Migrating Natives and Foreign Immigration: Is There a Preference for Ethnic Residential Homogeneity?[a]

Henrik Andersson[b] Heléne Berg[c] Matz Dahlberg[d]

October 27, 2020

Abstract

We study the migration behavior of the native Swedish population following refugee immigration, with a particular focus on examining whether there is support for an ethnically based migration response. Using rich geo-coded Swedish data, we account for possible endogeneity problems by combining policy-induced initial immigrant settlements with exogenous contemporaneous immigration as captured by refugee shocks. We find the same flight among all natives, irrespective of their parental foreign background. This suggests that "ethnic distance" to the new immigrants is not the dominant channel causing natives' flight behavior. Instead, refugee immigration seems to lead to more socio-economically segregated neighborhoods.

Keywords: Immigration; Native migration; Flight; Avoidance; IV estimation; Shift-share strategy
JEL classification: C26; J15; R23

1 Introduction

Over the last decades, many European and other Western countries have witnessed increased immigration, with a drastic culmination in 2015. In this year alone, UNHCR estimated that around 1 million individuals reached the shores of Europe after having crossed the Mediterranean. In the wake of this experience, heated discussions have emerged on how and where to

[a]We are grateful to Leah Platt Boustan, Mattias Engdahl, Jon Fiva, Florian Morath, Albert Saiz, Matti Sarvimäki, Håkan Selin, Susanne Urban, two anonymous referees, and seminar participants at UC Irvine, UCLA, ETH/KOF in Zurich, Uppsala University, University of Verona, Statistics Norway, the 2015 IIPF conference in Dublin, UCFS meeting at Krusenberg, UCLS meeting in Uppsala, the 2017 Urban Economic Association Meeting in Vancouver, and the Norface meetings at Rosersberg, Birmingham, and EUI in Florence for helpful comments and discussions.

[b]Department of Government, Uppsala University. henrik.andersson@statsvet.uu.se.

[c]Department of Economics, Stockholm University; CESifo. helene.berg@ne.su.se.

[d]Institute for Housing and Urban Research and Department of Economics, Uppsala University; CESifo; IEB; VATT; IFAU. matz.dahlberg@ibf.uu.se.

accommodate all these refugees. In particular, a major political concern is the emergence of ethnically segregated neighborhoods. Aside from immigrants tending to one another, such a development is reinforced if the native population reacts by leaving or avoiding neighborhoods that become more ethnically diverse. The extent to which natives do so is the topic of this paper.

We study the migration behavior of the native population—here, native Swedes—when new immigrants arrive. We hypothesize that this may be manifested either in the form of *native flight* (immigration inducing natives to move out of a neighborhood) or in the form of *native avoidance* (immigration inducing natives to avoid moving into a neighborhood where more immigrants tend to settle). Ultimately, the aim is to use estimated migration responses to deduce whether natives prefer *ethnically* homogeneous neighborhoods. We approach this task by developing the so-called "shift-share" method into a much-improved identification strategy.

In order to create effective policies to combat segregation, it is important to know both *if* natives change their behavior following immigration and, if so, *why* they do so. The common hypothesis in the literature on "white flight" is that migration responses are due to preferences for ethnically homogeneous neighborhoods (see, for instance, Saiz, 2007; Boustan, 2010; Saiz and Wachter, 2011; Sá, 2015). However, newly arrived immigrants exhibit a number of different characteristics other than their ethnicity. For example, the average refugee typically has a lower level of education and income than the native population. Which trait do the natives actually react to? Do they react to the ethnicity of the immigrants, as typically hypothesized in the earlier literature, or their socio-economic composition?[1]

We reach three main conclusions. First, we do not find any evidence of either native flight or native avoidance when studying the entire population.

Second, we find it important to acknowledge that, due to various constraints, far from all individuals are able to react based on their residential preferences. To that end, we want to distinguish between households with a high/low ability to move to a preferred neighborhood following an increase in immigration. In our setting, we define mobile actors as homeowners

[1] The data used in the paper allows us to observe country of birth and country of emigration. We do not, however, have any data on self-proclaimed ethnicity, which is why we use source country to proxy for ethnicity.

rather than renters, as the rental market in Sweden is characterized by long (sometimes extreme) queues, and renters frequently compete over the same apartments as the newly arrived refugees.[2]

A key result is that, when studying the group of natives identified as having a high ability to move to a preferred neighborhood, we estimate significant flight responses. We further find that the flight in this group goes to neighborhoods *within* rather than *between* municipalities.[3] In contrast, there are no effects of increased immigration on the migration behavior among natives identified as having a low ability to move to a preferred neighborhood. The pronounced flight effect in the sub-sample of mobile natives is in accordance with assumptions made in theoretical models on the effects of immigration on native migration. However, it is a previously neglected aspect in empirical work, which means that our results potentially have implications for the interpretations made in existing, related studies that have only looked at aggregate (average) effects.

Third, we conclude that a preference for ethnically homogeneous neighborhoods is not the dominant channel causing flight. Instead, our analyses consistently indicate that natives have preferences for socio-economically homogeneous (or "better") neighborhoods.[4] In particular, we find that the flight effect is the same across different groups of natives with different parental foreign backgrounds. This, in turn, suggests that the "ethnic distance" to the new immigrants is not the dominant channel causing natives' flight behavior.

The paper brings several value-adding features to the literature on native flight and avoidance. Thanks to comprehensive, detailed register data, a contribution of the paper is to examine the validity of the presumed ethnicity channel in a way that earlier literature was unable to do. Our data allows us to identify natives with different parental foreign backgrounds. And since many native-born individuals with non-Western parents are ethnically quite similar to current immigrants (in terms of country of origin),

[2]The rationale for this division is discussed in detail in section 3.1 and in the online Appendix. The online Appendix also provides additional facts on the rental queuing system in Sweden and sensitivity analyses for the results provided in the paper.

[3]We find evidence of native flight, but not native avoidance. A possible interpretation is that natives mostly notice and consequently react to increased immigration into the neighborhood where they currently reside.

[4]The fact that high-income individuals prefer to live with high-income neighbors has been shown in a US context by, for example, Boustan (2013).

yet in many cases socio-economically more similar to native-born individuals with Swedish-born parents, we can use the parental information to explore the validity of the ethnicity channel. Consequently, we are able to examine whether there is support for the hypothesis that residential preferences are formed along an ethnic dimension by estimating the migration response of natives conditioned on their parents' country of birth.

The geographic level of detail of our data allows us to examine where those who move following increased immigration go. Do they move to neighborhoods within their municipality (short-distance moves) or do they move to another municipality (long-distance moves)? Our finding that these moves mainly occur between neighborhoods within a municipality points to the importance of having fine-grained geographic data when examining native flight, which is very rare in the previous literature. Consequently, papers only having studied out-migration from large geographical areas, such as municipalities or metropolitan areas, might have underestimated the flight effect of increased immigration.

This paper also contains several methodological improvements. Our data contains information on each individual immigrant's reason for applying for a residence permit—whether or not he or she arrived as a refugee, tied family migrant, or labor migrant. This is a unique feature enabling us to make a distinct methodological improvement to related studies. In particular, we focus on refugees. Not only is this a highly topical and interesting group to study, but from a methodological point of view, refugee migration is arguably more exogenous to the characteristics of the receiving city or neighborhood, as compared to labor, student or family migration.[5]

To see how the focus on refugee migration constitutes a methodological improvement, note that much of the previous literature on white flight has focused on the US[6] and has frequently categorized all immigration as one common treatment. Besides the fact that generalizing all immigration into one concept implies that nuanced mechanisms may be lost, it may also cause endogeneity problems. A large proportion of immigrants to the US

[5] *Refugees* in our paper include all asylum-related residence permits, most importantly "Geneva convention" refugees (in which case there is an individual reason for asylum) as well as those given protection due to conflicts and war.

[6] See, in particular, Farley et al. (1978), Farley et al. (1994), Boustan (2010), Saiz and Wachter (2011), Wang (2011), and Shertzer and Walsh (forthcoming). An exception is Accetturo et al. (2014) who conduct an analysis on Italian data.

are *pulled* to specific places[7], whereas refugees tend to have been *pushed* from their native country by wars and other catastrophes.[8] Hence, increases in US-type immigration could to a larger extent than the refugee immigration we focus on be a function of regional chocks and pull factors, which affect both immigrants but potentially also the behavior of natives.[9]

We identify the causal effect of foreign immigration on the residential choice of natives by combining (i) contemporary refugee migration to Sweden with (ii) previous immigrant settlement patterns resulting from a refugee placement policy that was in place in the early years of our study period. In short, the policy meant that refugees were not allowed to decide for themselves where to settle but were assigned a municipality by the Migration Board. We argue that this policy-generated settlement is yet another methodological improvement to existing studies. The rationale for this is that immigrant settlement patterns from the early 1990s, which subsequently attracted more recent push-driven refugee migrants, are more likely to be uncorrelated with neighborhood characteristics that matter for natives' residential preferences compared to settlement patterns in the absence of such a policy.

Our panel data also allows us to incorporate neighborhood-fixed effects. Ultimately, we construct an instrumental variable for *changes* in immigration based on the interaction of, on the one side, immigrant settlements during the placement policy era and, on the other, the timing of contemporary, refugee-driven immigrant shocks. Arguably, this results in an improvement of the typical shift-share instrument used previously in the literature, where both initial immigrant settlement as well as contemporary immigration are added directly to the analysis and are thus likely to be endogenous to the outcome (see, for instance, Altonji and Card, 1991; Card and DiNardo, 2000; Saiz, 2007; Basten and Siegenthaler, 2013; Chalfin and Levy, 2013;

[7]According to the Migration Policy Institute, only 13 percent of all new US green card holders in 2016 were refugees, while almost half of all new permanent residents were refugees in Sweden (see `https://www.migrationpolicy.org/article/frequently-requested-statistics-immigrants-and-immigration-united-states` and `https://www.migrationsverket.se/English/About-the-Migration-Agency/Facts-and-statistics-/Statistics/Overview-and-time-series.html`).

[8]Zimmermann (1996) presents a stylized definition of push and pull migration.

[9]Consider, for example, a case where native US citizens increasingly appreciate Japanese food and culture. This could attract more Japanese into the States, while also making natives more inclined to live in Japanese-dense neighborhoods.

Sá, 2015).[10]

The data used in this paper originates from the GeoSweden database, covering the entire Swedish population since 1990. Several of the value-adding features and important sets of results (as laid out above) are only made possible thanks to the detailed information in this database. In particular, we are able to, on a year-to-year basis, identify immigrants granted a residence permit in Sweden based on refugee status, as well as natives who rent or own their homes. Furthermore, for each immigrant living in Sweden, there is information on his or her country of origin. Finally, all variables are presented as an annual panel covering relatively small neighborhoods. While the panel structure enables the fixed effects mentioned above, the high geographical resolution means that we can capture more nuanced residential preferences, as we are able to observe relatively short moves that are likely less costly than moving between larger units, such as metropolitan areas.

