JOHANNES KEPLER
UNIVERSITY LINZ

DEPARTMENT OF ECONOMICS
JOHANNES KEPLER UNIVERSITY OF LINZ

Do co-ethnic commuters disseminate labor market information? Evidence from geocoded register data

by

Johan KLAESSON
Özge ÖNER
Dieter PENNERSTORFER

Working Paper No. 2316
December 2023

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

dieter.pennerstorfer@jku.at

Do co-ethnic commuters disseminate labor market information? Evidence from geocoded register data*

Johan Klaesson[†] Özge Öner[‡] Dieter Pennerstorfer[§]

December 12, 2023

Abstract

This article provides causal evidence of the significant role ethnic networks play in facilitating labor market integration by reducing information frictions. Using full population geocoded employer-employee matched Swedish register data, we investigate how co-ethnic commuters can influence the work location of immigrants for their initial employment. We argue that these ethnic peers transmit job specific information from their places of work to fellow ethnic peers within the same residential neighborhood who seek jobs. We find that a new immigrant's likelihood of securing their first job at a certain location increases with the presence of co-ethnic commuters from their residential neighborhood: Each additional commuter of the same ethnic network increases the probability of finding employment in a specific neighborhood by 2.3 %. This effect is more pronounced for women, co-ethnic commuters with similar education levels, and immigrants who land their first jobs in larger firms.

Keywords: Co-ethnic commuters, information frictions, ethnic networks, labor market integration, ethnic enclaves

JEL codes: F22, J61, J64, O18, R23

*We thank late Edward Lazear, David Albouy, Johannes Hagen, Shade Shutters, José Lobo, conference participants of the 13th Meeting of the Urban Economics Association in New York (USA) and the Annual Meeting of the Austrian Economic Association in Graz (Austria), and workshop and seminar participants at the University of Cambridge (UK), the University of Birmingham (UK), the Max-Planck-Institute in Munich (Germany), Wageningen University (Netherlands), Jönköping International Business School (Sweden), the University of Innsbruck (Austria), the University of Vienna (Austria), and the University of Linz (Austria) for helpful comments on earlier versions of this paper.

[†]Centre for Entrepreneurship and Spatial Economics, Jönköping International Business School (JIBS)

[‡]University of Cambridge, Department of Land Economy and Centre for Entrepreneurship and Spatial Economics (CEnSE)

[§]Department of Economics, Johannes Kepler University Linz.

1 Introduction

The influx of immigrants and refugees into developed nations often sparks debates about their integration into society. One critical aspect of this discourse centers on the formation of ethnic enclaves,[1] which may pose challenges to integration. Responding to these concerns, Denmark (Damm, 2009, 2014), Sweden (Edin et al., 2003; Aslund et al., 2011; Lindgren et al., 2022; Andersson et al., 2021), Switzerland (Auer, 2018; Marten et al., 2019) and other countries have adopted policies that regulate refugees' settlement patterns. These policies, aiming to disperse immigrant populations, are not new. Historical precedents include Germany's approach to location assignment for guest workers in the 1960s and 1970s (Danzer et al., 2022) and the settlement policy for ethnic German immigrants from the late 1980s onwards (Chakraborty et al., 2019). The societal impact of these policies is considerable, fueling ongoing debates. At the heart of these discussions is whether ethnic enclaves facilitate or impede the integration of immigrants, particularly concerning their economic integration into the labor market.

The empirical studies examining ethnic enclaves' impact on labor market integration yield mixed results.[2] Cutler and Glaeser (1997) observe negative effects for African-Americans in the U. S., whereas Collins and Margo (2000) find no such impact in earlier periods. In Europe, however, Edin et al. (2003) and Damm (2009) identify positive effects of immigrant enclaves on labor market integration, while Xie and Gough (2011) report negative consequences. Studies like Marten et al. (2019) and Battisti et al. (2022) suggest initial positive effects for refugees living with co-nationals, but these benefits reduce (Marten et al., 2019) or disappear over time (Battisti et al., 2022). Damm (2014), focusing on enclave size, finds no

[1] The co-location of immigrants was first documented in a seminal study by Bartel (1989) and confirmed by a large number of empirical articles (see, among others, Åslund, 2005; Beine et al., 2011; Damm, 2009; Gross and Schmitt, 2003; Nowotny and Pennerstorfer, 2019; Pedersen et al., 2008; Zavodny, 1999; Zorlu and Mulder, 2008). Several articles have formulated hypotheses explaining migrant concentrations theoretically (see, e.g., Carrington et al., 1996; Chiswick and Miller, 2005; Gross and Schmitt, 2003; Massey et al., 1993).

[2] For a discussion of neighborhood effects in general (i.e., outside the context of international migration and enclaves), see Chyn and Katz (2021) for an excellent survey and the literature cited there for more details.

significant effects but highlights the importance of enclave quality. She notes that skill levels and employment rates within immigrant communities facilitate labor market integration, a conclusion supported by Klaesson et al. (2021). However, Kristiansen et al. (2022) find no significant correlation with the employment rates of ethnic peers. Andersson et al. (2014, 2019) document gender specific variations in these effects, while Klaesson and Oner (2021) report mixed results based on immigrant group and type of economic activity (employment or entrepreneurship). In summary, despite substantial research on the effects of enclaves on labor market integration, the empirical literature is still far from reaching a consensus.

The complexity and contextual nature of neighborhood effects make generalizations challenging (Katz et al., 2001; Kling et al., 2007; Galster, 2012). When investigating enclave effects, it is often difficult to identify specific mechanisms, resulting in a net effect that encompasses various channels. Some mechanisms can facilitate economic integration, while others can hinder it. Which of the two effects predominates depends on the specific context, which could explain the different and contradictory findings.

The mechanism most emphasized in the literature on the influence of ethnic enclaves on labor market integration is the role of information dissemination within these communities: Immigrants with common ethnic or cultural origins living in the same neighborhood often form social (ethnic) networks that facilitate the sharing of job information (Bertrand et al., 2000; Damm, 2009; Battisti et al., 2022). Through these networks, information about job vacancies or job referrals can be disseminated.[3] However, empirical evidence directly attributing labor market success to this information channel is notably limited because of data and identification challenges.

There are two alternative approaches presented in the literature to isolate this information channel. The first approach utilizes survey data that contain direct information about the channel through which employees find jobs. In the context of ethnic enclaves, this type of data is used by Dustmann et al. (2016) and Battisti et al. (2022). Their findings indicate a

[3]See Granovetter's (1973) seminal work on the importance of "weak ties" within social networks to disseminate information.

higher likelihood of immigrants securing their first job through personal contacts when the presence of co-nationals in the firm is higher (Dustmann et al., 2016) or the enclave size in the district of residence is larger (Battisti et al., 2022). Beyond the challenge of reconciling revealed versus stated preferences, surveys often suffer from a constrained sample size. This limitation narrows the analytical scope when examining specific subsamples, restricting the extent of inference that can be drawn from such data.

The second approach uses data on the composition of both the network and the workplace. The underlying hypothesis is that members of a network have better access to job related information about their workplace (e.g., the establishment itself or the neighborhood it is located), which they can then pass on to fellow network members seeking jobs. The higher the number of members of a network working in a particular place, the easier it gets for the job seekers in the same network to access information on that workplace, consequently increasing the likelihood of finding a job there. Only a handful of papers rely on this approach: Bayer et al. (2008) and Hellerstein et al. (2011) analyze cross-sections of U. S. employee data that contain information on both place of residence and place of work. Bayer et al. (2008) discover that individuals from the same census block are more likely to be employed in the same location (i.e., census block) compared to those from adjacent blocks. Similarly, Hellerstein et al. (2011) observe clustering of workers from the same neighborhoods in specific establishments.

A limitation of such approach is that the positive correlation between an individual's place of work and place of residence cannot be interpreted as evidence of neighborhood networks disseminating job information. Such a correlation could be symptomatic of reverse causality, where information about housing shared among colleagues leads to residential clustering. Moreover, individual characteristics could drive simultaneous selection into specific neighborhoods and workplaces, suggesting that observed labor market outcomes might not derive directly from network-driven information spillovers but rather from neighborhood choice. Bayer et al. (2008) confront this potential confound by incorporating data on residential tenure and prior labor market conditions, often finding stronger correlations in scenarios suggestive of active sharing of job related information. Nonetheless, the degree to which this correlation can be ascribed to an exchange of labor market information rather than reverse causality remains unresolved.

We contribute to the literature on ethnic enclaves and labor market integration by applying a similar dyad analysis that links residential and workplace location. However, we are able to isolate the causal impact of information spillovers from residential ethnic networks on the location of the found job, a novelty in this domain. Leveraging geocoded, matched employer-employee register data for the entire population of Stockholm metropolitan area spanning from 1992 to 2013, we concentrate on two refugee immigration waves stemming from armed conflicts.[4] Richness of the data encompassing details on ethnic background, residential address and workplace enables us to compute the commuter flows of ethnic peers and the broader population across all neighborhoods annually. This approach allows for an examination of immigrants at a fine spatial resolution regarding their places of residence and employment, while capitalizing on the longitudinal nature of the data.

We estimate the likelihood of an individual taking up their first job at a specific location against all possible alternatives in the pertinent local labor market. This likelihood is determined by the prevalence of co-ethnic commuters living in the same neighborhood of residence prior to the transition into the labor market. For job seekers, information about a potential workplace is relayed by co-ethnic commuters who already work there. This analysis carries two consequential implications: Firstly, it enables us to identify a distinct mechanism by which ethnic networks aid labor market integration—specifically through the channel of information transfer, given that other influences are tied to the place of residence. Secondly, by accounting for individual heterogeneity, we lay the groundwork for a causal interpretation of our findings.

Our results show that a new immigrant's likelihood of securing their first job at a partic-

[4] We study the immigration waves following the Bosnian War (1992–1995) and the Second Iraq War (Third Gulf War, 2003–2011).

ular workplace increases with the number of co-ethnic commuters from their neighborhood. This effect is not only statistically significant but also quantitatively substantial: We find that each additional commuter from the same ethnic background increases the chance of employment at a specific job location by 2.3 % in our main specification. The effect is more pronounced for women, co-ethnic commuters with similar education levels, and immigrants securing their initial employment in larger firms. The impact of co-ethnic commuters remains constant, regardless of whether immigrants have changed their place of residence before obtaining their first job, how old they are, or how quickly they transition into employment. This indicates that ethnic networks play a critical role in alleviating information frictions, thereby smoothing the pathway for new immigrants into the workforce.

The remainder of this article is organized as follows: Section 2 describes the institutional and historical background of immigration to Sweden. Section 3 discusses the research design and identification strategy and describes the data used for the empirical analysis. Section 4 presents the main results and Section 5 shows the robustness of our findings across different subgroups and model designs. Section 6 concludes.

2 Institutional background

The literature suggests that immigrants, by virtue of living in close proximity and sharing ethnic or cultural ties, form networks that act as channels for information exchange. These networks not only inform new immigrants about employment opportunities, such as job openings, but also allow network members to introduce the newly arrived immigrants to potential employers, for instance, through job referrals. Understanding this process is critical, as it has profound implications for urban planning and the integration of immigrants within the local labor markets, both of which can be influenced by policy design.

