

# Migrating Natives and Foreign Immigration\*

Henrik Andersson<sup>†</sup>    Heléne Berg<sup>‡</sup>    Matz Dahlberg<sup>§</sup>

July 13, 2016

## Abstract

In this paper we investigate the extent of native flight and native avoidance. Specifically, we study Sweden and ask if native-borns tend to move away from (native flight) or avoid moving in to (native avoidance) high-immigration neighborhoods. We contribute to the existing literature with better and more detailed data, as well as with improved identification methods: We use rich geo-coded register data spanning over 20 years, and to account for possible endogeneity problems, we apply a shift-share strategy which combines policy-induced initial immigrant settlement with exogenous contemporaneous shocks in the source country. While the results show no indications of native avoidance, we see indications of native flight. This outflow seems to be driven by second-generation immigrants, and is concentrated among the high-income earners and the highly educated. An increased segregation following immigration hence seems to take place along socioeconomic lines rather than along ethnic lines.

**Keywords:** Immigration; Native migration; Flight; Avoidance; IV estimation

**JEL classification:** C26; J15; R23

---

\*We are grateful to Leah Platt Boustan, Florian Morath, Susanne Urban and seminar participants at UC Irvine, UCLA, ETH/KOF in Zurich, Uppsala University, the 2015 IIPF conference in Dublin, Norface meeting at Rosersberg castle, UCFS meeting at Krusenberg, and UCLS meeting in Uppsala for helpful comments and discussions.

<sup>†</sup>Institute for Housing and Urban Research and Department of Economics, Uppsala University. henrik.andersson@nek.uu.se.

<sup>‡</sup>Department of Economics, Stockholm University; CESifo. helene.berg@ne.su.se.

<sup>§</sup>Institute for Housing and Urban Research and Department of Economics, Uppsala University; CESifo; IEB; IFAU. matz.dahlberg@ibf.uu.se.

# 1 Introduction

Many western countries have long faced increasing levels of immigration. In 2015 alone, UNHCR estimated that about 1 million individuals reached the shores of Europe by crossing the Mediterranean. Such an historic inflow has raised heated discussions on the geographical distribution of refugees and the fear of increasing levels of ethnic segregation. Although increased immigration *need not* imply increased ethnic segregation, there are several reasons why we might observe such a development. First, immigrants might face housing discrimination or resource constraints, second, immigrants could be attracted to co-ethnics already in the country, and third, an inflow of immigrants could induce the native population to leave or avoid ethnically diverse neighborhoods.

The extent to which natives react in such a way is the focus of this paper. Specifically, we take the fact that new immigrants are attracted to older ones as given, and use this fact for identification purposes. Our aim is to study the migration behavior of the native population when new immigrants arrive. We hypothesize that this reaction can take place in one of two ways; either in the form of *native flight* (natives migrating out of their neighborhood as more immigrants settle), or in the form of *native avoidance* (natives choosing not to migrate into a neighborhood where more immigrants settle).

Our study covers immigration into Sweden during the past two decades, a setting which we—for several reasons—think is interesting and highly appropriate for the question at hand. Much of the previous literature on white flight has focused on the US.<sup>1</sup> Like the US, Sweden is a country that, at least since the mid 1980s, has experienced increasing immigration inflows. But—and this will turn out to be beneficial for our proposed identification strategy—the nature of the immigration to Sweden is quite different. In particular, being the “land of opportunity”, immigrants are *pulled* from their home country to the US, whereas Swedish immigrants tend to have been *pushed* from their home country by wars and other catastrophes.<sup>2</sup> From an empiricist’s point of view, the US-type pull-migration comes with potential problems that the Swedish-type push-migration does not: In the case of pull-migration, increases in immigration are likely to be endogenous in the sense that shocks in the destination country that attract (pull) more immigrants also affect the behavior of natives directly.<sup>3</sup>

We identify the causal effect of foreign immigration on the residential

---

<sup>1</sup>See in particular Farley et al. (1978), Farley et al. (1994), Boustan (2010), Saiz and Wachter (2011) and Wang (2011).

<sup>2</sup>Zimmermann (1996) provides a stylized economic definition of push and pull migration.

<sup>3</sup>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.

choice of natives by combining these push features of contemporary migration into Sweden with previous immigrant settlement patterns generated partly by 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 means that the settlement patterns of immigrants from the early 1990s, who subsequently attracted more recent push-driven immigrants, 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. Relying on this assumption, our identification strategy is based on the shift-share approach (see, e.g., Altonji and Card, 1991; Card and DiNardo, 2000; Saiz, 2007; Carl and Siegenthaler, 2013; Chalfin and Levy, 2013; Sá, 2014) that departs from the idea above that new immigrants tend to be drawn to the familiar turf offered by their countrymen. Specifically, we construct an instrumental variable for immigrant inflows into a neighborhood based on the *interaction* of immigrant settlements during the placement policy era and the timing of contemporary immigrant shocks into Sweden due to events in the source countries.<sup>4</sup>

Our shift-share approach becomes particularly appropriate when applied to our detailed data from GeoSweden, a database that covers the full Swedish population. There are three valuable features of the data that we would like to stress. First, for each immigrant living in Sweden, there is information on the country of origin. The data therefore tells us in which neighborhoods immigrants from specific countries previously settled. Second, rather than using decennial censuses, the annual frequency of our data means that we rather precisely can relate the timing of new immigrants to shocks around the world that are arguably exogenous to Swedish neighborhood characteristics.

Finally, our empirical approach is strengthened by the fine geographical resolution in our data. While most previous studies have studied segregation across relatively large units (such as metropolitan statistical areas in the case of the US), our unit of analysis is much smaller neighborhoods within a municipality. This is important because it allows us to control for fixed effects at the more aggregate municipal level, and thus rely on variation in immigration across neighborhoods within a municipality.<sup>5</sup>

Aside from the methodological improvement, the quality of our register data, with very detailed information at the individual level, also makes it possible to increase our understanding of the mechanisms behind the natives' migration behavior. For example, by separating between different socio-economic groups (defined by income and education level) and between

---

<sup>4</sup>The push-driven immigration we identify and use in the analyses is from Iraq, Syria and Somalia; see section 4.2.1 for further details.

<sup>5</sup>A similar approach is taken in Saiz and Wachter (2011).

natives with both parents born in Sweden (“native-natives”) and natives with at least one parent born outside Sweden (“second-generation immigrants”), we can study the flight and avoidance behavior for these different groups separately. This allows us to get a better grasp of whether individuals react on ethnic or on socio-economic characteristics.

An influential, related literature 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). Here, much thanks to superior data, we can instead approach the heart of the issue *directly* by estimating the extent to which the residential choices of natives are affected by immigration. Admittedly, we are not the first to do so.<sup>6</sup> But we argue that the data at hand combined with the previous refugee placement program and the subsequent push-type Swedish immigration as described above allow us to add a somewhat unique study.

The paper is also closely related 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 look at continuous native migration. Finally, complementing the studies of the effects *of* residential segregation (Edin et al., 2003), our focus is on effects of immigration *on* residential segregation.

The analyses provide three main results. First, while we do not find any indications of native avoidance, we find some signs of native flight. Second, the outflow (flight) that we do observe is primarily driven by second-generation immigrants, rather than by Swedish-born individuals with Swedish-born parents. Third, heterogeneity results suggest that it is the high-income earners and the highly educated individuals that move out, while the low-educated individuals and the low-income earners have a tendency to move in and to stay. These results indicate that there is a sorting process along socio-economic lines rather than along ethnic lines that takes place following immigration.

In the following, 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, 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.