Apart from the literature *directly* estimating the extent to which the residential choices of natives are affected by immigration,[11] our paper is related to an influential set of literature having *indirectly* studied the response of natives to increased immigration by estimating effects on house prices (Saiz, 2007; Saiz and Wachter, 2011; Sá, 2015; Accetturo et al., 2014; Balkan et al., 2018; Braakmann, 2019) and wages (Card, 1990; Altonji and Card, 1991).

This paper is also closely linked to the so-called tipping point literature estimating at which potential proportion of immigrants in a neighborhood or a city the native population disproportionately starts to leave (Schelling, 1971; Card et al., 2008; Aldén et al., 2015). We instead focus on continuous native migration. Finally, complementing the studies of the effects *of* residential segregation (Edin et al., 2003), our focus is on the effects of immigration *on* residential segregation.

In the next section, we describe recent immigration patterns to Sweden. Section 3 then discusses the theoretical mechanisms through which we hypothesize that these patterns affect natives' migration responses and, in particular, describes our idea for examining whether there is any support for

[10] Jaeger et al. (2018) suggest a set of improvements to the shift-share instrument. As a robustness test, we apply their version in the Appendix, which yields virtually the same results.

[11] In addition to the papers in the field of economics referred to above (e.g., Card, 1990; Altonji and Card, 1991; Saiz, 2007; Boustan, 2010; Saiz and Wachter, 2011; Sá, 2015), a substantial body of literature in sociology and geography studies this phenomenon; see Rathelot and Safi, 2014 and the references therein.

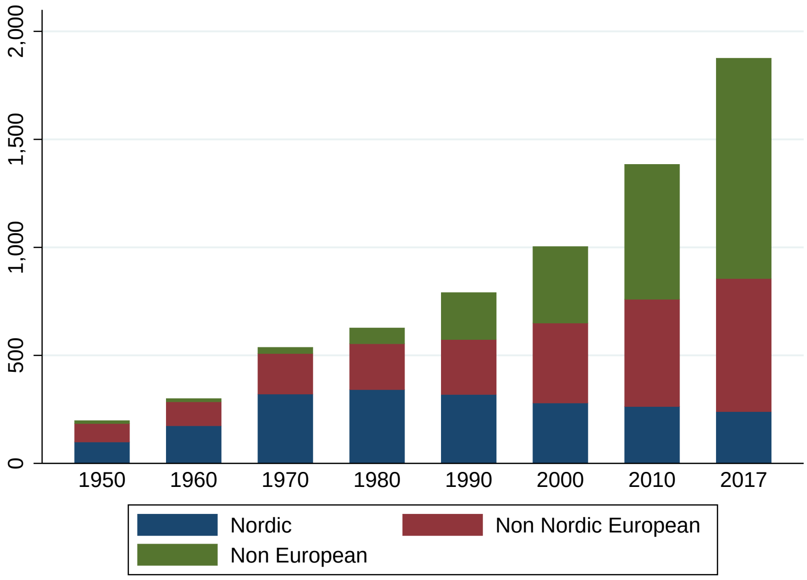
the ethnicity-based mechanism. While Section 4 lays out the strategy used to estimate these responses empirically, Section 5 presents the data used to obtain the main results, which are provided in Section 6. In the final section, we conclude the paper.

2 Immigration to Sweden

The size and character of immigration to Sweden have changed over the last decades. In 1970, less than seven percent of the Swedish population was born in another country.[12] Of those, a large majority had arrived as labor immigrants from another Nordic or European country in the 1950s and 1960s. From the late 1970s/early 1980s, immigration changed character in terms of going from being mainly labor-induced migration to more refugees starting to arrive. Consequently, there has since then been a drastic change in both the number and the origin of the foreign-born population in Sweden. The changing pattern of the foreign-born population is clear in Figure 1. While the proportion with origins in the Nordic countries decreases over time, the proportion originating in non-European countries is increasing. In 1950, the approximately 200,000 foreign-born individuals living in Sweden constituted around 2.8 percent of the total population of around 7 million. By the end of 2017, the approximately 1,900,000 foreign-born individuals constituted more than 18 percent of the total population of around 10 million. More than half of these individuals are born outside of Europe.

[12]Statistics Sweden, Yearbook of Sweden 2012, table 4.30 "Population by country of birth."

Figure 1: Number of foreign-born individuals in Sweden by region of origin, 1950–2017.



Notes: Y-axis in units of thousands.

Source: Statistics Sweden.

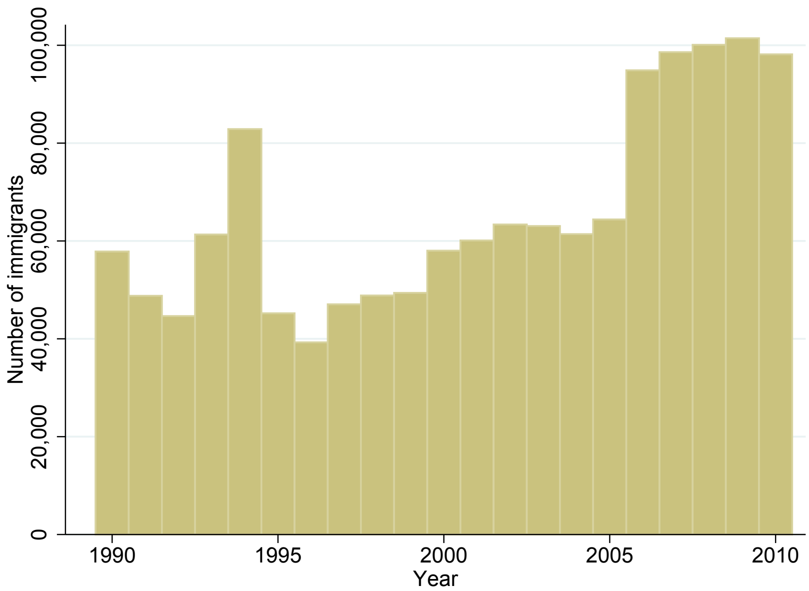
Compared to a majority of European countries, Sweden has a relatively large proportion of foreign-born individuals. According to statistics from Eurostat,[13] in 2010, 47 million individuals in the EU27 were not born in the country in which they resided. This amounted to almost ten percent of the total population. The majority of these, slightly more than 31 million, were born outside the European Union. There is, however, a large variation in these numbers across the union, ranging from Poland (with 1.2 percent foreign-born individuals), the Czech Republic, Hungary and Finland (with around 4 percent foreign-born individuals) to Austria (15.2 percent), Sweden (14.3 percent), Spain (14 percent) and Germany (12 percent).

Switching our focus from stocks to flows, the annual immigration to Sweden during the period we study, 1990–2010, is shown in Figure 2. Up until 2006, typically 50–60,000 individuals arrived each year.[14] Then, from 2006

[13]The figures in this section originate from issues 98/2008, 27/2010, 45/2010, and 34/2011 of Eurostat's Statistics.

[14]The spike in the early 1990s was due to increased refugee immigration following the Balkan war, and the increase in 2006 was primarily related to an escalation of the Iraqi war.

Figure 2: Total immigration to Sweden, 1990–2010



Source: GeoSweden (see Section 5 for further details).

and onward, there has been a discrete increase in the number of immigrants, with a yearly average of around 100,000.

3 Potential reactions of natives

The literature on residential segregation typically studies two types of reactions from the majority population in relation to an immigration of minorities: *flight* (immigration inducing the majority population to move out of a neighborhood) and *avoidance* (immigration inducing the majority population to avoid moving into a neighborhood).[15] For the analysis in this paper, it is necessary to distinguish between "native" and "white." The concepts of *native* flight and avoidance differ from *white* flight and avoidance. The latter stems from a US tradition of research on the effects of racial diversity. Primarily due to a different data practice in terms of how to classify individuals' backgrounds, rather than focusing on racial diversity, we study flight and avoidance due to increased diversity with regard to country of origin.

[15]For a complete set of potential reactions, one would additionally consider the concept of native *attraction*, referring to a scenario where, opposite to native flight and avoidance, immigration induces natives to move into or stay in an area.

Consequently, we refer to the potential reaction of the majority population as *native* flight and avoidance.

In this section, we discuss three things. First, we discuss the important issue of which households, in the Swedish context, can be considered having a high ability to actually react to an "immigration shock" by moving (see section 3.1). Second, we discuss our thinking in terms of our definition of natives. Since our definition of a native is everyone born in Sweden, our native group is quite heterogeneous with regard to individual parental foreign backgrounds. This is a feature that we will use in an attempt to disentangle the mechanisms behind the observed migration responses of natives (see Section 3.2 on preference-based mechanisms). Finally, we end this section with a discussion on non-behavioral mechanisms (see section 3.3).

3.1 Ability to react to a shock by moving

A prerequisite for deducing residential preferences from flight and avoidance estimates due to *any* mechanism is that people are indeed mobile. That is, that they are able to move to a preferred place. We recognize that this is far from true for everyone, meaning that some groups may not be able to react on the basis of their residential preferences. In the Swedish context, there is a dividing line between those who own and those who rent their housing. There are two reasons why the ability to move to a preferred place may differ between these two groups. One concerns differences in resources and the other is related to the institutional details of the Swedish housing market.[16]

Regarding differences in resources, it is clear that the more human capital and the more financial resources an individual possesses, the easier it is to move: the chances of finding a new job is higher, and the budget constraint is less strict.[17] In Sweden, those who rent their housing on average have smaller financial and human capital resources than those who own. Table 1 illustrates this aspect for the population of native-born Swedes: compared to owners, renters have significantly lower disposable income, the proportion with a university education is significantly lower, and the proportion receiv-

[16]See the online Appendix for a more detailed description of the Swedish rental market, also including a sensitivity analysis based on rental market tightness.

[17]Large increases in house prices during the past decades imply that a lack of financial resources is a significant obstacle for buying a home in Sweden.

Table 1: Financial and human capital resources among renting and owning native Swedes, respectively

	Renters	Owners	Difference
Disposable income	1609.1	2271.5	-662.4***
	(1213.6)	(4741.6)	(5.75)
Univ. education	0.219	0.322	-0.103***
	(0.414)	(0.467)	(0.001)
Social assistance	0.086	0.012	0.074***
	(0.280)	(0.108)	(0.002)
N	687,923	4,322,989	

*** p<0.01, ** p<0.05, * p<0.1. Disposable income is measured in 100 SEK, university education and social assistance are defined as shares. Variables are defined in 2010.

ing social assistance is significantly higher. This indicates that renters may find it more difficult to react to an "immigrant shock" in the neighborhood by moving to a preferred neighborhood.