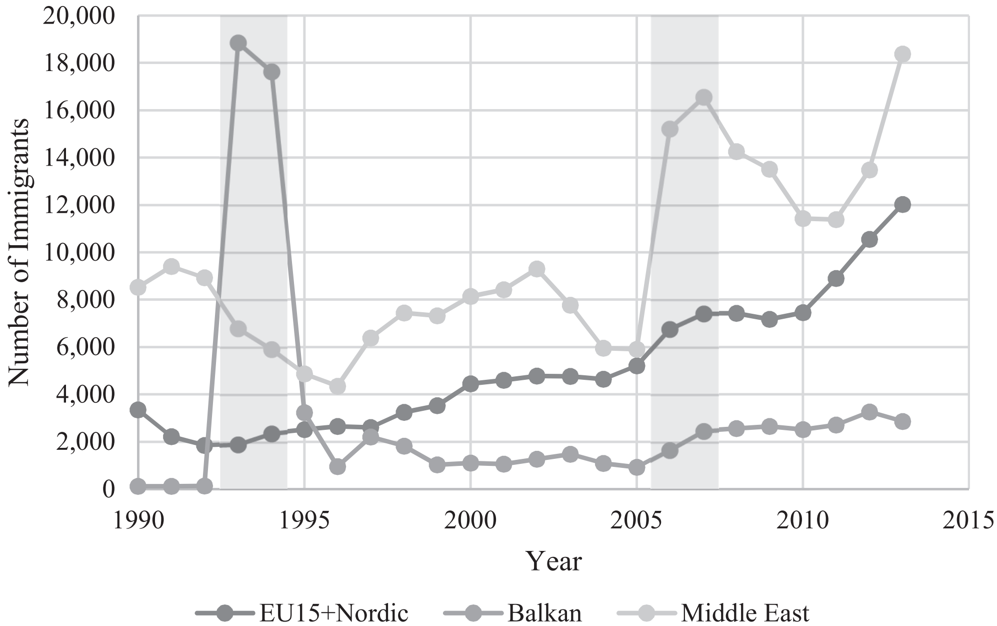
What makes Sweden an ideal case to study ethnic enclave effects—in addition to the availability of high-quality register data, see Section 3—is the country's immigration history.

Over the past fifty years, Sweden has witnessed a substantial increase and transformation in its immigration patterns. In 1950, foreign-born individuals constituted less than three percent of Sweden's population, predominantly originating from other Nordic countries. However, by the close of 2017, this share increased to 1.8 million, accounting for approximately 18 percent of the population, with the majority originating from countries outside Europe (Andersson, 2021).

The rapid increase in refugee numbers over the last 20 years makes the country particularly relevant for our empirical design. Sweden has experienced several migration waves during this period following exogenous shocks, two of which we investigate in our analysis: Immigrants from the Balkans who arrived between 1993 and 1994, and those from the Middle East between 2005 and 2006. The conflicts in these regions, notably the Bosnian War (1992–1995) as part of the broader Yugoslavian wars, and the Second Iraq War (2003–2011), precipitated a significant rise in refugee migration to Sweden. Figure 1 shows the migration patterns to Sweden of these two immigrant groups. Focusing on these two particular waves of immigration allows us to identify a cohort that is predominantly made up of refugees and represents relatively homogeneous groups—a crucial feature in the context of enclaves. While data constraints prevent us from indentifying the exact national origins of these immigrants at the neighborhood level, aggregated figures indicate that the entirety of Balkan refugees to Sweden between 1992 and 1993 originated from Bosnia and Herzegovina, and those from the Middle East between 2005 and 2006 primarily migrated from Iraq.

Concerning the spatial distribution of these refugees, Sweden implemented a centralized placement program from the mid-1980s to mid-1994, encompassing the arrival period of the Balkan immigrants. This program, administered by the Migration Board, allocated refugees to municipalities in a manner exogenous to individual characteristics. A previous study by Andersson et al. (2021) show that the characteristics of these immigrants and municipal characteristics are by and large orthogonal to each other. However, this centralized placement policy did not dictate the intra-city spatial sorting, thereby limiting its usefulness

Figure 1: Annual inflow of immigrants from three regions to Sweden between 1990 and 2015



Data source: Statistics Sweden, figure made by authors. Balkan countries include Albania, Kosovo, Bosnia and Herzegovina, Macedonia, Montenegro, and Serbia. The Middle East includes Bahrain, Egypt, the United Arab Emirates, Iraq, Iran, Israel, Yemen, Jordan, Kuwait, Lebanon, Oman, Qatar, Saudi Arabia, Syria, and Turkey. The EU15+Nordic include Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Iceland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, and the United Kingdom.

for analyzing neighborhood specific patterns. While no central placement policy was active during the influx of Middle Eastern immigrants, the pre-existing Iraqi population was too small to enable self-selection of new refugees into established ethnic enclaves within Swedish cities, given the absence of such enclaves at the onset of their arrival. Andersson et al. (2021) use the two immigration spikes as a part of their identification strategy, because these are exogenous immigration shocks similar to our empirical design.

In our study, we define both neighborhoods of residence and workplace locations as grid cells of 1 km $\times$ 1 km size. We restrict the analysis to the Stockholm local labor market region. To define local labor markets we follow Sweden's Growth Analysis Agency that consolidates Sweden's 290 municipalities into 72 functional regions based on commuting patterns. Therefore, municipalities within the same functional region are considered part of a local labor market area, with only small commuter flows between different labor markets. The Stockholm local labor market area comprises 36 municipalities, as shown in Figure A.1 in the Appendix.

It is also important to mention the urban structure in this local labor market area. Sweden is known to consist of monocentric municipalities, which allows us to identify a central business district for each municipality and for the entire labor market region. Previous work has shown that the accessibility of the center of a region by different means of commuting correlates strongly with the Euclidean distance to the center, so that both distance measures can be used interchangeably (Johansson et al., 2002). As outlined later in the empirical design (see Section 3), we control for the intensity of commuters between any two locations. Therefore, if there exists an urban infrastructure that makes some places more connected to one another, this is captured by the total number of commuters in a dyad (i.e., from a neighborhood of residence to a place of work). Immigrant communities disproportionately live in neighbourhoods located at the peripheries of the municipalities. The buildings in these neighborhoods were often constructed between 1965 and 1974 as part of an ambitious constructed program known as the Million Program (Miljonprogrammet). Therefore, it is important to control for the proximity of a residential area to the urban core in order to avoid a systematic bias that could be caused by these settlement patterns.

3 Empirical framework

3.1 Research design

The basic idea of the research design is that we investigate *where* (in which neighborhood) immigrants find their first job, but not *whether* or *when* they are integrated into the labor market. We therefore study individual immigrants from their arrival in Sweden until they have found their first job and analyze in which neighborhood they start to work. We relate an immigrant's first work location to the commuter flows originating from the immigrant's neighborhood of residence. We interpret ethnic peers living in the same neighborhood as an indicator of an immigrant's social (ethnic) network[5] and argue that co-ethnic commuters pass information between the neighborhoods where they work and the neighborhoods where they live. Co-ethnic commuters can provide new immigrants with information about jobs in their place of work (e.g., job vacancies). In addition, ethnic peers can provide the company with information about the new immigrant and, for example, make a job referral. Ethnic peers can therefore facilitate the transfer of labor market related information in both directions—from the new immigrant to the workplace and vice versa. Co-ethnic commuters can therefore help to integrate immigrants into the labor market by reducing information frictions (commonly labeled as "information channel", see Bertrand et al., 2000).

Although there are many mechanisms how ethnic enclaves can influence a new immigrant's labor market integration, our research design allows us to isolate this information channel. Alternative mechanisms discussed in the literature comprise the following channels: Living in an enclave can reduce incentives to acquire host country specific human capital (e.g., language skills, see Lazear, 1999) and thus labor market success. New immigrants can benefit or be harmed by spillovers of neighbors' human capital, depending on the quality of the enclave (human capital externalities, see Borjas, 1995). Furthermore, social norms may evolve within an ethnic network (Bertrand et al., 2000) that can influence labor market success directly (e.g., work ethics) and indirectly (e.g., attitudes towards education). It is important to note that all these alternative mechanisms of how ethnic peers in the neighborhood can affect the labor market integration of new immigrants depend only on the neighborhood in which the immigrant lives (i.e., they are neighborhood and not dyad specific).

In this article, we exploit the variation of co-ethnic commuters (living in the same neighborhood) to different workplace locations by comparing the labor market success of a given

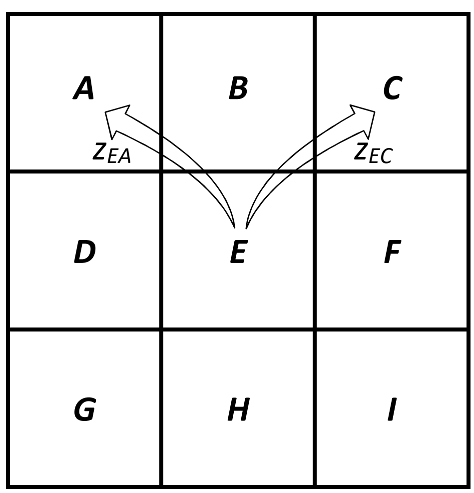
[5]In contrast to family networks (see, e.g., Kramarz and Skans, 2014), social networks are difficult to observe directly. It is therefore common to use people who live in the same neighborhood as a proxy variable for social networks (e.g., Bertrand et al., 2000; Damm, 2009; Edin et al., 2003; Nowotny and Pennerstorfer, 2019). See Bayer et al. (2008) and the references therein for a discussion of the extent to which individual social networks are local in a spatial sense.

immigrant in different (potential) job locations. This research design allows us to control for all channels that operate through the place of residence and we can therefore isolate information spillovers, since information is the only channel that differs across potential workplace locations. Moreover, this approach allows us to control for (unobserved) individual heterogeneity, which enables a causal interpretation of the results. Our research design differs from Bayer et al. (2008) and Hellerstein et al. (2011) because we use panel data, which allows us to identify senders (already employed ethnic peers) and receivers (new and unemployed immigrants) of information. In this way, we can isolate the effect of neighbors sharing labor market information from co-workers sharing residential information (reverse causality).

By investigating information frictions on labor market outcomes, our work is related to the existing literature on job search. It is important to recognize that individuals cannot freely choose among all alternative workplace locations in their choice set, because "job seeker cannot choose an employer who does not offer him a job" (Horowitz, 1991, p. 1237). Standard models of job search view the job search as a sequential process in which each individual decides how long to continue the search. The search process ends as soon as a satisfactory offer is received. The decision to end the search depends on the search costs, the reservation wage, and the distribution from which the potential job types are drawn (Faggian, 2021). In our empirical analysis, we control for the reservation wage by including individual-time specific fixed effects and the length of the commute. The costs associated with the search process are influenced by the quality of the ethnic social network, because information about a particular workplace location is acquired by ethnic peers who are employed there, which reduces search costs. While we only observe market equilibria (i.e., accepted job offers), a potentially positive effect of ethnic peers is likely to stem from an increase in the job offer arrival rate rather than a higher probability of a job offer being accepted.

We first describe the intuition of our identification strategy and then provide a formal treatment of this issue. The identification strategy is illustrated in Figure 2 below for a (reduced) set of potential workplace locations[6] $W = \{A, B, ..., I\}$. Whether an individual i living in neighborhood E finds a job in, for example, neighborhood C, depends on (i) the individual's characteristics (e.g., education, sex, and ability), (ii) on the peculiarities of the residential neighborhood E (like size and quality of the ethnic enclave), (iii) on the characteristics of the potential workplace location C, $\boldsymbol{X}_C$ (e.g., the number of jobs available at that location), and (iv) on dyad specific variables depending on both residential and workplace location, $\boldsymbol{Z}_{EC}$. Dyad specific variables include indicators of connectivity between the two neighborhoods E and C, like the distance or the overall number of commuters between these locations. Additionally—and most importantly—this type of variables also includes the number of ethnic peers commuting between neighborhoods E and C, as an indicator of the information spillovers between these two locations disseminated through the ethnic network. If more ethnic peers living in neighborhood E commute to workplace location C, we expect that a prospective immigrant worker (who also lives in neighborhood E) will receive more labor market relevant information about potential jobs in workplace location C, and therefore be more likely to find a job there, ceteris paribus.

