WORKING PAPER 2018:12

Migrating natives and foreign immigration:

Is there a preference for ethnic residential homogeneity?

Henrik Andersson
Heléne Berg
Matz Dahlberg

IFAU INSTITUTET FÖR ARBETSMARKNADS- OCH UTBILDNINGSPOLITISK UTVÄRDERING

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala.

IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

More information about IFAU and the institute's publications can be found on the website www.ifau.se

ISSN 1651-1166

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

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

September 21, 2018

Abstract

In this paper we investigate the migration behavior of the native population following foreign (refugee) immigration, with a particular focus on examining whether there is support for an ethnically based migration response. If ethnicity is the mechanism driving the change in natives' migration behavior, our maintained hypothesis is that native-born individuals who are ethnically similar to arriving refugees should not change their migration behavior to the same extent as native-born individuals with native-born parents (who are ethnically quite different from refugees). Using rich geo-coded register data from Sweden, spanning over 20 consecutive years, we account for possible endogeneity problems with an improved so-called "shift-share" approach; in particular, our strategy combines policy-induced initial immigrant settlements with exogenous contemporaneous immigration as captured by refugee shocks. We find no evidence of neither native flight nor native avoidance when studying the full population. We do, however, find native flight among individuals who are expected to be more mobile, and within this group, we find that all natives, irrespective of their parents' foreign background, react similarly to increased immigration. Our results therefore indicate that preference for ethnically homogeneous neighborhoods may not be the dominant channel inducing flight. The estimates instead indicate that immigration leads to more socio-economically segregated neighborhoods. This conclusion may have implications for the ethnically based tipping point literature.

Keywords: Immigration; Native migration; Flight; Avoidance; IV estimation
JEL classification: C26; J15; R23

---

[a]We are grateful to Leah Platt Boustan, Mattias Engdahl, Jon Fiva, Florian Morath, Albert Saiz, Matti Sarvimäki, Håkan Selin, Susanne Urban 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 three Norface meetings at Rosersberg castle, 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.

Contents

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 accommodate all 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 settle). Ultimately, the aim is to deduce from estimated migration responses whether natives prefer *ethnically* homogeneous neighborhoods. We approach this task by developing the so-called "shift-share" method into, in several ways, 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 maintained hypothesis in the literature on "white flight" is that migration responses are due to preferences for ethnically homogeneous neighborhoods (see, e.g., Saiz, 2007; Boustan, 2010; Saiz and Wachter, 2011; Sá, 2014). But newly arrived immigrants hold a number of different characteristics other than their ethnicity; the average refugee does, for example, typically have a lower education level and lower income than the native population. Which trait do the natives actually react on? Do they react on the ethnicity of the immigrants, as typically hypothesized in the earlier literature, or on the socio-economic part?[1]

Thanks to comprehensive, detailed register data, we contribute in this paper by, aside from studying the *if*, examining the validity of the pre-

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

sumed ethnicity channel in a way that earlier literature has not been able to do. In particular, our data allows us to identify natives with different parental foreign background. Because many native-born individuals with non-Western parents are ethnically quite similar (in terms of country of origin) to current immigrants, yet in many cases socio-economically more similar to native-born individuals with Swedish-born parents, we use the parental information to explore the validity of the ethnicity channel. By estimating the migration response of natives conditioning on their parents' country of birth, we can examine whether there is support for the hypothesis that residential preferences are formed along an ethnic dimension.

The paper also contains several methodological improvements. Our data holds information on each individual immigrant's reason for residence permit—whether or not he or she arrived as a refugee, a tied mover or a labor migrant. This is a unique feature that allows us to make a distinct methodlogical improvement to related studies. In particular, we focus on refugees, which, besides being a highly topical and interesting group to study, arguably is more exogenous to the characteristics of the receiving city or neighborhood, as compared to labor, student or family migration.[2]

Much of the previous literature on white flight has focused on the US[3], in which often all immigration has been chategorized as one common treatment. The generalization of all immigration as one concept imply that nuanced mechanisms may be lost. It can also be problematic from an empirical point of view. A large share of immigrants to the US are *pulled* to specific places[4], whereas refugees, tend to have been *pushed* from their home country by wars and other catastrophes.[5] Increases in the US type immigration could therefore, to a larger extent than the refugee immigration that we focus on, be a function of regional chocks and pull factors, which affect

[2] *Refugees* in our paper includes 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.

[3] See in particular Farley et al. (1978), Farley et al. (1994), Boustan (2010), Saiz and Wachter (2011) and Wang (2011).

[4] 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`).

[5] Zimmermann (1996) provides a stylized economic definition of push and pull migration.

both immigrants but potentially also the behavior of natives.[6]

We identify the causal effect of foreign immigration on the residential choice of natives by combining (i) contemporary refugee migration into Sweden with (ii) previous immigrant settlement patterns resulting from a refugee placement policy that was in place in the earliest 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 to a municipality by the Migration Board. We argue that this policy-generated settlement is yet another improvement to existing studies. The rationale for this is that settlement patterns of immigrants from the early 1990s, who subsequently attracted more recent push-driven refugee migrants, are more likely to be uncorrelated with neighborhood characteristics that matter for natives' residential preferences than what would have been the case in the absence of the policy.

Our panel data 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 to the typical shift-share instrument used earlier in the literature, where both initial immigrant settlement as well as contemporary immigration are taken directly into the analysis and therefore likely to be endogenous to the outcome (see, e.g., Altonji and Card, 1991; Card and DiNardo, 2000; Saiz, 2007; Carl and Siegenthaler, 2013; Chalfin and Levy, 2013; Sá, 2014).[7]

A final contribution is that we acknowledge that, due to various constraints, far from all individuals are able to react on their residential preferences following an increase in immigration. Consequently, we focus the analysis on households characterized as having a high possibility to move.[8] This is in accordance with assumptions made in theoretical models on the effects of immigration on native migration—but is a previously neglected aspect empirically—and it turns out to matter greatly for the results.

[6]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.

[7]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; this yields very similar results.