Aside from being more constrained in terms of resources, an additional reason why Swedish renters cannot easily react by moving is due to the institutional setting. In particular, features of the Swedish housing markets imply that renters' mobility constraints are likely to be particularly pronounced as a consequence of increased immigration. Municipalities are responsible for accommodating newly arrived refugees unable to find a place on their own. This is typically done through municipality-owned rental apartments.[18] These apartments make up a majority of the rental market and, in turn, a relatively large part of the total housing market. Access to these public rentals requires queuing, in many municipalities for several years (or even decades, as in the case of Stockholm; see the online Appendix). This is also true for existing tenants, as well as for many private rentals.[19]

These two facts imply that the competition for rental apartments is accentuated in high-immigration municipalities (given fixed short-run housing supply). Ultimately, following increased immigration, moving to a new neighborhood within the municipality will be particularly difficult for individuals living in rentals.[20] Note that this is not to say that renters *in*

[18]As documented in Andersson et al. (2010).

[19]Although under certain circumstances, so-called "switching contracts" where two renters change apartments with one another may be approved.

[20]Leaving the municipality is not subject to this problem. But long-distance moves are instead significantly more costly, not least from a labor market point of view. Additionally,

general are less mobile. Rather, this follows from the combination of most immigrants occupying rental apartments and that the non-renter market is inaccessible to (budget-constrained) renters.[21]

To take these mobility and resource constraints into account, we examine whether the migration behaviors differ between owners and renters. A contribution of our paper is that, to our knowledge, we are the first to take these kinds of mobility constraints into account when analyzing natives' migration responses to increased immigration.[22]

3.2 Preference-based mechanisms

Why would increased immigration affect natives' decisions where to live? Scholars in the fields of sociology, economics, and geography have highlighted several potential mechanisms, where the dominant one is related to preferences for racial and/or ethnic homogeneity. Primarily sociologists have used attitude surveys to document racial and ethnic preferences. These might take the form of strict preferences for living with co-ethnics or an aversion to perceived social unrest (Farley et al., 1978, 1994). Economists have incorporated this notion into their models by introducing a parameter capturing "distaste for immigrants" (or, analogously, "preference for homogeneity"). An illustrative example is the set-up in Sá (2014), where the preferences of the native population are modeled as:[23]

$$U_{n,i} = V_{n,i} + f(h,x) - \delta I, \tag{1}$$

where $V_{n,i}$ measures the value individual n attaches to the local amenities in neighborhood i, $f(h,x)$ is a function measuring utility from the consumption of housing services (h) and other goods (x), and δ captures natives'

moving to a new municipality often implies lost points in the queue.

[21]In the online Appendix, we provide more facts underpinning our argument and, in addition, we show that the degree of rental market tightness in a municipality is important for understanding our results.

[22]Again, the online Appendix presents evidence in support of our argument that owners have greater possibilities to react to an increased immigration by moving to a preferred neighborhood. But owners and renters may also differ in dimensions unrelated to this type of mobility. Thus, we cannot be entirely sure that the difference in results we recover is only explained by the differences in mobility that we highlight.

[23]See equation (9) in Sá (2015).

preferences for immigrants I. The mobility response of natives to immigration is derived by maximizing the utility function in (1) subject to the relevant budget constraint. This yields the intuitive prediction that native flight will increase if natives have a preference for homogeneity/a distaste for immigration (i.e., in terms of the model, if $\delta > 0$).

But what is the interpretation of the preference parameter δ? Does it measure natives' preferences for ethnicity or their preferences for other traits presented by the newly arrived immigrants? In Sweden, newly arrived immigrants are to a large extent refugees. Particularly during his or her first years in the country, the average refugee has a lower income and is less educated than the native population in the neighborhoods in which refugees locate. If natives have preferences for neighborhoods with homogeneous (high) levels of income and/or education, the change in the socio-economic composition in the neighborhood resulting from, in particular, refugee immigration may drive native out-migration. In other words, if natives experience that the neighborhood status is dropping due to increased immigration, then observed native flight/avoidance might in fact be economic flight/avoidance.[24]

That immigrants' socio-economic status plays a role in relation to natives' locational decisions has obviously been discussed earlier in the literature (see, for instance, Boustan (2010); Saiz and Wachter (2011); Rathelot and Safi (2014); Sá (2015)). Probably due to data restrictions, however, this has never really been examined. Here, we contribute by disentangling this socio-economic channel from the commonly assumed ethnic channel by using the detailed information in the Swedish register data regarding the foreign background of the parents of the native-born individuals.

Native-born Swedes represent many different ethnic backgrounds on the parental side; some have Swedish-born parents, others have parents born in another Western country, and still others have parents born in non-Western countries who mostly arrived as refugees (or as family members of refugees) before having children.[25] Assume that ethnicity is the only characteristic

[24]We refer to this channel as preferences *for homogeneity* along the socio-economic dimension. Since refugees generally have a lower socio-economic status, this is (empirically) equivalent to preferences *against a lower composition* of socio-economic traits. As an illustration of refugees generally having a lower socio-economic status, we note from our GeoSweden data that the median refugee did not have any earned income in the first year after his or her arrival.

[25]See `https://www.migrationsverket.se/English/About-the-Migration-Agency/Facts-and-statistics-/Statistics/Overview-and-time-series.html` for information

among the new immigrants that matters for the migration decision of the natives—that is, natives have a strong preference for ethnically homogeneous neighborhoods—so that δ only captures this dimension (call it $\delta^{Ethnicity}$). We would then expect the following hypotheses to hold:

$$\delta^{Ethnicity}_{Swedish\_Parents}, \delta^{Ethnicity}_{Western\_Parents} > \delta^{Ethnicity}_{Non-Western\_Parents} \tag{2}$$

That is, the mobility response within the group of natives on average more ethnically dissimilar to the newly arrived refugees (native-born individuals with Swedish- and other Western-born parents) will be greater than the response within the group of natives on average more ethnically similar to the newly arrived refugees (native-born individuals with parents born in a non-Western country). If there is a strong preference for ethnic homogeneity, we thus expect δ to be the smallest among natives with non-Western parents. By relating our empirical results to the different δ-coefficients in equation (2), we can examine the validity of the ethnicity-based channel vs. the socio-economic channel.

3.3 Non-behavioral mechanisms

Aside from the two preference-based channels discussed in section 3.2, there are non-behavioral mechanisms to consider. First, immigration may lead to changes in house prices that, in turn, may induce native flight and avoidance. Boustan (2010) explains this clearly. When studying historical white flight within the US, she uses a model where house prices are a function of the number of inhabitants. Assuming an inelastic housing supply, immigration will initially cause prices to rise. Since locational decisions are likely to be affected by house prices, this will induce movement from the current population. Under such a scenario, part of the observed flight is thus due to price increases rather than behavioral effects induced by the preferences or perceptions of the native majority. A similar line of reasoning can be found in, for example, Saiz (2007).

There is also the possibility of a reverse price effect if the neighborhood status is (perceived to be) dropping with increased immigration. This could

on number and type of residence permits per country of origin from 1980 and onward.

induce homeowners concerned about falling house prices to leave. However, the housing stock in high-immigration neighborhoods is typically characterized by a large proportion of rental apartments (see Section 5), and since the Swedish rental market is highly regulated, immigration cannot affect rental prices, regardless of whether they move up or down. This is particularly true in the short-run perspective adopted in our analysis (we consider native migration within one year of additional foreign immigration). Ultimately, we thus expect these non-behavioral mechanisms via house price changes to be rather small in the current setting. At the very least, they should not differ between the groups of natives with a different parental background, meaning that the relative importance of preferences along the ethnic vs. socio-economic dimension can be assessed as laid out above.

In addition to price effects, given that the housing supply is not perfectly elastic, there is also a "mechanical effect" to consider. In the extreme case when the housing supply is perfectly *inelastic*, irrespective of residential preferences, a person can only move into a neighborhood if someone else has moved out. Thanks to the high frequency in our data, we are more or less able to rule out this mechanical effect for the case of flight. We know the place of residence on December 31 for each individual living in Sweden in that particular year. We thus observe immigrants as well as natives registered in a particular neighborhood on that very date and can therefore with fairly good precision measure only native outflow occurring after the arrival of new immigrants. This means that our measure of native flight is net of any such potential mechanical effect.

For the case of avoidance, however, regardless of data frequency, it is not possible to completely rule out that measured native avoidance is mechanically driven by a fixed housing supply. Specifically, when a person moves into a neighborhood where the housing supply is fixed, there is one less apartment/house available for everybody else. Even if a native was contemplating moving there, this possibility might then not exist. This, however, should at most imply a (negative) 1:1 relation, meaning that we can rule out larger negative effects than that as solely driven by such a mechanical effect.

To sum up the discussion in section 3, if we observe substantial native flight among those with a high ability to react to an immigrant shock (those who own their housing), this is most likely driven by preferences against living in an ethnically diverse neighborhood and/or in a socio-economic diverse neighborhood. The same is true for observed native avoidance larger than a (negative) 1:1 relation. Furthermore, if natives with varying parental foreign backgrounds react to a similar extent, this suggests that preferences are formed along socio-economic dimensions and thus that preferences for ethnically homogeneous neighborhoods are (at most) of second order.

4 Econometric strategy

This section covers our econometric approach. Here, we discuss the general set-up, identification strategy, and our improvements compared to the previous literature.

4.1 General set-up

Let us begin by defining native outflow, $outflow_{i,t}$, as the number of natives who leave neighborhood i in year t. Analogously, we define native inflow, $inflow_{i,t}$, as the number of natives who move into i in year t. In other words, $outflow_{i,t}$ is the number of natives who lived in i in $t-1$ but live in another neighborhood in t, whereas $inflow_{i,t}$ is the number of natives who did not live in i in $t-1$ but do so in t.[26] The two variables $outflow_{i,t}$ and $inflow_{i,t}$ are our main outcome variables, and our two parameters of interest in the following two equations are β^{out} and β^{in}:

$$outflow_{i,t+1} = \alpha^{out} + \beta^{out} im_{i,t} + \epsilon^{out}_{i,t+1} \quad (3)$$

$$inflow_{i,t+1} = \alpha^{in} + \beta^{in} im_{i,t} + \epsilon^{in}_{i,t+1} \quad , \quad (4)$$

where $im_{i,t}$ is the number of new immigrants in neighborhood i in year t. Returning to the discussion from the previous section, we predict the following of β^{out} and β^{in}:

Empirical predictions. *If increased immigration cause…*

… *native* flight*, then* $\beta^{out} > 0$.

[26]Note that for the natives' responses, we only consider migration within the country (i.e., not emigration responses).

