

Linguistic Enclaves, Sorting, and Language Skills of Immigrants

Agnieszka Kanas, Yuliya Kosyakova & Ehsan Vallizadeh

To cite this article: Agnieszka Kanas, Yuliya Kosyakova & Ehsan Vallizadeh (2024) Linguistic Enclaves, Sorting, and Language Skills of Immigrants, Journal of Immigrant & Refugee Studies, 22:4, 847-861, DOI: [10.1080/15562948.2022.2132572](https://doi.org/10.1080/15562948.2022.2132572)

To link to this article: <https://doi.org/10.1080/15562948.2022.2132572>



© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC.



[View supplementary material](#)



Published online: 14 Oct 2022.



[Submit your article to this journal](#)



Article views: 3007



[View related articles](#)






[View Crossmark data](#)



Citing articles: 2 [View citing articles](#)

Linguistic Enclaves, Sorting, and Language Skills of Immigrants

Agnieszka Kanas^a , Yuliya Kosyakova^{b,c}  and Ehsan Vallizadeh^{b,d} 

^aDepartment of Public Administration and Sociology, Erasmus University Rotterdam (EUR), Rotterdam, the Netherlands; ^bDepartment Migration and International Labour Studies, Institute for Employment Research (IAB), Nuremberg, Germany; ^cChair of Sociology, Area Societal Stratification, Otto-Friedrich University of Bamberg, Bamberg, Germany; ^dDepartment of Economics, Otto-Friedrich University of Bamberg, Bamberg, Germany

ABSTRACT

The literature argues that linguistic enclaves negatively affect immigrants' language proficiency by reducing their exposure and incentives to learn destination language. This negative association may, however, be spurious, arising due to the self-selection of immigrants into regions with larger enclaves. Exploiting the natural experiment of the German residential policy, this paper analyses the influence of linguistic enclaves on refugees' language proficiency. We find no evidence that enclaves hamper German language learning among refugees. Our results are robust to various measurements and model specifications. We conclude that the negative relationship reported in previous research is mainly driven by immigrant's residential sorting.


KEYWORDS


Destination language acquisition; linguistic enclaves; residential policy; refugees; Germany

Introduction

In most countries, immigrants tend to live in spatially concentrated residential areas that are usually seen as ethnic or linguistic enclaves (Borjas, 1995; Damm, 2009a).¹ This is also true for Germany, where 50 percent of immigrants tend to concentrate in 53 out of 401 counties (own calculation based on the DESTATIS, 2019). While numerous studies have recognized the importance of immigrants' spatial concentration, empirical evidence is yet inconclusive regarding whether living in such enclaves benefit or hinders immigrants' integration in the host country. Some studies have shown that enclaves improve the labor market integration of immigrants due to lower job-search costs and better job-referral networks (Chiswick & Miller, 1996; Damm, 2009b; Martén et al., 2019; Mouw & Chavez, 2012; Portes & Sensenbrenner, 1993). Other studies, however, suggest that enclaves have an adverse effect on immigrants' integration process, for example, by impairing destination language skills (Bauer et al., 2005; Danzer & Yaman, 2016; Dustmann & van Soest, 2004; Espenshade & Fu, 1997; van Tubergen & Kalmijn, 2005).²

One key argument is that immigrants living in residential areas with larger linguistic enclaves accumulate less destination language skills due to low learning incentives—an effect also known as the *learning effect* (Bauer et al., 2005; Kónya, 2007; Lazear, 1999; Monteiro, 2021). Following this premise, a large body of migration research has studied the importance of ethnic/linguistic enclaves for immigrants' language acquisition (Bauer et al., 2005; Chiswick & Miller, 2001; Danzer & Yaman, 2016; Dustmann & van Soest, 2004; Espenshade & Fu, 1997; van Tubergen & Kalmijn, 2005). Although these studies have often argued that immigrant enclaves are detrimental for the

CONTACT Agnieszka Kanas  kanas@essb.eur.nl  Department of Public Administration and Sociology, Erasmus University Rotterdam (EUR), Rotterdam, the Netherlands.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/15562948.2022.2132572>.

© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

learning of the destination language, the causal effect of enclaves on language acquisition has been less often addressed by this literature.

Estimating the causal effect of enclaves on destination language proficiency is challenging due to immigrants' residential sorting. Immigrants who are less able or have lower incentives to learn the destination language are more inclined to move to areas where many others speak their own language due to lower expected communication costs and greater support from the ethnic community (Bauer et al., 2005; Lazear, 1999). Consequently, immigrants' residential choices are not random, and the selection into or out of an enclave depends importantly on the immigrants' ability and incentives to learn the destination language—an effect also known as the *sorting effect* (Lazear, 1999). In this paper, we are able to address immigrants' residential sorting and bring to the forefront the question of how linguistic enclaves impair destination language proficiency from the longitudinal perspective. We define linguistic enclaves as group of people who are similar in terms of languages used. Immigrants' access to information and social interaction in the host country, especially upon arrival, is commonly characterized by interacting with other immigrant groups that speak a similar origin language (Chiswick & Miller, 1996; Portes & Sensenbrenner, 1993).

To estimate the causal effect of linguistic enclaves on immigrants' language development over time, we exploit a unique natural experiment distilled from two important policies in Germany: (a) the initial allocation policy of recently arrived refugees³ (*Erstverteilung der Asylsuchenden*, EASY) and (b) the residency obligation policy (*Wohnsitzauflage*). The first policy implies that upon arrival, refugees are exogenously assigned to German counties, conditional on regional economic strength and population size. Note, that similar initial allocation policies of refugees are a common practice in many Western countries (e.g., Australia, Denmark, Ireland, Norway, Sweden, Switzerland, the UK, and the U.S.; see Kosyakova & Kogan, 2022). The second policy implies a residency obligation for approved refugees for additional three year. This policy, introduced by the Integration Act in August 2016, has essentially prohibited the residential mobility of refugees. As such, the two policies enable us to fully control for immigrants' sorting in and out of initial linguistic enclaves, allowing us to draw causal inferences about the impact of linguistic enclaves on language acquisition among immigrants. Unlike previous studies, where immigrants are free to choose their place of residence, refugees in Germany are exogenously allocated to their first residence place and restricted to stay there for a substantial time. Correspondingly, regional disparities refugees are exposed to in their first residence place could have consequences for their integration process. For instance, previous research has revealed the importance of economic conditions for refugees' integration (e.g., Aksoy et al., 2020; Åslund & Rooth, 2007; Kanas & Kosyakova, 2022; Kanas & Steinmetz, 2021). Likewise, local opportunities for coethnic networks seems to be an essential resource fostering labor market integration of refugees (e.g., Damm, 2009b; Martén et al., 2019).

The initial period of language acquisition is more crucial for refugees than for migrants arriving for economic or family motives. Because of their more abrupt and unprepared migration, refugees seldom arrive with destination language proficiency (Brücker et al., 2020; Kosyakova et al., 2022), resulting in persisting linguistic gaps with other immigrant groups (Chiswick et al., 2006; Kosyakova et al., 2022). By examining the effects of linguistic enclave on refugees' language proficiency from a longitudinal perspective, we contribute to the social networks' literature regarding the role of linguistic enclaves for the integration process of immigrants in the host country. We also contribute more generally to the migration literature concerned with the importance of local reception context for immigrants' life chances. In addition to this substantive interest, the current study addresses the heterogenous impact of linguistic enclaves on immigrants' destination language skills.

