Revisiting the Relationship between Ethnic Diversity and Preferences for Redistribution*

Lena Nekby* and Per Pettersson-Lidbom

JEL: D64, I30, Z18, J15, J18

Key words: ethnic diversity, income redistribution, re-analyzes

This draft: August 23, 2012

Abstract

In this paper, we return to the question raised in Dahlberg et al. (2012) concerning a causal relationship between ethnic diversity and preferences for redistribution. A re-analysis of their study indicates that results are based on an endogenous instrument and severe sample attrition bias. Correcting for either of these two problems reveals that there is no relationship between ethnic diversity and preferences for redistribution. More generally, we provide results that put into question the conventional description of the Swedish refugee placement policy.

* We thank Thina Carlsson at the Archives of the Swedish Board of Immigration for generous help with data collection. We also thank Torsten Persson, David Strömberg, Mårten, Palme, Peter Skogman Thoursie, and Peter Fredriksson for useful comments.
* Department of Economics, Stockholm University, E-mail: lena.nekby@ne.su.se
* Department of Economics, Stockholm University, E-mail: pp@ne.su.se
1. Introduction

Numerous papers have analyzed the relationship between increased ethnic heterogeneity and the size of the welfare state or preferences for redistribution (e.g. Alesina et al., 2001 and Alesina and Glaeser, 2004). In a recent study, Dahlberg, Edmark and Lundqvist (2012) add to this literature by using a refugee placement policy which, they argue, provides exogenous variation in the number of refugees placed in Swedish municipalities from 1985 to 1994.

Here, we revisit this question and show (i) that their instrument (actual refugee settlement) does not provide exogenous variation in refugee settlement and (ii) that their results are compromised by extensive sample attrition bias stemming from unwarranted sample restrictions not discussed in the paper.

The Swedish refugee placement policy, discussed in greater detail below, was in place from 1985 to 1994. During this time, the Swedish Board of Immigration (SIV) and local municipal governments negotiated and signed contracts concerning the number of refugees any given municipality commits to annually for a given period of time. Thereafter, as refugees arrived in Sweden, SIV and respective contracted municipality came to agreements on actual refugee placements.1 Dahlberg et al. (2012) use actual refugee settlement as an instrument variable for their main (endogenous) explanatory variable of interest, the share of immigrants in a municipality.2 We collected data on the contracts established between SIV and municipalities and find that there is little or no correlation between contracted levels and actual refugee settlement. This implies that Dahlberg et al. use one endogenous variable, actual refugee settlement, to instrument for another endogenous variable, share of immigrants.

Another serious problem with the Dahlberg et al. study concerns attrition bias resulting from unwarranted sample restriction choices. We show that this leads to a very large sample attrition rate of 66 percent leading to a large bias (more negative) in their estimated coefficients. In contrast, when estimation is based on samples with considerably less sample attrition, the estimated coefficients are reduced considerably and are no longer significantly different from zero in the sample with the smallest attrition rate (33 percent). Even more importantly, when IV estimation is based on the correct definition of the policy (contracted number of refugees), there is no statistical relationship regardless of estimation sample used.

1 There are two types of policies; discretionary and rule-based. Discretionary policies, by definition cannot be used as an exogenous source of variation while rule-based policies, under certain conditions, can (Besley and Case, 2000).
2 See also Dahlberg and Edmark (2008) for a similar use of this instrument as an exogenous source of variation in welfare benefit levels.
In this study, a critical analysis of the Swedish refugee placement policy is undertaken in order to provide new information concerning the degree to which this policy can be described as providing conditionally random placement of refugees to municipalities, i.e. exogenous variation in refugee settlement. Thereafter, the robustness of results reported in Dahlberg et al. (2012) to use of the contracted levels of refugee assignment as an instrument variable in the analysis is tested. In addition, numerous sample restrictions made in Dahlberg et al. are discussed and the sensitivity of results to these restrictions tested.

The rest of this paper is organized as follows. In Section 2, the exogeneity of the Swedish refugee placement policy is discussed and tested. In Section 3, the Dahlberg et al study is re-analyzed focusing particularly on instrument endogeneity and sample selection issues. Results are presented in Section 4 and concluding remarks are presented in Section 5.

2. Exogeneity of the refugee placement policy
In this section, we analyze whether the Swedish refugee placement policy provides an exogenous source of variation in refugee shares allowing it to be used as an instrumental variable to identify the causal impact of immigrant shares on preferences for redistribution. More formally, Dahlberg et al. (2012) define their first-stage relationship as

\[ x_{jt} = \alpha_j + \lambda_t + \beta z_{jt} + \epsilon_{jt} \]

where \( x_{jt} \) is the share of immigrants residing in municipality \( j \) at time period \( t \), \( z_{jt} \) is the instrument, i.e., the Swedish refugee placement policy, \( \alpha_j \) is a municipality fixed-effect and \( \lambda_t \) is a time fixed-effect. In other words, only the within-municipality variation in the policy instrument is used to identify the causal relationship of interest—the effect of immigration on preferences for redistribution.

Dahlberg et al. define the instrument as actual refugee settlement in municipalities rather than the contracted levels determined by the formal agreements between municipalities and the government authority in charge of refugee settlements, the Swedish Immigration Board (SIV). Surprisingly, the issue of how to correctly characterize the Swedish refugee policy has not been properly discussed, or empirically investigated, in the literature to date. Edin et al. (2003), the first study to exploit the refugee placement policy, write “Unfortunately, there is very little documentation about the practical implementation of the placement policy. Therefore, part of the information is based on interviews with placement
officers and other officials of the Immigration Board.” Similarly, Dahlberg et al. also largely rely on interviews with program officials.³

There are however a number of studies that have now documented how the Swedish refugee placement policy was implemented, many of which have not earlier been cited in the economics literature. This body of work provides a more varied and nuanced view on the workings of the refugee placement policy.