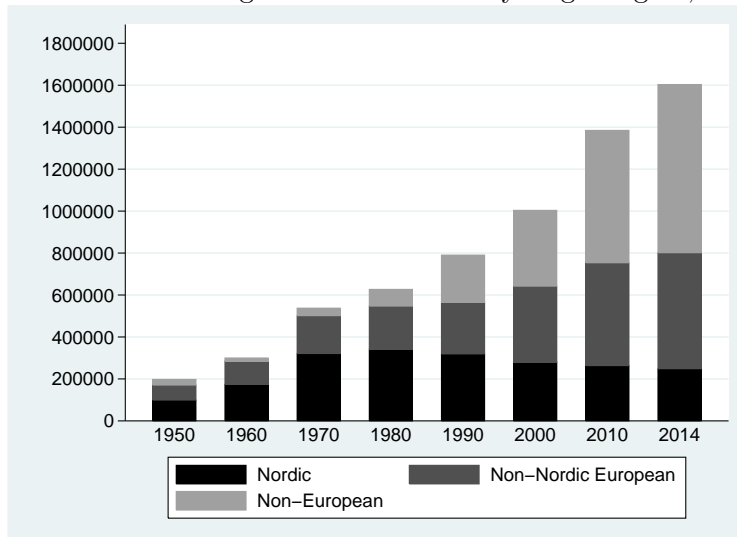
---

<sup>6</sup>In addition to the papers in the economics literature referred to above (e.g., Card and DiNardo, 2000 and Saiz and Wachter, 2011), a substantial body in the sociology literature studies this phenomenon; see Rathelot and Safi (2014) and the references therein.

## 2 Immigration to Sweden

The size and character of immigration to Sweden has changed over the last decades. In 1970, less than seven percent of the Swedish population was born in another country<sup>7</sup>, and of those the large majority had arrived as 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 been a drastic change over the last three to four decades in both the number and the origin of the foreign-born population in Sweden. The changing pattern of the foreign born-population in Sweden is clear from Figure 1. While the share of the foreign-born with roots in the Nordic countries is decreasing over time (the black bars), the share originating from non-European countries is increasing (the light gray bars). 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. In the end of 2014, the approximately 1,600,000 foreign-born individuals living in Sweden constituted a little bit more than 16 percent of the total population of close to 10 million. About half of these are born outside of Europe.<sup>8</sup>

Figure 1: Number of foreign-born in Sweden by origin region, 1950–2014.



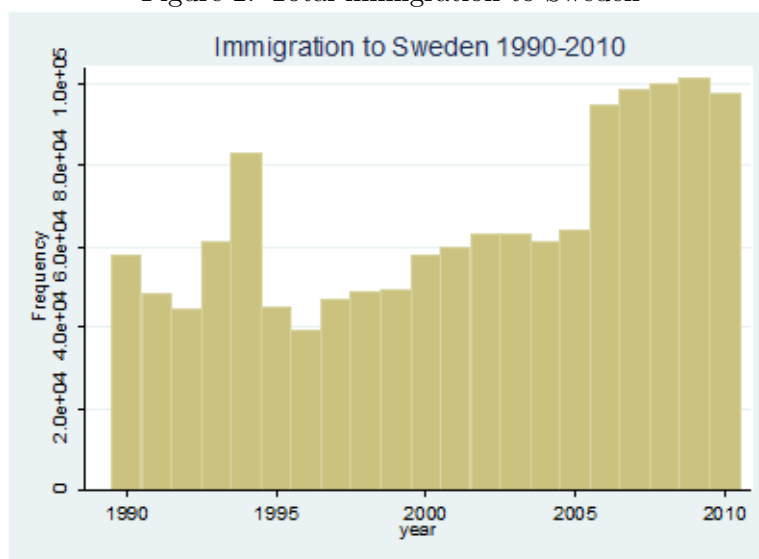
Source: Blind and Dahlberg (2015).

<sup>7</sup>Statistics Sweden, Yearbook of Sweden 2012, table 4.30 "Population by country of birth".

<sup>8</sup>See also Statistics Sweden/Statistics database/Population/Population statistics/Foreign-born persons/Foreign-born persons in Sweden by country of birth, age and sex. Year 2000–2014.

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.<sup>9</sup> Then, from 2006 and onward, there has been a discrete increase in the number of immigrants, with a yearly average of around 100,000. The bulk of the increase is explained by non-western refugee migration, which we return to in later sections.

Figure 2: Total immigration to Sweden



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

Compared to most other European countries, Sweden has a substantially larger share of foreign-borns. According to statistics from Eurostat,<sup>10</sup> in 2010, 47 million individuals in the EU 27-countries were not born in the country in which they resided. This amounted to almost ten percent of the total population. The majority of these, slightly more than 31 million, were born outside of the European Union. There is however a large variation in these numbers across the EU, 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).

<sup>9</sup>The spike in the early 1990s is due to increased refugee immigration following the Balkan war.

<sup>10</sup>The figures in this section comes from the following issues of Eurostat’s Statistics in focus: 98/2008, 27/2010, 45/2010, and 34/2011.

### 3 Potential reaction 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). Before discussing potential mechanisms behind these reactions, two things deserve notice. First, the concepts of *native* flight and avoidance are somewhat 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 the terms of country of origin. Hence, we refer to the potential reaction of the majority population as *native* flight and avoidance. We provide results where native is defined either as being born in Sweden irrespectively of where the parents are born, as being born in Sweden and having both parents Swedish-born, or as being born in Sweden with at least one parent being born abroad.<sup>11</sup>

Why, then, might we expect increasing immigration to affect natives' location decisions? Scholars within sociology, economics and geography have lifted several potential mechanisms.

First, 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, Farley et al., 1994). Second, native mobility might be induced by changing socio-economic characteristics. To the extent that refugees often are relatively poorer, the neighborhoods where they locate will experience reduced socio-economic status. In such cases, native movement might be driven by perceived unfavorable conditions in the area, such as high unemployment and low levels of security. In other words, if natives experience that the neighborhood status is dropping due to increased immigration, then native flight/avoidance might in fact be economic flight/avoidance. A third potential mechanism, related to the second, is the care for one's children. If parents believe that school quality is dropping due to increased minority presence an exodus from schools could occur. Furthermore, if the place of residence is directly linked to the school of the child, actions to change school of ones child might require moving.

All of these three mechanisms can be expected to cause both flight and avoidance. However, aside from such behavioral effects, there are also possible price effects of increased immigration to consider. Boustan (2010) explains this clearly; in investigating historical white flight within the US,

---

<sup>11</sup>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.

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 house prices, this will induce movement from the current population. Under such a scenario, part of the observed flight is therefore due to such 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).

Last, 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 able to rule out this mechanical effect for the case of flight. For the case of avoidance, however, this mechanism cannot be completely ruled out. We discuss both cases below.

### 3.1 Main mechanisms in the Swedish context

Above we listed five potential mechanisms through which we can expect effects on native migration from foreign immigration: Ethnic preferences, socio-economic preferences, school-quality preferences, price effects and purely mechanical effects due to fixed supply of housing. Given the Swedish institutional context and the frequency of our data, we argue that three of these can largely be ruled out for our application.

First, we expect the price effect to be small. The rental market in Sweden is highly regulated, implying that immigration cannot affect rental prices, at least not in the short-run perspective that our analysis takes (we mostly consider native migration within one year of additional foreign immigration). Further, to buy a house or an apartment, there are fairly strict rules on capital requirements and income. Since our focus is on refugee immigration, it is highly unlikely that many will purchase houses during their first year in Sweden. Still, it is possible that property prices increase via a substitution effect; as immigrants occupy more rental housing, demand for buying a property in the neighborhood could increase among some natives, consequently inducing other natives to leave (then because of higher prices, not because of immigration). We assess the possible extent of such price substitution effects by characterizing the typical type of housing stock in high-immigration neighborhoods; see Section 5.

Second, we believe that the school mechanism is likely to be of limited relevance for migration behavior in the Swedish context. Since 1992, the place of schooling is not determined by place of residency, but rather, there is a voucher system where parents can choose where to put their children to school. Thus, we expect (perceived) worse school quality to have at most a minor effect on the locational decisions of native Swedes.