Figure 2: Illustration of identification strategy



Notes: The grid cells illustrate neighborhoods. The term $\boldsymbol{Z}_{EA}$ ($\boldsymbol{Z}_{EC}$) summarizes variables that depend on the neighborhood of residence E and the potential workplace location A (C), such as the distance or commuter flows between these neighborhoods.

[6] We use the terms neighborhood and location interchangeably. We will define these terms in Section 3.3.

We restrict the analysis to the year in which the immigrant finds the first job. Consequently, we examine whether the immigrant is more likely to find the first job in a particular workplace location (e.g., in neighborhood C) compared to another location (e.g., neighborhood A). For both potential workplace locations A and C, the characteristics of the immigrant and their residential neighborhood E are the same. Comparing the probability of finding a job in different workplace locations therefore allows us to control for all (observed and unobserved) individual-level and neighborhood-level characteristics.

Formally, we estimate the probability p_{irwt} that a person i living in neighborhood (of residence) r finds the first job at workplace location w at time t (i.e., $y_{irwt} = 1$), summarized by the following equation:

$$p_{irwt} \equiv Pr[y_{irwt} = 1 | \alpha_{ir,t-1}, \boldsymbol{Z}_{rw,t-1}, \boldsymbol{X}_{w,t-1}] = F(\alpha_{ir,t-1} + \boldsymbol{Z}_{rw,t-1}\beta^d + \boldsymbol{X}_{w,t-1}\beta^w), \quad (1)$$

$\alpha_{ir,t-1}$ denotes an individual-time specific fixed effect that captures individual and residential characteristics (both time varying and time fixed). The matrix $\boldsymbol{Z}_{rw,t-1}$ captures dyad specific variables that vary between both the neighborhood of residence r of individual i and the potential workplace location w. This category of variables includes the distance between neighborhoods of residence r and work w, the overall number of commuters, as well as the number of co-ethnic commuters. The matrix $\boldsymbol{X}_{w,t-1}$ summarizes workplace location specific variables, and β^d and β^w are the respective vectors of parameters to be estimated. The explanatory variables are lagged by one year to ensure that these characteristics are observed before the start of the first employment and to rule out reverse causality. Regression Equation (1) can be estimated by a logit model, with $F(\cdot)$ as the cumulative distribution function of the logistic distribution (see Cameron and Trivedi, 2005, p. 465 ff.).

Including the individual-time specific fixed effects $\alpha_{ir,t-1}$ is crucial for our analysis and has important implications: In the period between an immigrant's arrival to Sweden and the year prior to finding a job there is no variation in outcomes, i.e., the person does not find a job in any location in these years. The individual-time specific fixed effects perfectly predict the outcomes in these years, and the respective observations drop out of the estimation, leaving only observations in the sample if an individual finds their first job in a given year. We therefore model the workplace location of the first job, conditional on finding the first job in the respective year, which can be captured by a conditional logit (CL) model (McFadden, 1974). The regression equation can be stated as:

$$\begin{aligned} p_{irwt} &\equiv Pr[y_{irwt} = 1 | \boldsymbol{Z}_{rw,t-1}, \boldsymbol{X}_{w,t-1}, \sum_{v \in W^i} y_{irvt} = 1] = \\ &= \frac{\exp(\alpha_{ir,t-1} + \boldsymbol{Z}'_{rw,t-1}\beta^d + \boldsymbol{X}'_{w,t-1}\beta^w)}{\sum_{v \in W^i} \exp(\alpha_{ir,t-1} + \boldsymbol{Z}'_{rv,t-1}\beta^d + \boldsymbol{X}'_{v,t-1}\beta^w)} = \\ &= \frac{\exp(\boldsymbol{Z}'_{rw,t-1}\beta^d + \boldsymbol{X}'_{w,t-1}\beta^w)}{\sum_{v \in W^i} \exp(\boldsymbol{Z}'_{rv,t-1}\beta^d + \boldsymbol{X}'_{v,t-1}\beta^w)} \end{aligned} \tag{2}$$

with W^i as the choice set, i.e., the set of potential workplace locations. The estimated probabilities range from zero to one and sum to one across all workplace locations. As can be seen in Equation (2), the fixed effects $\alpha_{ir,t-1}$ drop out of the equation: While we do not explicitly estimate the fixed effects $\alpha_{ir,t-1}$, the model nevertheless controls for individual-time specific characteristics that do not vary across alternative workplace locations.

3.2 Identification and interpretation

We are interested in finding causal evidence of information spillovers, and are therefore particularly interested in the estimated parameter for the number of co-ethnic commuters. Our identification strategy is to compare the labor market success of an immigrant in different (potential) job locations. We thus offer an alternative approach to the existing empirical literature that usually compares the labor market integration of different immigrants assigned to different (more or less segregated) neighborhoods. Our approach allows us to control for (observed and unobserved) individual and neighborhood characteristics in the framework of a conditional logit (CL) model and is thus not plagued by an omitted variables bias due to unobserved individual characteristics.[7]

An omitted variables bias may nevertheless occur if important dyad specific variables or relevant variables at the workplace location level are left out. The parameter estimates on co-ethnic commuters will be inconsistent if parts of the immigrant group have similar preferences regarding both place of residence and place of work, resulting in similar commuting patterns of old and new immigrants. An immigrant group specific transportation network, for example, could constitute a preference for immigrants to find both a place of residence and a work location along this transportation network, resulting in a correlation between the commuting routes of new immigrants and ethnic peers that is independent of information spillovers. However, since Stockholm has a dense, efficient and heavily subsidized (and thus quite cheap) public transport system, this argument seems unlikely in the present context.[8]

We find it plausible to interpret the effect of co-ethnic commuters as information spillovers, since the individual-time dependent fixed effects included in all model specifications control for all (residential) neighborhood heterogeneity that does not vary across alternatives (i.e., across potential workplace locations). In this way, we take into account all the other potential enclave effects discussed in Section 3.1, like enclave or immigrant group specific social norms, reduced incentives to acquire host-country specific skills, or human capital externalities, since all these mechanisms operate at the neighborhood level. While we are able to isolate the information channel from other mechanisms how enclaves may affect immigrants' labor market success, we acknowledge that ethnic enclaves may also provide general (i.e., not location specific) information about the labor market to newly arrived immigrants, which we cannot account for due to our research design. Our results will thus be rather conservative

[7]It is a common problem in empirical research on the effects of ethnic enclaves (or neighborhood effects in general) that (unobserved) individual heterogeneity is correlated with neighborhood characteristics (labeled as "sorting by ability"), leading to inconsistent results of neighborhood characteristics due to an omitted variables bias. Edin et al. (2003) and Damm (2009) convincingly address this issue by evaluating the random assignment of refugees to neighborhoods in the context of spatial dispersal policies on refugees in Sweden and Denmark, respectively.

[8]Although we do not directly observe public transportation infrastructure, a high level of connectivity between two neighborhoods should be reflected in the Euclidean distance and total number of commuters, variables that are included in all model specifications.

estimates of all information spillovers.

Interpreting the results as evidence of the information channel would be invalid if prospective workers gained utility from commuting together with their ethnic peers.[9] We think that this is an unlikely mechanism, because starting and end time of workdays may differ between ethnic peers, impeding to commute together even if the commuting routes are similar. Furthermore, the gains of commuting together are rather small when using public transportation, which is the main commuting mode in the Stockholm metropolitan region.[10]

Regarding the heterogeneity of workplace locations, we choose a parsimonious model for the main specification and include only a small number of variables indicating the attractiveness of a workplace location (like the number of jobs). In the sensitivity analysis, we control for the heterogeneity of the workplace locations to a greater extent by including variables on industry structure, labor market dynamics, and even location specific dummy variables.

3.3 Data

In our empirical analysis, we can draw on a full sample of employer-employee matched micro data for the entire population of Sweden for the years between 1992 and 2013. The data set is a panel and we are able to observe every individual over the entire sample period. All observations are geocoded, which allows us to match residential and workplace location (at the establishment level) of all employed individuals. The data includes information on wage income, employer characteristics (both at the firm and the establishment level), and a vector of individual characteristics comprising age, sex, marital status, number of children, and education levels. For foreign-born individuals the sample includes information on the region of birth (18 different regions) and the year of arriving in Sweden.

We restrict our sample in two ways: First, we focus on two particular immigrant groups,

[9] We argue that the number of co-ethnic commuters reduces information frictions and therefore increases the job offer arrival rate. If immigrants have a preference for commuting with their ethnic peers, the observed positive relationship could stem from an increase in the probability of accepting a job rather than a higher job offer arrival rate.

[10] The Metro is the main mode of commuting in the Stockholm region. The Metro share of inner-city commuter trips is about 75 % at peak hour (Borjesson et al., 2014).

namely immigrants who arrived in Sweden in 1992 and 1993 from the Balkan countries following the Yugoslavian War, and Middle Eastern immigrants arriving in Sweden in 2005 and 2006 following the Iraq War.[11] The immigrants' year of arrival is defined as their first year with a residence permit in Sweden, which allows them to obtain work. Second, we only include immigrants residing in the Stockholm metropolitan region in the year of arrival. The sample of immigrants consists of 13,066 individuals. People are discarded if they did not find a job within the sample period, if they left the Stockholm region before finding their first job, if they found their first job in the year of arrival (and thus location characteristics of the previous year are not available), or if the location of the establishment could not be precisely identified. This leaves 3,841 individuals who found their first job in the Stockholm region within the sample period (see Table A.1 in the Appendix for details).

To evaluate neighborhood or enclave effects, we define both residential neighborhoods and workplace locations as (square) grid cells with a size of $1\,\text{km} \times 1\,\text{km}$, resulting in neighborhoods with population sizes similar to U. S. census tracts used by Cutler and Glaeser (1997) and Cutler et al. (1999, 2008). The grid cells are exogenously defined by Statistics Sweden and partition the entire Stockholm region. All variables included in the analysis are based on individual data that are aggregated at the neighborhood (i.e., grid cell) level. We define an immigrant's ethnic peers as all individuals who share the same region of birth, and an ethnic enclave as all ethnic peers living in the immigrant's neighborhood. The ethnic enclave is considered a proxy variable for the immigrant's ethnic network.

Explanatory variables can be classified as dyad specific and workplace specific variables, as indicated in Equations (1) and (2). Dyad specific variables include the (Euclidean) distance between the (centroid of the) residential neighborhood and the (centroid of the) workplace lo-

[11] Balkan countries include Albania, Kosovo, Bosnia and Herzegovina, Macedonia, Montenegro and Serbia, and the Middle East includes Bahrain, Egypt, the United Arab Emirates, Iraq, Iran, Israel, Yemen, Jordan, Kuwait, Lebanon, Oman, Qatar, Saudi Arabia, Syria and Turkey. Data on the source countries of immigrants are not available at the neighborhood level. Most immigrants from Balkan countries in 1992 and 1993 come from Bosnia and Herzegovina, and most immigrants from the Middle East in 2005 and 2006 from Iraq. At the end of the sample period, immigrants from Bosnia and Herzegovina and Iraq constitute the largest groups, accounting for about 40 % of the total stock of immigrants from the respective regions.

cation (*Distance*), the total number of commuters ($Commuters_{total}$), and the number of commuters of an immigrant's ethnic network ($Commuters_{ethnic}$). The variable $Commuters_{total}$ may capture information spillover effects, but can also reflect the connectivity or accessibility between two locations (as a higher connectivity will increase the number of commuters, ceteris paribus). The variable $Commuters_{ethnic}$ comprises a subset of the total number of commuters included in the variable $Commuters_{total}$. The parameter estimate on the variable $Commuters_{ethnic}$ therefore indicates the additional effect of a commuter on an immigrant's probability of finding the first job in the commuter's workplace location, if this commuter and the immigrant are ethnic peers. We interpret this variable as a measure of information sharing among ethnic peers about a potential job location.