[8]In our setting, we will define mobile actors as home owners rather than renters, as the rental market in Sweden is characterized by long (sometimes extreme) queues, and renters in many cases compete for the same apartments as the newly arrived refugees.

Many of the value-added features in this paper can only be implemented thanks to the detailed information in our data from the GeoSweden database. This is a database that covers the full Swedish population since 1990. Some of the valuable aspects of the data have already been covered; in particular, that we on a year-to-year basis can identify immigrants that are granted a residence permit in Sweden based on refugee reasons, as well as natives who indeed are mobile. Second, for each immigrant living in Sweden, there is information on the country of origin. Last, all variables come as an annual panel covering relatively small neighborhoods. While the panel structure allows for the fixed effects as mentioned above, the fine geographical resolution means that we can capture more nuanced residential preferences, as we are able to observe relatively short moves which are likely less costly than migration between larger units such as metropolitan areas.

Apart from the literature that *directly* estimates the extent to which the residential choices of natives are affected by immigration[9], our paper is related to an influential literature that has *indirectly* studied the response of natives to increased immigration by estimating effects on house prices (Saiz, 2007; Saiz and Wachter, 2011; Sá, 2014) and wages (Card, 1990; Altonji and Card, 1991).

The paper is also closely linked to the tipping-point literature that estimates at which potential share 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. However, to relate the results to previous work, we do also provide a set of tipping point-type estimates where we condition on initial share of immigrants. Finally, complementing the studies of the effects *of* residential segregation (Edin et al., 2003), our focus is on effects of immigration *on* residential segregation.

We reach four main conclusions. First, we do not find any evidence of neither native flight nor native avoidance when studying the full population—that is, irrespectively of potential mobility.

Second, we find that distinguishing between households with high/low possibility to move following an increased immigration is important; when

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

studying the group of natives identified as having high possibilities to move, we estimate significant flight responses.[10] In contrast, there are no effects of increased immigration on the migration behavior among natives identified as having a low possibility to move. The pronounced flight effect in the sub-sample of mobile natives potentially has implications for the interpretation in existing, related studies that—much due to data limitations—only have looked at aggregate (average) effects.

Third, we find that all natives, irrespective of their parents' foreign background, react similarly to increased immigration. The likely interpretation of this is that a preference for ethnically homogeneous neighborhoods is not the dominant channel causing flight. Instead, our analyses indicate that natives have preferences for socio-economically homogeneous (or, "better") neighborhoods.

Finally, conditioning on the initial immigrant share and thereby relating to tipping point estimates, we again find similar patterns irrespective of the natives' parental foreign background. This is thus further evidence against the ethnicity channel, indicating that the tipping point literature might have focused on the wrong trait.

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 response, and, in particular describes our idea for examining whether there is any support for 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. Finally, we conclude.

2 Immigration to Sweden

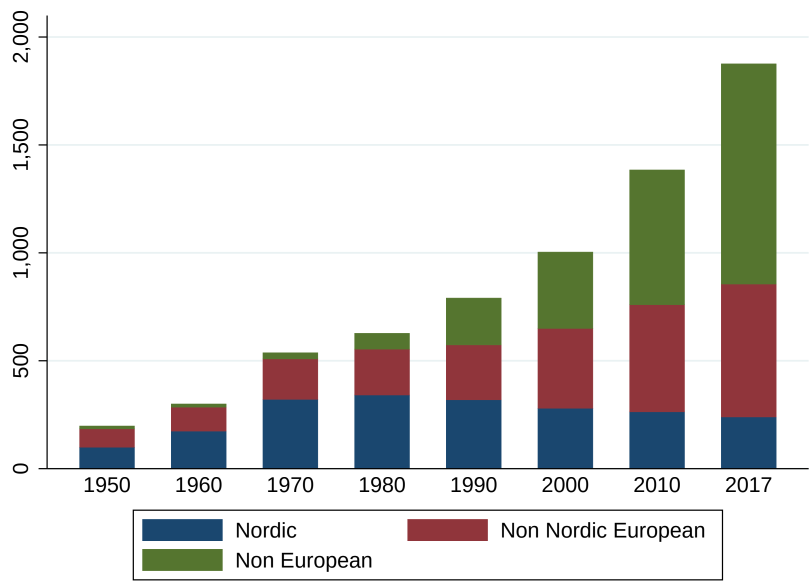
The size and character of immigration to Sweden have changed over the last decades. In 1970, less than seven percent of the Swedish population were born in another country[11], and of those the large majority had arrived as

[10]It is noteworthy that we find evidence of native flight, but not native avoidance. A possible interpretation is that natives mostly notice and consequently react on increased immigration into the neighborhood where they currently live.

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

labor immigrants from another Nordic or European country in the 1950s and 1960s. From the late 1970s/early 1980s, the immigration changed character; going from being mainly labor-induced, more refugees started to come. 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 from Figure 1. While the share with roots in the Nordic countries is decreasing over time, the share originating from 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 are born outside of Europe.

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



Notes: Y-axis in units of thousands.

Source: Statistics Sweden.

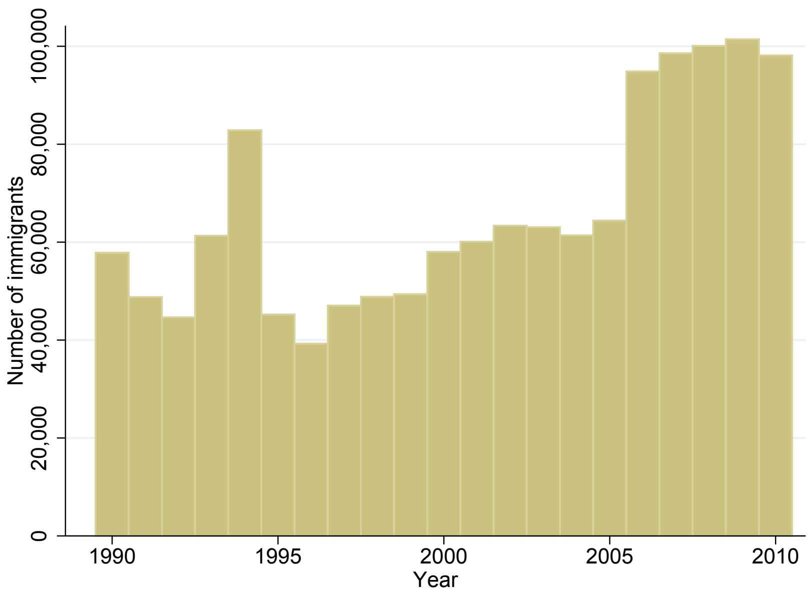
Compared to most other European countries, Sweden has a relatively large share of foreign-borns. According to statistics from Eurostat,[12] in 2010, 47 million individuals in the EU27 were not born in the country in which