Regarding the mechanical effect due to fixed supply of housing, our register data allows us to see the place of residence on December 31<sup>st</sup> 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. In contrast, 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.

To sum up, if we observe substantial native flight, 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. Due to the nature of the immigration that we study, distinguishing between ethnic diversity and socio-economic diversity is intrinsically difficult, but we provide an extensive heterogeneity analysis that is suggestive of a more pronounced socio-economic channel.

## 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 left neighborhood  $i$  in year  $t$ . Analogously, we define native inflow,  $inflow_{i,t}$ , as the number of natives who moved into  $i$  in year  $t$ . In other words,  $outflow_{i,t}$  is the number of natives who lived in  $i$  in  $t - s$  but lives in another neighborhood in  $t$ , whereas  $inflow_{i,t}$  is the number of natives who did not live in  $i$  in  $t - s$  but does so in  $t$ , where  $s$  defines the lag with which we assume that natives' response occurs. For most of our analysis,  $s$  will be one year.<sup>12</sup> The two variables  $outflow_{i,t}$  and  $inflow_{i,t}$  are our main

---

<sup>12</sup>Note that for the natives' responses, we only consider migration within the country (i.e., not emigration responses).

outcome variables, and our two parameters of interest are  $\beta^{out}$  and  $\beta^{in}$  in the following two equations:

$$outflow_{i,t+s} = \alpha^{out} + \beta^{out} im_{i,t} + \epsilon_{i,t+s}^{out} \quad (1)$$

$$inflow_{i,t+s} = \alpha^{in} + \beta^{in} im_{i,t} + \epsilon_{i,t+s}^{in} \quad , \quad (2)$$

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  $\beta^{out}$  and  $\beta^{in}$ :

**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  $\beta^{out}$  and  $\beta^{in}$ , we will use what we consider to be an improved version of the 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 push migration, arguably providing more exogenous variation in immigration than when combined with pull 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: A shift-share approach

The shift-share strategy has been used in several papers to solve the endogenous location choice of immigrants (see, e.g., Altonji and Card, 1991; Card and DiNardo, 2000; Saiz, 2007; Sá, 2014). The strategy builds on the fact 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  $i\tilde{m}_{i,t}$ , defined as (exemplified by immigration to Sweden):

$$i\tilde{m}_{i,t} = \sum_c i\tilde{m}_{c,i,t} = \sum_c \left( \phi_{c,i,t^0} \times im_{c,SWE,t} \right), \quad (3)$$

where

$$\phi_{c,i,t^0} = \frac{im_{c,i,t^0}}{im_{c,SWE,t^0}} \quad (4)$$

is the fraction of immigrants from source country  $c$  that arrived in Sweden and settle in neighborhood  $i$  in some baseline period  $t^0$ .  $im_{c,SWE,t}$  represents total immigration to Sweden from source country  $c$  in year (or period)  $t$ . The instrument  $\tilde{im}_{i,t}$  defined in equation (3) thus measures the contemporary immigration that would have been the result had the settlement of these immigrants and those who came in the baseline period been the same.

But how, then, to choose source country  $c$  and baseline period  $t^0$ ?

#### 4.2.1 Definition of source country

In the earlier studies discussed in the introduction, which mainly focus on US and UK data, typically *all* immigration has been used in the analyses. The majority of the immigration to these countries (as for most) is pull-driven. That is, most of the immigrants have already beforehand decided to go to a certain location in a specific country for a specific reason (something that might be particularly true for the immigration to the US and the UK). This type of selection worsens the endogeneity problem, since these specific reasons are likely to be directly related to natives' residential preferences and thus to native migration.

In contrast, we argue that the push-driven migration, i.e., migration that is the consequence of an exogenous shock (such as a war or famine) that pushes people to migrate from their country of residence, makes the selection effect less pronounced and thereby the endogeneity problem less severe.

Refugee migration constitutes a large part of the immigration to Sweden, implying that much of it is indeed push-driven. Still, a non-negligible fraction of total immigration is likely to be endogenous, meaning that without further caution, we would not get rid of the identification problem that studies which mainly relies on pull-migration likely have. However, thanks to the detailed information in our data—we know the exact year in which an individual migrates to Sweden, and from which country—we are able to restrict our analysis to immigration that is likely driven by exogenous shocks in the source country.

As was clear from Figure 2, there was a discrete increase in the number of individuals immigrating to Sweden in 2006, thus generating substantial variation that can potentially be used for identification purposes. Furthermore, a closer look at the data shows that this increase is largely driven by im-

migration from non-OECD countries: see Figure 3.<sup>13</sup> For a non-European, non-OECD country with substantial immigration to Sweden in 2006 to qualify for our what we argue is a causal measure of foreign immigration, we require two things: (1) That there is an exogenous shock in the source country in 2006, implying that immigration is likely push-driven; and (2) that “sufficiently many” individuals from those countries also immigrated to Sweden in the baseline period, in order to obtain more precise variation in the instrument (in particular, in  $\phi_{c,i,t^0}$  in equation (3))

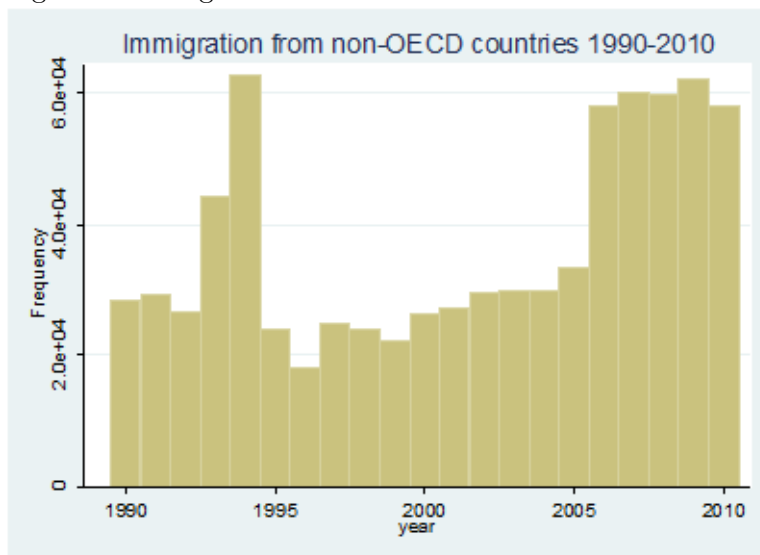
Three countries fulfill these criteria; Iraq, Somalia and Syria. As is clear from Figure 4, there was a sharp increase in immigration from these countries in 2006, and they all had a history of migration to Sweden. Furthermore, all these three countries experienced events that, from a Swedish perspective, exogenously pushed people to flee: There was an escalation of the Iraqi war in 2006, as witnessed, for instance, by a sharp increase in the number of individuals killed in Iraq due to the war in 2006<sup>14</sup>; Somalia experienced a severe drought in early 2006, and also came into conflict with Ethiopia later that year; and in Syria, there was a severe drought that started in 2006 (and that lasted until approximately 2011). Based on this, our decision of source countries  $c$  is thus Iraq, Somalia and Syria.

---

<sup>13</sup>Of the non-European non-OECD countries, most immigrants in 2006 came from Iraq (9,684 individuals), Thailand (2,871 individuals), Somalia (2,360 individuals), China and Taiwan (2,225 individuals), Iran (2,118 individuals), Lebanon (1,947 individuals), and Syria (1,680 individuals).

<sup>14</sup>See figure 6 in the Appendix for a description of the rise in number of casualties from the Iraqi war in 2006.