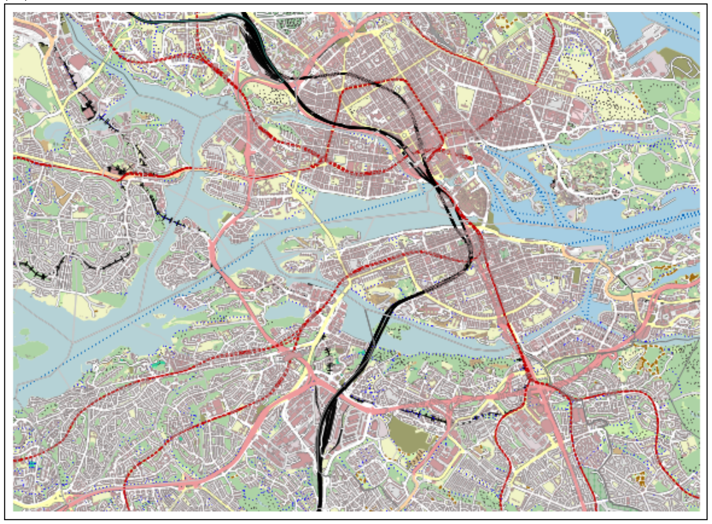
The definition of neighborhoods of residence, workplace locations, and commuter flows are shown in Figure 3 for a map section of Stockholm. Panel (c) shows the commuter flows of all employed persons living in the southwesternmost grid cell of the map section. The number in this cell (in italics) indicates the number of employed persons residing in this neighborhood. The numbers in the other cells denote how many of them work in that cell. For example, of the 12,035 employed people living in the southwesternmost cell, 197 people work in the cell directly east of it. Panel (d) shows commuter flows of ethnic peers living in the southwesternmost neighborhood in the same way.[12]

Workplace specific variables include the potential workplace location's (Euclidean) distance to the central business district (CBD) of the respective municipality (*Distance CBD*). The total number of individuals, $Employees_{total}$, and the number of ethnic peers, $Employees_{ethnic}$, working in a given location are included to indicate the labor market opportunities in that neighborhood. As argued by Damm (2009), some areas may provide better job opportunities for a specific immigrant group that may not be adequately captured by the total number of jobs. The workplace specific variables are included to control for work-

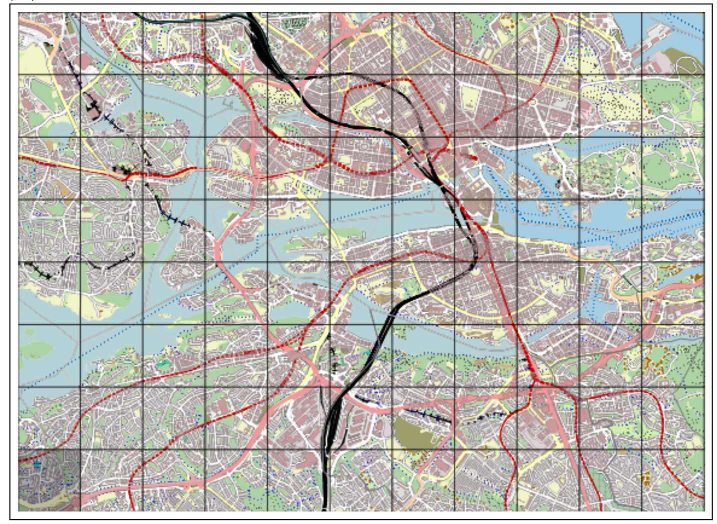
[12] We include Figure 3 to illustrate the definition of neighborhoods and the construction of commuter flows. The specific numbers are fictitious, as we are not allowed to report individual-level data for data protection reasons.

Figure 3: Neighborhood definition and commuter flows

(a) Map section of Stockholm



(b) Map section of Stockholm with grid cells



(c) Commuter flows of all residents

21	37	52	8	82	92	102	116	98	3	8
19	19	0	14	76	154	167	80	157	5	0
72	28	25	28	23	83	111	89	186	7	0
34	36	18	0	4	52	62	70	18	4	8
24	35	8	72	24	27	83	35	43	0	12
0	67	190	173	61	37	72	53	86	39	19
83	32	105	120	28	81	37	25	54	8	96
12035	197	83	167	43	73	56	83	73	27	42

(d) Commuter flows of ethnic peers

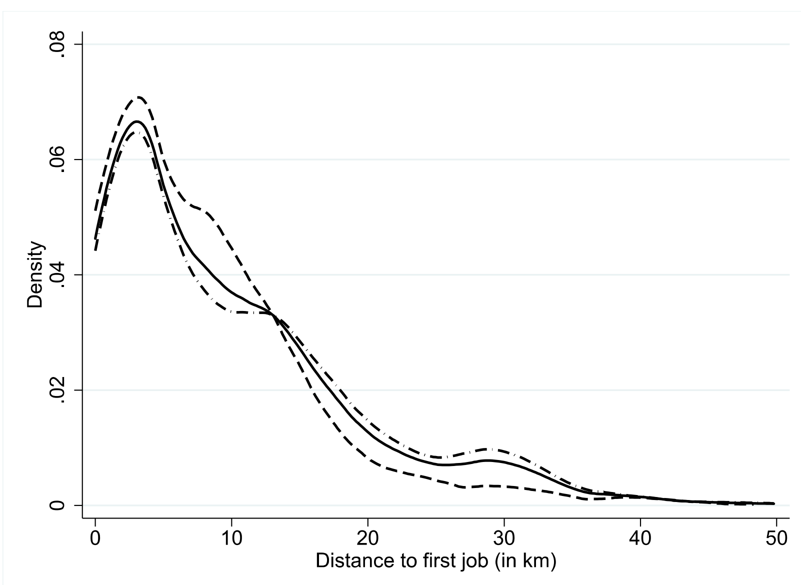
3	0	6	0	1	9	7	5	7	0	0
6	0	0	0	2	3	3	6	8	0	0
1	1	4	3	2	4	10	8	11	0	0
1	2	2	0	0	2	4	4	0	0	0
0	0	0	7	0	5	8	7	1	0	0
0	7	3	1	2	7	1	0	2	4	1
6	6	8	7	7	9	3	7	1	2	2
468	3	12	3	5	12	2	0	3	2	3

Notes: All four figures show the same map section of Stockholm. The grid lines delineate 1 km × 1 km cells. The bottom panels illustrate commuter flows from the grid cell in the southwest of the map (indicated by dotted hatching) to the respective grid cell. The figures in the cell in the southwest corner of the maps (in italics) show the number of all employed persons (panel (c)) and the number of all employed ethnic peers (panel (d)) living in that neighborhood (i.e., grid cell). The numbers in all other grid cells in panel (c) show the places of work of these employed persons and thus the commuter flows. The figures in panel (d) report commuter flows of ethnic peers only. Larger commuter flows are indicated by darker colors. The specific numbers are fictitious, as we are not allowed to report individual-level data for data protection reasons.

place location heterogeneity so that we can consistently estimate the effect of the number of commuters from an immigrant's ethnic network on the location of their first job.

Descriptive statistics on all variables, both for the entire sample and for the chosen alternatives (i.e., the locations where the immigrants found their first jobs) are summarized in Table 1. One observation is a combination of individual immigrant and potential workplace location, conditional on finding the first job in the respective year (i.e., $\sum_{v \in W^i} y_{irvt} = 1 \,\forall\, irt$). The choice set is restricted to potential workplace locations within a (Euclidean) distance of up to 50 kilometers from the place of residence, since the probability of finding the first job at a greater distance is very low (see Figure 4). Table 1 indicates that the chosen workplace location w (with $y_{irwt} = 1$) is characterized—relative to all potential workplaces—by a smaller commuting distance, by a larger number of commuters, by being relatively close to the CBD of the municipality, and by providing a large number of jobs, as expected.

Figure 4: Commuting distance to first job by immigrant groups



Notes: The figure shows kernel density estimates for the Euclidean distance between immigrants' neighborhoods of residence and places of first employment ($N = 3{,}841$). The solid line represents all immigrants in our sample. The dashed (dashed-dotted) line represents immigrants from Balkan countries (from the Middle East). The figure is based on an Epanechnikov kernel with a bandwidth of 2.

Table 1: Descriptive statistics

Variable	Description	Sample	Obs	Mean	Std. Dev.	Min	Max
Dyad-specific variables							
Distance	Distance between residential and potential workplace location (in km)	all	14,781,242	30.55	12.35	0	50.00
		$y_{irwt} = 1$	3,841	9.41	8.90	0	49.40
$Commuters_{total}$	Total number of commuters	all	14,781,242	0.37	4.35	0	1,259.00
		$y_{irwt} = 1$	3,841	26.99	51.64	0	1,259.00
$Commuters_{ethnic}$	Number of co-ethnic commuters	all	14,781,242	0.04	0.63	0	85.00
		$y_{irwt} = 1$	3,841	3.27	8.66	0	80.00
Location characteristics							
Distance CBD	Distance between potential workplace location and municipality CDB	all	14,639,356	10.51	6.82	0	50.79
		$y_{irwt} = 1$	3,840	4.23	3.95	0	43.04
$Employees_{total}$	Total number of employees in potential workplace location	all	14,781,242	213.17	1,367.39	0	41,678.00
		$y_{irwt} = 1$	3,841	4,894.67	7,749.91	0	41,678.00
$Employees_{ethnic}$	Number of co-ethnic employees in potential workplace location	all	14,781,242	6.10	35.94	0	1,084.00
		$y_{irwt} = 1$	3,841	128.09	182.64	0	1,084.00

Note: One observation is a combination of individual immigrant i (living in neighborhood r) and workplace location w, conditional on finding the first job in the respective year (i.e., $\sum_{v \in W^i} y_{irvt} = 1 \ \ \forall \ \ irt$).

4 Main results

Table 2 reports the odds ratios of the regression results of the conditional logit (CL) estimates. The most sparsely specified model [1] only includes the distance between the neighborhoods of residence and work, as well as the total number of jobs at the potential workplace location ($Employees_{total}$) and its distance to the CBD (of the respective municipality). We find a strong negative effect of the distance between the neighborhood of residence and the potential workplace location. An odds ratio of 0.24 suggests that the odds of finding a job decrease by about three fourth if the distance from the residential neighborhood to a potential workplace location doubles.[13] If the potential workplace is more attractive (i.e., the workplace location is close to the CBD of the municipality and offers many employment opportunities), the probability of finding the first job there increases, as expected. Adding the number of commuters ($Commuters_{total}$) as an additional explanatory variable—as in specification [2]—shows that it is more likely to take up a job if the number of commuters is larger: An odds ratio of 1.045 means that the probability of finding a job in that location increases by about 4.5 % if the number of people living in the immigrant's neighborhood who commute to that location increases by 10.[14]

The immigrant group specific variables are included as additional regressors in specification [3]. If many ethnic peers from an immigrant's neighborhood commute to a potential workplace location ($Commuters_{ethnic}$), the prospective worker is much more likely to find the first job there: The probability of finding the first job in a particular workplace location increases by 23 % if the number commuters of the ethnic network increases by 10. This effect is significantly positive at the 1 % significance level. The odds ratio for the overall number of commuters, $Commuters_{total}$, is much smaller compared to specification [2] and takes a value of only 1.02, presumably because the commuters who are members of the ethnic network are

[13]If an immigrant finds the first job in the neighborhood of residence, then $distance = 0$ and $\log(distance)$ is undefined. Individuals finding their first jobs within their residential neighborhoods are thus excluded throughout the analysis.