... *native* avoidance, *then* $\beta^{in} < -1$

The geographic location of immigrants is not random but might rather be correlated—either directly or via some unobserved neighborhood characteristic—with our outcome of interest: native migration. In other words, an endogeneity problem exists that needs to be solved. To identify β^{out} and β^{in}, we use an instrumental variable we argue substantially improves the instruments typically used earlier in the literature (the so-called shift-share instrument; see Altonji and Card, 1991 for the first use of this instrument). In short, this improvement is mainly attributed to two factors. First, we only consider refugee migration, arguably exhibiting more exogenous variation in immigration than when conflated with other migration. Second, we use a Swedish refugee placement policy that was in effect in the early part of the period we study, arguably generating a more exogenous historical allocation of immigrants than when they self-select their place of residence.

In the following, we discuss the general shift-share approach and how we improve it.

4.2 Identification: Interaction between push-driven immigration and a historical placement policy

The instruments used in the previous literature to solve the endogenous location choice of immigrants typically follow the shift-share strategy (see, for instance, Altonji and Card, 1991; Card and DiNardo, 2000; Saiz, 2007; Sá, 2015). This strategy is based on the observation that new immigrants tend to be drawn to places where former immigrants sharing their background have already settled. The idea is to instrument $im_{i,t}$ with the prediction $\tilde{im}_{i,t}$, defined as (exemplified by immigration to Sweden):

$$\tilde{im}_{i,t} = \sum_c \tilde{im}_{c,i,t} = \sum_c \Big(\phi_{c,i,t^0} \times im_{c,SWE,t} \Big), \tag{5}$$

where

$$\phi_{c,i,t^0} = \frac{im_{c,i,t^0}}{im_{c,SWE,t^0}} \tag{6}$$

is the fraction of immigrants from source country c who arrived in Sweden and settled in neighborhood i in some baseline period t^0. $im_{c,SWE,t}$ represents the total refugee immigration to Sweden from source country c in year (or period) t. The instrument $\tilde{im}_{i,t}$ defined in equation (5) thus measures the contemporary refugee immigration that would have been the result had the settlement of these refugees and those who came during the baseline period been the same.

To implement the shift-share approach, source country c and baseline period t^0 must be chosen, and these two decisions are key for our methodological improvement. We discuss these two aspects in turn.

4.2.1 Definition of source country

In previous research, which mainly focuses on US and UK data, typically *all* immigration has been used in the analyses. Departing from this approach enables us to make significant contributions. For one thing, the immigrants' source country plays a major role in our aim to separate between ethnically and socio-economically induced flight and avoidance. The mechanism is likely different in a scenario where native flight occurs due to an increase of individuals from geographically and culturally distant countries but not due to immigration from more similar countries.

Furthermore, a unique feature of our data is the inclusion of the immigrants' *reason for immigration*[27]. This enables us to focus our analysis on refugee immigration, which is advantageous from an identification perspective. As noted, we argue that the settlement of refugees is less driven by pull factors *of the neighborhood.* In particular, for other forms of migration (e.g., labor and student migration), pull factors are to a larger extent city or neighborhood features in the destination country. Even though pull factors are not entirely irrelevant for refugee migration, they are national rather than local in nature, such as how liberal the asylum policies are. Consequently, by singling out refugees, we can restrict the analysis to push-type immigration driven by exogenous shocks.

Focusing on refugee migration also has a technical, methodological advantage. As we use neighborhood-fixed effects, identification in our shift-

[27] *Grund för bosättning* in Swedish.

share setting comes from variation within neighborhoods over time. By construction, the distribution of immigrants in the baseline years is constant. This means that identification over time stems from variation in the country-specific annual inflow of immigrants, which needs to be substantial in order to separate the predicted neighborhood level of immigration in t from that in $t + 1$. Now, country-specific flows of *refugees* indeed change heavily from year to year (e.g., due to conflict escalation). On the contrary, labor and student migration is more consistent over time.[28]

The information on the reason for immigrating is available from 1997, and our period of analysis is 1997–2010. Individuals entering Sweden with refugee status during this period arrive from all source countries. However, we drop those from OECD countries, since it is less likely that we observe flight based on migration from, for example, Germany or Denmark. Many of these individuals, furthermore, are likely Dublin cases with citizenship from other countries. We further drop Egypt and Eritrea. There are no/only 30 individuals arriving from Egypt/Eritrea in the baseline period,[29], thus implying that ϕ_{c,i,t^0} in equation (5) is not defined/will be highly imprecise. From the remaining source countries, at least 100 individuals or more arrived during the baseline period. The full list of these 34 countries and the frequency of refugees arriving in 1997–2010 are available in Table 15 in the Appendix.

4.2.2 Definition of baseline period

As seen in equation 5, the yearly national inflow of refugees from country c is scaled by the neighborhood proportion of immigrants from the same country during the baseline year. Since this scaling is based on historical behavior, it represents a problem for identifying whether historical immigrant settlement patterns were guided by (unobserved) sticky or fixed factors that are correlated with natives' migration decisions still to this day.[30]

[28] This is at least the case in the Swedish setting, where large spikes or changes over time are generally related to changes in refugee migration (see, for example, Figure 2).

[29] For a definition of baseline period, see the next section.

[30] This is different from the problem of long-term effects accumulating over time. Such dynamic effects arise if immigration causes flight in the baseline period, which, in turn, sets a long-term response in motion that might still be in the process of evolving in the year of the migration response in focus. This problem has been discussed and addressed by Jaeger et al. (2018). We estimate our model with their suggested solution in the Appendix, yielding no alternations to the main results presented in the paper.

This is a problem left unsolved in the existing migration literature applying the shift-share approach, and one of our methodological improvements is to exploit a refugee placement policy that was in effect in Sweden from 1985 to mid-1994. During this period, refugees could not decide for themselves where to settle but were assigned to a municipality through municipality-wide contracts, coordinated by the Immigration Board.[31] The number of municipalities with such a contract increased over the years, and by 1991, 277 out of 286 municipalities were part of the program.

One of the main aims of the refugee placement program was to break the concentration of immigrants to larger cities (mainly Stockholm, Gothenburg, and Malmö) to instead achieve a more even distribution of refugees over the country. This aim was successfully fulfilled, as illustrated in, for example, Figure 3B in Dahlberg et al. (2012) and Table 1 in Edin et al. (2004).

Motivated by this, for our baseline period t^0, we choose the early years in our data in which the refugee placement program was in place, 1990–93 (our data starts in 1990). We think that this adds credibility to the instrument since, thanks to the placement program, the immigrant settlement pattern across neighborhoods back then is less likely to be driven by endogenous factors also affecting the migration pattern of natives following contemporary immigration increases (compared to a situation in which the policy had not existed). This is especially true conditional on neighborhood-fixed effects and a set of neighborhood characteristics that we include in our estimation model. That is, we argue that the placement program can pick up possible time-varying unobservables not picked up by the fixed effects or the included time-varying covariates. Note that we do not require the program-generated placement of refugees across municipalities to be random.[32] What we argue is rather that, since the refugees received by the municipalities were effectively assigned to a specific apartment rather than they themselves choosing where to live, conditional on a set of characteristics, the variation in immigration to a neighborhood *within* a given municipality is likely to be exogenous to contemporaneous native flight and avoidance.[33]

[31] They were, however, allowed to move after the initial placement.

[32] In fact, it was not entirely random, but rather determined by, for example, available housing (Dahlberg et al., 2012) and even party constellation in the municipal council (Folke, 2014). For a lengthier discussion on the exogeneity of the placement program with respect to municipal characteristics, we refer to Dahlberg et al. (2012).

[33] A couple of caveats are to be noted here: First, for the years 1990–93, we have no information on the reason for immigration. Instead, we use all immigrants from the

We now proceed by specifying the details of our proposed estimation model, including the neighborhood characteristics upon which we condition the exogeneity assumption.

4.3 Estimation model

We analyze panel data where the year of refugee immigration, t in equations (3) and (4), refers to the years 1997–2009, while the migratory response by natives takes place in $t+1$, implying that the effects are estimated for the years 1998–2010.[34]

Besides instrumenting $im_{i,t}$ with $\tilde{im}_{i,t}$, our final estimation model differs from the basic equations in (3) and (4) in a few ways. First and foremost, the panel structure of the data means that we can include neighborhood-fixed effects,[35] μ_i, and thereby exploit *changes* in immigration shocks within neighborhoods over time. Second, we include linear, quadratic and cubic controls for population size (pop) in $t-1$. The purpose of these is to flexibly control for the fact that, in absolute terms, larger neighborhoods typically experience larger immigration inflows as well as larger population turnover in general. Third, since refugee immigration could be correlated with immigration for other reasons, which, in turn, could lead to further migratory responses, we control for all non-refugee immigration from the refugees' source countries in year $t-1$.[36] Fourth, we include time-fixed effects to control for aggregate shocks affecting all neighborhoods the same way in a given year.

countries defined as refugee countries in the later time period t. Second, the placement program became less strict after 1992, mainly due to an unexpected and large increase in immigration from former Yugoslavia. For efficiency reasons, we still include 1993 so as to increase the number of observations in our baseline period. It is also worth noting that when we apply the IV-design suggested in Jaeger et al. (2018)—an approach that does not rely on the exogeneity of the initial settlement—we still get the same results (see Appendix A).

[34] We focus on the short-term perspective of one year since, at least in a quantitative sense, the estimated effects of immigration become less reliable the longer the native response is allowed to take. The reason is that immigration during and after year t is likely to be correlated, implying that native migration measured later may either be longer-run responses to immigration in year t or short-run responses to immigration after year t.

[35] A neighborhood is defined as a so-called *SAMS*; see the following section.

[36] The main concern is that tied family migrants move to the same neighborhoods as the refugees, causing an additional effect on native migration. Since we primarily worry about tied family migration, we only control for other types of migration from the refugee countries we use to construct $im_{i,t}$. However, we have estimated a model with *all* other immigration as a covariate, with no important alterations to the baseline estimates. These results are available upon request.

Finally, we control for a set of time-varying socio-economic characteristics of the neighborhood (measured in $t-1$): average disposable income, number of students, per capita cost of social assistance, and number of public rental estates.[37]

Letting the vector $\mathbf{X}$ include the variables for non-refugee immigration and the socio-economic characteristics, the first stage in our IV approach is:

$$im_{i,t} = \gamma \tilde{im}_{i,t} + \sum_{p=1}^{3} \phi^p pop^p_{i,t-1} + \Gamma\mathbf{X} + \mu_i + \tau_t + \epsilon_{i,t} \tag{7}$$

The prediction $\widehat{im}_{i,t}$ from this first stage is then used in the two equations capturing the migratory response of the native population:

$$outflow_{i,t+1} = \beta^{out}\widehat{im}_{i,t} + \sum_{p=1}^{3} \delta^p pop^p_{i,t-1} + \Pi\mathbf{X} + \mu_i + \tau_t + \varepsilon^{out}_{i,t+1} \tag{8}$$

and

$$inflow_{i,t+s} = \beta^{in}\widehat{im}_{i,t} + \sum_{p=1}^{3} \delta^p pop^p_{i,t-1} + \Pi\mathbf{X} + \mu_i + \tau_t + \varepsilon^{in}_{i,t+s} \tag{9}$$

Our approach thus estimates effects of immigration on native migration within neighborhoods and over time. The identifying variation in immigration stems from contemporary year-to-year changes in the inflow of refugees from specific countries, weighted by the placement policy-induced immigrant settlement from several years before.[38]

[37]The reason we date all variables in $t-1$ is to avoid a bad control problem—that is, we control for things in fact constituting responses to/implications of immigration.