Figure 3: Immigration to Sweden from non-OECD countries



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

#### 4.2.2 Definition of baseline period

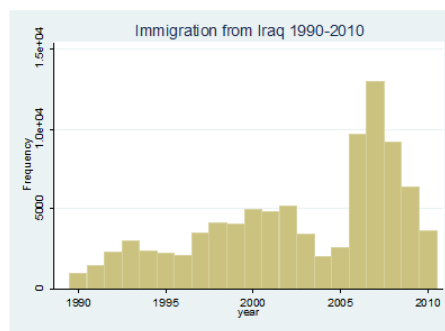
The push-driven migration generates an exogenous increase in immigration to Sweden. But even though the push-driven migration makes the selection problem less severe, it is still the case that once the refugee-migrants enter Sweden, they can still choose where to settle, and hence there is still a potential selection problem. The shift-share strategy uses the fact that new immigrants tend to be drawn to places where former immigrants sharing their background have already settled, the argument being that this generates an exogenous allocation of the newly arrived immigrants. However, a remaining problem for identification is if the historical immigrant settlement patterns were guided by (unobserved) factors that are correlated with natives' migration decisions still today.

This is a problem that is left unsolved in the existing papers in the migration literature applying the shift-share approach, and part of our contribution is to exploit a refugee placement policy that was in effect in Sweden from the beginning of 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.<sup>15</sup> The number of municipalities that had such a contract increased over time, and by 1991, 277 out of 286 were part of the program.

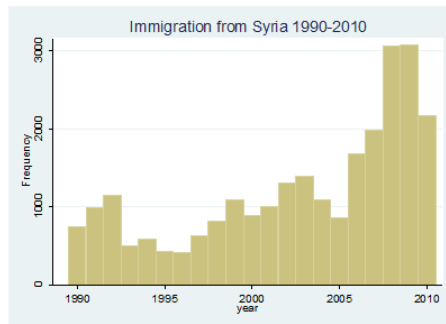
One of the main aims of the refugee placement program was to break the concentration of immigrants to larger towns (mainly Stockholm, Gothen-

<sup>15</sup>They were, however, allowed to move after the initial placement.

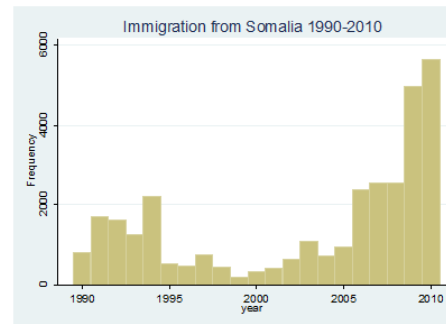
Figure 4: Immigration to Sweden from Iraq, Syria and Somalia



(a) Immigration to Sweden from Iraq



(b) Immigration to Sweden from Syria



(c) Immigration to Sweden from Somalia

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

burg and Malmö) and, consequently, 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 shift-share approach 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 the immigration increase in 2006 (compared to a situation in which the policy had not existed). This is especially true conditional on municipality fixed effects and a set of neighborhood characteristics that we include in our estimation model. Put differently, it is not necessary that the program-generated placement of refugees across municipalities was random.<sup>16</sup> What we argue is rather that, since the refugees that the municipalities received 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.

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

Our analysis will be cross-sectional, where the year of foreign immigration ( $t$  in equations (1) and (2)) refers to year 2006 when immigration from Iraq, Syria and Somalia increased significantly, while the year of migratory response by natives ( $t+s$ ) for most of the analysis refers to the year 2007. We will, however, also provide a longer term analysis where we study the natives' migratory response in 2014, eight years after the 2006 immigration.<sup>17</sup>

Besides instrumenting  $im_{i,2006}$  with  $\tilde{im}_{i,2006}$ , our final estimation model differs from the basic equations in (1) and (2) in a few ways. First, as just indicated, we include municipality fixed effects,  $\mu_j$ . This means that we compare neighborhoods with different levels of immigration that are located in

---

<sup>16</sup>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).

<sup>17</sup>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 2006 is likely to be correlated, implying that native migration measured later may either be longer-run responses to immigration in 2006, or short-run responses to later immigration.

the same municipality. Second, we include linear, quadratic and cubic controls for population size (*pop*) in 2006. The purpose of these is to flexibly control for the fact that, in absolute terms, larger neighborhoods typically experience larger immigration inflows as well as larger population turnover in general. Third, to account for the fact that different neighborhoods differ also in relative population turnover, we add variables capturing historical in- and outmigration. Specifically, we control for the average number of natives who moved out of as well as the number of natives who moved into the neighborhood during the period 1990-2000.<sup>18</sup> We consider these variables to be important additions to the model, since they may well be correlated with the variation in immigration stemming from the refugee placement program, and since population turnover may be rather persistent. Fourth, since immigration from Iraq, Syria and Somalia could be correlated with immigration from other countries, which in turn could lead to further outflow, we control for immigration in 2006 from all other non-OECD countries. Fifth, we add controls for the neighborhood fraction of individuals born in Syria, Somalia, Iraq and all other non-OECD countries in 2005.<sup>19</sup> We know that immigration is correlated over time (this is the very logic behind the shift-share instrument). Hence, any out-migration of natives following the 2006 immigration could be a result of immigration in earlier years. By controlling for the share in 2005 we make sure that the out-migration is only due to additional immigration after 2005.<sup>20</sup> Finally, we control for a set of socio-economic characteristics of the neighborhood (also measured in 2006); average disposable income, the number of individuals with at least some university education (or equivalent), the number of individuals living in public rentals, and the per capita cost of social assistance.<sup>21</sup>

Letting the vector  $\mathbf{X}$  include the variables for historical turnover, other

---

<sup>18</sup>The exact definition is:  $Outflow_{his} = \sum_{t=1990}^{2000} \frac{Outflow_{ij,t}}{11}$  and  $Inflow_{his} = \sum_{t=1990}^{2000} \frac{Inflow_{ij,t}}{11}$ . The results are not sensitive to alterations of the period used for obtaining a variable that controls for historical, neighborhood-specific, population turnover. In particular, using the period 1990-2005 does not affect the results.

<sup>19</sup>Note that this is different from the variation used in the instrument, were we use anyone who immigrated *from* from Iraq, Syria and Somalia. When using the stock in 2005 we have no other option but to use those born in these countries, since information on which country a certain individual immigrated from, is not known to us prior to 1990.

<sup>20</sup>Since the level of individuals born in Syria, Somalia, and Iraq is highly correlated with the instrumental variable, we are not able to control for these variables in levels (it would kill all variation in the first-stage estimation) but have to use them as shares. We do however consider this as an improvement compared to the literature using decennial data; in this type of literature it is only possible to control for variables dated ten years back, which increases the probability that natives' migration today is a function of immigration in intervening years.

<sup>21</sup>One potential concern is that defining the covariates in the same year as when immigration is measured creates a bad control problem—that is, that we control for things that are in fact responses to immigration. However, re-running the models with all controls based in 2005 instead of 2006 does not alter the baseline coefficients of the paper.



non-OECD immigration, the fraction born in the three chosen source countries as well as in other non-OECD countries and the socio-economic characteristics, the first stage in our shift-share approach hence looks as follows:

$$im_{ij,2006} = \gamma \tilde{im}_{ij,2006} + \sum_{p=1}^3 \phi^p pop_{ij,2006}^p + \Gamma \mathbf{X} + \mu_j + \epsilon_{ij,2006} \quad (5)$$

The prediction  $\widehat{im}_{ij,2006}$  from the first stage is then used in the two equations capturing the migratory response of the native population:

$$outflow_{ij,2006+s} = \beta^{out} \widehat{im}_{ij,2006} + \sum_{p=1}^3 \delta^p pop_{ij,2006}^p + \Pi \mathbf{X} + \mu_j + \epsilon_{ij,2006+s}^{out} \quad (6)$$

and

$$inflow_{ij,2006+s} = \beta^{in} \widehat{im}_{ij,2006} + \sum_{p=1}^3 \delta^p pop_{ij,2006}^p + \Pi \mathbf{X} + \mu_j + \epsilon_{ij,2006+s}^{in}, \quad (7)$$

with  $s = 1$  for most of the analysis, and  $s = 8$  for the longer-run analysis.