[14]Note that we measure commuters and workers in 10 persons to facilitate the presentation of the results.

Table 2: Regression results on location of first job: Main results

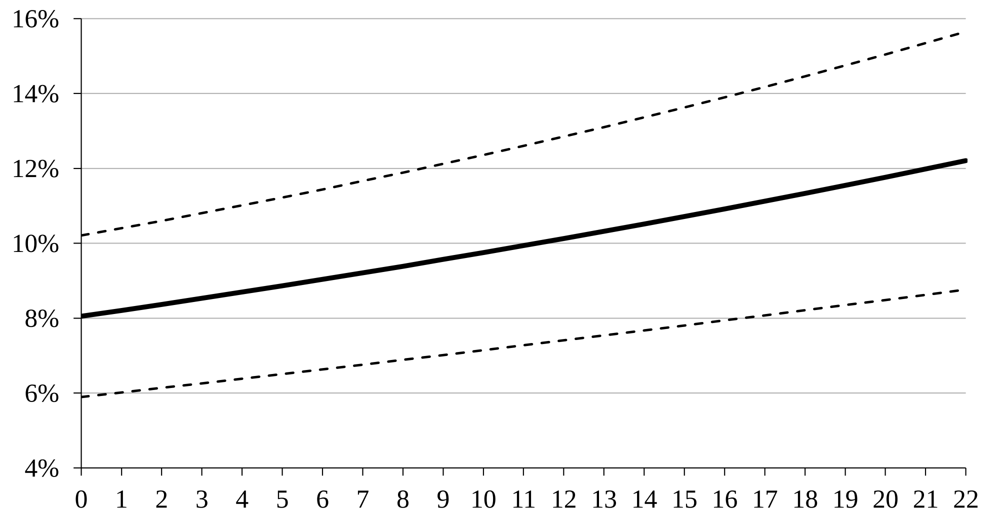
Variables	Model [1]	Model [2]	Model [3]	Model [4]	Model [5]
Dyad-specific variables					
Distance (in log)	0.23948***	0.25029***	0.25598***	0.25078***	0.25880***
	(0.00439)	(0.00477)	(0.00493)	(0.00772)	(0.00645)
$Commuters_{total}$ (in 10s)		1.04482***	1.01984***	1.02348**	1.01902***
		(0.00534)	(0.00545)	(0.00954)	(0.00677)
$Commuters_{ethnic}$ (in 10s)			1.23392***	1.24866***	1.19663***
			(0.04259)	(0.06677)	(0.05492)
Characteristics of workplace location					
Distance CDB (in log)	0.39922***	0.40587***	0.42011***	0.42881***	0.41602***
	(0.00925)	(0.00941)	(0.00978)	(0.01605)	(0.01237)
$Employees_{total}$ (in 10s)	1.00036***	1.00028***	0.99997	1.00014**	0.99979***
	(0.00003)	(0.00003)	(0.00004)	(0.00006)	(0.00006)
$Employees_{ethnic}$ (in 10s)			1.02074***	1.01129***	1.03110***
			(0.00196)	(0.00264)	(0.00282)
Sample	Full sample	Full sample	Full sample	“Movers”	“Stayers”
Number of observations	13,554,864	13,554,864	13,554,864	5,403,521	8,151,343
Number of individuals	3,565	3,565	3,565	1,421	2,144
Log-likelihood	–20,742	–20,704	–20,559	–8,205	–12,315
Pseudo-R^2	0.293	0.295	0.299	0.299	0.302

Notes: All regressions are estimated by a conditional logistic model and include fixed effects at the municipality level. The term “movers” (“stayers”) refers to immigrants who have changed (did not change) their neighborhood of residence after arriving in Sweden and before taking up their first job. Estimated coefficients are reported in odds ratios and standard errors in parenthesis. Asterisks denote statistical significance in a t-test at 1 % (***), 5 % (**) or 10 % (*) level.

now considered in a separate variable. Note that the number of co-ethnic commuters are included in both variables, $Commuters_{total}$ and $Commuters_{ethnic}$. The effect of commuters who are members of the ethnic network on the odds ratio of finding a job in a particular location is thus about 12 times higher than the influence of commuters with a different ethnic background. To illustrate the magnitude of the effect, Figure 5 shows the estimated probabilities of an immigrant finding their first job in a hypothetical workplace location, depending on the number of ethnic peers commuting to that workplace. The probability of finding a first job in that location increases steadily with the number of co-ethnic commuters from about 8 % (if no member of the ethnic network commutes to this place) to more than 12 % (if the number of ethnic peers increases to 22).[15]

[15] We choose 22 as the upper bound because this is approximately equal to the mean plus twice the standard deviation of the number of commuters to the chosen workplace location. See Table 1 for details.

Figure 5: Illustration of the size of the effect of co-ethnic commuters



Notes: The figure illustrates the predicted probabilities of finding the first job in a hypothetical workplace location (vertical axis), depending on the number of co-ethnic commuters (horizontal axis). The solid curve indicates the point estimates and the dotted curves indicate the 95 % confidence interval. The predicted values are based on the regression results of model [3] reported in Table 2. The distance between the place of residence and the workplace location is set to 2 km. The dummy variables for the municipality where the workplace is located are set to their mean values, and the variables for commuters, employment level at the workplace location and distance to the CBD are set to the mean values of the chosen alternatives (as reported in Table 1).

The number of ethnic peers working in a particular location ($Employees_{ethnic}$) also has a positive and statistically significant influence on the immigrant's labor market success in that area. The estimated odds ratio of the total number of jobs in the workplace location ($Employees_{total}$) is not significantly different from one anymore. Parameter estimates on the variables related to distance, on the other hand, are hardly affected by including these additional explanatory variables.

To investigate the issue of (negative) sorting of immigrants into enclaves, we analyze immigrants who have changed their neighborhood of residence after arriving in Sweden and before taking up their first job ("movers") and immigrants who did not move before becoming integrated into the labor market ("stayers") separately. The regression results on the subsample of movers are reported in column [4] and on stayers in specification [5] in Table 2. The parameter estimates of all dyad specific variables of the two subsamples take the same sign as in the main specification [3] and are significantly different from zero at least at the 5 % significance level. The point estimates vary only slightly by a statistically insignificant amount compared to the regression results based on the entire sample (model [3]). In particular, the odds ratios on co-ethnic commuters, $Commuters_{ethnic}$, are slightly larger for movers (1.25) than for stayers (1.20), but these odds ratios are not significantly different from each other.

The differences in the parameter estimates of workplace specific variables between these two groups are much larger: While the total number of jobs in a potential workplace location, $Employees_{total}$, is positively correlated with the probability of finding a job for movers, the relationship is significantly negative for stayers. Additionally, while the estimated odds ratio of $Employees_{ethnic}$ is significantly larger than one for both groups, the positive effect is much larger for stayers: An increase in the number of ethnic peers working in a particular location by 10 increases the odds for finding a job in that location by 3.1 % for stayers, but only by 1.1 % for movers. The regression results based on the subsamples of movers and stayers confirm that these groups are different. For example, stayers are more dependent on immigrant group specific employment opportunities, while general labor market opportunities are less important for them than for movers. This can be interpreted as an indication of negative sorting. However, the results clearly show that the effects of all dyad specific variables—in particular the number of co-ethnic commuters, $Commuters_{ethnic}$—are not sensitive to the mobility of immigrants.

5 Sensitivity analysis

In this section, we provide a wide range of sensitivity analyses to show the robustness of our findings, but also to elaborate on the heterogeneity of the results with respect to the immigrants' individual characteristics (Subsection 5.1) and workplace locations (Subsection 5.2). Subsection 5.3 discusses issues related to the spatial definition of neighborhoods and potential spillover effects across neighborhoods. Finally, Subsection 5.4 investigates the functional form of the relationship between the commuting patterns of ethnic peers and new immigrants' labor market integration. This last subsection also discusses the so-called independence of irrelevant alternatives (IIA) assumption, a necessary assumption in a conditional logit framework. In describing and discussing the results, we mainly focus on our main variable of interest, namely the number of commuters of the respective ethnic network, $Commuters_{ethnic}$, and will relate the findings of the sensitivity analysis to the results of the main specification [3], reported in Table 2.

5.1 Individual heterogeneity

To examine heterogeneity in the size of the effects of co-ethnic commuters on the location of the first employment of new immigrants, we split the sample by sex, education level, age, and the time span between the immigrants' arrival in Sweden and their integration into the labor market. For men and women as well as for different education levels we split both the sample and the commuter flows along these lines, as summarized in Table 3. Looking at the overall number of commuters, the results show that men's locations of first jobs are positively related to male commuters, but not significantly associated with female commuters (see model [6]). While we see the same pattern for commuters of the ethnic network, the size of the effects are much larger. Women's first workplace location, on the other hand, is positively associated with female commuters only, but negatively with male commuters. Parameter estimates for commuters of their ethnic networks are generally larger than for men (see specification [7]), but women appear to benefit more from male than from female co-ethnic commuters.

If immigrants are analyzed according to their level of education (models [8] to [10]), it is found that immigrants benefit exclusively from ethnic peers with a similar level of education: Both low and highly educated commuters of the ethnic network exert only a significantly positive impact on new immigrants with low or high education, respectively. Similarly, commuters with medium education have large positive effects for the labor market integration of

Table 3: Regression results on location of first job: Heterogeneity of commuter flows

Variables	Model [3]	Model [6]	Model [7]	Model [8]	Model [9]	Model [10]
Dyad-specific variables						
Distance (in log)	0.25598***	0.29233***	0.20630***	0.27460***	0.25013***	0.26015***
	(0.00493)	(0.00733)	(0.00648)	(0.00889)	(0.00889)	(0.00879)
$Commuters_{total}$ (in 10s)	1.01984***					
	(0.00545)					
$Commuters_{total}$ (men; in 10s)		1.04210**	0.92772**			
		(0.02169)	(0.02922)			
$Commuters_{total}$ (women; in 10s)		0.98471	1.12401***			
		(0.02047)	(0.02723)			
$Commuters_{total}$ (low education; in 10s)				1.46197***	1.84625***	1.30158***
				(0.12662)	(0.19139)	(0.12706)
$Commuters_{total}$ (medium education; in 10s)				1.06274	0.94582	1.05964
				(0.05121)	(0.05864)	(0.05307)
$Commuters_{total}$ (high education; in 10s)				0.91262***	0.90026**	0.93414***
				(0.02634)	(0.04451)	(0.02318)
$Commuters_{ethnic}$ (in 10s)	1.23392***					
	(0.04259)					
$Commuters_{ethnic}$ (men; in 10s)		1.37260***	1.55759***			
		(0.11644)	(0.21574)			
$Commuters_{ethnic}$ (women; in 10s)		0.88478	1.29294*			
		(0.09901)	(0.18705)			
$Commuters_{ethnic}$ (low education; in 10s)				1.75388***	0.75341	0.92340
				(0.37257)	(0.23131)	(0.21440)
$Commuters_{ethnic}$ (medium education; in 10s)				0.58000**	1.64240	0.73476
				(0.12541)	(0.50534)	(0.17520)
$Commuters_{ethnic}$ (high education; in 10s)				0.69167	0.84068	1.80612***
				(0.17992)	(0.28116)	(0.37004)
Characteristics of workplace location						
Distance CDB (in log)	0.42011***	0.43044***	0.40068***	0.45699***	0.39686***	0.43405***
	(0.00978)	(0.01262)	(0.01601)	(0.01762)	(0.01764)	(0.01785)
$Employees_{total}$ (in 10s)	0.99997	0.99985**	1.00011	0.99980**	1.00006	0.99992
	(0.00004)	(0.00006)	(0.00007)	(0.00008)	(0.00007)	(0.00008)
$Employees_{ethnic}$ (in 10s)	1.02074***	1.02850***	1.01045***	1.03349***	1.01453***	1.02368***
	(0.00196)	(0.00264)	(0.00290)	(0.00344)	(0.00336)	(0.00335)
Sample	Full sample	Men	Women	Low educ.	Medium educ.	High educ.
Number of observations	13,554,864	8,539,040	5,015,824	5,189,805	3,901,044	4,464,015
Number of individuals	3,565	2,245	1,320	1,356	1,034	1,175
Log-likelihood	–20,559	–13,405	–7,045	–7,836	–5,887	–6,682
Pseudo-R^2	0.299	0.275	0.352	0.299	0.308	0.309