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

they resided. This amounted to almost ten percent of the total population. The majority of these, slightly more than 31 million, were born outside of the European Union. There is however a large variation in these numbers across the union, ranging from Poland (with 1.2 percent foreign-born), Czech Republic, Hungary and Finland (all with around 4 percent foreign-born) to Austria (15.2 percent), Sweden (14.3 percent), Spain (14 percent) and Germany (12 percent).

Switching focus from stocks to flows, the annual immigration to Sweden during the period that we study, 1990–2010, is shown in Figure 2. Up until 2006, typically 50–60,000 individuals came each year.[13] Then, from 2006 and onward, there has been a discrete increase in the number of immigrants, with a yearly average of around 100,000.

Figure 2: Total immigration to Sweden, 1990–2010



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

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

3 Potential reactions of natives

The literature on residential segregation typically studies two types of reactions of the majority population to 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).[14] For the analysis in this paper, it is necessary to distinguish between "native" and "white". The concepts of *native* flight and avoidance are different 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 how to classify individuals' background, rather than focusing on racial diversity, we will study flight and avoidance due to increased diversity in terms of country of origin. Consequently, we refer to the potential reaction of the majority population as *native* flight and avoidance.

Our main definition of native is everyone born in Sweden. This means that our native group is quite heterogeneous in terms of their parental foreign background, a feature which we use in an attempt to disentangle the mechanisms behind the observed migration responses of natives. We continue with explaining this in more detail.

3.1 Preference-based mechanisms

Why would increasing immigration affect natives' location decisions? Scholars within sociology, economics and geography have lifted several potential mechanisms, where the dominating 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 of aversion against perceived social unrest (Farley et al., 1978, 1994). Economists have incorporated this thought 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:[15]

[14]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.

[15]See equation (9) in Sá.

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

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 consumption of housing services (h) and of other goods (x), and δ captures natives' 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 that the newly arrived immigrants carry? In Sweden, newly arrived immigrants are to a large extent refugees. Particularly in the first years in the country, the average refugee has lower income and is less educated than the native population in the neighborhoods in which they 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 especially 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.[16]

That immigrants' socio-economic status might matter for natives' locational decisions has of course been discussed earlier in the literature, see e.g. Boustan (2010); Saiz and Wachter (2011); Rathelot and Safi (2014); Sá (2014). Probably due to data restrictions, it has however 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 about the foreign background of

[16] We refer to this channel as preferences *for homogeneity* along the socio-economic dimension. Because refugees generally have 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 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 arrival.

the parents of the native born.

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 tied family members to refugees) before having children.[17] Assume that ethnicity is the only characteristic 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 captures this dimension only (call it $\delta^{Ethnicity}$). Then we would 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 who on average are ethnically more dissimilar to the newly arrived refugees (the native-born individuals with Swedish- and other Western-born parents) will be greater than the response within the group of natives who on average are ethnically more similar to the newly arrived refugees (the native-born individuals with parents born in a non-Western country). If there is a strong preference for ethnic homogeneity, we therefore expect δ to be 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 one.

3.2 Non-behavioral mechanisms

Aside from the two preference-based channels, 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; in investigating historical white flight within the US, she sets up 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

[17] See `https://www.migrationsverket.se/English/About-the-Migration-Agency/Facts-and-statistics-/Statistics/Overview-and-time-series.html` for information on number and type of residence permits per country of origin from 1980 and onward.

house prices, this will induce movement from the current population. Under such a scenario, part of the observed flight is therefore due to price increases rather than to behavioral effects induced by the preferences or the perceptions of the native majority. A similar 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 induce home owners who are worried about falling house prices to leave. However, the housing stock in high-immigration neighborhoods is typically characterized by a large share of rental apartments (see Section 5), and because the Swedish rental market is highly regulated, immigration cannot affect rental prices, neither up nor down. This is particularly true in the short-run perspective that our analysis take (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 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 housing supply is not perfectly elastic, there is also a "mechanical effect" to consider. In the extreme case when housing supply is perfectly *inelastic*, irrespectively 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^{st} of the year for each individual living in Sweden at that point in time. 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 that takes place 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, no matter the 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 housing supply is fixed, there is one less apartment/house available for everybody else. Even if a native was contemplating moving there, the possibility might then not exist. This should, however, at most imply a (negative) 1:1 relation, meaning that we can rule out larger negative effects than that as being solely driven by such a mechanical effect.

3.3 Possibility to move

A prerequisite for deducing residential preferences from flight and avoidance estimates due to *any* mechanism is that people indeed are mobile. We recognize that this is far from true for everyone, meaning that some groups may not be able to react on their residential preferences. A contribution of our paper is that, to our knowledge, we are the first to take such mobility constraints into account.

The particular reason why some individuals cannot easily move depends on the institutional setting. In the current context, mobility constraints of individuals renting rather than owning their homes are likely to be especially pronounced as a consequence of increased immigration. This is because, first, renters are often resource constrained. Many renters are therefore constrained to other rental apartments, should they wish to move. Second, municipalities are responsible for accommodating newly arrived refugees who are not able find a place on their own. Usually this is 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). This is true also for existing tennants, 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 thus be particularly difficult for individuals living in rentals.[20] To take these mobility/budget constraints into account, we focus much of the empirical analysis on the group owning

[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 can be approved.

