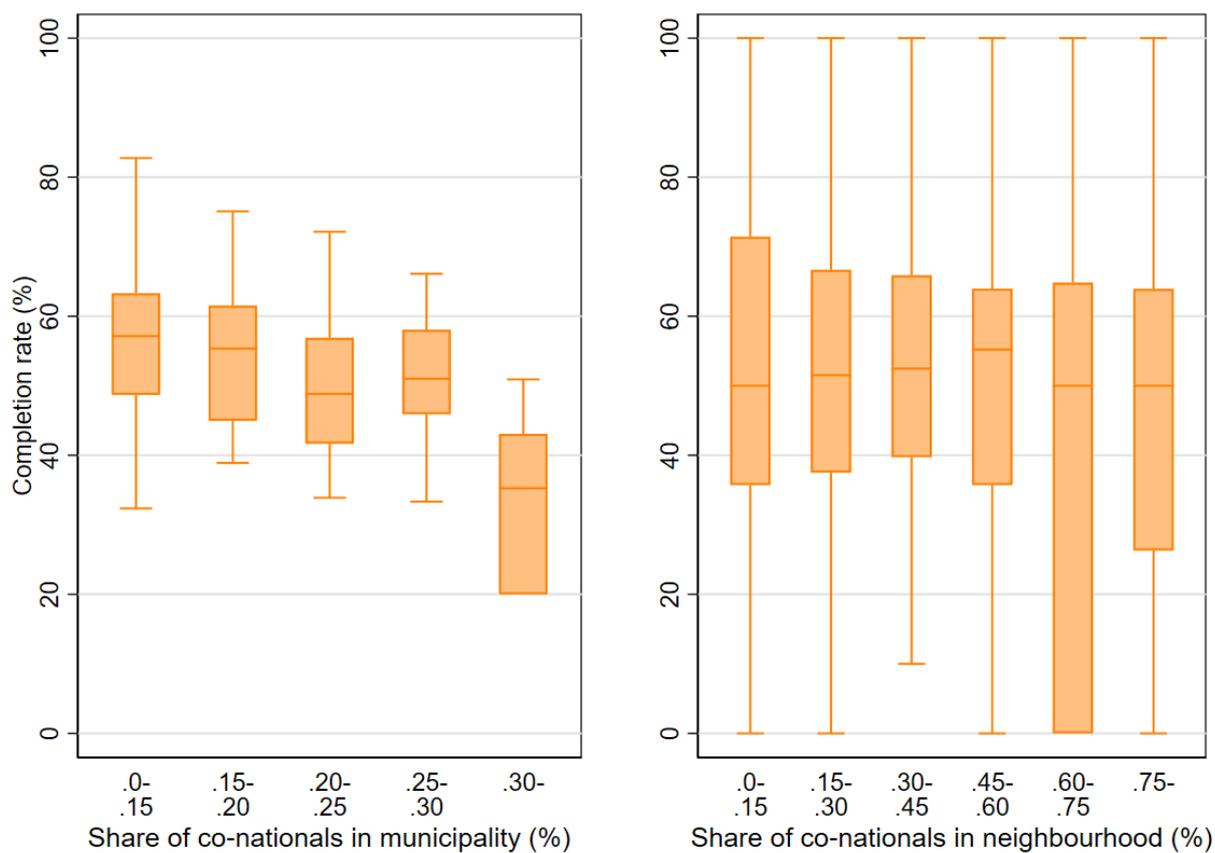
**Appendix Figure A3.** Share of co-nationals in the local area upon settlement by educational level.



*Source:* Authors' calculations based on administrative register data from Statistics Denmark, 2004–2020, linked with register on neighbourhood of residence constructed by Damm et al. (2019a).

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). The bottom and top of the box form the first and third quartile, while the line in the box shows the median and the whiskers illustrate the lower and upper adjacent values.

**Appendix Figure A4.** Share of refugees completing a language course within four years of asylum according to share of co-nationals in the municipality and neighbourhood.



*Source:* Authors' own calculations based on administrative register data from Statistics Denmark, 2004–2020, linked with register on neighbourhood of residence constructed by Damm et al. (2019a).

*Notes:* Sample: Adult settled refugees in Denmark 2004–2015 (N=20,390). The bottom and top of the box form the first and third quartile, while the line in the box shows the median and the whiskers illustrate the lower and upper adjacent values.

## Appendix 1: Identifying Language Enclaves

The aim of this appendix is to show the way in which we identify language enclaves. In particular, our goal is to count the number of persons in each municipality and neighbourhood who share the same official language(s) with a person arriving as a refugee. The period under consideration is 1986–2019. Denmark is divided into 5 regions, 98 municipalities, and Damm et al. (2019a) divided 476,171 inhabited hectare cells of Denmark into 1,961 macro-neighbourhoods (hereafter *neighbourhoods*).

### 1. Sources for construction of datasets

#### 1.1. Official languages of countries

To identify the official languages spoken in each country, we use the document from the United Nations (2017) stating the name of countries and the official languages (national official) of those countries.<sup>35</sup> We considered the official languages of countries rather than any other languages spoken because we have information only about the country the refugee left and not about the refugee's area (or region). 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-country. 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 immigrants (refugees) to speak not only with their co-nationals (in the municipality or neighbourhood) but also with persons from other countries with the same official language. Overall, there are 222 countries and 109 unique official languages.