Our empirical analyses are based on a longitudinal data from the representative survey of refugees in Germany, the IAB-BAMF-SOEP Survey of Refugees (Bruecker et al., 2017; Kühne et al., 2019), launched in the aftermath of the surge of refugee migration to Europe in 2015.

The data provide detailed information on self-reported German language proficiency in reading, writing, and speaking—important measures that have become standard practice in studies on language skills (van Tubergen & Kalmijn, 2005, p. 1248). Moreover, the survey includes information on the initial county of assignment of refugees. A rich set of information in the data regarding sociodemographic characteristics, pre- and postmigration educational attainment, and premigration German language proficiency allow us to account for potential confounders in our natural experiment.

The allocation and residency obligation policies in Germany

Concerns about the capacity of local labor markets to integrate large groups of refugees, displacement effects on housing markets and the concentration of immigrants in ethnic ‘ghettos’ are widespread in many Western countries (Fasani et al., 2022; Lastrapes & Lebesmuehlbacher, 2020; Triandafyllidou, 2018). As a result, many destination countries enforce allocation and residency obligation policies requiring refugees to settle in the specific locations within the country (Bansak et al., 2018; Fasani et al., 2022; Martén et al., 2019).

In Germany, refugees are distributed across counties (*Kreis*) according to allocation policies, which aim to ensure an equal distribution of refugees across German regions without considering individual preferences (Schneider, 2012). This distribution occurs in two steps: first across and then within German federal states (*Bundesländer*). More specifically, newly arrived refugees are registered in the computer-based system EASY, which then assigns them to the initial reception center (typically nearby) by a quota (*Königssteiner Schlüssel*). This quota is updated annually and is based on the tax revenues (2/3 weight) and population size (1/3 weight) of a federal state. In the second step, the authorities of the federal states are responsible for the further distribution of the assigned refugees within their territory (based either on similar quota regulations or solely on population size, see Geis & Orth, 2016).

Refugees’ first residential allocation is binding. Refugees with pending or rejected asylum applications face very strict residency obligation, including travel bans (*Residenzpflicht*, §56 Residence Act). The obligation to reside in the county of the initial allocation can be abolished upon the official approval of refugee status. These rules have been modified with the Integration Act since August 2016. Since then, approved refugees have been obliged to take residence in the federal state in which their asylum application was processed for an additional three years after approval (§12a, Residence Act).⁴

Note, that the Immigration Act from 2016 delegates the application of the residency obligation policy to the authorities at the Federal State level. While mobility across the boundaries of the sixteen Federal States was generally restricted, it remained the decision of the Federal States to restrict mobility at the more local level. Correspondingly, in seven out of 16 federal states, the implementation of the residency obligation policy is more restrictive, such that refugees are even obliged to take up residence not only within the federal state, but also within counties (*Kreis*) and sometimes even within municipalities (*Gemeinde*). Some (urban) federal states—Berlin, Bremen, and Hamburg—represent only one county each, hence, may be considered more restrictive after the implementation of the reforms as well. Consequently, the residency obligation policy prolongs the initial distribution of refugees for a substantial period even after the approval of their asylum applications. In terms of timing, the federal law became effective in August 2016 and had retroactive effect from January 2016 for approved refugees and refugees with a resident permit. Moreover, in states which enforce more restrictive residency obligation policy (i.e., at the county or municipal level) the implementation took place at different points in time (see Brücker et al., 2020; Kanas & Kosyakova, 2022; see also Table A3 in the Online Appendix). Because of the retroactive nature and variation across states in the implementation of residency obligation, it is rather unlikely that refugees could anticipate this policy.

Linguistic enclaves and immigrants' destination language proficiency

Previous studies have constantly observed a negative association between immigrant ethnic or linguistic enclaves and destination language skills, which has been supported in a number of countries and among various groups of immigrants (Bauer et al., 2005; Chiswick & Miller, 2001; Danzer & Yaman, 2016; Dustmann & van Soest, 2004; Espenshade & Fu, 1997; van Tubergen & Kalmijn, 2005). To infer the nature of this negative association, we build upon the conventional model of destination language acquisition (Chiswick & Miller, 2001). According to this model, immigrants improve their language proficiency depending on their exposure to the new language, their incentives to learn the destination language, and the individual efficiency—the extent to which a given amount of destination-language exposure produces language fluency. Following this model, immigrants invest in destination language skills if the economic returns to their investments in destination language skills are high (incentives). The investment depends also on whether refugees can hear and use the destination language frequently (exposure) and if they find it easier to learn a new language (individual efficiency) (Chiswick & Miller, 2001).

Correspondingly, living and working among many immigrants who speak the same language may lower the exposure to the destination language and reduce the incentives to acquire destination language skills (Chiswick & Miller, 1996). This is because linguistic enclaves reduce opportunities to hear and speak the destination language, thus slowing down the rate of language acquisition. Linguistic enclaves also reduce the intensity of immigrants' exposure to the destination language indirectly by lowering the intensity of social interactions with natives (Chiswick & Miller, 1996; van Tubergen, 2010). Likewise, living and working among other immigrants speaking the same language increase opportunities for social interaction and work as well as give access to ethnic goods and services (such as newspapers and employment agencies), reducing the necessity and importance of investing in destination language skills (Espenshade & Fu, 1997; van Tubergen & Kalmijn, 2005).

Although the third determinant of destination language acquisition—individual efficiency—is unlikely to be affected by the presence of linguistic enclaves, it is likely to play an important role in the self-selection of immigrants in and out of enclaves (Bauer et al., 2005; Chiswick & Miller, 1996). Thus, while previous arguments infer that linguistic enclave can impair the *learning* of the destination language, in the following, we will argue that this negative association can be spurious and may arise due to immigrants' residential *sorting*. In particular, immigrants with higher efficiency to learn new language—commonly measured by years of schooling (e.g., Chiswick & Miller, 2001)—are said to be more capable in converting exposure to the destination language into greater language skills. More educated immigrants have often greater abilities to learn different school subjects, which might extend to the learning of foreign languages. Likewise, they may have had more exposure to foreign languages at school before migration. Higher-educated immigrants usually have more economic incentives to invest in destination language acquisition because such skills are required in many better-paid jobs (Bauer et al., 2005; Chiswick & Miller, 1996).

However, how is the individual efficiency to learn the destination language related to immigrants' decision to sort into and out of linguistic enclaves? The decision to live in a region with a local population speaking the same language is not made at random—it is rather a deliberate decision to self-select into a preferred residential setting. With regard to individual efficiency, selection into linguistic enclaves is argued to be unfavorable for learning the destination language (Bauer et al., 2005; Lazear, 1999). That is, immigrants who are less capable or willing to learn the destination language select themselves into linguistic enclaves to reduce communication costs and benefit from the support of the linguistic community.