[20] Moves out of the municipality are not subject to this problem. But long-distance moves are instead significantly more costly, not the least from a labor market point of view. Additionally, moving to a new municipality often implies lost queuing points.

their homes. Note that this is not to say that renters in general are less mobile. Rather, this follows from the combination of most immigrants occupying rental apartments, and that the non-renter market is inaccessible for (budget constrained) renters.

To sum up the discussion in section 3, if we observe substantial native flight among those with a high possibility to move, 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 background react to a similar extent, this suggest 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; we discuss the general set-up, the identification strategy, and our improvement compared to the earlier 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 lives in another neighborhood in t, whereas $inflow_{i,t}$ is the number of natives who did not live in i in $t-1$ but does so in t.[21] The two variables $outflow_{i,t}$ and $inflow_{i,t}$ are our main outcome variables, and our two parameters of interest are β^{out} and β^{in} in the following two equations:

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

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

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

where $im_{i,t}$ is the number of new immigrants in neighborhood i in year t. Recalling 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* $\beta^{out} > 0$.

... *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, there is an endogeneity problem that must be solved. To identify β^{out} and β^{in}, we will use an instrumental variable that we consider substantially improves on 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, the improvement is mainly attributed to two factors. First, we only consider refugee migration, arguably providing more exogenous variation in immigration than when conflated with other migration. Second, we make use of a Swedish refugee placement policy that was in effect in the early part of the period that we study, arguably generating a more exogenous historical allocation of immigrants than when they self-select the place of residency.

In the following, we discuss the general shift-share approach and our improvements to it.

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

The instruments used in the earlier literature to solve the endogenous location choice of immigrants typically follow the shift-share strategy (see, e.g., Altonji and Card, 1991; Card and DiNardo, 2000; Saiz, 2007; Sá, 2014). The strategy builds 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 that arrived in Sweden and settled in neighborhood i in some baseline period t^0. $im_{c,SWE,t}$ represents 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 in the baseline period been the same.

To implement the shift-share approach, source country c and baseline period t^0 must be chosen, and it is in these two decisions that our methodological improvement lies. 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 allows 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 nations, 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*[22]. This allows us to focus the analysis on refugee immigration, which is advantegous from an identification point of view. 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

[22] *Grund för bosättning* in Swedish.

(e.g., labor and student migration), pull factors are to a larger extent city or neighborhood features in the destination country. 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 will use neighborhood fixed effects, identification in our shift share setting comes from variation within neighborhoods over time. By construction, the distribution of immigrants in the baseline years is constant. Thus, identification over time stems from variation in the country-specific annual inflow of immigrants, which needs to be stubstanital in order to be able 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, for example due conflict escalation. On the contary, labor and student migration is more consistent over time.[23]

The information on reason for immigration 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, but we drop those from OECD countries, since it is less likely that we observe flight from migration from for example Germany or Denmark. Also, many of these 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,[24] 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 in the baseline period. The full list of these 34 countries and the frequency of refugees arriving in 1997–2010 are available in Table 13 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 share of immigrants from the same country in the baseline year. Since the scaling is based on historical behavior, it is

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

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

a problem for identification if the historical immigrant settlement patterns were guided by (unobserved) sticky or fixed factors that are correlated with natives' migration decisions still today.[25]

This is a problem that is 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 themselves where to settle, but were assigned to a municipality through municipality-wise contracts, coordinated by the Immigration Board.[26] The number of municipalities that had such a contract increased over the years, and by 1991, 277 out of 286 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) and, instead, to achieve a more even distribution of refugees over the country. This aim was successfully fulfilled, as illustrated for example in Figure 3B in Dahlberg et al. (2012) and Table 1 in Edin et al. (2004).

Motivated by this, we choose for our baseline period t^0 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 that also affect 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 that the program-generated placement of refugees across municipalities was ran-

[25]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 of interest. 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.

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

dom.[27] What we argue is rather that, since the refugees received by the municipalities were effectively assigned to a specific apartment rather than choosing themselves 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.[28]

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 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.[29]

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 most importantly, the panel structure of the data means that we can include neighborhood fixed effects,[30] 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 are 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 immigration of refugees could be correlated

[27] 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 of the exogeneity of the placement program with respect to municipal characteristics, we refer to Dahlberg et al. (2012).

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

[29] We focus on the short-term perspective of one year because, 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 post 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.

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

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 $t-1$.[31] Fourth, we include time fixed effects to control for aggregate shocks that affect all neighborhoods in the same way in a given year. Finally, we control for a set of time-varying socio-economic characteristics of the neighborhood (measured in $t-1$); average disposable income, the number of students, the per capita cost of social assistance and the number of public rental estates.[32]

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 on native migration of immigration, within neighborhoods, 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

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

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

settlement from several years before.

5 Data and descriptive statistics

In this section we present the data, which is obtained from the GeoSweden database, and our defintion of a "neighborhood". All data is collected and made anonymous 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, covers all individuals living in Sweden and is very comprehensive. It contains variables from several different registers such as the education, the income and the employment registers, and it contains information on individual characteristics such as year and country of birth, marital status, the number of children in the household, as well as 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 of extra importance for this paper is the detailed geographical information on where the individuals live, information on the date, from which country, and for what 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 Areas for Market Statistics). A SAMS is a geographical unit that Statistics Sweden has defined 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 tractable 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 1 provides summary statistics of the variables used in the analysis, along with a clarifying description. The SAMS-neighborhoods span between very small places with only a couple of individuals to large neighborhoods in inner Stockholm with around 20,000 inhabitants. Around 85 natives on average 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 because 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 can be seen, 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, and we show below that this is indeed the case.