[38]A possible problem with the identifying variation used is that immigrants may sort into a few districts in Sweden. If so, it would be difficult to separate national and local immigration shocks from each other. In Table 11 in the Appendix, we show the situation for the five largest countries of origin in our sample—Iraq, Somalia, Yugoslavia, Bosnia, and Iran—which together make up more than 3/4 of the sample. Refugees from these countries spread across many neighborhoods (for example, Iranians were placed in 1,800 different neighborhoods, and no neighborhood took in more than 1.3 percent of the incoming Iraqis.

5 Data and descriptive statistics

In this section, we present the data, which is obtained from the GeoSweden database, as well as our definition of a "neighborhood." All data is collected and anonymized by Statistics Sweden and administered by the Institute for Housing and Urban Research at Uppsala University.

5.1 The GeoSweden database

The GeoSweden database is collected on a yearly basis, it is very comprehensive, and covers all individuals living in Sweden. It contains variables from several different registers, such as the education, income, and employment registers, as well as information on individual characteristics such as year and country of birth, marital status, number of children in the household, in addition to the individuals' level and type of education. It also contains pre-tax income from different sources, disposable income as well as various variables concerning the individual's employment.

What is particularly significant for this paper is the detailed geographical information on where individuals live, information on the date, from which country, and for which reason an individual immigrates to Sweden, as well as annual information on migration patterns within Sweden.

We define a neighborhood to be a so-called *SAMS* (Small Area for Market Statistics). A SAMS is a geographical unit defined by Statistics Sweden to obtain a countrywide division of municipalities into homogeneous areas. Sweden consists of approximately 9,200 SAMS with an average population of around 1,000 individuals. In our sample, we have excluded SAMS that were not traceable throughout the study period or that lack population at some point in time. This leaves us with 8,723 neighborhoods. The average number of SAMS per municipality is around 30 and the number of neighborhoods per municipality is highly correlated with the population of the municipality. We analyze the sensitivity of the first stage to the type of SAMS in Section 6.1.

5.2 Descriptives

Table 2 presents summary statistics of the variables used in the analysis, along with a clarifying description. The SAMS neighborhoods span between very small locations with only a couple of individuals to large neigh-

Table 2: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Key variables:					
Outflow	114,477	85.2	118	0	2,352
Inflow	114,470	85.2	121	0	2,716
Immigration (main)	114,478	0.82	4.7	0	313
Predicted immigration (instrument)	113,503	0.82	3.6	0	253
Control variables:					
Population	114,478	1019	1236	1	20,285
Students	114,478	53.1	107.5	0	2,642
Disposable income	114,478	155,870	53,849	-107,050	5,688,067
Social assistance	114,478	8,742	22,452	0	108,200
Other non-OECD imm.	114,478	2.4	9.4	0	588
Public rentals	113,681	2.1	5.2	0	408

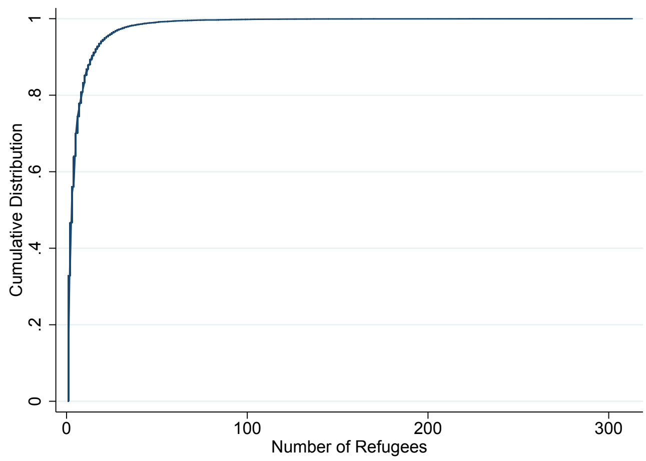
Outflow and Inflow measure the number of natives moving out of and into a given neighborhood in a given year. Immigration (main) is the main endogenous independent variable, measuring the annual number of refugees, and Predicted immigration is the instrument for this variable. Population denotes total SAMS population and students the number who receive some student contributions (majority of Swedish students). Disposable income and Social assistance are measured in SEK, other non-OECD immigration shows the number of non-refugee immigrants and Public rentals is the number of public rental estates. The unit of observation is SAMS-by-year, and the time span is 1997–2010. Note that disposable income is measured net of taxes, and from 2004; includes capital losses.

borhoods in the center of Stockholm with around 20,000 inhabitants. On average, around 85 natives move out of a neighborhood in a given year, which represents about 8 percent of the neighborhood population.

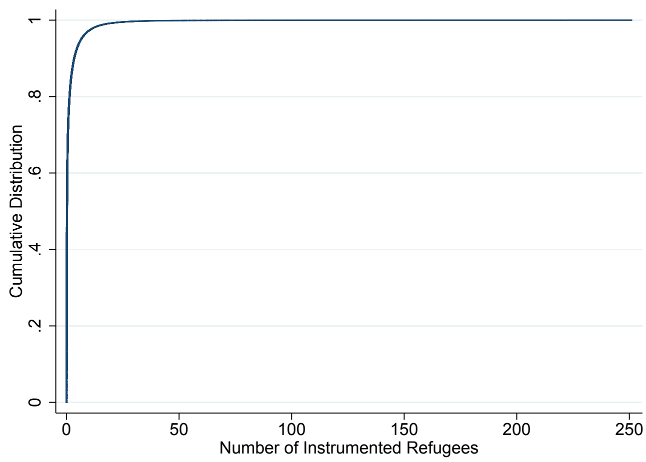
For the main endogenous immigration variable as well as its instrument (corresponding to $im_{i,t}$ and $\hat{im}_{i,t}$ in the above equations), the standard deviations are large relative to their means. This reflects the fact that roughly 85 percent of the observations contain zeros, which, in turn, is due to the fact that many SAMS are very small. To get a better sense of the variation in the data, Figure 3 shows the distribution of these two immigration variables, conditional on positive migration. As we can see, the majority of neighborhoods have a fairly low level of immigration. Half the neighborhoods received 3 people or less, while 90 percent received 14 or less. The figures also suggest that the two distributions are highly correlated. This is indicative of a strong instrument, which we below show to indeed be the case.

Figure 3: Distribution of actual and predicted immigration

(a) Actual number of immigrants



(b) Predicted number of immigrants



Note: The figures show the cumulative distribution of immigration, actual (panel a) and as predicted by the instrument (panel b), conditional on positive immigration. The unit of observation is SAMS-by-year, and the time span is 1997–2010.

Source: GeoSweden.

6 Results

We now turn to the results. After establishing in Section 6.1 that our instrumental variable works well in the first-stage regression, we provide the IV estimates of the effects of foreign immigration on native migration in Section 6.2. By focusing on homeowners, we study households that are indeed relatively able to move following increased immigration. Thoughts concerning mechanisms are discussed and tested in Section 6.3. The more constrained group of renters, finally, is studied in section 6.4.

6.1 First stage

Table 3 shows the baseline estimation of the first stage as specified in equation (7). For the years 1997–2010, the inflow of refugees to neighborhood i in year t is regressed on the inflow as predicted by equation (5). An estimate of 1 implies a perfect correlation; that is, a prediction based on the interaction of previous settlement patterns and current shocks of one more immigrant into neighborhood i in year t corresponds to an actual inflow of one more immigrant to that very neighborhood in that year. Since treatment is defined at the level of SAMS-by-year, our default approach is to cluster the standard errors at SAMS.[39]

Column 1 presents raw correlations, while column 2 adds fixed effects and control variables according to the preferred model based on the discussion in Section 4.3. In the latter column, we see that, conditional on last years' population, socio-economic and demographic characteristics of the neighborhood, non-refugee immigration, as well as year- and neighborhood-fixed effects, one additional predicted immigrant is associated with 0.6 actual immigrants. The coefficient is highly significant, meaning that the instrument clearly fulfills the relevance condition. The model is also very stable, as adding all control variables, including the fixed effects, does not affect the estimate much (from 0.67 to 0.60; cf. columns 1 and 2).

[39]For robustness, we have also re-estimated the model clustering at municipality.

Table 3: First-stage estimates

	(1) No controls	(2) Baseline specification
$\tilde{im}$	0.669***	0.603***
	(0.080)	(0.085)
Observations	113,503	104,251
Number of SAMS	8,731	8,710
F-stat	69.49	50.01
Adjusted R-squared	0.172	0.171

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on SAMS-level. Column 1 estimates the unconditional correlation, while column 2 estimates the first-stage according to the preferred specification as described in Section 4.3; it includes year and neighborhood fixed effects as well as linear, quadratic and cubic controls for population size, non-refugee immigration from the refugees' source countries, average disposable income, the number of students, the per capita cost of social assistance and the number of public rental estates. All covariates are at the neighborhood level and measured in year $t-1$.

As a robustness check, in Table 4 we run the first stage for several different subsamples. First, we remove the 10 percent of neighborhoods with the smallest population (fewer than 113 individuals) and the largest population (more than 2,043 individuals), respectively (see columns 1–2). Clearly, the estimations are more dependent on the larger neighborhoods. This is expected, as immigration is more consistent over time when it comes to larger neighborhoods. However, the coefficient is highly statistically significant in both subsamples.

Gothenburg (the second-largest city in Sweden) with its almost 800 neighborhoods is a clear outlier, as very few municipalities have over 100 neighborhoods, and Stockholm (the capital) has fewer than 200. We thus exclude Gothenburg, resulting in no major change in either power or significance (see column 3). Finally, it is interesting to see how the first stage depends on the number of immigrants. Since the majority of neighborhoods in a typical year did not receive any refugees, dropping the top 10 percent of the distribution of immigrated refugees (as in columns 1–2 for population) would be too much of a restriction. Instead, we drop the top 10 percent of the sample, *given positive immigration.* In practice, this implies any neighborhood receiving more than 14 immigrants. Just as when dropping neighborhoods with large populations, the first stage drops in power. Once

Table 4: Robustness of the first-stage estimate over different subsets of the sample

	(1) Excl. least pop. n'hoods	(2) Excl. most pop. n'hoods	(3) Excl. Gbg	(4) Excl. n'hoods with most im
$\tilde{im}$	0.611***	0.250***	0.603***	0.117***
	(0.086)	(0.029)	(0.089)	(0.011)
Observations	94,762	93,635	95,327	102,792
Number of SAMS	8,010	7,927	7,960	8,707
F-Stat	50.26	74.12	46.04	117.00
Adj. R-squared	0.180	0.0433	0.174	0.0400

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on SAMS-level. Column 1 excludes neighborhoods in the bottom decile of the population size distribution, column 2 excludes neighborhoods in the top decile of the population size distribution, column 3 excludes Gothenburg (Gbg) and column 4 excludes neighborhoods in the top decile in the distribution of received immigrants (given positive immigration). See Table 3 for details of the estimated model.

again, however, it is still highly significant (see column 4).[40]

The first stage may be concluded as being strong. The baseline estimate implies that an increase of 1 predicted refugee to a neighborhood is associated with 0.6 more actual refugees to the very same neighborhood. It is highly stable for the inclusion of fixed effects as well as several control variables. It is also robust to the exclusion of segments of the sample, although the prime part of the variation is identified through larger neighborhoods.