Combining these arguments, if the self-selection of immigrants into linguistic enclaves is systematic with respect to their efficiency and incentives to learn and acquire a new language specifically, the effect of linguistic enclaves on language acquisition established in prior studies is likely to be biased due to the sorting of immigrants. Moreover, there is likely to be a variation in immigrant sorting over time (Mouw & Chavez, 2012). That is, as the length of stay increases,

some immigrants will improve their language proficiency and—as relying on family and coethnic connections become less necessary—select themselves out of linguistic enclaves. This idea of (self-)selection out of enclaves has been also supported by Bauer et al. (2005) who examined the effect of English language proficiency on the location choice among Mexican immigrants in the U.S. The authors showed that immigrants choose areas with smaller enclaves as their English language proficiency improves.

A crucial omission in previous literature on destination language acquisition is not being able to account properly for immigrants' sorting into and out of enclaves, mainly due to data availability. As argued above, immigrants' characteristics, such as their efficiency to learn the destination language, are likely to influence a decision to move into or out of a linguistic enclave and confound thereby the effect of linguistic enclaves on immigrants' language skills (Bauer et al., 2005; Chiswick & Miller, 1996). In a first attempt to address these sorting mechanisms, Lazear (1999) showed that immigrants with poor English language proficiency before migration to the U.S. tend to sort into linguistic enclaves because they are less able to interact with the native population. However, whether linguistic enclaves affect the destination language fluency of immigrants remains unanswered.

Closely related to our study, Danzer and Yaman (2016) exploited dispersal policy, under which guest worker immigrants were exogenously allocated to German counties. While Danzer and Yaman (2016) observed a negative impact of coethnic enclaves on guest workers' German language proficiency, they had only information on immigrants' residence counties during the survey period—on average 14 years after immigration. This implies that the authors could neither fully control for residency changes and residential sorting in and out of enclaves nor observe the development of immigrants' language skills over the 14-year period since migration, leaving the questions about the causal effect of enclaves on language learning rather unresolved.

In contrast, refugees in Germany (and in our data) are not only assigned to their first residence place but also restricted to stay there for a period of three years. Moreover, we are able to capture refugees' learning trajectories early since arrival—a very dynamic stage that is crucial for individuals' language proficiency (Stevens, 1999). The longitudinal nature of our data allows us to depict the development of language skills over time, taking into account initial skill levels. In doing so, we provide empirical evidence for the dynamic effects of linguistic enclaves on immigrants' language proficiency. We improve previous research by exploiting the allocation and residency obligation policies in Germany to investigate how the exogenous variation in the size of linguistic enclaves at refugee's residency upon arrival affects refugees' destination language acquisition. At the core of our empirical analysis, we test the hypothesis whether larger linguistic enclaves have a negative effect on refugees' language proficiency. If this hypothesis is supported, the conclusion is that linguistic enclaves are detrimental to the learning of destination language. However, if the negative effect of linguistic enclave established by previous studies on language acquisition is spurious—i.e., due to less favorable sorting into linguistic communities—we should find no negative effect of larger linguistic enclaves on refugees' language proficiency.

Data and method

Data and sample

The IAB-BAMF-SOEP Refugee Survey is a nationally representative longitudinal household survey of adult asylum-seekers and refugees in Germany launched in 2016 (Bruecker et al., 2017; Kühne et al., 2019). The findings presented here refer to the first five waves of the survey, conducted between 2016 and 2020. The data were drawn from the Central Register of Foreign Nationals (*Ausländerzentralregister*, AZR). The sampling frame targeted the population of refugees who entered Germany between 2013 and 2016 (irrespective of their factual legal status by the sampling date). The response rate in the first wave amounted to approximately 50 percent, which in comparison to other surveys of individuals with a migration background is quite high (Kroh

et al., 2017). The survey itself covers the respondent and all household members of the respondent. Interviews were conducted in a computer-assisted personal interview (CAPI) mode and were supported by interpreters if needed. The questionnaires were available in seven languages (Arabic, English, Farsi/Dari, German, Kurmanji, Pashtu, and Urdu) and included auditory instruments for survey participants who were illiterate.

For our analyses, we restrict initial sample to refugees aged 18-55 at arrival who were first surveyed in their first two years of residence in Germany and had valid information on German language proficiency and country of origin (to measure linguistic enclaves, see below). To identify refugees subject to restrictive residency obligation policy, we considered month and year of asylum application and decision, and the type of decision, as well as the county of assignment.⁵ Hence, respondents with missing or inconsistent information on those parameters were excluded. In line with the German residential policy rules, we consider following refugee groups being subject to the restrictive residency obligation policy. First, this refers to refugees with a pending or negative decision on their asylum application. Second, approved refugees assigned to the restrictive federal state who received their positive decision on the asylum application after the restrictive residency obligation policy was implemented in the corresponding federal state and who were interviewed within 36 months since approval. For example, consider a refugee assigned to Bavaria (the implementation of restrictive residency obligation policy in Bavaria was effective in September 2016) whose asylum application was approved in October 2016 and who was interviewed each December between 2016 and 2019. In such a case, we include observations of this respondent in 2016 and 2017, while the observations in 2018 and 2019 are excluded.

Applying restrictions criteria specified above, the analytical sample reduced to 5732 person-year observations (2677 persons) (see Table A1 in the Online Appendix for detailed description on the sample construction). Table A2 includes a detailed overview of the Federal States with the restrictive residency obligation policy and information on the date when the policy was implemented for each region.

Variables

To analyze the *German language proficiency* of refugees in Germany, we rely on self-reported German language skills—a conventional measure used in studies relying on survey data (e.g., Espenshade & Fu, 1997; van Tubergen & Kalmijn, 2005). Respondents rated their reading, speaking and writing skills on a scale from 0 “Not at all” to 4 “Very well”. For German language proficiency, we compiled an additive index of these three measures, which ranges from 0 to 12, with a mean (mode) value of 5.8 (6) (weighted results). The internal consistency, as indicated by Cronbach’s alpha, is very high and amounts to 0.873. In our sample, the share of respondents with (very) good self-reported German language proficiency amounts to 38 percent with a mean proficiency index of 8.8. The remaining 62 percent had an average language index of 4.0 points (weighted results).

To measure the size of *initial linguistic enclave* in the assigned county in the year of arrival, we augment the IAB-BAMF-SOEP Refugee Survey with the county data on immigrants from the Federal Statistical Office (DESTATIS, 2020a, 2020b). To calculate the linguistic enclaves, we use the information on linguistic proximity between the respondent and the immigrants in the county of the assignment. The linguistic proximity was developed by ethnolinguistics and ethnobiologists and measures the similarities between a fixed number of words in every pair of languages (Melitz & Toubal, 2014). For our measure of linguistic enclaves, we consider the total of respondent’s linguistic proximities in the county of assignment to each non-German resident in this county. This measure is then divided by local population size (to account for the potential diffusion of an enclave across the county). Formally, we define the linguistic enclave, $ENCLAVE_{iact_0}$, in arrival year (t_0) of an individual i from country of origin a in county c as follows:⁶

$$\text{ENCLAVE}_{iact_0} = \sum_{b \in B} \frac{S_{ab} \times N_{bc}}{N_c}, \quad (1)$$

where S_{ab} is the measure of language proximity between origin country a and origin country b (B indicates a set of all origin countries in a county), N_{bc} is the number of individuals from origin country b in county c , and N_c is the total population in county c . To infer on the origin country, we consider respondents' country of birth.