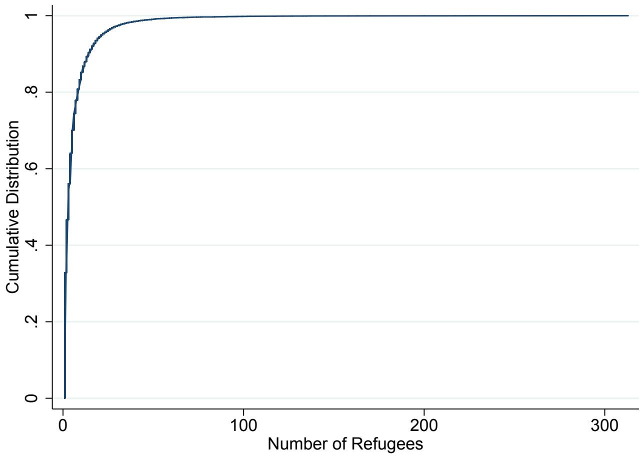
Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Key variables:					
Outflow	114,477	85.2	118	0	2,352
Inflow	114,477	85.2	121	0	2,716
Immigration (main)	114,478	0.82	4.7	0	313
Predicted immigration (instrument)	113,503	0.81	3.6	0	251
Control variables:					
Population	114,478	1019	1236	1	20,285
Students	114,478	53.1	107.5	0	2,642
Disposable income	114,478	155,838	538	-107,050	5,688,067
Social assistance	114,478	8,700	22,500	0	108,200
Other non-OECD imm.	114,478	2.4	9.4	0	590
Public rentals	114,478	2.1	5.2	0	408

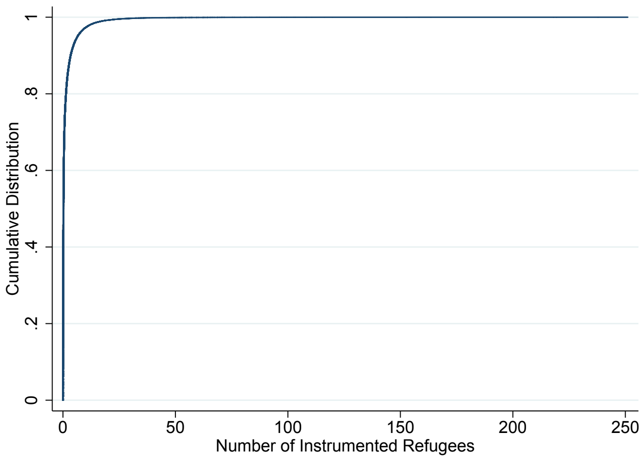
Outflow and Inflow measure the number of natives moving out of and into a given neighborhood in a given year. Immigration (main) is the main endogenous independent variable, measuring the annual number of refugees, and Predicted immigration is the instrument for this variable. Population denotes total 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.

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 home owners, we study households that indeed have a fair possibility to move following increased immigration. Ideas about mechanisms are discussed and tested in Section 6.3. The more constrained group of renters is studied in section 6.4, while Section 6.5 relates the results to the tipping point literature.

6.1 First stage

Table 2 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 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. Because treatment is defined at the level of SAMS-by-year, our default is to cluster the standard errors at SAMS.[33]

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. We see in the latter 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, such that the instrument clearly fulfills the relevance condition. The model is also very stable; adding all the control variables, including the fixed effects, does not affect the estimate much (from 0.67 to 0.61; cf. columns 1 and 2).

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

Table 2: First-stage estimates

	(1) No controls	(2) Baseline specification
$\tilde{im}$	0.674***	0.608***
	(0.0808)	(0.0859)
Observations	113,503	104,251
Number of SAMS	8,731	8,710
F-stat	69.55	49.98
Adjusted R-squared	0.170	0.170

*** 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 3 we run the first stage for several different subsamples. First, we remove the 10 percent of neighborhoods with the smallest population (less than 322 individuals) and the largest population (more than 2043 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 to larger neighborhoods. The coefficient is however highly statistically significant in both subsamples.

Gothenburg (the second largest city in Sweden) with its almost 800 neighborhoods is a clear outlier; very few municipalities have over 100, and Stockholm (the capital) has less than 200. We therefore exclude Gothenburg, with no big change in either power or significance; see column 3. Last, it is interesting to see how the first stage depends on the number of immigrants. Because 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 but, again, it is still highly significant; see column 4.[34]

[34]Because the placement program became less strict after 1992, we have also estimated

The first stage can be concluded as 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.

Table 3: 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
$\tilde{im}$	0.622***	0.251***	0.608***	0.140***
	(0.0882)	(0.0292)	(0.0896)	(0.0134)
Observations	78,997	93,635	95,327	103,168
Number of SAMS	6,770	7,927	7,960	8,709
F-Stat	49.84	74.13	45.96	108.16
Adj. R-squared	0.186	0.0422	0.173	0.0392

*** 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 2 for details of the estimated model.

6.2 Native flight and avoidance: Average effects

Moving 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 (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.

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 4: Second-stage estimates of native flight and avoidance

	(1) All natives	(2) Home owners
OUTFLOW		
im	0.0645	0.347**
	(0.158)	(0.156)
INFLOW		
im	-0.0847	0.170
	(0.183)	(0.137)
Observations	104,250	104,250
Number of SAMS	8,710	8,710
Mean of dep. variable	85	39

*** 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 2 for details of the estimated model.

The left column of Table 4 includes all natives and shows neither signs of flight nor 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 their possibilities of moving. Capturing residential preferences through flight and avoidance is however only possible if people indeed are 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—and especially so 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, for much of the remaining analyses we therefore restrict the sample to natives owning their home. In the right column of Table 4, it becomes clear that the insignificant aggregate effects mask interesting heterogeneity. In particular, the estimated outflow effect among native home owners is a statistically significant 0.35.[35] The interpretation of this coefficient is that, when a neighborhood receives one more immigrant than on average, 0.35 additional natives move out.

In contrast to outflow, there is no statistically significant inflow effect of increased immigration among home-owning natives. A possible explana-

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

tion is that home owners mostly notice and consequently react on increased immigration into the neighborhood where they currently live. 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 that the moving natives sell.