6.2 Native flight and avoidance: Average effects for high-mobility households

Moving on to the estimated native flight and avoidance effects, Table 4 presents results from estimating the second-stage equations of outflow and inflow, as specified in equations (8) and (9), respectively. Any native residing in neighborhood i on the last day of t but living in another neighborhood $-i$ on the last day of $t+1$ is counted as outflow from i. Analogously, any native residing in neighborhood i on the last day of $t+1$ but in another neighborhood $-i$ on the last day of t is counted as inflow into i.

[40]Since the placement program was made less strict after 1992, we have also estimated the first stage using only the years 1990 and 1991 for the baseline period. This still yields significant point estimates, but the instrument is not as powerful in terms of F-statistics.

Table 5: Second-stage estimates of native flight and avoidance

	(1) All natives	(2) Home owners
OUTFLOW		
im	0.0637	0.346**
	(0.158)	(0.156)
INFLOW		
im	-0.0847	0.170
	(0.183)	(0.136)
Observations	104,250	104,250
Number of SAMS	8,710	8,710

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on SAMS-level. Column 1 includes all natives and column 2 is restricted to native home owners. See Table 3 for details of the estimated model.

For comparative reasons, the left-hand column in Table 5 includes the results when all natives are included in the analysis and does not show signs of flight or avoidance, as both coefficients are small and statistically insignificant. This null result is interesting, as it differs from previous literature despite the clear analogy of being based on the full population irrespective of people's ability to move. Capturing residential preferences through flight and avoidance, however, is only possible if people are, in fact, mobile. This is true in any institutional setting, although what determines mobility varies. In Sweden, as explained above, renting rather than owning your home constitutes a significant obstacle to moving in a situation when the municipality has received and accommodated many immigrants.

As a way of getting closer to residential preferences and reactions to increased immigration, and following the discussion in section 3.1, we restrict the sample to natives owning their home. In the right-hand column in Table 5, it becomes clear that the insignificant aggregate effects conceal interesting heterogeneity. In particular, the estimated outflow effect among native homeowners is a statistically significant 0.35.[41] The interpretation of this coefficient is that, when a neighborhood receives one more immigrant than on average, 0.35 additional natives move out.[42]

[41] Results for natives renting their homes are provided in section 6.4.

[42] The mean outflow among homeowners is 39. As a sensitivity check, we have also estimated the outflow equations in Table 5 with more parsimonious specifications. When

The institutional context that leads us to separate our sample in renters and owners is thoroughly discussed and explained in the online appendix. In this appendix we also provide a robustness check, by splitting the Swedish housing market along local rental market "tightness". We use data on municipality housing queues and availability of rental apartments in 1998. We then run the outflow analysis in Table 5 separately for neighborhoods depending on the share of the rental stock available in the municipality. In support of our argument, the flight coefficient is larger for owners than for renters *only* in municipalities with few rental apartments available.

In contrast to outflow, there is no statistically significant inflow effect of increased immigration among homeowning natives. A possible explanation is that homeowners mostly notice and consequently react to increased immigration into the neighborhood they currently inhabit. Furthermore, a likely interpretation of the difference between the estimated flight and avoidance effects is that other immigrants and/or current renters are (at least partly) the buyers of the houses and apartments sold by the moving natives.

Long-run moves are more costly than short-run moves, not least as the former may require finding a new job. This implies that the observed flight estimate in Table 5 for homeowners ought to be mainly driven by moves to nearby neighborhoods. Table 6 shows that this is indeed the case; for moves within the same municipality, the estimated outflow effect among homeowners is similar to the overall counterpart in Table 5. In contrast, for moves out of the municipality, the estimated effect is closer to zero and statistically insignificant. This illustrates the importance of having data with a high geographical resolution in order to capture more nuanced residential preferences. The question is whether these preferences are formed along ethnic or socio-economic lines. This is the topic we now turn to.

including only neighborhood-fixed effects, we get a point estimate of 0.29 (standard error 0.137). When adding the population controls (but not the other socio-economic and demographic covariates), we get a point estimate of 0.39 (standard error 0.158). The results are also stable to including year×labor market region-specific fixed effects (to control for year- and region-specific shocks). When these fixed effects are added, we obtain a point estimate of 0.32 (standard error 0.153).

Table 6: Second-stage estimates of native flight within and out of the municipality, respectively, for home owners

	(1) Out of municipality	(2) Within municipality
im	0.0892	0.257**
	(0.067)	(0.110)
Observations	104,250	104,250
Number of SAMS	8,710	8,710

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on SAMS-level. Column 1 shows the results when conditioning the out-migration from a neighborhood on a move out of the municipality and column 2 shows the results when conditioning the out-migration from a neigfhborhood on a move to another neighborhood within the same municipality. See Table 3 for details of the estimated model.

6.3 Is native flight determined by ethnically based preferences?

The pronounced flight effect in the subsample of homeowners is interesting in its own right, in part as it potentially has implications for previous studies mostly having looked at aggregate flight effects—which have nevertheless been fairly in line with the effects in the group characterized as mobile above. We now make additional use of our data to consider the mechanism behind the estimated effects within this group.

Refugees come from a different ethnic background and typically also from groups located lower on the socio-economic ladder than the average native. If natives move due to increased immigration, they may thus do so either because they prefer ethnically homogenous neighborhoods and/or if they have preferences for socio-economic homogeneity. As discussed in Section 3.2, we want to examine whether the commonly assumed ethnic channel is supported by the data by grouping individuals according to their parental foreign background. While previous works have speculated as to whether the ethnic or the socio-economic channel is the driving force, (see Saiz and Wachter, 2011; Sá, 2015; Rathelot and Safi, 2014), to our knowledge, we are the first to explicitly approach this question with relevant data.

As a group, native Swedes with non-Western parents are on average ethnically more similar to the current immigrants, yet socio-economically more similar to natives with Swedish-born parents. This is the rationale for why we expect the relationship in equation (2) to apply if natives indeed do react based on the immigrants' ethnicity. Under such a scenario, estimated

Table 7: Second-stage estimates of native flight and avoidance among home owners with different parental background

	(1)	(2)	(3)	(4)
	All	Parental background:		
	natives	Native	Western	Non-Western
OUTFLOW				
im	0.009**	0.008**	0.017***	0.008**
	(0.004)	(0.004)	(0.006)	(0.004)
INFLOW				
im	0.004	0.004	0.005	0.008
	(0.004)	(0.004)	(0.005)	(0.005)
Observations	104,248	104,248	104,250	104,248
Number of SAMS	8,710	8,710	8,710	8,710

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors clustered on SAMS-level. Column 1 includes all native home owners, column 2 is restricted to native home owners with Swedish-born parents, column 3 is restricted to native home owners with at least one parent born in another Western country, and column 4 is restricted to native home owners with at least one parent born in a non-Western country. The dependent variables are standardized with its respective mean. See Table 3 for details of the estimated model.

native flight would be higher among natives with native parents than among natives with non-Western parents. If, on the contrary, flight is observed to a similar extent among all natives irrespective of parental background, then the main mechanism is more likely to be socio-economic.

We continue to focus on homeowners and construct three groups of native homeowners based on their foreign/ethnic background and provide outflow and inflow effects for these respective groups in columns 2–4 in Table 7 (column 1 reproduces the above average effect among homeowners). Column 2 contains those with native-born parents, column 3 contains those with at least one parent born in another Western country[43], and column 4 contains those with at least one parent born in a non-Western country. The average number of movers differs between the groups, and in order to facilitate comparisons, the dependent variable is standardized with its mean.

As seen in Table 7, all flight estimates are positive and statistically significant. When comparing the magnitude across columns, it is clear that the relative magnitude is very similar across the groups of natives with different

[43] Countries that are members of the OECD are defined as Western.

parental backgrounds. In particular, the estimated effects for natives with native-born and non-Western-born parents are strikingly similar. In both cases, the point estimates are around 0.008, which means that one more refugee into the neighborhood causes an additional outflow of 0.8 percent of the average annual number of movers in the respective groups. Meanwhile, there are no statistically significant effects on inflow. The point estimates are generally positive and thus, just as in Table 5, there is no evidence of avoidance.

Treatment (increased immigration) is defined at the level of SAMS-by-year, and as noted above, we thus cluster the standard errors by SAMS. Yet, the first phase of the placement program that defines our baseline period placed refugees in municipalities. Hence, we have re-estimated the model clustering at the municipality level. The number of clusters then decreases substantially, from around 8,700 to 290. Still, the first stage is hardly affected and remains statistically significant at conventional significance levels. For the second stage, the change in statistical significance varies. While the standard errors hardly change for the group of natives with parents born in other Western and non-Western countries, the estimate for natives with native-born parents is no longer significant with the municipality clusters.[44] All in all, we trust the estimates obtained from clustering at the SAMS level and note that despite some loss in precision, the comparison across groups of natives with a different parental background mainly holds also when clustering at the municipality level.

The fact that natives with different ethnic parental backgrounds display similar flight behaviors quite strongly suggests that their residential preferences are not mainly shaped along the ethnic dimension. Rather, to the extent that immigrants on average are (or are perceived to be) less educated and poorer, the socio-economic diversity of the neighborhood may be the dimension along which natives' residential preferences are shaped.[45] To study

[44]These results are available upon request. Natives with non-native parents are more likely to be concentrated to a few SAMS within a given municipality. Allowing for within-municipality correlation thus has little impact on the standard errors for these groups.