Notes: All regressions are estimated by a conditional logistic model and include fixed effects at the municipality level. Educational attainment of immigrants is classified into low (9 years of schooling or less), medium (secondary education) or high (post-secondary education). Estimated coefficients are reported in odds ratios and standard errors in parenthesis. Asterisks denote statistical significance in a t-test at 1 % (***), 5 % (**) or 10 % (*) level.

immigrants with medium education, but the effect is imprecisely estimated and therefore not statistically significant. These findings are consistent with Bayer et al. (2008), who document stronger referral effects for workers with more similar sociodemographic characteristics, and can be explained in two ways: First, more similar ethnic peers may provide labor market information that is more relevant for the prospective worker. Second, we do not observe the ethnic network directly, but proxy this network by the immigrants from the same source region living in the same neighborhood. Using educational attainment (in addition to region of birth) might simply describe the immigrants' social networks more accurately, as social networks are in general often characterized by similar education levels. Regarding the total number of commuters $Commuters_{total}$, the results show that only low educated commuters influence the new immigrants' labor market successes positively. It appears that of all commuters (regardless of their ethnic background, i.e., predominantly natives), those with a low level of education provide more relevant labor market information than commuters with the same level of education. This result is consistent with findings of Larsen et al. (2018) and Kracke and Klug (2021), who show that new immigrants often pick jobs where they are overqualified.

Splitting the sample by the age of immigrants in the year of their arrival (under or over 30) or by whether they found their first job more quickly or more slowly (within the first two years or later) does not lead to any major differences. For all subsamples, summarized in models [11] to [14] in Table 4, co-ethnic commuters influence the odds ratios of an immigrant's first job location positively. The respective parameter estimates are significantly different from zero at the 1 % level and are of similar size as (and not significantly different from) the respective parameter estimates of our main model [3].

5.2 Workplace and firm heterogeneity

While we account for a small number of workplace characteristics in the main specification, the regressions reported in Table 5 control for a wider range of heterogeneity. In specification

Table 4: Regression results on location of first job: Individual heterogeneity

Variables	Model [3]	Model [11]	Model [12]	Model [13]	Model [14]
Dyad-specific variables					
Distance (in log)	0.25598***	0.26552***	0.25036***	0.27767***	0.23836***
	(0.00493)	(0.00806)	(0.00628)	(0.00779)	(0.00646)
$Commuters_{total}$	1.01984***	1.01485*	1.02389***	1.00356	1.04733***
(in 10s)	(0.00545)	(0.00869)	(0.00702)	(0.00725)	(0.00897)
$Commuters_{ethnic}$	1.23392***	1.25293***	1.22535***	1.16287***	1.23050***
(in 10s)	(0.04259)	(0.06534)	(0.05682)	(0.06125)	(0.05772)
Characteristics of workplace location					
Distance CDB (in log)	0.42011***	0.43231***	0.41185***	0.43638***	0.41437***
	(0.00978)	(0.01530)	(0.01275)	(0.01402)	(0.01402)
$Employees_{total}$	0.99997	0.99997	0.99995	0.99969***	1.00009
(in 10s)	(0.00004)	(0.00007)	(0.00006)	(0.00007)	(0.00006)
$Employees_{ethnic}$	1.02074***	1.02464***	1.01777***	1.04357***	1.00890***
(in 10s)	(0.00196)	(0.00298)	(0.00263)	(0.00324)	(0.00239)
Sample	Full sample	< 30 years old	30 years or older	job within 2 years	job after 2 years
Number of observations	13,554,864	6,028,868	7,525,996	6,939,711	6,615,153
Number of individuals	3,565	1,576	1,989	1,815	1,750
Log-likelihood	–20,559	–8,970	–11,550	–10,589	–9,881
Pseudo-R^2	0.299	0.309	0.294	0.292	0.314

Notes: All regressions are estimated by a conditional logistic model and include fixed effects at the municipality level. In model [11] and [12] individuals are classified depending on their age in the year of arrival in Sweden. Estimated coefficients are reported in odds ratios and standard errors in parenthesis. Asterisks denote statistical significance in a t-test at 1 % (***), 5 % (**) or 10 % (*) level.

[15], the dummy variable *Same municipality* indicates whether the residential neighborhood and the potential workplace are located in the same municipality, as neighborhoods within the same municipality may be better connected, ceteris paribus. An estimated odds ratio significantly larger than one for this variable confirms this conjecture. However, the parameter estimate for co-ethnic commuters, $Commuters_{ethnic}$, remains almost unaffected by this modification. Model [16] includes a large number of workplace specific controls, such as industry structure (employment shares of the 13 different industries) and variables indicating industry variety and job growth over the past two years. As expected, both industry variety and job dynamics increase the probability of finding a job in this location. In the model specification [17], we include workplace location specific fixed effects to control for all (unobserved) heterogeneity between potential workplace locations. Note that the sample size is substantially

reduced by the inclusion of fixed effects, since in many potential job locations none of the immigrants in our sample found their first job. These locations are perfectly explained by the fixed effects and therefore drop out of the analysis. While the effects of $Commuters_{ethnic}$ on the odds of finding a job decline somewhat when using these additional control variables, the estimated odds ratio remain significantly larger than one at the 1 % level in both models [16] and [17].

Table 6 reports subsample results depending on whether individuals find their first jobs in small (1-9 employees), medium (10-49 employees) or large firms (50 or more employees). Interestingly, the larger the company where individuals find their first job, the stronger the effect: If the prospective worker finds the first job in a small firm, an increase in the number of co-ethnic commuters by 10 increases the probability of finding a job at the respective location by merely 3.9 %, which is not significantly different from zero. For immigrants who find a job in a medium-sized company, the probability increases by 23.3 % (which is similar to the main results), and for immigrants who find the first job in a large company, the probability increases by as much as 41.9 %. Both effects are significantly positive at the 1 % level. The results in the main model [3] do not seem to be driven by new immigrants finding jobs in businesses run by ethnic peers, as these are predominantly small service enterprises, while we find the strongest effects for individuals getting employed in large firms.

5.3 Definition of neighborhoods

In this sensitivity analysis, we address spatial aspects of the definition of neighborhoods. First, we investigate whether the use of 1 km $\times$ 1 km grid cells is an appropriate definition of residential neighborhoods. A drawback of our approach to delineate neighborhoods in this way is that they have neither natural (such as rivers) nor artificial (e.g., infrastructure) boundaries. Consequently, our neighborhood definition might be inaccurate for individuals living close to the neighborhood's border. As each 1 km $\times$ 1 km grid can be divided in 16 cells of 250 m size, we first restrict our sample to immigrants living in the four 250 m cells in

Table 5: Regression results on location of first job: Workplace heterogeneity

	Model [3]	Model [15]	Model [16]	Model [17]
Dyad-specific variables				
Distance (in log)	0.25598***	0.30295***	0.25583***	0.28825***
	(0.00493)	(0.00720)	(0.00504)	(0.00623)
$Commuters_{total}$ (in 10s)	1.01984***	1.02021***	1.00566	0.98631**
	(0.00545)	(0.00547)	(0.00564)	(0.00648)
$Commuters_{ethnic}$ (in 10s)	1.23392***	1.21637***	1.15189***	1.11670***
	(0.04259)	(0.04215)	(0.04107)	(0.04541)
Same municipality		1.97569***		
		(0.10680)		
Characteristics of workplace location				
Distance CDB (in log)	0.42011***	0.42047***	0.75917***	
	(0.00978)	(0.00979)	(0.02221)	
$Employees_{total}$ (in 10s)	0.99997	0.99993*	0.99961***	
	(0.00004)	(0.00004)	(0.00005)	
$Employees_{ethnic}$ (in 10s)	1.02074***	1.02252***	1.00402**	
	(0.00196)	(0.00197)	(0.00200)	
Industry variety			1.04155***	
			(0.00124)	
Employment dynamics			1.00010***	
			(0.00002)	
Municipality level FE	Yes	Yes	Yes	No
Workplace location FE	No	No	No	Yes
Industry structure	No	No	Yes	No
Sample	Full sample	Full sample	Full sample	Full sample
Number of observations	13,554,864	13,554,864	13,554,864	1,915,793
Number of individuals	3,565	3,565	3,565	3,612
Log-likelihood	–20,559	–20,481	–19,724	–17,890
Pseudo-R^2	0.299	0.302	0.328	0.203

Notes: All regressions are estimated by a conditional logistic model. *Same municipality* is a dummy variable and equals 1 if place of residence and place of work are located in the same municipality. *Industry variety* counts the number of different industries that can be found in the respective workplace location (at the three-digit level) and *employment dynamics* denotes the net change in employees within the last two years in the workplace location. The industry structure comprises the employment shares of the individual industries (13 categories) at the workplace location. Estimated coefficients are reported in odds ratios and standard errors in parenthesis. Asterisks denote statistical significance in a t-test at 1 % (***), 5 % (**) or 10 % (*) level.