6.3 Is native flight determined by ethnically based preferences?

The pronounced flight effect in the subsample of home owners is interesting as such, in part because it potentially has implications for previous studies that mostly have looked at aggregate flight effects—which yet have 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 lower socio-economic groups than the average native. If natives move due to increased immigration, they may therefore do so either because they prefer ethnically homogenous neighborhoods, and/or if they have preferences for socio-economic homogeneity. As discussed in Section 3.1, we wish to examine if the commonly assumed ethnic channel is supported by the data, by grouping individuals according to their parental foreign background. While earlier work have speculated about whether the ethnic or the socio-economic channel is the driving one, (see Saiz and Wachter, 2011; Sá, 2014; 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 react on the immigrants' ethnicity. Under such a scenario, estimated 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 their parental background, then the main mechanism is more likely to be socio-economic.

We continue the focus on home owners and construct three groups of native home owners based on their foreign/ethnic background, and provide outflow and inflow effects for these respective groups in columns 2–4 of Table 5 (column 1 reproduces the average effect among home owners from above); column 2 contains those with native-born parents; column 3 contains those with at least one parent born in another Western country[36]; and column 4 contains those with at least one parent born in a non-Western country. The average number of movers differ between the groups, and in order to facilitate comparisons, the dependent variable is standardized with its mean.

As can be seen in Table 5, all flight estimates are positive and statistically significant. In comparing the magnitude across columns, it is clear that the relative magnitude is very similar across the groups of natives with different parental background. In particular, the estimated effects for natives with native-born and non-Western born parents are strikingly similar. In both cases close, the point estimates are around 0.008, which says 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.

Regarding inflow, the only (weakly) statistically significant effect is observed for the group in column 4—that is, for native home owners with at least one parent born in a non-Western country. But the point estimate is positive and thus, as in Table 4, there is no evidence of avoidance.

Treatment (increased immigration) is defined at the level of SAMS-by-year, and as noted above, we therefore cluster the standard errors by SAMS. Yet, the first phase of the placement program that defines our baseline period placed refugees to municipalities. We have therefore reestimated 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; whereas 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.[37] All in all, we trust the estimates obtained from clustering at the SAMS level, and note that despite some loss in precision, the com-

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

[37]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 will then have little impact on the standard errors for these groups.

parison across groups of natives with different parental background mainly holds also when clustering at the municipality level.

Table 5: 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	0.00881**	0.00787**	0.0172***	0.00841**
	(0.00396)	(0.00393)	(0.00636)	(0.00420)
INFLOW				
im	0.00432	0.00375	0.00503	0.00799*
	(0.00347)	(0.00353)	(0.00507)	(0.00447)
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 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 2 for details of the estimated model.

That natives with different ethnic parental background display similar flight behavior thus 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, increased immigration creates socio-economically more diverse neighborhoods, and may be the dimension along which natives' residential preferences are shaped. To study this further, we use our model to study how immigration affects the income and education level in the neighborhood, respectively, among home owners in the same sub-groups as analyzed in Table 5. 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

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

That is, the outcome variable is replaced by the average disposable income, $income_{i,t+1}$, and the share of university-educated, $university_{i,t+1}$, in the neighborhood. We define and estimate equations (10) and (11) for all home owners (that is, both native and non-native) as well as separately by the same sub-groups as analyzed in Table 5. And as above, to facilitate comparisons across groups, the dependent variable is standardized with its mean.[38]

The results, provided in Tables 6 and 7, show that the effect of increased immigration is that the income as well as the educational level decrease among all home owners irrespectively 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 home owners and all native home owners (cf. columns 1 and 2) imply that the 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.

[38]Note that when estimating the effect on the education level in the neighborhood (equation 11, Table 7) we add a lagged control for the number of university educated in $t-1$.

Table 6: 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**	-0.00130*	-0.00149	-0.000494
	(0.000845)	(0.000874)	(0.000785)	(0.00113)	(0.000687)
Obs.	90,825	90,748	90,555	87,025	87,946
N'hoods	8,358	8,350	8,331	8,159	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 2 for details of the estimated model.

Table 7: 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	-0.00358***	-0.00344***	-0.00361***	-0.00359***	-0.00173***
	(0.000888)	(0.000873)	(0.000923)	(0.00124)	(0.000584)
Obs.	95,551	95,551	95,551	95,551	95,551
N'hoods	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 neighborhood share of home owners in the respective groups that has at least some university education. 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. Covariates are the same as described in Table 2, 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 analysis above focusing on home owners, who in the current setting, as argued, are those who indeed can react on increased immigration. Table 8 instead presents flight and avoidance estimates for natives in publicly provided rental apartments, again grouped according to parental foreign background. As can be 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 argument above (see Section 3.3), 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 of Table 8). One possible explanation to this result is that this group starts a housing career when new immigrants arrive. To test this, we look specifically at the outflow of native renters *becoming home owners*. Table 9 presents the results, showing that the only positive, significant effect is again among natives with non-Western parental background. The magnitude is also similar to the overall effect for this group. Thus, the results are indeed consistent with this group making a housing career. And again, the estimated flight and avoidance behavior among renters are difficult to reconcile with an ethnically based mechanism.

Table 8: 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	-0.00391	-0.0126***	-0.0127	0.0503***
	(0.00496)	(0.00487)	(0.0107)	(0.0127)
INFLOW				
im	-0.00426**	-0.00618***	-0.00152**	0.0104
	(0.00205)	(0.00207)	(0.000768)	(0.00992)
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 2 for details of the estimated model.

Table 9: Second-stage estimates of native flight among renters with different parental background becoming home owners

	(1)	(2)	(3)	(4)
	All	Parental background:		
	Natives	Native	Western	Non-Western
im	0.00412	-0.00785*	0.0184	0.0843***
	(0.00571)	(0.00462)	(0.0159)	(0.0246)
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-home owners, column 2 is restricted to native renters-home owners with Swedish-born parents, column 3 is restricted to native renters-home owners with at least one parent born in another Western country, and column 4 is restricted to native renters-home owners with at least one parent born in a non-Western country. The dependent variables are standardized with its respective mean. See Table 2 for details of the estimated model.