Our approach thus identifies effects on native migration of immigration stemming from the fact that placement policy-induced immigrant settlement differs across neighborhoods in a municipality, and following an international shock, different neighborhoods will experience different immigrant inflows.

## 5 The GeoSweden data

In this section we present our data, which is obtained from the GeoSweden database, and how we define our key concepts of “neighborhoods”, “foreign immigration” and “natives”. All data is collected and made anonymous by Statistics Sweden.

The data used for the analysis comes from the comprehensive database GeoSweden, which is administered by the Institute for Housing and Urban Research at Uppsala University. The database, which 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 that the database includes detailed geographical information (given by geographic coordinates) on where the individuals live, information on the date and from which country an individual immigrates to Sweden, along with annual information on migration patterns within Sweden. The geographical coordinates make up administrative units such as municipalities, parishes, etc., which we use to define a “neighborhood”. In particular, 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 all those that were not tractable across our time period, or that lack population at some point in time. This leaves us with 8,723 SAMS (neighborhoods). The average number of SAMS per municipality is 130, although excluding Gothenburg (a clear outlier), the number is 69. Naturally, the number of neighborhoods per municipality is highly correlated with the population. We analyze the sensitivity of the first stage to the type of SAMS in Section 6.1.

The information on source country is used to define immigrants from Iraq, Syria and Somalia (and other non-OECD immigrants that we control for). Our main definition of “native” is anyone born in Sweden. We also provide a subset of results separately for second-generation immigrants and those born in Sweden with both parents being Swedish-born.

## 5.1 Descriptives

Table 1 provides summary statistics of the variables used in the analysis, along with a clarifying description. As can be seen, for the main endogenous immigration variable as well as its instrument (corresponding to  $im_{ij,2006}$  and  $\hat{im}_{ij,2006}$  in the above equations), the standard deviations are large relative to their means. This reflects the fact that roughly 80 percent of the neighborhoods did not receive any immigrants from our three source countries in 2006 ( $im_{ij,2006}$ ) or in the baseline period ( $\hat{im}_{ij,2006}$ ).<sup>22</sup> To get a sense of the variation in the data conditional on having positive immigration in these respective periods, Figure 5 shows the remaining conditional distribution of these two immigration variables. As can be seen, the majority of neighborhoods have a fairly low level of immigration. 80 percent of the neighborhoods received 10 people or less, while 90 percent received 20 or less.

---

<sup>22</sup>The fact that most of the variables have minimum value zero is because there is a substantial number of really small SAMS.

Figure 5: Immigration from Iraq, Somalia and Syria, actual and predicted.

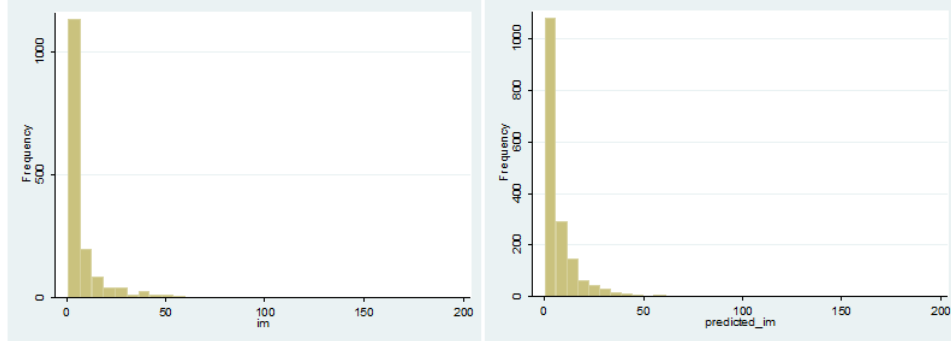


Figure 5 also suggests that the two distributions are highly correlated.<sup>23</sup> This is indicative of a strong instrument, and we show below that this is indeed the case. But before turning to the estimations, following up on the potential mechanisms behind observed native migration as discussed in Section 3.1, we assess from the data the extent of potential price substitution effects.<sup>24</sup> In particular, we estimate the correlation coefficient between the amount of public rental housing and the main immigration variable to be 0.55, whereas the corresponding correlation with the predicted immigration (the instrument) is 0.50.<sup>25</sup> The share of public housing is in turn a substantial part of the rental apartments. In the 3,467 neighborhoods with at least 1 public rental inhabitant, public renters constitute almost 70 percent of the people renting. Thus, these relatively high correlations confirm the notion that increasing property prices due to substitution effects between rental and owner-occupied housing are small, and are likely to be well captured by controlling for the amount of public rentals in the regressions.

## 6 Results

We now turn to the results. After establishing in Section 6.1 that our shift-share instrument works well in the first stage regression, we estimate the second stage effects of foreign immigration on native migration in Section 6.2. In section 6.3 we provide separate analyses for the effects on natives with both parents born in Sweden and for natives with at least one parent born outside the country, and in section 6.4 we provide heterogeneity analyses based on socio-economic characteristics (income and education level).

<sup>23</sup>The raw correlation between the two variables is 0.63

<sup>24</sup>Recall from above that we expect direct price effects to be negligible, as immigrants tend to move into price-regulated rental apartments.

<sup>25</sup>Net of the municipality fixed effects – thus using the same variation as in the empirical model – these correlations are instead 0.54 and 0.48, respectively.

Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
<i>Key variables:</i>					
Outflow	8,723	88.0	122.7	0	2294
Inflow	8,723	87.8	126.8	0	2545
Immigration (main)	8,723	1.5	9.3	0	352
Predicted immigration (instrument)	8723	1.5	7.3	0	273
<i>Control variables:</i>					
Population	8,723	1,036	1,250	1	19,783
Disposable income	8,723	1,738	525.3	-175.5	25,119
Highly educated	8,723	225.0	394.2	0	9,045
Public rentals	8,683	144.8	449.5	0	9,158
Social assistance	8,723	7.9	18.4	0	595
Historic inflow	8,714	84.1	114.1	0	2,133
Historic outflow	8,713	84.1	109.6	0	2,030
Other non-OECD immigration	8,723	5.0	16.8	0	697
Percentage Syrians 2005	8,723	0.1	0.5	0	12.5
Percentage Iraqis 2005	8,723	0.4	1.6	0	38.8
Percentage Somalis 2005	8,723	0.1	0.6	0	19.0
Percentage from other non-OECD countries	8,723	3.8	5.1	0	100

Outflow and Inflow measure the number of natives moving out of and into a given neighborhood between 2006 and 2007, respectively. This goes for historic inflow and outflow as well, only the historic variable is based on the average number between 1990 and 2000. Immigration (main) is the main (endogenous) independent variable, measuring the number of Iraqis, Somalis and Syrians who came to a neighborhood in 2006, and Predicted immigration is the instrument for this variable. Population denotes total SAMS population, Disposable income is measured in 100 SEK, Highly educated denotes the number of individuals with a university education, Public rentals denotes the number of individuals staying in publically owned rental properties, all measured in 2006. Percentage Syrians, Iraqis and Somalis denotes the percentage of the neighborhood population born in the respective country, measured in 2005. The unit of observation is SAMS area.

Finally, we examine the long run effects in section 6.5.

## 6.1 Relevance and stability of the first stage

Table 2 shows the baseline estimations of the first stage in equation (5). The first column regresses the inflow of immigrants from Syria, Iraq and Somalia in 2006 to neighborhood  $i$  and municipal  $j$  on the predicted inflow to that very neighborhood. The covariates included as well as the detailed specification is discussed in Section 4.3. A coefficient of 1 implies perfect correlation; that is, a prediction based on previous settlement patterns of one more immigrant into a neighborhood  $i$  corresponds to an actual inflow of one more immigrant to that very neighborhood.