Accordingly, the measure varies across origin-country and counties of arrival, and for each refugee, it is fixed to the value in the year of arrival. There are 401 counties in Germany, with a mean (median) of 65,801 (43,643) migrants per county. Our sample of refugees is distributed across 275 assignment counties. Our measure of the linguistic enclave has an average size of 0.005 with a standard deviation of 0.008 (weighted results). In our sample, the refugees with the highest value of the average linguistic enclave size are those from Azerbaijan (0.020), followed by refugees from Turkey (0.019), Yemen (0.016), Sudan (0.15), Kuwait Morocco and Libya (0.14), Montenegro and Egypt (0.012), Iraq and Syria (0.011) (weighted results). For the empirical analyses, we standardize the linguistic enclave variable to have a mean of zero and a standard deviation of one. Higher values indicate higher linguistic concentration.

Note, that our measure of linguistic enclaves is exogenous to refugee's individual characteristics, i.e., regional sorting decisions. Likewise, the natural experiment design as ours ensures that other county characteristics are randomly distributed across refugees. Yet, it does not exclude a possibility that the relationship between the linguistic enclaves and refugees' language proficiency is due to unobserved county characteristics that correlate with a local linguistic enclave. For instance, local labor market conditions in the past could have attracted more immigrants. At the same time, there could be heterogeneous preferences among previous immigrants regarding regional characteristics driven by historical immigration patterns in Germany (Tanis, 2020). Hence, we account for potential confounders by augmenting our data with the information on spatial classification (*initial rural-type county*), *initial population density*, *initial unemployment rate* (in percent), *initial share of foreigners* (in percent), and fixed effects for *origin country* (grouped to Syria, Afghanistan, Iraq, Eritrea, Iran, rest of MENA, former USSR, West Balkan, rest Africa, and other states). Note that country of origin fixed effect wave out potential issues related to accelerated asylum procedure for some countries of origin and, hence, faster access to integration courses (Kosyakova & Brenzel, 2020).

Moreover, a natural experiment such as ours allows for more noise than a controlled laboratory or field experiment. For this reason, we include a set of pre- and post-migration individual-level variables as controls (e.g., Kónya, 2007; Kosyakova et al., 2022; Monteiro, 2021). Among the sociodemographic variables, we account for being *female*, *age at arrival*, *partnership* and *children* status, and *duration of stay* measured via months since arrival in Germany. Socioeconomic confounders include a *premigration German language proficiency*, *premigration years of education*, indicators for having *premigration work experience*, *country-of-origin literacy*, *(very) good health*, living in *shared accommodation*, *asylum application status*, and *language course participation*, as well as *weekly contacts with Germans*. To absorb any systematic differences related to the survey design, we control for the survey sample (M3, M4, M5, M6). Survey year fixed effects account for time trends. Table A3 in the Online Appendix provides information about the unweighted distributions of all variables and presents their definitions.

Analytical strategy

To estimate the effect of linguistic enclaves on language proficiency process, we first estimate a linear regression model with standard errors clustered at the person level to account for the fact that some refugees are surveyed repeatedly. Our model is specified in equation (2):

$$Y_{it} = \alpha_0 + \alpha_1 ENCLAVE_{iact_0} + \alpha_2 X_i + \alpha_3 X_{it} + \alpha_4 X_{ct_0} + \lambda_t + \lambda_f + \varepsilon_{ict_0t} \quad (2)$$

where Y_{it} denotes German language proficiency of respondent i in survey year t , $ENCLAVE_{iact_0}$ denotes the linguistic enclave the person is exposed to in assigned county c in the year of arrival t_0 (see Section *Variables*), vector X_i denotes time-invariant individual-level characteristics, vector X_{it} denotes time-variant individual-level characteristics, and vector X_{ct_0} denotes the characteristics of the assigned county in the year of arrival. λ_t denotes fixed effects for year of interview t . λ_f denotes fixed effects for assigned federal state f . ε_{ict_0t} is the error term. Our main interest is in the effect of time-invariant variable – linguistic enclave measured at t_0 of the panel. The coefficients of the linguistic enclave indicate whether individuals assigned to counties with larger linguistic enclaves become on average less fluent in German than those assigned to counties with smaller linguistic enclaves.

In the second step we add an interaction effect between linguistic enclave and the duration of stay in Germany of respondent i in survey year t ($DURSTAY_{it}$), to measure whether the effect of linguistic enclave on language proficiency changes with the progressed time since arrival in Germany.

$$Y_{it} = \alpha_0 + \alpha_1 ENCLAVE_{iact_0} + \alpha_2 DURSTAY_{it} + \alpha_3 ENCLAVE_{iact_0} \times DURSTAY_{it} + \alpha_4 X_i + \alpha_5 X_{it} + \alpha_6 X_{ct_0} + \lambda_t + \lambda_f + \varepsilon_{ict_0t}, \quad (3)$$

Results

Refugees' language proficiency over duration of stay

Figure 1 describes the initial language proficiency level and its development over the duration of stay for the recent refugees in Germany. In line with previous research, virtually no refugee possessed German language skills upon their arrival in Germany (Kosyakova et al., 2022). However, there is a notable increase in German language proficiency since arrival with refugees achieving proficiency at around roughly 7.9 points five years after arrival. We also observe the typical process of language acquisition with initially more rapid progress (reflected in a steeper learning curve) that flatters over longer duration of stay (Hartshorne et al., 2018).

The effect of linguistic enclaves on refugees' language proficiency

Table 1 investigates whether and to what extent linguistic enclaves are detrimental for refugees' language proficiency using linear regression models.

Our empirical results indicate several novel insights. In particular, we find no evidence that living in an area with larger linguistic enclave detracts refugees' German language proficiency. Specifically, the beta coefficient of linguistic enclave is negative but not statistically significant (Model 1.1: $p=0.343$). This insignificant effect also remains after controlling for county-level and individual-level confounders (Model 1.2: $p=0.227$ and Model 1.3: $p=0.381$, respectively). In support of our descriptive results, we find increasing German language proficiency over refugees' duration of stay: a one-month increase in stay duration increases language proficiency by 0.05 points (Model 1.3).⁷ However, the dynamics of German language learning is not affected by the size of linguistic enclaves where refugees were initially assigned to (Model 1.4). That is, neither the main effect of linguistic enclave nor the interaction with duration of stay are statistically significant. In sum, the negative effects of enclaves on destination language acquisition established in previous research seem to be at least partially attributed to refugees' sorting into and out of linguistic enclaves.