[45]Natives with native parents may differ in various ways from natives with parents born in non-Western countries. For example, as illustrated in Table 16 in the Appendix, those with a non-Western parental background have more limited financial and human capital resources. It is interesting that, despite having lower resources, their reaction to increased immigration is similar to that of natives with a Swedish parental background (cf. Table 7).

this further, we use our model to see how immigration affects the income and education level in the neighborhood, respectively, among homeowners in the same subgroups as analyzed in Table 7. The first stage is the same as before, while the second stage is now given by:

$$income_{i,t+1} = \beta^{inc}\widehat{im}_{i,t} + \sum_{p=1}^{3} \delta^{p} pop^{p}_{i,t-1} + \Pi'\mathbf{X} + \mu_i + \tau_t + \varepsilon^{inc}_{i,t+1} \qquad (10)$$

and

$$university_{i,t+1} = \beta^{univ}\widehat{im}_{i,t} + \sum_{p=1}^{3} \delta^{p} pop^{p}_{i,t-1} + \Pi'\mathbf{X} + \mu_i + \tau_t + \varepsilon^{univ}_{i,t+1}. \qquad (11)$$

That is, the outcome variable is replaced by the average disposable income, $income_{i,t+1}$, and the proportion of university-educated individuals, $university_{i,t+1}$, in the neighborhood. We define and estimate equations (10) and (11) for all homeowners (i.e., both native and non-native) as well as separately by the same subgroups as analyzed in Table 7. And, just as above, the dependent variable is standardized with its mean to facilitate comparisons across groups.[46]

The results, presented in Tables 8 and 9, show that the effect of increased immigration is that income as well as level of education decrease among all homeowners irrespective of their own or their parents' foreign background (although the income estimates for natives with foreign-born parents are statistically insignificant). Interestingly, the similar magnitudes of the effects for all homeowners and all native homeowners (cf. columns 1 and 2) imply that socio-economic segregation is driven by natives from higher socio-economic groups moving out rather than by immigrants from lower socio-economic groups moving in. These results further strengthen the conjecture that preferences are formed along socio-economic rather than ethnic lines.[47]

[46]Note that when estimating the effect on the level of education in the neighborhood (equation 11, Table 9), we add a lagged control for the number of university-educated individuals in $t-1$.

[47]However, note that we are not able to sort out the relative importance of socio-economic preferences versus budget constraints.

Table 8: Second-stage estimates of the average disposable income among different types of home owners

	(1)	(2)	(3)	(4)	(5)
		All	Natives by parental background:		
	All	natives	Native	Western	Non-Western
im	-0.00186**	-0.00192**	-0.00129*	-0.00146	-0.000562
	(0.000842)	(0.000871)	(0.000782)	(0.00113)	(0.000682)
Obs.	90,825	90,748	90,555	87,025	87,954
N'hoods	8,358	8,350	8,331	8,158	8,179

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on SAMS-level. The outcome variable is the neighborhood average disposable income in the respective groups. Column 1 includes all home owners, column 2 includes all native home owners, column 3 is restricted to native home owners with Swedish-born parents, column 4 is restricted to native home owners with at least one parent born in another Western country, and column 5 is restricted to native home owners with at least one parent born in a non-Western country. The dependent variables are standardized with its respective mean. See Table 3 for details of the estimated model.

Table 9: Second-stage estimates of the share university educated among different types of home owners

	(1)	(2)	(3)	(4)	(5)
		All	Natives by parental background:		
	All	natives	Native	Western	Non-Western
im	-0.00346***	-0.00330***	-0.00348***	-0.00349***	-0.00156***
	(0.000859)	(0.000842)	(0.000892)	(0.00122)	(0.000567)
Observations	95,551	95,551	95,551	95,551	95,551
Number of sams	8,709	8,709	8,709	8,709	8,709

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on SAMS-level. The outcome variable is the share of the n'hood population that are home owners and has at least some university education. Column 1 includes all home owners, column 2 counts only native home owners, column 3 is restricted to native home owners with Swedish-born parents, column 4 is restricted to native home owners with at least one parent born in another Western country, and column 5 is restricted to native home owners with at least one parent born in a non-Western country. The dependent variables are standardized with its respective mean. Covariates are the same as described in Table 3, with the addition of the lagged number of university educated.

6.4 Flight and avoidance among renting natives

The conclusion that changing socio-economic characteristics rather than ethnic heterogeneity seems to be the primary channel explaining natives' migration behavior pertains to the above analysis focusing on homeowners. In the current setting, as argued, they are the ones who are indeed able to react to increased immigration. Table 10 instead presents flight and avoidance estimates for natives in publicly provided rental apartments, again grouped

Table 10: Second-stage estimates of native flight and avoidance among renters with different parental background

	(1)	(2)	(3)	(4)
	All	Parental background:		
	Natives	Native	Western	Non-Western
OUTFLOW				
im	-0.00394	-0.0126***	-0.013	0.050***
	(0.00496)	(0.00487)	(0.0108)	(0.013)
INFLOW				
im	-0.00425**	-0.00616***	-0.00152**	0.010
	(0.00205)	(0.00207)	(0.00076)	(0.009)
Observations	104,250	104,250	104,250	104,250
Number of SAMS	8,710	8,710	8,710	8,710

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on SAMS-level. Column 1 includes all native renters, column 2 is restricted to native renters with Swedish-born parents, column 3 is restricted to native renters with at least one parent born in another Western country, and column 4 is restricted to native renters with at least one parent born in a non-Western country. The dependent variables are standardized with its respective mean. See Table 3 for details of the estimated model.

according to parental foreign background. As seen, the effects among these groups are generally negative. That is, increased immigration leads to fewer renters moving out, which, in turn, leaves less room for others to move in. This is in line with the above argument (see Section 3.1) that increased competition for public rentals in the wake of an inflow of immigrants causes lock-in effects among the initial renters.

An exception to the negative coefficients is the effect for natives with non-Western parents, who instead react by moving out of the neighborhood (cf. column 4 in Table 10). One possible explanation for this result is that this group begins a housing career when new immigrants arrive. Regardless, the estimated flight and avoidance behaviors among renters are difficult to reconcile with an ethnically based mechanism.

7 Concluding remarks

In this paper, we have applied detailed and comprehensive register data to a refined shift-share method to answer whether native flight and avoidance constitute important phenomena in Sweden. In particular, by using information in the data enabling us to identify native-born individuals who to different degrees are ethnically close to the newly arrived refugee immigrants (as defined via their heterogeneous parental background), we have examined whether there is support for the hypothesis that natives prefer to live in ethnically homogeneous neighborhoods. Our study spans the period 1990–2010, which is an important and interesting period to study based on at least two factors: first, there was a large increase in refugee-based immigration to Sweden during this time period and, second, in the early part of the period, there was a refugee placement policy in Sweden that may arguably be used to improve the shift-share instrument.

By using push-driven refugee immigration to Sweden interacted with a settlement pattern of their countrymen in the early 1990s partially generated by this state-run placement policy, we reach three main conclusions.

First, we find no evidence of either native flight or native avoidance when studying the entire population.

Second, a conclusion distinguishing between mobile/immobile households when examining the effects of immigration on native migration is important but is something the previous literature has not been able to do. When we look specifically at households with a high ability to move following increased immigration (homeowners in the Swedish context), we do detect native flight. That is, homeowning natives move out of neighborhoods experiencing an increase in immigration.[48] Among natives identified as having a low ability to move following increased immigration (here, renters), we do not find any such effects. Another interesting finding is that, for mobile households, the flight occurs to neighborhoods within the same municipality. This points to the importance of having granular geographic data when examining the effects of immigration on natives' migration behavior and that studies on larger geographical units may have missed important flight behaviors.

Third, we find that the "ethnic distance" between native-born individuals and newly arrived refugees does not matter for observed flight behaviors: all Swedish-born individuals react in a very similar fashion to increased immigration, irrespective of their parental foreign background. Hence, preferences for ethnically homogeneous neighborhoods do not seem to be the

[48] We do not find evidence of native avoidance, meaning that natives do not move into these neighborhoods to a lesser extent. A possible interpretation of this discrepancy is that natives mostly notice and consequently react to increased immigration into the neighborhood where they currently reside.

main channel causing flight. Rather, our analyses consistently indicate that natives have preferences for socio-economically homogeneous (or "better") neighborhoods.

If political decision-makers want to initiate policies to combat segregation, it is important to identify the mechanisms behind observed changes in natives' migration behavior, as successful policies will likely differ depending on whether the main channel is ethnically or socio-economically based. This paper suggests the former. However, more research is needed before any firm policy conclusions can be drawn.

Several future extensions are of interest. First, while the one-year lag allows us to identify more precise quantitative causal effects, we acknowledge that this focus potentially misses flight behaviors occurring after a longer period of time. Studying longer time lags thus constitutes an interesting follow-up. Second, our results indicate that the ethnically based tipping point literature might have focused on the wrong trait. This calls for studying tipping points along dimensions other than ethnicity, such as socio-economic traits. Finally, an alternative way of channeling any preferences for homogeneity is via school choices (rather than via residential choices). If parents perceive that school quality is dropping due to an increased minority presence, an exodus from the neighborhood school could occur. Hence, increasing school segregation also represents an interesting topic for future research.

References

Accetturo, A., F. Manaresi, S. Mocetti, and E. Olivieri (2014): "Don't stand so close to me: The urban impact of immigration," *Regional Science and Urban Economics*, 45, 45–56.

Aldén, L., M. Hammarstedt, and E. Neuman (2015): "Ethnic Segregation, Tipping Behavior, and Native Residential Mobility," *The International Migration Review*, 49, 36–69.

Altonji, J. G. and D. Card (1991): "The Effects of Immigration on the Labor Market Outcomes of Less-Skilled Natives," in *Immigration, Trade, and the Labor Market*, University of Chicago Press, 201–234.

Andersson, R., L. Magnusson Turner, and E. Holmqvist (2010): "Contextualising Ethnic Residential Segregation in Sweden: Welfare, Housing and Migration-Related Policies," Country report for sweden; norface program on migration, KOF.

Balkan, B., E. O. Tok, H. Torun, and S. Tumen (2018): "Immigration, Housing Rents, and Residential Segregation: Evidence from Syrian Refugees in Turkey," Discussion Paper 11611, IZA.

Basten, C. and M. Siegenthaler (2013): "Do Immigrants Take or Create Residents' Jobs? Quasi-Experimental Evidence from Switzerland," Working Paper 335, KOF.

Boustan, L. P. (2010): "Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration," *The Quarterly Journal of Economics*, 125, 417–443.

——— (2013): "Local Public Goods and the Demand for High-income Municipalities," *Journal of Urban Economics*, 76, 71–82.

Braakmann, N. (2019): "Immigration and the Property Market: Evidence from England and Wales," *Real Estate Economics*, 47, 509–533.

Card, D. (1990): "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial & Labor Relations Review*, 43, 245–257.

Card, D. and J. DiNardo (2000): "Do Immigrant Inflows Lead to Native Outflows?" *American Economic Review*, 90, 360–367.

Card, D., A. Mas, and J. Rothstein (2008): "Tipping and the Dynamics of Segregation," *The Quarterly Journal of Economics*, 123, 177–218.

Chalfin, A. and M. Levy (2013): "The Effect of Mexican Immigration on the Wages and Employment of U.S. Natives: Evidence from the Timing of Mexican Fertility Shocks," Mimeo UC Berkeley.