Looking at the estimates we see that, given population, socio-economic characteristics of the neighborhood, historic migration of natives, the share of non OECD inhabitants in the prior year and other non OECD immigration, the prediction explains more than 25 percent of the variation. The coefficient is also highly significant. We consider the model in the first column to be our preferred specification.

In order to assess the stability of the first stage, we alter the baseline first stage regression in various ways (see columns 2–5 in Table 2). First, neighborhoods differ with respect to the size of the population. Taking away the smallest decile (neighborhoods with less than 120 inhabitants; see column 2), however, only decreases the coefficient marginally. The opposite is perhaps more likely, that is that the results are driven by *more* populated neighborhoods. We expect this since larger neighborhoods are likely to have a more consistent flow of immigration.<sup>26</sup> We therefore drop the top decile, which consists of all neighborhoods with more than 2,065 inhabitants. As can be expected, this exclusion affects the size of the estimate more, with a drop from 0.26 to 0.037 (see column 3). The coefficient is still significant, but going forward, it is worth keeping in mind that more populated neighborhoods make up most of the variation in the instrument.

Second, municipalities differ with respect to how many SAMS there are. For example, Gothenburg is a clear outlier; 758 SAMS in our sample belong to Gothenburg, whereas only two other municipalities have more than 200 neighborhoods.<sup>27</sup> We therefore run the baseline specifications without Gothenburg. The change is, however, negligible (c.f. column 4). Lastly, we exclude the decile of municipalities with the fewest SAMS areas (less than 12). Again, this does not affect the estimate (c.f. column 5).

All in all, we deem our first stage estimations to be stable, but note

---

<sup>26</sup>Recall that the majority of neighborhoods did not experience any immigration at all from our three source countries. Furthermore, since our model is specified in absolute levels (controlling for population) rather than in population shares, more weight is put on more populated areas.

<sup>27</sup>Only Malmö with 324 and Uppsala with 223.

Table 2: First stage and robustness of baseline estimation

VARIABLES	(1) First stage Baseline	(2) First stage > Pop 120	(3) First stage < Pop 2065	(4) First stage no Gothenburg	(5) First stage no small municipals
$\tilde{im}_{ij,2006}$	0.269*** (0.0574)	0.253*** (0.0605)	0.037* (0.0204)	0.268*** (0.0541)	0.279*** (0.0571)
Observations	8,674	7,851	7,801	7,932	7,875
Municipality Fixed effects	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors clustered on municipality. See Section 4.3 for description and explanation of the covariates included.

that the more populated neighborhoods make up a significant part of the identifying variation.

## 6.2 Native flight and avoidance: Baseline model

Moving to the 2SLS estimations of equations (6) and (7), Table 3 shows the results for flight (outflow) and avoidance (inflow) of natives. For neighborhood  $i$ , any native that resides in  $i$  on the last day of 2006, but lives in another neighborhood  $-i$  on the last day of 2007 is counted as outflow, while any native residing in area  $i$  on the last day of 2007 but in another neighborhood  $-i$  on the last day of 2006, is counted as inflow. All columns include standard errors clustered at the municipality level as well as municipality fixed effects.

Starting with the flight estimates, the results are given in the upper panel of Table 3. From column 1 we note that the outflow coefficient is approximately 0.6, implying that, on average, 0.6 natives leave the neighborhood in 2007 for each immigrant that arrived in 2006. As discussed in Section 3.1, the observed flight is likely driven by residential preferences rather than by price and mechanical effects operating through the housing market. We therefore consider the size of the effect to be quite substantial.

Under the hypothesis that natives react to the arrival of new immigrants because of preferences for homogeneity, those moving out of a neighborhood should move to a neighborhood with a lower share of immigrants (i.e., to a more native-dense neighborhood). To examine this, we separately study migration to more native-dense neighborhoods. The results, presented in column 2, are in line with the hypothesized mechanism; the statistically significant point estimate indicates that one immigrant to a neighborhood in 2006 cause, on average, 1.6 more natives to move to a neighborhood with a lower share of immigrants (compared to the number of natives moving to such neighborhoods from places that did not receive any immigrants). This targeted outflow is thus much larger than the aggregate outflow, imply-

Table 3: Second stage baseline estimations: Native flight (outflow) and native avoidance (inflow) in 2007

	(1)	(2)
	All moves	To native dense areas
VARIABLES	Outflow	Outflow
$im_{ij,2006}$	0.636** (0.283)	1.614** (0.792)
VARIABLES	Inflow	-
$im_{ij,2006}$	1.163** (0.496)	-
Observations	8,674	8,674
Municipality Fixed effects	YES	YES
Controls	YES	YES

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors clustered on municipality. Column (1) looks at all moves, column (2) looks at moves to neighborhoods with a larger share of natives. See Section 4.3 for description and explanation of the covariates included.

ing that fewer natives in immigrant-receiving neighborhoods move to other immigrant-dense places (again, compared to in neighborhoods that did not experience as much immigration).

Turning to the avoidance (inflow) estimates (see the results in the lower panel of Table 3), immigration has a positive and statistically significant effect on natives' in-migration behavior; for every immigrant arriving in a neighborhood, approximately one native moves into the neighborhood. This result is somewhat surprising, and might actually suggest native attraction rather than native avoidance. The result is even more surprising considering that we observe a positive outflow coefficient. As we will argue later on (see Section 6.4), both the observed outflow and the observed inflow of natives are likely part of a socio-economic sorting pattern.

The results in Table 3 hence indicate native flight but not native avoidance. As indicated, a central question regarding the interpretation of these results is if individuals react on increased ethnic diversity, or if they react on some other characteristics correlated with newly arrived immigrants (such as education level, income level, or some other socio-economic characteristics). This question is relevant for all related papers in the literature, but is rarely addressed. However, with our very detailed individual-level data, we are in a good position for examining these issues a bit further. To this aim, in the next section, we start to dig into this question by disentangling the migratory behavior of native-born individuals with native-born parents

from the migratory behavior of native-born individuals with at least one foreign-born parent.

### **6.3 Flight and avoidance among “native-natives” and “second-generation immigrants”**

The estimates in Table 3 are based on the definition of native as anyone born in Sweden. Included in this definition is a non-negligible group of “second-generation immigrants” (here defined as individuals born in Sweden with at least one parent born abroad). In this section, we disentangle the flight and avoidance behavior among this group from that of the rest (which we refer to as “native-natives”). Being able to make this distinction is very valuable, since it helps to understand whether the observed flight behavior is a reaction to the ethnic or to the socio-economic dimension of immigration; if people born in Sweden with Swedish-born parents react more strongly, it is more likely that migration decisions are taken on ethnic or racial grounds.

The results in Table 4 yields a pattern that indicates that the division into “native-natives” and second-generation immigrants is important for understanding the baseline results. While the positive inflow coefficient in Table 3 seems to be entirely driven by “native-natives” (c.f. the coefficients in columns 1 and 3 in the lower panel of Table 4), the outflow (flight) coefficient in Table 4 is mainly driven by the out-migration of second-generation immigrants; the outflow coefficients for those born in Sweden with Swedish-born parents are statistically insignificant while the same coefficients for second-generation immigrants are positive and statistically significant (see columns 1–4 in the upper panel in Table 4). These results hence indicate that ethnicity might not be the main driving characteristic for the migration-decision among those born in Sweden.

To further our understanding of the role of socio-economic characteristics, we next examine if the migratory behavior of high-income earners and the highly educated differ from that of low-income earners and the low-educated.



Table 4: Second stage estimations: Separate flight and avoidance results in 2007 for Swedish-born with Swedish-born parents ("native-natives") and Swedish-born with at least one foreign-born parent (second generation immigrants).