Table 6: Regression results on location of first job: Firm heterogeneity

Variables	Model [3]	Model [18]	Model [19]	Model [20]
Dyad-specific variables				
Distance (in log)	0.25598***	0.28455***	0.23295***	0.25496***
	(0.00493)	(0.01015)	(0.00790)	(0.00808)
$Commuters_{total}$ (in 10s)	1.01984***	1.03405***	1.01373	1.01194
	(0.00545)	(0.01240)	(0.00981)	(0.00771)
$Commuters_{ethnic}$ (in 10s)	1.23392***	1.03889	1.23357***	1.41870***
	(0.04259)	(0.06895)	(0.07464)	(0.07997)
Characteristics of workplace location				
Distance CDB (in log)	0.42011***	0.41115***	0.45049***	0.39987***
	(0.00978)	(0.01691)	(0.01891)	(0.01533)
$Employees_{total}$ (in 10s)	0.99997	0.99954***	1.00009	1.00011*
	(0.00004)	(0.00009)	(0.00008)	(0.00007)
$Employees_{ethnic}$ (in 10s)	1.02074***	1.03357***	1.01941***	1.01599***
	(0.00196)	(0.00386)	(0.00348)	(0.00295)
Sample	Full sample	Small firms	Medium firms	Large firms
Number of observations	13,554,864	4,284,178	4,345,121	4,925,565
Number of individuals	3,565	1,125	1,135	1,305
Log-likelihood	–20,559	–6,700	–6,440	–7,318
Pseudo-R^2	0.299	0.277	0.311	0.318

Notes: All regressions are estimated by a conditional logistic model and include fixed effects at the municipality level. Small (medium) [large] firms are firms with 1-9 (10-49) [50 or more] employees. Estimated coefficients are reported in odds ratios and standard errors in parenthesis. Asterisks denote statistical significance in a t-test at 1 % (***), 5 % (**) or 10 % (*) level.

the middle of the 1 km grids. This restriction ensures that the individuals in this subsample do not live close to the border of their 1 km × 1 km neighborhoods. The results given in model [21] in Table 7 show that the odds ratios for $Commuters_{ethnic}$ (but also for all other variables) are hardly affected by this modification.[16]

Second, we examine whether the spatial delineation of the residential neighborhood or the workplace location is too small. In addition to commuter flows from an immigrant's place of residence to a potential workplace location, we also consider commuters living in the proximity of the immigrant's neighborhood in specification [22]. All grid cells are considered to be close-by if they are adjacent (horizontal, vertical or diagonal) to the respective residential neighborhood. In model [23], we consider commuter flows from the immigrant's residential

[16]Note that the odds ratios are estimated less precisely: The standard errors in model [21] are about twice as high as in model [3] because the sample size has been quartered.

Table 7: Regression results on location of first job: Definition of neighborhoods

	Model [3]	Model [21]	Model [22]	Model [23]
Dyad-specific variables				
Distance (in log)	0.25598***	0.26722***	0.25595***	0.24503***
	(0.00493)	(0.01040)	(0.00497)	(0.00490)
$Commuters_{total}$ (in 10s)	1.01984***	1.01977*	1.02138***	1.05623***
	(0.00545)	(0.01063)	(0.00618)	(0.00724)
$Commuters_{total}$ *around place of residence* (in 10s)			0.99881	
			(0.00215)	
$Commuters_{total}$ *around workplace location* (in 10s)				0.98480***
				(0.00201)
$Commuters_{ethnic}$ (in 10s)	1.23392***	1.22491***	1.23117***	1.18890***
	(0.04259)	(0.07594)	(0.04640)	(0.04822)
$Commuters_{ethnic}$ *around place of residence* (in 10s)			1.00585	
			(0.03826)	
$Commuters_{ethnic}$ *around workplace location* (in 10s)				0.98181
				(0.01485)
Characteristics of workplace location				
Distance CDB (in log)	0.42011***	0.46702***	0.42017***	0.40272***
	(0.00978)	(0.02220)	(0.00978)	(0.00959)
$Employees_{total}$ (in 10s)	0.99997	0.99996	0.99998	0.99999
	(0.00004)	(0.00009)	(0.00005)	(0.00004)
$Employees_{ethnic}$ (in 10s)	1.02074***	1.02477***	1.02053***	1.02037***
	(0.00196)	(0.00402)	(0.00209)	(0.00194)
Sample	Full sample	Residence in center of neighborhood	Full sample	Full sample
Number of observations	13,554,864	3,405,478	13,554,864	13,554,864
Number of individuals	3,565	893	3,565	3,565
Log-likelihood	–20,559	–5,184	–20,559	–20,516
Pseudo-R^2	0.299	0.295	0.299	

Notes: All regressions are estimated by a conditional logistic model and include fixed effects at the municipality level. Neighborhoods around the place of residence or work, respectively, denote all neighborhoods adjacent to the respective location, applying the concept of queen contiguity. Estimated coefficients are reported in odds ratios and standard errors in parenthesis. Asterisks denote statistical significance in a t-test at 1 % (***), 5 % (**) or 10 % (*) level.

neighborhood to neighborhoods in the proximity of a potential workplace location (using the same concept of proximity as before). These specifications investigate whether residential neighborhoods are too narrowly defined, and whether co-ethnic commuters can provide labor market information in a larger area around their place of work. The parameter estimates for $Commuters_{ethnic}$ are significantly positive and similar in magnitude to our main specification [3], while neither co-ethnic commuters living near an immigrant's residential neighborhood nor commuters working in neighborhoods close to a potential workplace location have a significant influence on the immigrant's labor market integration. These results suggest that our definitions of residential neighborhoods and workplace locations are spatially appropriate and not too small.

5.4 Alternative functional form and independence of irrelevant alternatives

To address the issue of a potential non-linear relationship between ethnic peers commuting to a particular neighborhood and the associated information spillovers, we include—in addition to the number of commuters—a dummy variable indicating if there is at least one commuter between these two neighborhoods. These results are reported in specification [24] in Table 8. While the parameter estimates on the number of co-ethnic commuters, $Commuters_{ethnic}$, are hardly affected by this modification, the odds ratio of the dummy variable are significantly positive, suggesting that the first commuter of the ethnic network is particularly important.

As information spillovers may be different for workplace locations close to the neighborhood of residence, we exclude immigrants who found their first jobs in their residential neighborhood throughout our empirical analysis. In specifications [25] and [26], we also exclude those who found their first jobs within 2 km or within 5 km distance of their place of residence, respectively. In both cases, the effects of commuters of the ethnic network are significantly positive. When restricting the sample to individuals who found their first jobs at least 5 km away, the estimated effect is substantially larger. This may also point in

Table 8: Regression results on location of first job: Alternative functional form

	Model [3]	Model [24]	Model [25]	Model [26]	Model [27]
Dyad-specific variables					
Distance (in log)	0.25598	0.39564***	0.23872***	0.23297***	0.25670***
	(0.00493)	(0.00863)	(0.00627)	(0.01105)	(0.00508)
$Commuters_{total}$ (in 10s)	1.01984***	1.02507***	1.02209***	1.08057***	1.03575***
	(0.00545)	(0.00527)	(0.00669)	(0.01893)	(0.00702)
$Commuters_{total} > 0$		9.10010***			
		(0.58579)			
$Commuters_{ethnic}$ (in 10s)	1.23392***	1.24386***	1.16437***	1.74282***	1.19054***
	(0.04259)	(0.04346)	(0.04413)	(0.21170)	(0.04812)
$Commuters_{ethnic} > 0$		1.73403***			
		(0.08489)			
Characteristics of workplace location					
Distance CDB (in log)	0.42011***	0.58645***	0.41442***	0.43795***	0.42335***
	(0.00978)	(0.01461)	(0.01059)	(0.01367)	(0.01014)
$Employees_{total}$ (in 10s)	0.99997	1.00001	0.99996	0.99992	0.99995
	(0.00004)	(0.00004)	(0.00005)	(0.00005)	(0.00005)
$Employees_{ethnic}$ (in 10s)	1.02074***	1.01305***	1.02018***	1.01788***	1.02218***
	(0.00196)	(0.00191)	(0.00204)	(0.00237)	(0.00204)
Sample	Full sample	Full sample	*Distance* $\geq$ 2km	*Distance* $\geq$ 5km	1,000 randomly sampled alternatives
Number of observations	13,554,864	13,554,864	11,891,132	8,551,797	3,528,812
Number of individuals	3,565	3,565	3,126	2,259	3,565
Log-likelihood	–20,559	–19,658	–18,411	–13,668	–15,848
Pseudo-R^2	0.299	0.330	0.285	0.265	0.356

Notes: All regressions are estimated by a conditional logistic model and include fixed effects at the municipality level. $Commuters_{total} > 0$ ($Commuters_{ethnic} > 0$) is a dummy variable and equals 1 if the number of commuters (number of commuters of the same immigrant group) is larger than 0. In model [27], the choice set is restricted to 1,000 randomly sampled non-chosen alternatives (in addition to the chosen workplace location). Estimated coefficients are reported in odds ratios and standard errors in parenthesis. Asterisks denote statistical significance in a t-test at 1 % (***), 5 % (**) or 10 % (*) level.

the direction of a potential non-linear relationship, with the first commuter being especially important, as discussed in the previous paragraph.

One restriction inherent to the conditional logit model is the so-called independence of irrelevant alternatives (IIA) assumption, stating that the ratio of choice probabilities between two alternatives (e.g., between potential workplace locations A and C from the example illustrated in Figure 2) depends only on the characteristics of these two alternatives, but not on the availability or characteristics of other alternatives. A violation of this assumption would render the CL model invalid. This assumption may be restrictive in some settings, but if this assumption holds, then the parameters can be consistently estimated when the number of alternatives is reduced by randomly dropping non-chosen alternatives (i.e., workplace locations) altogether (see Train, 2009, p. 48 f.). Therefore, the estimated coefficients whenusing a subset of alternatives should not differ significantly from the parameter estimates obtained when using the full set of alternatives.

We follow this suggestion and randomly reduce the number of non-chosen potential workplace locations for each immigrant to 1,000 neighborhoods. The results are reported in the final specification [27] in Table 8. The odds ratios on the number of co-ethnic commuters remain significantly positive at the 1 % level and are not significantly different from the respective results of the main model specification [3]. This result suggests that the IIA assumption is not violated in the empirical analysis.

6 Conclusions

In this article, we investigate *where* (in which neighborhoods) immigrants take up their first jobs and relate the immigrants' workplace location to the number of co-ethnic commuters. Matched employer-employee individual-level data of the entire population of the Stockholm labor market region for over 20 years (including information on the place of residence, place of work and ethnic background) allow us to calculate commuting paths of every single individual and to isolate ethnic peers from the overall commuter flows. We consider all ethnic peers living in the same neighborhood as a proxy variable for the immigrant's ethnic network. This network can influence the labor market integration of new immigrants in various ways, for example by influencing (immigrant group specific) social norms or by reducing incentives to acquire host-country specific (e.g., language) skills. Additionally, ethnic peers can help reducing information frictions for new immigrants, for example by providing their employer with information about a new immigrant (by making job referrals) or by disseminating information about job openings.

The key point is that ethnic peers are able to provide better labor market information about their workplace location than than about other neighborhoods. We exploit this (information) heterogeneity by estimating the workplace locations of new immigrants as a function of co-ethnic commuters in a conditional logit (CL) framework. This empirical design allows us to control for all (observed and unobserved) individual-level and neighborhood-level heterogeneity. We are therefore able to interpret the impact of co-ethnic commuters on the new immigrants' workplace locations as causal evidence of information spillovers for the following reasons: First, in the CL framework, we control for (potentially unobservable) individual heterogeneity that could lead to inconsistent parameter estimates of neighborhood characteristics due to omitted variables bias (often referred to as "sorting by ability"). Second, we account for all other channels through which the ethnic network might influence the labor market integration of new immigrants, as other mechanisms operate through the place of residence. Third, by exploiting the longitudinal dimension of the data, we can rule out reverse causality and thus ensure that the result is evidence that neighbors transmit labor market information and not work colleagues share residential information.

Our results show that new immigrants are more likely to find their first job in a particular workplace location if more of their ethnic peers commute there. This effect is both statistically robust and sizeable, as one additional co-ethnic commuter increases the probability of finding the first job at a particular workplace location by 2.3 % (in our main specification). This effect is larger for women, co-ethnic commuters with similar education levels, and for immigrants who find their first jobs in larger firms. The impact of co-ethnic commuters are similar regardless of whether immigrants changed residence before finding a first job, how old they are, or whether they found their first jobs quickly.