6.5 Native flight and tipping points

As is common in the native/white flight and avoidance literature, we have estimated a continuous model of migration behavior. This contrasts with the tipping point literature, which is based on a model in which natives' migration behavior changes abruptly when the share of immigrants reaches a certain level (see, e.g., Schelling, 1971; Card et al., 2008). A question is then how our results relate to this latter literature hypothesizing that there is a neighborhood tipping point? To get a sense of this, we analyze how the flight coefficients vary with the initial share of immigrants. To this aim, the model is estimated separately depending on the neighborhood share of immigrants in the first year of our study period (1997). We again focus on home owners and distinguish between natives of different parental foreign background.

Figure 4 shows the resulting estimates for each 1-percentage interval of the initial immigrant share from 4 to 25 percent (the minimum and maximum value, respectively). Although the estimates across the different initial levels are not significantly different from one another, there are signs of a discrete increase in the parameter estimate at an initial share of 18 percent or above. This is something to examine further in future research (as is the U-shaped pattern of the point estimates, which is not readily reconciled with one unique tipping point).[39]

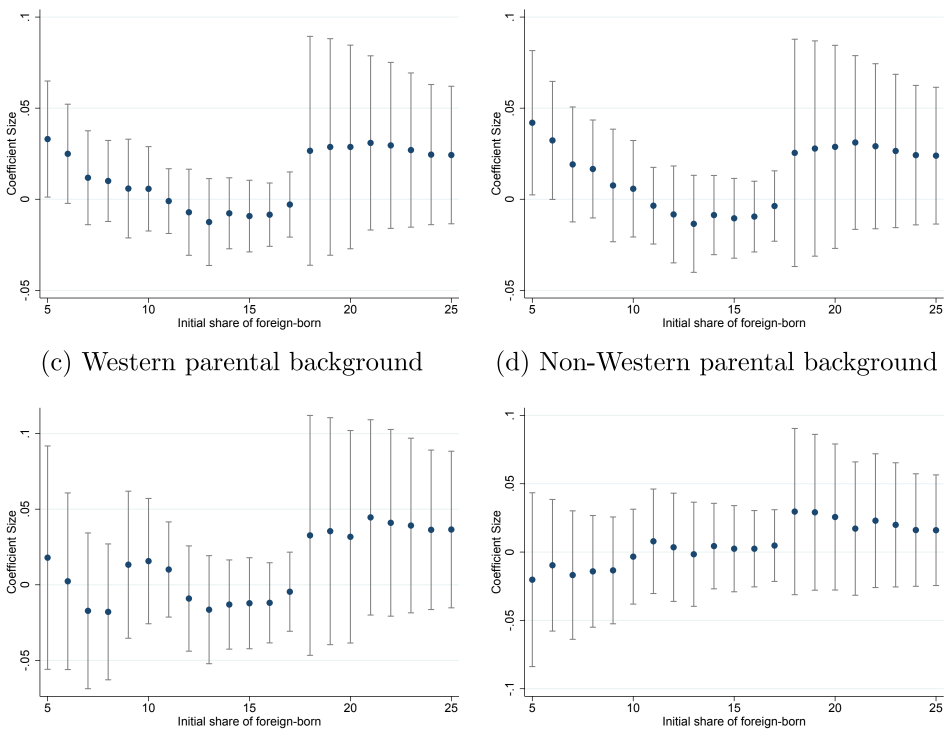
Furthermore, the three groups of natives with different parental foreign background display very similar patterns (cf. panels b-d). This is interesting and in line with the results found in previous sections. The conclusion thus pertains; residential preferences do not seem to be formed along ethnic lines, and the observed migration behavior of native-born individuals with varying foreign background is rather likely to be a reaction to changing socio-economic neighborhood characteristics. These results potentially have implications for the tipping point literature; significant tipping points for all native-born individuals irrespectively of their parental foreign background is suggestive of native tipping behavior based not, as commonly assumed, on ethnic/racial grounds, but rather on socio-economic grounds.

[39] Compare also with the tipping points found in Böhlmark and Willén (2017) and in Aldén et al. (2015) using Swedish data.

Figure 4: Second-stage estimates of native flight among home owners with different parental background, by initial immigrant share

(a) All natives (b) Native parental background

(c) Western parental background (d) Non-Western parental background



Note: The estimates are obtained from separate regressions depending on the share of foreign-born in the neighborhood in the initial year of the analysis (1997), for each 1-percentage interval between 4 and 25 percent. The dependent variable is standardized with its mean. Figure 4a includes all native home owners, Figure 4b is restricted to native home owners with Swedish-born parents, Figure 4c is restricted to native home owners with at least one parent born in another OECD country, and Figure 4d is restricted to native home owners with at least one parent born in a non-Western country.

Source: GeoSweden.

7 Concluding remarks

In this paper, we have applied detailed, comprehensive, register data to a refined shift-share method to answer whether native flight and avoidance are important phenomena in Sweden. In particular, using information in the data that allows us to identify native-born individuals that to different degrees are ethnically close to the newly arrived refugee immigrants (as defined via the heterogeneous parental background), we have examined if 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 for at least two reasons; first, there was a large increase in refugee-based immigration to Sweden over this time period and, second, in the early part of the period, there was a refugee placement policy in Sweden which arguably can be used to improve the shift-share instrument.

Using push-driven refugee immigration to Sweden interacted with a settlement pattern of their countrymen in the early 1990s that was partly generated by this state-run placement policy, we reach four main conclusions.

First, we find no evidence of neither native flight nor native avoidance when studying the full population.

Second, when we look specifically at households with high possibilities to move following increased immigration (home owners in the Swedish context), we do detect native flight. That is, home-owning natives move out of neighborhoods experiencing an increase in immigration. We do however not find evidence of native avoidance; 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 on increased immigration into the neighborhood where they currently live. We do not find any flight or avoidance effects among natives identified as having low possibilities to move following increased immigration (here renters). Hence, distinguishing between mobile/immobile households when examining the effects of immigration on native migration seems important, but is something the earlier literature has not been able to do.