While the claim that linguistic enclaves obstruct destination language learning is clearly rejected, the standard framework of language acquisition (Chiswick & Miller, 2001) is well

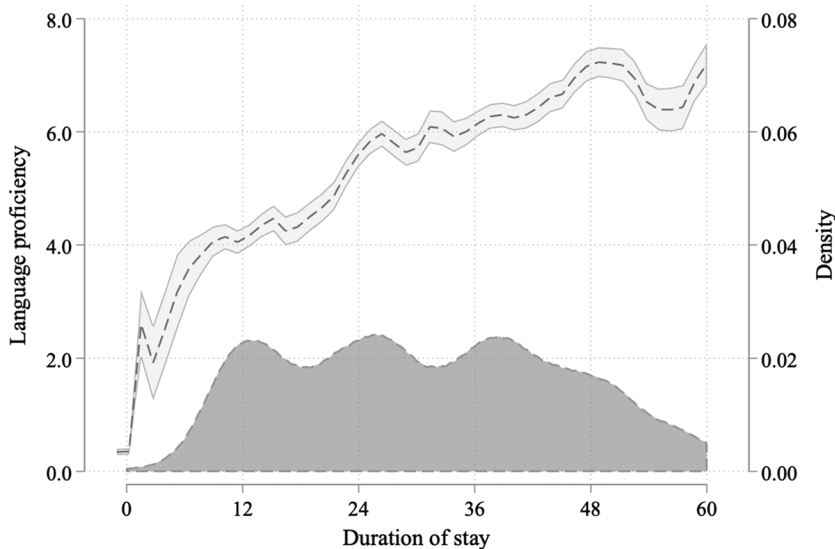


Figure 1. Development of language proficiency and density distribution of the duration of stay. *Data source:* IAB-BAMF-SOEP Survey of Refugees 2016–2020, weighted.

supported by our data. Refugees who arrived at the younger age, had more years of schooling, are literate in the origin country language, reported better health and arrived with some German language proficiency indicate higher German language proficiency than those who are older, less educated, less literate in origin language, less healthy and have no initial German proficiency (Model 1.3). We also find that family situation affects the language learning process. For example, being women or having children in the household is negatively associated language fluency. Likewise, individuals living in shared accommodation indicate lower German language proficiency. Compared to refugees with pending asylum application, those with approved application report better German skills while rejected refugees show no statistically significant differences. Finally, those who are more exposed to German language via language course participation or weekly contact with German natives have better German language proficiency.

We also explore the heterogeneity effects of linguistic enclaves on refugees' language proficiency. It can be argued that linguistic enclaves are particularly detrimental for those refugees who are less able and incentivized to learn destination language, which should be particularly the case for immigrants who arrived at the older age, reported worse health status, and acquired less years of education (Chiswick & Miller, 2001). Given the gendered division of labor and traditional gender expectations among recent refugees (Kosyakova & Kulic, 2022), the linguistic enclaves likely reduce incentives to invest in the German language stronger among female than male refugees (Kanas & Müller, 2021; Roder & Muhlau, 2014). Likewise, the negative effects of linguistic enclaves on learning destination language are likely to be larger for those who otherwise have little exposure to contact with natives. Here, we focus on two factors which are associated with increased opportunities to hear and speak the destination language: language course participation and contacts with German natives (Kosyakova et al., 2022; van Tubergen & Kalmijn, 2005). We find almost no evidence for the heterogeneous impacts of linguistic enclaves on refugees' language proficiency (Table A4 in the online Appendix). In all model specifications, the effect of linguistic enclaves on language proficiency remains statistically insignificant regardless of refugees' gender, age at migration, health status, education, and language course participation. One exception is the positive interaction effect between the linguistic enclave and weekly contacts with German variables, while the main effect of linguistic enclave on language acquisition remains negative and insignificant. A closer examination revealed that the positive association between frequent contacts with Germans and German language proficiency is stronger when refugees reside in

Table 1. Linear regression of language proficiency among refugees.

	Model 1.1		Model 1.2		Model 1.3		Model 1.4	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
<i>Duration of stay</i>	0.06***	(0.01)	0.06***	(0.01)	0.05***	(0.01)	0.05***	(0.01)
Linguistic enclaves								
<i>Initial linguistic enclave, std.</i>	-0.08	(0.09)	-0.12	(0.10)	-0.07	(0.08)	-0.04	(0.12)
<i>x Duration of stay</i>							-0.00	(0.00)
Regional-level confounders								
<i>Initial rural-type county</i>			-0.39	(0.23)	-0.30	(0.18)	-0.30	(0.18)
<i>Initial unemployment rate, std.</i>			0.05	(0.15)	-0.08	(0.11)	-0.08	(0.11)
<i>Initial share of foreigners, std.</i>			0.01	(0.22)	0.16	(0.17)	0.16	(0.17)
<i>Initial population density, std.</i>			-0.01	(0.23)	-0.13	(0.18)	-0.13	(0.18)
Individual-level confounders								
<i>Female</i>					-0.65***	(0.15)	-0.65***	(0.15)
<i>Age at immigration</i>					-0.06***	(0.01)	-0.06***	(0.01)
<i>Single</i>					0.12	(0.12)	0.12	(0.12)
<i>Child aged < 16</i>					-0.46***	(0.14)	-0.46***	(0.14)
<i>Premigration German language proficiency</i>					0.15***	(0.04)	0.15***	(0.04)
<i>Premigration years of education</i>					0.14***	(0.02)	0.14***	(0.02)
<i>Premigration work experience</i>					-0.18	(0.14)	-0.18	(0.14)
<i>Country-of-origin literacy</i>					1.04***	(0.17)	1.03***	(0.17)
<i>Good health</i>					0.28*	(0.13)	0.28*	(0.13)
<i>Shared accommodation</i>					-0.43***	(0.11)	-0.43***	(0.11)
<i>Status of asylum application (ref. pending)</i>								
<i>Approved</i>					0.44**	(0.16)	0.44**	(0.16)
<i>Rejected</i>					-0.41**	(0.15)	-0.41**	(0.15)
<i>Language course participation</i>					1.27***	(0.13)	1.27***	(0.13)
<i>Weekly contact with Germans</i>					0.87***	(0.09)	0.87***	(0.09)
<i>_cons</i>	4.44***	(0.25)	4.54***	(0.34)	2.24***	(0.43)	2.24***	(0.43)
<i>Country of origin FE</i>	YES		YES		YES		YES	
<i>Federal state FE</i>	YES		YES		YES		YES	
<i>Survey year FE</i>	YES		YES		YES		YES	
<i>N observations</i>	5732		5732		5732		5732	

Notes: *p < 0.05, **p < 0.01, ***p < 0.001 (two-sided tests). FE = Fixed effects. All models control for sample fixed effects and missing values indicators. Standard errors are clustered at the person level.

Data source: IAB-BAMF-SOEP Survey of Refugees 2016–2020, weighted.

larger enclaves. In fact, foreigners tend to concentrate in metropolitan areas (Chiswick & Miller, 2004; Semyonov & Raijman, 2021), which is also true for Germany (Heider et al., 2020; Tanis, 2020). Apparently, our measure of linguistic enclaves is generally larger in bigger cities and metropolitan areas. In these areas, opportunities for jobs, social and cultural activities for immigrants are more pronounced and, thus, there are more options for interaction with natives. Hence, our findings imply that refugees assigned to larger enclaves may have also greater access to networks with Germans, reliance on which may be particularly beneficial for German language proficiency.

Robustness checks

We perform a series of robustness checks using our benchmark model (Models 1.3 and 1.4 in Table 1). The results of all the robustness checks are presented in the Online Appendix.

First, we test various specification of enclaves (Table A5), considering country of citizenship instead of country of birth to identify linguistic enclaves (Models 2.3 and 2.4). Alternatively, we measure enclaves based on the share of people with the same reported country of birth (co-ethnics; Models 3.3 and 3.4) or country of nationality (conationals; Models 4.3 and 4.4). Replicating our benchmark models using any specification of enclaves did not alter our conclusions: enclave has no significant effect on German language proficiency.