Our analysis contributes to the literature on the effects of ethnic enclaves and ethnic networks on the (labor market) integration of new immigrants. This relationship is complex and ethnic networks influence the labor market success of its members in many ways. The effects of some of the mechanisms identified by the literature are positive (like information dissemination, Bertrand et al., 2000), some are negative (lower incentives to acquire language skills, see Lazear, 1999), and some are ambiguous (e.g., human capital externalities, Borjas, 1995). The empirical literature usually estimates an overall effect by comparing different immigrants living in more or less segregated neighborhoods (e.g., Edin et al., 2003; Damm, 2009, 2014), because it is difficult to discriminate between different mechanisms. Thus, the estimated impacts are the net effects of various possible channels. We contribute to this literature by isolating the effects of one important mechanism, namely that immigrants disseminate labor market information through their ethnic network ("information channel").

Our research design is not limited to isolating labor market relevant information spillovers disseminated through ethnic networks, but can also be applied to other social networks in the neighborhood, given that the social networks can be identified. For example, one might investigate information sharing among college graduates or workers with specific training, rather than ethnic peers. Moreover, information dissemination is not limited to labor markets, and one can apply this research design when "commuters" provide location specific information. For example, patients can share information about the quality of hospitals they have previously consulted with new patients seeking a particular treatment, or students may provide prospective students with information about the schools they attend. Our research therefore also contributes to the broader literature on neighborhood effects, where "we know little about the relative importance of the different mechanisms that are typically 'bundled' together within a neighborhood" (Chyn and Katz, 2021, p. 216). A better understanding of the relative importance of specific neighborhood effects is key to improving policy responses to the challenges that segregation in general and ethnic enclaves in particular may pose to society.

References

Åslund, O. (2005). Now and forever? Initial and subsequent location choices of immigrants. *Regional Science and Urban Economics*, 35(2):141–165.

Andersson, H. (2021). Ethnic enclaves, self-employment, and the economic performance of refugees: evidence from a swedish dispersal policy. *International Migration Review*, 55(1):58–83.

Andersson, H., Berg, H., and Dahlberg, M. (2021). Migrating natives and foreign immigration: Is there a preference for ethnic residential homogeneity? *Journal of Urban Economics*, 121:103296.

Andersson, R., Musterd, S., and Galster, G. (2014). Neighbourhood ethnic composition and employment effects on immigrant incomes. *Journal of Ethnic and Migration Studies*, 40(5):710–736.

Andersson, R., Musterd, S., and Galster, G. (2019). Port-of-entry neighborhood and its effects on the economic success of refugees in Sweden. *International Migration Review*, 53(3):671–705.

Aslund, O., Edin, P.-A., Fredriksson, P., and Gronqvist, H. (2011). Peers, neighborhoods, and immigrant student achievement: Evidence from a placement policy. *American Economic Journal-Applied Economics*, 3(2):67–95.

Auer, D. (2018). Language roulette – The effect of random placement on refugees' labour market integration. *Journal of Ethnic and Migration Studies*, 44(3):341–362.

Bartel, A. P. (1989). Where do the new U. S. immigrants live? *Journal of Labor Economics*, 7(4):371–391.

Battisti, M., Peri, G., and Romiti, A. (2022). Dynamic effects of co-ethnic networks on immigrants' economic success. *Economic Journal*, 132(641):58–88.

Bayer, P., Ross, S. L., and Topa, G. (2008). Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy*, 116(6):1150–1196.

Beine, M., Docquier, F., and Özden, C. (2011). Diasporas. *Journal of Development Economics*, 95(1):30–41.

Bertrand, M., Luttmer, E. F. P., and Mullainathan, S. (2000). Network effects and welfare culture. *The Quarterly Journal of Economics*, 115(3):1019–1055.

Borjas, G. J. (1995). Ethnicity, neighborhoods, and human-capital externalities. *American Economic Review*, 85(3):365–390.

Borjesson, M., Jonsson, R. D., and Lundberg, M. (2014). An ex-post CBA for the Stockholm Metro. *Transportation Research Part A*, 70:135–148.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics. Methods and Applications*. Cambridge University Press, Cambridge.

Carrington, W. J., Detragiache, E., and Vishwanath, T. (1996). Migration with endogenous moving costs. *American Economic Review*, 86(4):909–930.

Chakraborty, T., Schueller, S., and Zimmermann, K. F. (2019). Beyond the average: Ethnic capital heterogeneity and intergenerational transmission of education. *Journal of Economic Behavior & Organization*, 163:551–569.

Chiswick, B. and Miller, P. W. (2005). Do enclaves matter in immigrant adjustment? *City & Community*, 4(1):5–35.

Chyn, E. and Katz, L. F. (2021). Neighborhoods matter: Assessing the evidence for place effects. *Journal of Economic Perspectives*, 35(4):197–222.

Collins, W. J. and Margo, R. A. (2000). Residential segregation and socioeconomic outcomes: When did ghettos go bad? *Economics Letters*, 69(2):239 – 243.

Cutler, D. M. and Glaeser, E. L. (1997). Are ghettos good or bad? *The Quarterly Journal of Economics*, 112(3):827–872.

Cutler, D. M., Glaeser, E. L., and Vigdor, J. L. (1999). The rise and fall of the American ghetto. *Journal of Political Economy*, 107(3):455–506.

Cutler, D. M., Glaeser, E. L., and Vigdor, J. L. (2008). When are ghettos bad? Lessons from immigrant segregation in the United States. *Journal of Urban Economics*, 63(3):759 – 774.

Damm, A. P. (2009). Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics*, 27(2):281–314.

Damm, A. P. (2014). Neighborhood quality and labor market outcomes: Evidence from quasi-random neighborhood assignment of immigrants. *Journal of Urban Economics*, 79(SI):139–166.

Danzer, A. M., Feuerbaum, C., Piopiunik, M., and Woessmann, L. (2022). Growing up in ethnic enclaves: Language proficiency and educational attainment of immigrant children. *Journal of Population Economics*, 35(3):1297–1344.

Dustmann, C., Glitz, A., Schönberg, U., and Brücker, H. (2016). Referral-based job search networks. *The Review of Economic Studies*, 83(2):514–546.

Edin, P.-A., Fredriksson, P., and Åslund, O. (2003). Ethnic enclaves and the economic success of immigrants – Evidence from a natural experiment. *The Quarterly Journal of Economics*, 118(1):329–357.

Faggian, A. (2021). *Job Search Theory*, pages 481–495. Springer, Berlin, Heidelberg.

Galster, G. C. (2012). *The Mechanism(s) of Neighbourhood Effects: Theory, Evidence, and Policy Implications*, pages 23–56. Springer Netherlands, Dordrecht.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380.

Gross, D. M. and Schmitt, N. (2003). The role of cultural clustering in attracting new immigrants. *Journal of Regional Science*, 43(2):295–318.

Hellerstein, J. K., McInerney, M., and Neumark, D. (2011). Neighbors and coworkers: The importance of residential labor market networks. *Journal of Labor Economics*, 29(4):659–695.

Horowitz, J. L. (1991). Modeling the choice of choice set in discrete-choice random-utility models. *Environment and Planning A: Economy and Space*, 23(9):1237–1246.

Johansson, B., Klaesson, J., and Olsson, M. (2002). Time distances and labor market integration. *Papers in regional science*, 81(3):305–327.

Katz, L. F., Kling, J. R., and Liebman, J. B. (2001). Moving to opportunity in Boston: Early results of a randomized mobility experiment. *Quarterly Journal of Economics*, 116(2):607–654.

Klaesson, J. and Oner, O. (2021). Ethnic enclaves and segregation – Self-employment and employment patterns among forced migrants. *Small Business Economics*, 56(3, SI):985–1006.

Klaesson, J., Oner, O., and Pennerstorfer, D. (2021). Getting the first job: Size and quality of ethnic enclaves and refugee labor market entry. *Journal of Regional Science*, 61(1):112–139.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

Kracke, N. and Klug, C. (2021). Social capital and its effect on labour market (mis)match: Migrants' overqualification in Germany. *Journal of International Migration and Integration*, 22(4):1–26.

Kramarz, F. and Skans, O. N. (2014). When strong ties are strong: Networks and youth labour market entry. *Review of Economic Studies*, 81(3):1164 – 1200.

Kristiansen, M. H., Maas, I., Boschman, S., and Vrooman, J. C. (2022). Refugees' transition from welfare to work: A quasi-experimental approach of the impact of the neighbourhood context. *European Sociological Review*, 38(2):234–251.

Larsen, E. N., Rogne, A. F., and Birkelund, G. E. (2018). Perfect for the job? Overqualification of immigrants and their descendants in the Norwegian labor market. *Social Inclusion*, 6(3):78–103.

Lazear, E. P. (1999). Culture and language. *Journal of Political Economy*, 107(6):95–126.

Lindgren, K.-O., Nicholson, M. D., and Oskarsson, S. (2022). Immigrant political representation and local ethnic concentration: Evidence from a Swedish refugee placement program. *British Journal of Political Science*, 52:997–1012.

Marten, L., Hainmueller, J., and 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.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegriono, A., and Taylor, J. (1993). Theories of international migration: A review and appraisal. *Population and Development Review*, 19(3):431–466.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In Zarembka, P., editor, *Frontiers in Econometrics*, chapter 4, pages 105–142. Academic Press, New York.

Nowotny, K. and Pennerstorfer, D. (2019). Network migration: Do neighbouring regions matter? *Regional Studies*, 53(1):107–117.

Pedersen, P. J., Pytlikova, M., and Smith, N. (2008). Selection and network effects – Migration flows into OECD countries 1990–2000. *European Economic Review*, 52(7):1160–1186.

Train, K. E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press, New York, second edition.

Xie, Y. and Gough, M. (2011). Ethnic enclaves and the earnings of immigrants. *Demography*, 48(4):1293–1315.

Zavodny, M. (1999). Determinants of recent immigrants' locational choice. *International Migration Review*, 33(4):1014–1030.

Zorlu, A. and Mulder, C. H. (2008). Initial and subsequent location choices of immigrants to the Netherlands. *Regional Studies*, 42(2):245–264.

Appendix Sample selection

The sample includes immigrants from the Balkans who arrived in Sweden in 1993 and 1994 and immigrants from the Middle East who arrived in 2006 and 2007. We only consider people who first arrived in the Stockholm metropolitan region when immigrating to Sweden, as illustrated in Figure A.1. The entire sample consists of 13,066 individuals (see Table A.1 below). Individuals are discarded if they dropped out of the sample before finding a job (because they moved or passed away), if they did not find jobs until the end of the sample period (in 2013), or if they found their first jobs in the year of arriving in Sweden (because their labor market success cannot be explained by the neighborhood characteristics of the previous year).

Figure A.1: Map of Stockholm metropolitan region

About 40 % (2,818) of the remaining 6,659 individuals had to be dropped because the location of the first job was either outside the Stockholm metropolitan region or further than 50 km away from their residential neighborhood, or because the workplace location could not be precisely identified. The reason for this inaccuracy is that information on wage income (and thus employment status) and on workplace location are surveyed differently: Having a job is defined as receiving a wage income larger than zero over the entire calendar year, whereas information on the workplace location is collected at a particular point in time during the year. In the first year in employment, immigrants often work only temporarily. If they do work some time during the year, but not at the specific date the data on workplace locations are surveyed, then they are employed (according to our definition), but we cannot identify their place of work. We discard these individuals, as we focus on finding the first job, because once immigrants have been integrated into the labor market, information frictions may become less important.

Table A.1: Sample for empirical analysis

	Number of individuals
Individuals in the sample (total)	13,066
Excluded because individuals ...	
dropped out of the sample before finding job	1,817
did not find a job until the end of the sample period (2013)	2,637
found first job in year of arrival	1,953
location of first job could not be precisely identified	2,818
Sample for the empirical analysis	**3,841**