	(1)	(2)	(3)	(4)
	All moves	To native dense areas	All moves	To native dense areas
	"Native-natives"		Second generation immigrants	
VARIABLES	Outflow	Outflow	Outflow	Outflow
$im_{ij,2006}$	0.176 (0.236)	0.851 (0.546)	0.460* (0.257)	0.762** (0.354)
VARIABLES	Inflow	-	Inflow	-
$im_{ij,2006}$	1.095*** (0.395)	- -	0.068 (0.177)	- -
Observations	8,674	8,674	8,674	8,674
Municipality Fixed effects	YES	YES	YES	YES
Controls	YES	YES	YES	YES

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Second stage estimations of native flight and avoidance. See Section 4.3 for description and explanation of the covariates included.

## 6.4 Flight and avoidance among high-income earners and the highly educated

We proceed by examining the natives' migratory behavior separately for, on the one hand, the highly educated and the low-educated and, on the other hand, the high-income earners and the low-income earners.

Starting with the education dimension, we group the natives into two categories: those with a university degree (highly educated) and those who did not complete upper secondary school (low-educated).<sup>28</sup> To make sure that we are not picking up any differences across neighborhoods in education levels, we add flexible controls for the number of individuals in a neighborhood with a university education, or who did not finish secondary school. The results for the two educational groups are provided in Table 5. Columns 1 and 2 show the results for the highly educated, and columns 3 and 4 show the results for the low-educated. While there is a positive and significant effect of immigration on the outflow of the highly educated, the low-educated have a tendency to stay. But it is also clear that immigration has a positive effect on the inflow of the highly educated.<sup>29</sup>

Turning to the income dimension, we define the native high-income earners as those with the 25 percent highest disposable income among the full population, and the native low-income earners as those with the 25 percent lowest. Again, to make sure that we are not picking up any difference across neighborhoods with respect to income levels, we add flexible controls for the number of individuals in each neighborhood who belong to the richest as well as the poorest category.<sup>30</sup> From the estimations, presented in Table 6, we note that the outflow-results are fairly similar for high- and low-income individuals; while the two groups do not display any flight effects on average, there is a tendency for both groups that increased immigration causes them to move to native dense areas. However, from the inflow results, there are no clear signs of avoidance for any of the groups, but there is a clear difference between the groups in the sense that low-income tend to move in to neighborhoods that have witnessed increasing immigration.<sup>31</sup>

All in all, the results obtained indicate sorting based on income and

---

<sup>28</sup>Equivalent to Swedish gymnasium.

<sup>29</sup>When we ran the same estimations for second-generation immigrants and for the Swedish-born with parents born in Sweden separately, we found that the significant effects found in Table 5 is in principle entirely driven by the "native-natives". These results are available upon request.

<sup>30</sup>Since we are adding controls in this as well as in the heterogeneity analysis along the education dimension, we also rerun the first stage estimations. These estimations, which are barely affected, along with the exact specification can be found in the Appendix.

<sup>31</sup>When we ran the same estimations for second-generation immigrants and for the Swedish-born with parents born in Sweden separately, the results show that the outflow of high-income earners are mainly driven by second-generation immigrants and the inflow of the low-income earners are mainly driven by "native-natives". These results are available upon request.

Table 5: Second stage estimations: "Flight" and "avoidance" in 2007 among high- and low-educated.

	(1)	(2)	(3)	(4)
	University educated All moves	University educated To native dense areas	No education All moves	No education To native dense areas
VARIABLES	Outflow	Outflow	Outflow	Outflow
$im_{ij,2006}$	0.863*** (0.257)	0.723** (0.311)	-0.210 (0.141)	0.137 (0.303)
VARIABLES	Inflow	-	Inflow	-
$im_{ij,2006}$	0.774*** (0.244)	-	0.426 (0.447)	-
Observations	8,674	8,674	8,674	8,674
Municipality Fixed effects	YES	YES	YES	YES
Controls	YES	YES	YES	YES

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Second stage estimations of native flight and avoidance of those with different levels of education. See Section 4.3 (equation 5) for description and explanation of the covariates included. Note that to the baseline specification, we add flexible controls for number of highly educated ( $\sum_{p=1}^3 \delta^p univ_{ij,2006}^p$ ) in (1) and (2), and flexible number of low educated in (3) and (4).

education rather than sorting on ethnicity. A likely explanation for the results in the heterogeneity analyses is simply that those with higher incomes and those with a higher education are, in general, less resource constrained and therefore have the option of moving. The outflow of the highly educated also decreases the status of the area, which gives economic incentives for low income groups not to move away, but rather to move in.<sup>32</sup> To get a final piece to the puzzle of understanding whether the natives' migration response to immigration is based on ethnicity or whether it is based on correlated socio-economic characteristics, we have also estimated migration responses conditioning on moving to a neighborhood with higher average income. The results indicate that both those with high income and the highly educated tend to move to areas with higher disposable income as a response to immigration.<sup>33</sup> The numbers are highly similar to those given in Table 3, again demonstrating the difficulty of separating socio-economic patterns from that of ethnic preferences.

<sup>32</sup>Theoretically, however, we could expect any sign on the education coefficient. If we accept that education can be a counter-force against xenophobia, we would expect less avoidance or flight when looking at the educated. However, as noted educated people also have a better possibility to move, due to more resources and labor market options.

<sup>33</sup>Results available upon request

Table 6: Second stage estimations: "Flight" and "avoidance" in 2007 by income groups.

	(1)	(2)	(3)	(4)
	High income All moves	High income To native dense areas	Low income All moves	Low income To native dense areas
VARIABLES	Outflow	Outflow	Outflow	Outflow
$im_{ij,2006}$	0.183 (0.201)	0.560** (0.250)	0.110 (0.113)	0.493** (0.209)
VARIABLES	Inflow	-	Inflow	-
$im_{ij,2006}$	-0.134 (0.187)	-	0.826*** (0.203)	-
Observations	8,674	8,674	8,674	8,674
Municipality Fixed effects	YES	YES	YES	YES
Controls	YES	YES	YES	YES

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Second stage estimations of native flight and avoidance of those with different levels of income. See Section 4.3 (equation 5) for description and explanation of the covariates included. Note that to the baseline specification, we add flexible controls for number of high income people ( $\sum_{p=1}^3 \delta^p highincome_{ij,2006}^p$ ) in (1) and (2), and flexible number of low income people in (3) and (4).

## 6.5 Long-run effects

Lastly, we look at the long-run migration response, measured as natives out- and in-migration to different neighborhoods between the years 2006 and 2014. Apart from the dependent variable in the second stage (the natives migration response), the model we estimate is the same as the one used for obtaining the short run estimates (i.e., same instrument and same controls, yielding exactly the same first stage).

These long-run estimates are similar in type to the ones estimated earlier in the literature, in which decennial or other infrequent data has been used. In these type of estimations, one should be very careful talking about causal effects. Since immigration is measured much earlier (in our case in 2006) than the natives migration response (in our case in 2014), the migration response can be the result of both the initial immigration (2006 in our case) and immigration in the intervening years (2007–2013 in our case). For this reason, we are reluctant to interpret the coefficient for the immigration variable in our long-run estimates solely as long-run effects of immigration (although this has been practice in the earlier literature)<sup>34</sup>.

Comparing the long-run estimates (see Table 7) with the short-run baseline estimates (see Table 3), two things can be noted. First, comparing the

<sup>34</sup>See the discussion in section 4.3

results in the upper panel in both tables, the coefficient in the outflow regression for all moves turns negative and becomes insignificant, indicating no native flight on average in the long run. However, the coefficient for out-migration to native dense areas becomes almost three times bigger. Second, comparing the results in the lower panel in both tables, the coefficient in the inflow regression is significant and the point estimate increases almost three times, indicating that neighborhoods that received more immigrants in 2006 have attracted more natives in the long run (but, again, we cannot say that this is an effect of the 2006 immigration only).