Second, the use of self-reported proficiency in language skills has become a common practice in studies on language proficiency, although it is often argued that a more objective assessment would be desirable (e.g., van Tubergen & Kalmijn, 2005). Our data provide us with an opportunity to validate the use of self-reported measures of refugees' German language proficiency by comparing self-reports with interviewers' reports. This allows us to reduce the measurement error due to wrong assessment by survey participants. The exact question was "How well could you conduct the interview in German?" with a five-point answer scale ranging from 1, "very well" to 5, "very bad", which we recoded so that higher values denote higher proficiency. The Persons' correlation between our measure of language proficiency and interviewers' report is statistically significant and amounts to 0.627 (weighted results). The altered additive index for refugees' language proficiency ranges from 0 to 16, with a mean (mode) value of 7.9 (8) (weighted results). Cronbach's alpha amounts to 0.836. Replicating our benchmark models are robust to corresponding specification of the dependent variable (Table A6).

Third, we regard various sample restrictions based on asylum application status, country of origin, duration of stay, arrival cohorts, and Federal state (Table A7). This restriction did not show any substantial differences from the original findings in Table 1.

Conclusion

For newly arrived immigrants, acquiring destination language skills is a key element for successful and sustainable integration into the host society. Insufficient destination language proficiency, although not the only responsible factor, often leads to various forms of inequality and discrimination in host societies. Linguistic enclaves have been often named one of the major forces detrimental to destination language acquisition. In particular, previous studies have stressed that living among many coethnics reduces migrants' incentives and exposure to the destination language and hence impairs the learning of the language (Bauer et al., 2005; Chiswick & Miller, 1996; Danzer & Yaman, 2016; Dustmann & van Soest, 2004; Espenshade & Fu, 1997; van Tubergen & Kalmijn, 2005). At the same time, researchers have acknowledged that enclaves may provide an important "stepping stone" in the host country to accommodate those who (for whatever reason) are unable or unwilling to learn the destination language (Mouw & Chavez, 2012). Correspondingly, enclaves may drive immigrants' sorting into and

out of enclaves (Damm, 2009a; Danzer & Yaman, 2016; Lazear, 1999), rendering the negative association between enclaves and language proficiency identified by prior studies to be spurious.

This paper contributes to the social network literature by examining the effect of linguistic enclaves on destination language proficiency. We utilized a unique natural experiment in which refugees are exogenously allocated to counties and their residential mobility across counties is constrained by German authorities. The exogenous assignment and mobility restrictions address the problem of immigrants' self-selection into and out of linguistic enclaves and provide a unique opportunity to estimate the effect of enclaves on German language proficiency from the causal perspective. Moreover, by using longitudinal data and following immigrants over time, we could explore the dynamic impact of linguistic enclaves on language proficiency over duration of stay.

Our results clearly show that living in an area with a large linguistic enclave has no detrimental effect on refugees' German language proficiency: Refugees who are exogenously assigned to counties with larger linguistic enclaves have no worse language skills than those assigned to counties with smaller enclaves. Likewise, the size of linguistic enclave has no significant effect on the rate of German language learning. These results are robust to different specifications of enclaves or sample restrictions. In addition, the effect of linguistic enclave on destination language learning remains insignificant independent on refugees' incentives and ability to learn German language and other sociodemographic variables. Taken together, these findings suggest that the negative relationship between enclaves and immigrant's language skills reported in previous research is spurious—i.e., due to less favorable sorting into communities where a large share of population speaks the same language.

Our findings also have important policy implications. Many policymakers argue that living in an area with a high concentration of people speaking the same language makes immigrants less likely to learn the destination-country language and, hence, slows down their integration process. As a result, policies placing restrictions on where newly arrived immigrants are to settle have been widely implemented in a majority of European countries (Rinne, 2013). Our results clearly reject the assumption of the negative impact of linguistic enclaves underlying the spatial allocation policies. This is in line with recent empirical evidence that enclaves are also not detrimental for immigrants' integration prospects when residential self-selection is appropriately accounted for. Instead, refugees living in locations with many coethnics are more likely to be employed and have higher earnings than those living in locations with fewer coethnics (Damm, 2009b; Martén et al., 2019).

Two limitations of our study call for further research. First, by focusing on a single country, Germany, we are unable to rule out the role of contextual factors that may shape the observed patterns in one direction or another. Thus, replication studies from other countries are important for the generalization of the results. Second, the insignificant effect of linguistic enclaves on destination language proficiency could be attributed to the fact that recent refugees as a group shows much smaller residential concentration than other well-established groups in Europe (e.g., Turkish immigrants) and in the U.S. (e.g., Mexican immigrants). Still, previous research using a related experimental design has shown that with respect to other integration outcomes, such as employment (e.g., Martén et al., 2019) and earnings (e.g., Damm, 2009b), the presumed negative effects of linguistic enclaves have been largely overestimated and deflate when immigrants' sorting is accounted for. Hence, by using the unique natural-experimental design, we add important methodological—and therefore theoretical—value to the predominantly observational knowledge about the importance of linguistic enclaves for immigrants' adjustment.

Notes

1. The literature refers to spatial concentration of immigrants in terms of either ethnicity or language. As discussed below, we refer to spatial concentration in terms of linguistic concentration, i.e., residential areas with a high concentration of immigrants speaking similar languages.

2. Linguistic enclaves could also obstruct integration by reducing connections with the native population (e.g., Portes & Sensenbrenner, 1993).
3. Henceforth, the term “refugees” is used colloquially and includes all persons who move to another country for humanitarian reasons, irrespective of their legal status (e.g., refugee, asylum-seeker, or other humanitarian migrants). Note that when we use the term “immigrants”, we refer to all immigrants, including refugees and other immigrants.
4. Exceptions apply for employees subject to social security contributions with at least 15 weekly working hours and monthly income of approximately 700 €.
5. The information on the county of the assignment is based on the self-reported information on the first place of residence (if moved only once) or the place where the individual resided the longest (if moved more than once) in Germany. We assume that the respondents were assigned to these reported places. For those who reported that they still lived in the first residence place, the current county was considered the county of assignment.
6. Consider a county, with a total population of 100 residents, and newly arrived five Syrians, two Russians, and five Ukrainians. The language proximity between Syria to Russia is 0.10 and between Syria to Ukraine 0.08, while the language proximity between Ukraine to Russia is 0.59. Hence, for the newly arrived Syrian refugee, the measure of the linguistic enclave in this county would be 0.056 ($= (2 \cdot 0.10 + 5 \cdot 0.08 + 5 \cdot 1) / 100$). For the migrants from Ukraine and Russia the measure of the linguistic enclave in this county would be 0.065 ($= (2 \cdot 0.59 + 5 \cdot 1 + 5 \cdot 0.08) / 100$) and 0.055 ($= (5 \cdot 0.59 + 2 \cdot 1 + 5 \cdot 0.10) / 100$), respectively.
7. An additional test showed that a squared term of the duration of stay is negative and significant ($p = 0.036$) whereas the interaction between a squared term of the duration of stay and linguistic enclave is negative but not statistically significant ($p = 0.938$). Hence, we decided in favor of more parsimonious models without a squared term of the duration of stay.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: This article uses the factually anonymous data of the IABBAMF-SOEP Survey of Refugees, waves 1–5 (2016–2020). The IAB-BAMF-SOEP Survey of Refugees in Germany is a representative longitudinal survey conducted jointly by the Institute for Employment Research (IAB) in Nuremberg, the Research Centre on Migration, Integration, and Asylum of the Federal Office for Migration and Refugees (BAMF-FZ) and the German Socio- Economic Panel (SOEP) at the DIW Berlin. Data access was provided via a Scientific Use File supplied by the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB). DOI: [10.5684/soep.iab-bamf-soep-mig.2020](https://doi.org/10.5684/soep.iab-bamf-soep-mig.2020). All documentation concerning the IAB-BAMF-SOEP Survey of Refugees and including questionnaires and data manuals are made available by the FDZ (https://fdz.iab.de/en/FDZ_Individual_Data/iab-bamfsoep.aspx) and DIW (https://www.diw.de/sixcms/detail.php?id?diw_01.c.814095.en https://www.diw.de/sixcms/detail.php?id?diw_01.c.814095.en). Due to the German Data Protection legislation, we cannot make the original data from the IAB-BAMF-SOEP Survey of Refugees or the dataset we generated available. Researchers can however apply for data access via the FDZ or DIW. Requests to access these datasets should be directed to https://fdz.iab.de/en/FDZ_Individual_Data/iab-bamfsoep.aspx. The computer code for the analysis is available at <https://osf.io/xz5g9/>.