Third, we find that the ethnic closeness between the native-born individuals and the newly arrived refugees does not matter for observed flight behavior; all Swedish-born individuals react in a very similar fashion to increased immigration. Preferences for ethnically homogeneous neighborhoods do therefore not seem to be the main channel causing flight. Rather, our analyses consistently indicate that natives have preferences for socioeconomically homogeneous (or, “better”) neighborhoods.

Fourth and finally, when conditioning on the initial share of foreign born, and thereby relating to the tipping point literature, we again find the same patterns regardless of the natives’ parental foreign background. This is thus further evidence against the ethnicity channel, which could indicate that the tipping point literature might have focused on the wrong trait.

If political decision-makers want to instigate policies to combat segregation, it is important to know what the mechanisms are for observed changes in natives' migration behavior; successful policies will likely differ depending on whether the main channel is ethnically or socio-economically based. More research is however 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 behavior that takes place after a longer period of time. Studying longer time lags is therefore an interesting follow-up. Second, the focus on small neighborhoods could be extended to larger areas, such as municipalities: this is not the least interesting as a comparison to the the earlier literature, which mostly has studied larger geographical units. Third, our results indicate that the ethnically-based tipping point literature might have focused on the wrong trait. This is important to investigate further, for example by extending the analysis for native-born individuals with different ethnic backgrounds. Lastly, a possible alternative way to channel any preferences for homogeneity is via school choices. If parents perceive school quality to be dropping due to increased minority presence, an exodus from the neighborhood school could occur. Increasing school segregation is thus also an interesting topic for future research.

References

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.

Böhlmark, A. and A. Willén (2017): "Tipping and the Effects of Segregation," *IFAU Working Paper Series*, 14.

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

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.

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

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 Berkely.

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. (2014): “Immigration and House Prices in the UK,” *The Economic Journal*, forthcoming.

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.

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

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

Assume an immigration shock to neighborhood i in 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 reversion of house prices. The shift share instrument uses correlation over time in immigrants' location patterns. Potentially, the instrument therefore 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 in t. Since both are endogenous, their solution implies two first stages:

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

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

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

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

$$inflow_{i,t+1} = \beta_1^{IV}\widehat{im}_{i,t} + \beta_2^{IV}\widehat{im}_{i,t-1} + \sum_{p=1}^{3}\delta_p^{IV} pop_{i,t-1}^{p} + \mathbf{X}\Pi^{IV\prime} + \eta_i + \lambda_t + \varepsilon_{i,t+1}^{IV} \quad (15)$$

Table 10 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 11, and for inflow in Table 12. Comparing these estiamtes to the baseline estimates ub Table 5, it is clear that the results are robust to this addition the model. In other words, dynamic effects do not appear to be a threat to our short-term estimates.

Table 10: First-stage estimates according to Jaeger et al. (2018)

	(1) im_{it}	(2) im_{it-1}
$\widehat{im_{it}}$	0.690***	-0.0152
	(0.0941)	(0.0178)
$\widehat{im_{it-1}}$	-0.180***	0.632***
	(0.0305)	(0.0898)
Observations	104,772	104,772
Number of SAMS	8,731	8,731
F-Stat	30.39	25.00
SAMS FE	YES	YES
Year FE	YES	YES
Covariates	YES	YES

*** 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 2 for details of the estimated model.

Table 11: 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}	0.00827***	0.00717**	0.0177***	0.00811**
	(0.00319)	(0.00312)	(0.00562)	(0.00378)
im_{it-1}	0.00144	0.00187	-0.00141	0.000804
	(0.00331)	(0.00325)	(0.00459)	(0.00440)
Observations	104,250	104,250	104,250	104,250
Number of SAMS	8,710	8,710	8,710	8,710
SAMS FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on SAMS-level. See Table 2 for details of the estimated model and Table 5 for details across columns.

Table 12: 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}	0.00399	0.00349	0.00466	0.00709
	(0.00280)	(0.00270)	(0.00472)	(0.00463)
im_{it-1}	0.000901	0.000686	0.00101	0.00241
	(0.00382)	(0.00378)	(0.00544)	(0.00444)
Observations	104,248	104,248	104,248	104,248
Number of SAMS	8,710	8,710	8,710	8,710
SAMS FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on SAMS-level. See Table 2 for details of the estimated model and Table 5 for details across columns.

B Source countries

Table 13: Frequency of refugees arriving 1997–2010.

Iraq	40,537	43.31%	43.31%
Somalia	11,597	12.39%	55.70%
Yugoslavia	8,345	8.92%	64.61%
Bosnia	6,727	7.19%	71.80%
Iran	5,105	5.45%	77.26%
Afghanistan	4,347	4.64%	81.90%
Syria	3,954	4.22%	86.12%
Russia	2,676	2.86%	88.98%
Lebanon	2,563	2.74%	91.72%
Thailand	1,225	1.31%	93.03%
Ethiopia	1,142	1.22%	94.25%
Croatia	887	0.95%	95.20%
Colombia	736	0.79%	95.98%
India	683	0.73%	96.71%
Peru	520	0.56%	97.27%
Bangladesh	469	0.50%	97.77%
Pakistan	468	0.50%	98.27%
China	269	0.29%	98.56%
Uganda	187	0.20%	98.76%
Romania	165	0.18%	98.93%
Bolivia	164	0.18%	99.11%
Vietnam	160	0.17%	99.28%
Algeria	125	0.13%	99.41%
Poland	108	0.12%	99.53%
Morocco	86	0.09%	99.62%
Tunisia	78	0.08%	99.70%
Latvia and Lithuania	71	0.08%	99.78%
Bulgaria	49	0.05%	99.83%
Estonia	38	0.04%	99.87%
Philippines	36	0.04%	99.91%
Gambia	31	0.03%	99.94%
Argentina	29	0.03%	99.98%
Slovenia	12	0.01%	99.99%
Brazil	11	0.01%	100.00%

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