Dahlberg, M., K. Edmark, and H. Lundqvist (2012): "Ethnic Diversity and Preferences for Redistribution," *Journal of Political Economy*, 120, 41–76.

Edin, P.-A., P. Fredriksson, and O. Aaslund (2003): "Ethnic Enclaves And The Economic Success Of Immigrants: Evidence From A Natural Experiment," *The Quarterly Journal of Economics*, 118, 329–357.

Edin, P.-A., P. Fredriksson, and O. Åslund (2004): "Settlement Policies and the Economic Success of Immigrants," *Journal of Population Economics*, 17, 133–155.

Farley, R., H. Schuman, S. Bianchi, D. Colasanto, and S. Hatchett (1978): "'Chocolate City, Vanilla Suburbs': Will the Trend Toward Racially Separate Communities Continue?" *Social Science Research*, 7, 319–344.

Farley, R., C. Steeh, M. Krysan, T. Jackson, and K. Reeves (1994): "Stereotypes and Segregation: Neighborhoods in the Detroit Area," *American Journal of Sociology*, 100, 750–780.

Folke, O. (2014): "Shades of Brown and Green: Party Effects in Proportional Election Systems," *Journal of the European Economic Association*, 12, 1361–1395.

Jaeger, D. A., J. Ruist, and J. Stuhler (2018): "Shift-Share Instruments and the Impact of Immigration," Working Paper 24285, National Bureau of Economic Research.

Rathelot, R. and M. Safi (2014): "Local Ethnic Composition and Natives' and Immigrants' Geographic Mobility in France, 1982–1999," *American Sociological Review*, 79, 43–64.

Sá, F. (2015): "Immigration and House Prices in the UK," *The Economic Journal*, 125, 1393–1424.

Saiz, A. (2007): "Immigration and Housing Rents in American Cities," *Journal of Urban Economics*, 61, 345–371.

Saiz, A. and S. Wachter (2011): "Immigration and the Neighborhood," *American Economic Journal: Economic Policy*, 3, 169–188.

Schelling, T. C. (1971): "Dynamic Models of Segregation," *Journal of Mathematical Sociology*, 1, 143–186.

Shertzer, A. and R. Walsh (forthcoming): "Racial Sorting and the Emergence of Segregation in American Cities," *Review of Economics and Statistics*.

Wang, Y. (2011): "White Flight in Los Angeles County, 1960–1990: A Model of Fuzzy Tipping," *The Annals of Regional Science*, 47, 111–129.

Zimmermann, K. F. (1996): "European Migration: Push and Pull," *International Regional Science Review*, 19, 95–128.

Appendix

A Distribution across neighborhoods

A possible issue related to the instrument used is that immigrants may sort into a few districts in Sweden. If so, it would be difficult to separate national and local immigration shocks from each other. In Table 11, we show the situation for the five largest sending countries in our sample—Iraq, Somalia, Yugoslavia, Bosnia, and Iran—which together make up more than 3/4 of the sample. The table has four columns. In the first, we count the number of neighborhoods (SAMS) that the different nationalities were settled in during the baseline years 1990–1993. As we can see, there is no tendency for any of the sending countries to sort into a few neighborhoods. Rather, the refugees are well-spread across many neighborhoods. The second and third columns show the mean and standard deviation of ϕ_{c,i,t^0} for the neighborhoods with any settlement during the baseline years. The last column shows the maximum value of ϕ_{c,i,t^0}. In the case of Iran, for example, the SAMS hosting the largest number of Iranians in 1990–1993 took in a mere 1.33 % of all Iranians arriving during the time period. This number never exceeds 2.38 %, which it does with regard to Iraqis.

Table 11: Distribution of nationalities over neighborhoods during the baseline years (1990–1993).

Country	(1) N	(2) Mean	(3) Std	(4) Max
Iraq	1,227	0.08%	0.14%	2.38%
Somalia	829	0.12%	0.18%	1.96%
Yugoslavia	1,855	0.05%	0.08%	1.10%
Bosnia	1,312	0.08%	0.14%	1.72%
Iran	1,827	0.05%	0.09%	1.33%

B Applying the IV estimator in Jaeger et al. (2018)

Jaeger et al. (2018) criticize the shift-share instrument for failing to account for dynamic effects. Their example is based on wages but extends to any setting including possible dynamics over time.

Assume an immigration shock to neighborhood i during the baseline period t_0. This triggers a short-term effect of native flight. This native flight may put other forces in motion, such as decreasing house prices, continued flight, or possibly a mean decline of house prices. The shift-share instrument uses correlation over time in immigrants' location patterns. Potentially, this instrument thus measures both the *short-term* effect of immigration, as well as the continued *dynamic process*. The resulting estimates are then difficult to interpret since they do not solely capture the effect of contemporary immigration.

The solution in Jaeger et al. (2018) is to add a lag to the model and estimate the effect of both immigration in $t-1$ and t. Since both are endogenous, their solution implies two first stages:

$$im_{i,t} = \gamma_{1,1}\tilde{im}_{i,t} + \gamma_{1,2}\tilde{im}_{i,t-1} + \sum_{p=1}^{3}\phi_p pop_{i,t-1}^{p} + \mathbf{X}\Gamma' + \mu_i + \tau_t + \epsilon_{i,t} \quad (12)$$

$$im_{i,t-1} = \gamma_{2,1}\tilde{im}_{i,t} + \gamma_{2,2}\tilde{im}_{i,t-1} + \sum_{p=1}^{3}\phi_p pop_{i,t-1}^{p} + \mathbf{X}\Gamma' + \mu_i + \tau_t + \epsilon_{i,t} \quad (13)$$

The two first stages in equations (12) and (13) then give the following second-stage equations for outflow and inflow, respectively:

$$outflow_{i,t+1} = \beta_1^{IV}\widehat{im}_{i,t} + \beta_2^{IV}\widehat{im}_{i,t-1} + \sum_{p=1}^{3}\delta_p^{IV} pop_{i,t-1}^{p} + \mathbf{X}\Pi^{IV'} + \eta_i + \lambda_t + \varepsilon_{i,t+1}^{IV} \quad (14)$$

$$inflow_{i,t+1} = \beta_1^{IV}\widehat{im}_{i,t} + \beta_2^{IV}\widehat{im}_{i,t-1} + \sum_{p=1}^{3}\delta_p^{IV} pop_{i,t-1}^{p} + \mathbf{X}\Pi^{IV'} + \eta_i + \lambda_t + \varepsilon_{i,t+1}^{IV} \quad (15)$$

Table 12 shows the results from running the first stages in equation 12 and 13. The F-statistics indicate that the instrument is strong in both cases. The second stage for outflow is seen in Table 13 and for inflow in Table 14. Comparing these estimates to the baseline estimates in Table 7, it is clear that the results are robust to this addition to the model. In other words, dynamic effects do not appear to be a threat to our short-term estimates.

Table 12: First-stage estimates according to Jaeger et al. (2018)

	(1) im_{it}	(2) im_{it-1}
$\tilde{im}_{it}$	0.685***	-0.0148
	(0.093)	(0.018)
$\tilde{im}_{it-1}$	-0.179***	0.627***
	(0.030)	(0.089)
Observations	104,772	104,772
Number of SAMS	8,731	8,731
F-Stat	30.47	24.99

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on SAMS-level. Column 1 estimates the first-stage equation 12 and column 2 estimates the first-stage equation 13. See Table 3 for details of the estimated model.

Table 13: Second-stage estimates of native flight among home owners with different parental background according to Jaeger et al. (2018)

	(1)	(2)	(3)	(4)
	All	Natives by parental background:		
	natives	Native	Western	Non-Western
im_{it}	0.00826***	0.00716**	0.0176***	0.0081**
	(0.00319)	(0.00311)	(0.0056)	(0.0038)
im_{it-1}	0.00143	0.00185	-0.0014	0.0008
	(0.00331)	(0.00325)	(0.0046)	(0.0044)
Observations	104,250	104,250	104,250	104,250
Number of SAMS	8,710	8,710	8,710	8,710

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on SAMS-level. See Table 3 for details of the estimated model and Table 7 for details across columns.

Table 14: Second-stage estimates of native avoidance among home owners with different parental background according to Jaeger et al. (2018)

	(1)	(2)	(3)	(4)
	All	Natives by parental background:		
	natives	Native	Western	Non-Western
im_{it}	0.00398	0.00348	0.00464	0.0070
	(0.00280)	(0.00270)	(0.0047)	(0.0046)
im_{it-1}	0.000900	0.000689	0.00099	0.0024
	(0.00383)	(0.00378)	(0.0054)	(0.00444)
Observations	104,248	104,248	104,248	104,248
Number of SAMS	8,710	8,710	8,710	8,710

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on SAMS-level. See Table 3 for details of the estimated model and Table 7 for details across columns.

C Additional tables

Table 15: Frequency of refugees arriving 1997–2010.

Iraq	40,537	43.26%	43%
Somalia	11,597	12.37%	56%
Yugoslavia	8,345	8.90%	65%
Bosnia	6,727	7.18%	72%
Iran	5,105	5.45%	77%
Afghanistan	4,347	4.64%	82%
Syria	3,954	4.22%	86%
Russia	2,676	2.86%	89%
Lebanon	2,563	2.73%	92%
Thailand	1,225	1.31%	93%
Ethiopia	1,142	1.22%	94%
Croatia	887	0.95%	95%
Colombia	736	0.79%	96%
India	683	0.73%	97%
Peru	520	0.55%	97%
Bangladesh	469	0.50%	98%
Pakistan	468	0.50%	98%
China	269	0.29%	98%
Uganda	187	0.20%	99%
Romania	165	0.18%	99%
Bolivia	164	0.17%	99%
Vietnam	160	0.17%	99%
Algeria	125	0.13%	99%
Sri Lanka	115	0.12%	99%
Poland	108	0.12%	100%
Morocco	86	0.09%	100%
Tunisia	78	0.08%	100%
Latvia	71	0.08%	100%
Bulgaria	49	0.05%	100%
Estonia	38	0.04%	100%
Philippines	36	0.04%	100%
Gambia	31	0.03%	100%
Argentina	29	0.03%	100%
Slovenia	12	0.01%	100%
Brazil	11	0.01%	100%

Number of refugees per emigration country who got a residence permit 1997–2010.

Table 16: Financial and human capital resources among native Swedes (home owners) with different parental background (2010)

	Parental background:			
	Native	Western	Non-Western	Total
Disposable income	2346.2	2308.6	1800.7	2271.5
	(5090.5)	(3206.1)	(2932.7)	(4741.6)
Univ. education	0.338	0.342	0.212	0.322
	(0.473)	(0.474)	(0.409)	(0.467)
Social assistance	0.012	0.021	0.009	0.012
	(0.107)	(0.142)	(0.092)	(0.108)

Disposable income is measured in 100 SEK, university education and social assistance are defined as shares. Variables are defined in 2010.