Acknowledgements

We would like to thank Sekou Keita for providing helpful advice on the earlier versions of this paper; Thijs Bol, Bram Lancee, Liliya Leopold, Thomas Leopold, and Beate Volker for engaging in insightful discussions at various stages of the research process; and all those who commented on the earlier presentations of this work at the following events: the Spring Meeting of the International Sociological Association (ISA) Research Committee on Social Stratification and Mobility (RC28) in Frankfurt (April 2019), the 34th Annual Conference of the European Society for Population Economics (ESPE) in Bath (June 2019), the 15th IMISCOE Annual Conference in Malmö (June 2019), the Immigrants’ Integration: Educational Opportunities & Life Chances conference in Monte Verità (June 2019), the Seminar in University of Amsterdam (September 2019), the Lunch Seminar at the Institute for Employment Research (November 2019), and European Consortium for Sociological Research annual conference in Amsterdam (July 2022). All authors contributed equally and are listed in alphabetical order.

ORCID

Agnieszka Kanas  <http://orcid.org/0000-0003-0057-2474>
 Yuliya Kosyakova  <http://orcid.org/0000-0002-9621-1755>
 Ehsan Vallizadeh  <http://orcid.org/0000-0002-9367-0374>

References

- Aksoy, C. G., Poutvaara, P., & Schikora, F. (2020). First time around: Local conditions and multi-dimensional integration of refugees. In *SOEP Papers on Multidisciplinary Panel Data Research*. DIW Berlin/SOEP. <https://doi.org/10.2139/ssrn.3738561>
- Åslund, O., & Rooth, D.-O. (2007). Do when and where matter? Initial labour market conditions and immigrant earnings. *The Economic Journal*, 117(518), 422–448. <https://doi.org/10.1111/j.1468-0297.2007.02024.x>
- Bansak, K., Ferwerda, J., Hainmueller, J., Dillon, A., Hangartner, D., Lawrence, D., & Weinstein, J. (2018). Improving refugee integration through data-driven algorithmic assignment. *Science (New York, N.Y.)*, 359(6373), 325–329. <https://doi.org/10.1126/science.aao4408>
- Bauer, T. K., Epstein, G. S., & Gang, I. N. (2005). Enclaves, language, and the location choice of migrants. *Journal of Population Economics*, 18(4), 649–662. <https://doi.org/10.1007/s00148-005-0009-z>
- Borjas, G. J. (1995). Ethnicity, neighborhoods, and human-capital externalities. *The American Economic Review*, 85(3), 365–390.
- Brücker, H., Hauptmann, A., & Jaschke, P. (2020). *Beschränkungen der Wohnortwahl für anerkannte Geflüchtete: Wohnsitzauflagen reduzieren die Chancen auf Arbeitsmarktintegration*. Vol. 3. IAB Kurzbericht.
- Brücker, H., Kosyakova, Y., & Vallizadeh, E. (2020). Has there been a “refugee crisis”? New insights on the recent refugee arrivals in Germany and their integration prospects. *Soziale Welt*, 71(1–2), 24–53. <https://doi.org/10.5771/0038-6073-2020-1-2-24>
- Bruecker, H., Rother, N., & Schupp, J. (2017). IAB-BAMF-SOEP-Befragung von Geflüchteten 2016: Studiendesign, Feldergebnisse sowie Analysen zu schulischer wie beruflicher Qualifikation, Sprachkenntnissen sowie kognitiven Potenzialen. In *IAB- Forschungsbericht, 13/2017*. DIW/SOEP.
- Chiswick, B. R., & Miller, P. W. (1996). Ethnic networks and language proficiency among immigrants. *Journal of Population Economics*, 9(1), 19–35.
- Chiswick, B. R., & Miller, P. W. (2001). A model of destination-language acquisition: Application to male immigrants in Canada. *Demography*, 38(3), 391–409. <https://doi.org/10.2307/3088354>
- Chiswick, B. R., & Miller, P. W. (2004). Where immigrants settle in the United States. *Journal of Comparative Policy Analysis: Research and Practice*, 6(2), 185–197. <https://doi.org/10.1080/1387698042000273479>
- Chiswick, B. R., Lee, Y. L., & Miller, P. W. (2006). Immigrants’ language skills and visa category. *International Migration Review*, 40(2), 419–450. <https://doi.org/10.1111/j.1747-7379.2006.00023.x>
- Damm, A. P. (2009a). Determinants of recent immigrants’ location choices: Quasi-experimental evidence. *Journal of Population Economics*, 22(1), 145–174. <https://doi.org/10.1007/s00148-007-0148-5>
- Damm, A. P. (2009b). Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics*, 27(2), 281–314. <https://doi.org/10.1086/599336>
- Danzer, A. M., & Yaman, F. (2016). Ethnic concentration and language fluency of immigrants: Evidence from the guest-worker placement in Germany. *Journal of Economic Behavior & Organization*, 131(4742), 151–165. <https://doi.org/10.1016/j.jebo.2016.08.012>
- DESTATIS. (2019). *Ausländische Bevölkerung. Ergebnisse des Ausländerzentralregisters 2018. Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 2*. Federal Statistical Office.
- DESTATIS. (2020a). *Statistisches Bundesamt. Ausländer: Kreise, Stichtag, Geschlecht, Ausgewählte Aufenthaltstitel, Staatsangehörigkeit. Code: 12521-0043*. Statistisches Bundesamt (DESTATIS, Federal Statistical Office). <https://www-genesis.destatis.de/>
- DESTATIS. (2020b). *Statistisches Bundesamt. Bevölkerung: Kreise, Stichtag. Code: 12411-0015*. Statistisches Bundesamt (DESTATIS, Federal Statistical Office). <https://www-genesis.destatis.de/>
- Dustmann, C., & van Soest, A. (2004). An analysis of speaking fluency of immigrants using ordered response models with classification errors. *Journal of Business & Economic Statistics*, 22(3), 312–321. <https://doi.org/10.1198/073500104000000163>
- Espenshade, T. J., & Fu, H. (1997). An analysis of English-language proficiency among U.S. immigrants. *American Sociological Review*, 62(2), 288–305. <https://doi.org/10.2307/2657305>
- Fasani, F., Frattini, T., & Minale, L. (2022). (The struggle for) refugee integration into the labour market: Evidence from Europe. *Journal of Economic Geography*, 22(2), 351–393. <https://doi.org/10.1093/jeg/lbab011>
- Geis, W., & Orth, A. K. (2016). *Flüchtlinge regional besser verteilen. Ausgangslage und Ansatzpunkte für einen neuen Verteilungsmechanismus*. Köln: Institut der deutschen Wirtschaft Köln.
- Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, 177, 263–277. <https://doi.org/10.1016/j.cognition.2018.04.007>
- Heider, B., Stroms, P., Koch, J., & Siedentop, S. (2020). Where do immigrants move in Germany? The role of international migration in regional disparities in population development. *Population, Space and Place*, 26(8), 1–19. <https://doi.org/10.1002/psp.2363>
- Kanas, A., & Kosyakova, Y. (forthcoming). Greater local supply of language courses improves refugees’ labor market integration. *European Societies*. <https://doi.org/10.1080/14616696.2022.2096915>
- Kanas, A., & Müller, K. (2021). Immigrant women’s economic outcomes in Europe: The importance of religion and traditional gender roles. *International Migration Review*, 55(4), 1231–1264. <https://doi.org/10.1177/01979183211008867>