#### 1.2. Name and code of countries

To identify the country of origin of each person in Denmark's Administrative Registers, Statistics Denmark defines a variable called "opr\_land"<sup>36</sup> consisting of 4-digit numbers, which makes a one-to-one link between country of origin and opr\_land variable.<sup>37</sup> To use the language dataset introduced in sub-section 1.1., we add opr\_land information to the language dataset. Then, we count the number of official languages in each country, as well as the frequency with which a language is official language across countries.

#### 1.3. Identifying the refugees' countries of origin and the countries' official languages

To identify the refugees' countries of origin, we used Residence Permit Register (OPHG) and restricted the sample to refugees whose residence permits were issued during 1997–2015.

To find the countries' official languages (which are also our target languages), we restricted the constructed dataset in sub-section 1.2 to refugee-origin countries. The new dataset covers refugee-origin countries and the related official languages.

---

<sup>35</sup> "11<sup>th</sup> 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](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 21 January 2021)

<sup>36</sup> Oprindelsesland

<sup>37</sup> <https://www.dst.dk/da/Statistik/dokumentation/Times/moduldata-for-befolkning-og-valg/opr-land> (accessed on 21 January 2021)

Overall, there are 123 refugee-origin countries and 78 official languages spoken in these countries. We classify those 123 countries in three groups:

- Group 1: Countries with one official language (84 countries).
- Group 2: Countries with more than one official language (28 countries).
- Group 3: Countries without an official language (11 countries).

By identifying target languages (i.e., official languages in refugee-origin countries), we restricted the languages to target languages. This allowed us to concentrate only on languages that are official in the refugees' origin countries, because our goal was to count the number of persons in each municipality and neighbourhood with whom a refugee could communicate using a common language (such as Arabic). Consequently, we considered all the countries but only official languages in refugee-origin countries.

## **2. Strategies for counting the number of persons with common target languages**

As mentioned in sub-section 1.3, we have 3 groups of countries based on the number of official languages spoken in those countries. In this section, we describe strategies for counting the number of persons (full population) with whom a refugee from a refugee-origin country can speak using the same language.

### **2.1. Group 1: Refugee-origin countries with one official language**

There are 84 refugee-origin countries with one official language. For countries in this group, counting the number of same-language persons in the municipality and neighbourhood of residence is straightforward. We merely need to identify the country of origin (`opr_land` variable) and the official language in that country. We can then count the number of persons in that municipality or neighbourhood speaking that specific language.

### **2.2. Group 2: Refugee-origin countries with more than one official language**

There are 28 refugee-origin countries with more than one official language. Counting the number of same-language persons (in the municipality or neighbourhood) for countries in this group is challenging because there could be countries (both refugee-origin and other countries) that share more than one language with a refugee-origin country. For example, suppose we have a refugee from refugee-origin country A with 3 official languages of Arabic, English, and French. In addition, suppose a person from country B (which can be from refugee-origin or other countries) with two official languages of Arabic and French lives in the same municipality (or neighbourhood) as the refugee lives. In this case, the refugee can talk with that person in either Arabic or French, or both. However, that person should be counted only once. Therefore, without considering this fact, it is possible to count a person two or more times, depending on the number of shared languages.

To deal with this issue, we introduced a new language variable for each of the refugee-origin countries in Group 2, and countries (both refugee-origin and other countries) take the value of one when they have *at least* one of those official languages as their official language. By doing so, for instance, persons from country B not only speak Arabic and French, but also the newly introduced language.

Nevertheless, we are interested in counting the number of persons speaking the newly introduced language. Accordingly, we counted them only once for refugees from country A.<sup>38</sup> This strategy helps avoid multiple counting. We then added these 28 newly introduced languages to the constructed dataset in sub-section 1.2. and made a new dataset.

### 2.3. Group 3: Refugee-origin countries without an official language

There are 11 refugee-origin countries without an official language. We assume that refugees belonging to one of the countries in this group can share a common language (only) with other persons from their own countries. That is, we simply count the number of co-nationals of each refugee from countries in this group.

## 3. Linking the constructed dataset to the Population Register during 1986–2016

The Danish Population Register (BEF) gives information about the country of origin (`opr_land`), municipality of residence, and unique housing address for each inhabitant of Denmark over time (in an encrypted version). Therefore, we can easily link the unique housing address of each person to the neighbourhood dataset constructed by Damm et al. (2019a) and find the neighbourhood of residence for each person. Consequently, we can obtain all the required information (i.e., country of origin, municipality and neighbourhood of residence). As a result, we linked the Population Register with the above-mentioned information to our language dataset using the country of origin (i.e., `opr_land` variable).

Finally, we counted the number of persons speaking each of the official languages in each municipality (neighbourhood) in each year.

---

<sup>38</sup> For instance, Sudan has two official languages, Arabic and English. Further, Eritrea has three official languages, Arabic, English, and Tigrinya. Suppose a refugee from Sudan lives in a municipality (or neighbourhood) where a person from Eritrea also lives. The refugee from Sudan can speak with the person from Eritrea in Arabic and/or English. What is important here is that there is one Eritrean in the municipality (or neighbourhood) with two official languages shared with a Sudanian refugee, which means that the person from Eritrea should be counted only once. Accordingly, we introduce a new language for persons from Sudan called “Sudan\_case”, and for persons coming from countries where Arabic or English (or both) is an official language, the `Sudan_case` variable takes the value of one and zero otherwise. In this example, for the person from Eritrea, `Sudan_case` takes the value one, irrespective of that person sharing two languages with the Sudanian refugee. Now, to count the number of same-language persons with a person from Sudan in the municipality (or neighbourhood), we count the number of persons for whom `Sudan_case` equals one (rather than counting the number of persons speaking Arabic and number of persons speaking English in the municipality and summing them up).