When we divide the natives into those with native-born parents (“native-natives”) and with at least one parent being foreign-born (second-generation immigrants), an interesting pattern emerges (see Table 8). In the long run, it seems as though the neighborhoods that received immigrants in 2006 have attracted native-born individuals with native-born parents (c.f. column 1); there has been both a significantly lower out-migration (upper panel) and a significantly higher in-migration of “native-natives” (lower panel) in the neighborhoods that received more immigrants in 2006. For the second-generation immigrants, the pattern is different; the coefficients in the out-migration regressions increase several times compared to the baseline estimates in Table 5, being strongly positive and highly significant (c.f. columns 3 and 4).

Dividing the sample further by income and educational level shows that it is the high-income second-generation individuals who drive these results.<sup>35</sup> The high-income second-generation immigrants both move away from and avoid moving into the immigrant-receiving neighborhoods in the long run. We also see a similar pattern regarding low income as in the short run results. Low income natives with Swedish born parents seem to move in to areas with higher immigration. A possible explanation for this is that the higher degree of immigration increases the demand for rental apartments. If more rental apartments are built, it can further attract a higher share of resource constrained individuals, including low income natives. It should however be noted that in the long run, also high income “native-natives”, although in a relatively smaller degree, are attracted to immigrant neighborhoods.

---

<sup>35</sup>Results available upon request.

Table 7: Second stage long run estimations: Native flight (outflow) and native avoidance (inflow) 2006-2014

	(1)	(2)
	All moves	To native dense areas
VARIABLES	Outflow 2006-2014	Outflow 2006-2014
$im_{ij,2006}$	-1.453 (1.049)	4.550*** (1.716)
VARIABLES	Inflow	-
$im_{ij,2006}$	3.510** (1.666)	- -
Observations	8,665	8,665
Municipality Fixed effects	YES	YES
Controls	YES	YES

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors clustered on municipality. Column (1) looks at all moves, column (2) looks at moves to neighborhoods with a larger share of natives. See Section 4.3 for description and explanation of the covariates included.

Table 8: Second stage long run estimations: Separate flight and avoidance results 2006-2014 for Swedish born with Swedish born parents ("native-natives") and second generation immigrants.

	(1)	(2)	(3)	(4)
	All moves	To native dense areas	All moves	To native dense areas
	"Native-natives"		Second generation immigrants	
VARIABLES	Outflow	Outflow	Outflow	Outflow
$im_{ij,2006}$	-3.562*** (1.247)	0.937 (1.403)	2.109*** (0.738)	3.614*** (0.973)
VARIABLES	Inflow	-	Inflow	-
$im_{ij,2006}$	3.921*** (1.371)	- -	-0.411 (0.488)	- -
Observations	8,665	8,665	8,665	8,665
Municipality Fixed effects	YES	YES	YES	YES
Controls	YES	YES	YES	YES

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Second stage estimations of native flight and avoidance. See Section 4.3 for description and explanation of the covariates included.

## 7 Concluding remarks

In this paper, we have applied detailed, comprehensive, register data to a refined shift-share methodological approach to answer whether native flight and avoidance are important phenomena in Sweden. Our study spans the period 1990–2014, 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 helps in improving the shift-share instrument.

Using push-driven immigration to Sweden, generated by wars and famines in Iraq, Syria and Somalia in 2006, in combination with a settlement pattern of their countrymen in Sweden in the early 1990s that was partly generated by a state-run placement policy, we reach four main conclusions.

First, while we do not find any indications of native avoidance, we find signs of native flight in the short run (in which the follow-up horizon is one year). The general mean effect is 0.6 natives moving out for each immigrant moving in.

Second, the outflow (flight) that we do observe is primarily driven by high-income “second-generation immigrants” (Swedish-born individuals with at least one parent born abroad), rather than by Swedish-born individuals with Swedish-born parents (“native-natives”). For individuals with low income and low education-level, we find in general small or no effects on flight.

Third, some of the inflow coefficients are positive, indicating that higher immigration causes more natives to move into the neighborhood. This type of migration behavior seems to be concentrated among the low-income “native-natives”, as well as among the highly educated “native-natives”.

Fourth, estimates over a longer time period (eight years) indicate that the socio-economic sorting observed in the short run is even more pronounced in the longer run.

The explanation for these results likely lies in resource constraints, where those with higher education and higher income have larger possibilities to migrate. All in all, our results indicate that immigration to Sweden in 1990–2014 caused a sorting process along socio-economic lines rather than along ethnic lines.

Even though the “native-natives” do not seem to react on immigration by moving out of the neighborhood, there might be an increased uncertainty regarding how the rising share of foreign-borns will affect the quality in the local school. As a result of this increased uncertainty, parents might choose to take advantage of the voucher system existing in Sweden and place their children in schools further away from home. This is a topic that is next on our research agenda.

## References

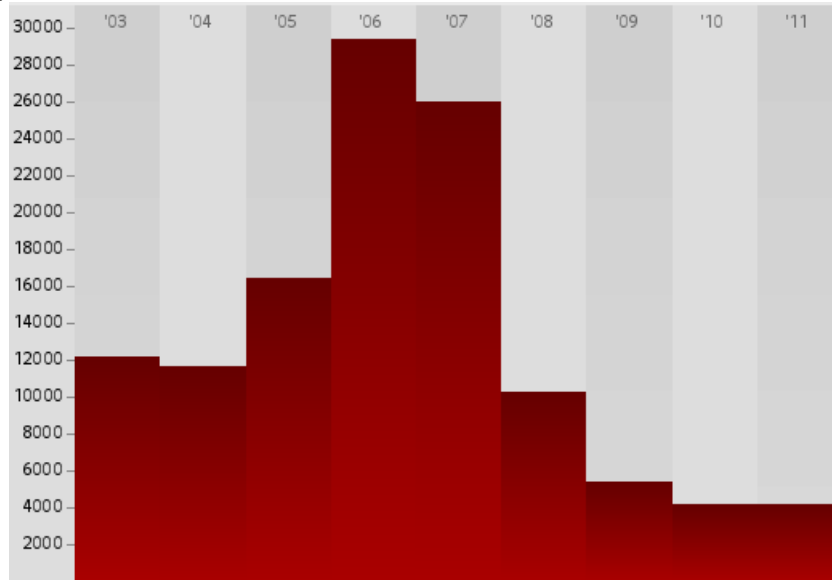
- ALDÉN, L., M. HAMMARSTEDT, AND E. NEUMAN (2015): “Ethnic Segregation, Tipping Behavior, and Native Residential Mobility,” *International Migration Review*, 49, 36–69.
- ALTONJI, J. G. AND D. CARD (1991): “The Effects of Immigration on the Labor Market Outcomes of Less-Skilled Natives,” in *Immigration, Trade, and the Labor Market*, University of Chicago Press, 201–234.
- BLIND, I. AND M. DAHLBERG (2015): “Immigration, New Religious Symbols, and the Dynamics of Neighborhoods,” Mimeo Uppsala University.
- 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.
- 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.
- SHELLING, 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.

## 8 Appendix

Figure 6: Maximum number of deaths from the Iraqi war according to Iraqi body count



Source: [www.iraqbodycount.org](http://www.iraqbodycount.org).

Table 9: First stage. Including controls for heterogeneity analysis

VARIABLES	(1) Univ	(2) low educ	(3) high inc	(4) low inc
$im_{ij,2006}$	0.257*** (0.0560)	0.181*** (0.0693)	0.208*** (0.0518)	0.220*** (0.0524)
Observations	8,674	8,674	8,674	8,674
Fixed effects	YES	YES	YES	YES
Controls	YES	YES	YES	YES

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors clustered on municipality. See Section 4.3 for description and explanation of the covariates included. Note that to the baseline specification, we add flexible controls for number of high educated people in (1), low educated in (2), high income people in (3) and low income people in (4).