- Kanas, A., & Steinmetz, S. (2021). Economic outcomes of immigrants with different migration motives: The role of labour market policies. *European Sociological Review*, 37(3), 449–464. <https://doi.org/10.1093/esr/jcaa058>
- Kónya, I. (2007). Optimal immigration and cultural assimilation. *Journal of Labor Economics*, 25(2), 367–391. <https://doi.org/10.1086/511378>
- Kosyakova, Y., & Kogan, I. (2022). Labor market situation of refugees in Europe: The role of individual and contextual factors. *Frontiers in Political Science*, 4. <https://doi.org/10.3389/fpos.2022.977764>
- Kosyakova, Y., & Kulic, N. (2022). Kinship, inter- and intraethnic social networks and refugees' division of housework. *Journal of Family Research*, 34(2), 802–822. <https://doi.org/10.20377/jfr-783>
- Kosyakova, Y., & Brenzel, H. (2020). The role of length of asylum procedure and legal status in the labour market integration of refugees in Germany. *Soziale Welt*, 71(1–2), 123–159. <https://doi.org/10.5771/0038-6073-2020-1-2-123>
- Kosyakova, Y., Kristen, C., & Spörlein, C. (2022). The dynamics of recent refugees' language acquisition: How do their pathways compare to those of other new immigrants? *Journal of Ethnic and Migration Studies*, 48(5), 989–1012. <https://doi.org/10.1080/1369183X.2021.1988845>
- Kroh, M., Kühne, S., Jacobsen, J., Siegert, M., & Siegers, R. (2017). *Sampling, nonresponse, and integrated weighting of the 2016 IAB-BAMF-SOEP survey of refugees (M3/M4) – revised version*. DIW/SOEP.
- Kühne, S., Jacobsen, J., & Kroh, M. (2019). *Sampling in times of high immigration: The survey process of the IAB-BAMF-SOEP survey of refugees*. Insights from the Field, pp. 1–9. <https://doi.org/10.13094/SMIF-2019-00005>
- Lastrapes, W. D., & Lebesmuehlbacher, T. (2020). Asylum seekers and house prices: Evidence from the United Kingdom. *Journal of Housing Economics*, 49, 101712. <https://doi.org/10.1016/j.jhe.2020.101712>
- Lazear, E. P. (1999). Culture and language. *Journal of Political Economy*, 107(S6), S95–S126. <https://doi.org/10.1086/250105>
- Martén, L., Hainmueller, J., & Hangartner, D. (2019). Ethnic networks can foster the economic integration of refugees. *Proceedings of the National Academy of Sciences of the United States of America*, 116(33), 16280–16285. <https://doi.org/10.1073/pnas.1820345116>
- Melitz, J., & Toubal, F. (2014). Native language, spoken language, translation and trade. *Journal of International Economics*, 93(2), 351–363. <https://doi.org/10.1016/j.jinteco.2014.04.004>
- Monteiro, S. (2021). Cultural assimilation: Learning and sorting. *Review of Economic Analysis*, 13(2), 115–156. <https://doi.org/10.15353/rea.v13i2.4045>
- Mouw, T., & Chavez, S. (2012). Occupational linguistic niches and the wage growth of Latino immigrants. *Social Forces*, 91(2), 423–452. <https://doi.org/10.1093/sf/sos186>
- Portes, A., & Sensenbrenner, J. (1993). Embeddedness and immigration: Notes on the social determinants of economic action. *American Journal of Sociology*, 98(6), 1320–1350. <https://doi.org/10.1086/230191>
- Rinne, U. (2013). The evaluation of immigration policies. In *International handbook on the economics of migration*, Amelie F. Constant & Klaus F. Zimmermann (Eds.), (Vol. 6369, pp. 530–552). Edward Elgar Publishing. <https://doi.org/10.4337/9781782546078.00038>
- Roder, A., & Muhlau, P. (2014). Are they acculturating? Europe's immigrants and gender egalitarianism. *Social Forces*, 92(3), 899–928. <https://doi.org/10.1093/sf/sot126>
- Schneider, J. (2012). The organisation of asylum and migration policies in Germany. In *Study of the German National Contact Point for the European Migration Network (EMN) Working*.
- Semyonov, M., & Rajjman, R. (2021). Immigration and immigrants in European countries. In *Companion to urban and regional studies*, Anthony M. Orum, Javier Ruiz-Tagle, Serena Vicari Haddock (Eds.), (pp. 420–449). Wiley. <https://doi.org/10.1002/9781119316916.ch20>
- Stevens, G. (1999). Age at immigration and second language proficiency among foreign-born adults. *Language in Society*, 28(04), 555–578. <https://doi.org/10.1017/S0047404599004030>
- Tanis, K. (2020). Regional distribution and location choices of immigrants in Germany. *Regional Studies*, 54(4), 483–494. <https://doi.org/10.1080/00343404.2018.1490015>
- Triandafyllidou, A. (2018). A “refugee crisis” unfolding: “Real” events and their interpretation in media and political debates. *Journal of Immigrant & Refugee Studies*, 16(1–2), 198–216. <https://doi.org/10.1080/15562948.2017.1309089>
- van Tubergen, F. (2010). Determinants of second language proficiency among refugees in the Netherlands. *Social Forces*, 89(2), 515–534. <https://doi.org/10.1353/sof.2010.0092>
- van Tubergen, F., & Kalmijn, M. (2005). Destination language proficiency in cross national perspective: A study of immigrant groups in nine Western countries. *American Journal of Sociology*, 110(5), 1412–1457. <https://doi.org/10.1086/428931>