2.1 Description of the Swedish refugee placement policy

Dahlberg et al. (2012) describes the refugee placement policy as follows: “Under the program, refugees arriving to Sweden were consequently not allowed to decide themselves where to settle but were assigned to a municipality through municipal contracts, coordinated by the Immigration Board (the refugees were, however, allowed to move after the initial placement). At the start of the program, only a fraction of the municipalities were contracted, but as the number of refugees soared in the late 1980s and early 1990s, so did the number of receiving municipalities. By 1991, as many as 277 out of the then 286 Swedish municipalities had agreed to participate.” They also write that it became “harder for municipalities to dismiss the refugee placement proposal from the Immigration Board; the refugees had to be placed somewhere, and it became necessary that all municipalities shared the responsibility. Second, refusals of refugee placement were in fact very rare, and those who at first did refuse got a lot of negative publicity.” Finally they write “housing availability seems to have become the more important factor” governing refugee placement.⁴ Given this description of the placement policy, and note that their description is based on the contracts signed between SIV and the municipalities, they argue that “the program is quite likely to provide exogenous variation conditional on a set of municipality-specific covariates.”

This description, however, differs in important respects from the ones provided by for example Andersson (2003), Andersson and Solid (2003), Hammar (1993) and Soininen (1992, 1993).⁵ Specifically, these studies raise doubts about the following claims; (i) that refugees had little or no ability to decide themselves where to settle (ii) that the formal agreements, the contracts, between SIV and the local governments were binding and (iii) that the availability of housing vacancies was the key factor in determining refugee placement in municipalities.

³ Dahlberg et al. also base their description of the placement policy on information provided in Bengtsson (2002) and Edin et al. (2003).
⁴ Here they cite Edin et al. (2003) and Bengtsson (2003), see footnote 13 in Dahlberg et al. (2012).
⁵ For other studies on the Swedish refugee placement policy, see for example; Sarstrand (2011), Similä (1992) Ålund and Schierup (1991) and Wickström (2008).
Regarding the first claim, that refugees had no say on where they could settle, Hammar (1993) notes that all refugee placement should be voluntary and that the wishes and demands of the refugees themselves should be taken into consideration. Moreover, Hammar argues that refugees could refuse an offer of municipal placement. Those that did either stayed in the asylum centers until offered a new placement or made their own arrangements. Andersson (2003) argues along the same lines that “it was nevertheless obvious that the State’s control [of refugee settlement] was only temporary, and perhaps illusionary, as the refugee could choose to migrate whenever (s)he wanted.” The only cost of doing so was delayed enrolment in language courses (or other introduction programs) (Åslund et al., 2011).

Regarding the claim that the Swedish Immigration Board (SIV) could largely determine the terms and conditions in the written contracts, Andersson (2003) writes “At the beginning of 1989, the SIV staff faced ever growing problems in their attempts to find enough municipal places for refugees, and despite the fact that almost all municipalities were engaged in the 1989 reception programme, the number of places was reduced to 17,800”, from a previous level of 22,000-24,000. Likewise, Andersson writes “The autumn of 1989 saw an even greater influx of new asylum seekers. Refugee camps were crowded, the municipalities became more reluctant to further increase receptions.”.

Finally, regarding the claim that municipal (public) housing vacancies was the key factor determining the placement of refugees, Soininen (1992, 1993) notes that a lack of housing vacancies was not considered a valid argument for a municipality to refuse signing contracts for refugee settlement during this period. At the beginning of 1989, SIV tried to convince some municipalities to establish temporary housing in institutional buildings that were not currently in use. This strategy was not initially successful but as open unemployment dropped to below two percent in 1990, municipalities became more interested in temporary solutions implying that refugees were housed in for example, hotels, military tents and cruise ships (Andersson, 2008).

These studies on the Swedish refugee placement policy suggest that the characteristics determining both contracted and actual refugee settlement varied over time and across municipalities. We now turn to an empirical assessment of the refugee placement policy and the degree to which it can be considered to generate conditionally random placement of refugees across municipalities.
2.2 An empirical analysis of the refugee placement policy

In this section, we perform an empirical analysis of the Swedish refugee placement policy and its determinants and provide new evidence showing that the instrument used by Dahlberg et al. is not likely to be exogenous with respect to municipal characteristics. As noted earlier, Dahlberg et al. define their instrument as actual refugee settlement which is determined after contracts between municipalities and the government authority SIV have been negotiated. Using actual refugee settlement may be of little concern if the formal agreements correspond closely to actual settlement. However, to date no one has empirically analyzed if this is the case. We therefore compiled information on the formal written agreements between SIV and the municipalities for the time period 1986-1994 from the Archives of the Swedish Immigration Board.

By studying the written contracts, we see that the negotiated contracts varied across municipalities in both contract length and stipulations. Almost all contracts were formulated such that the municipality in question committed to preparing for a given number of refugee settlements per year during the contract period. Most contracts included clauses stating that SIV must consult with the municipality before each refugee placement on factors such as date of arrival, nationality, language skills, family composition, education and labor market experience of the refugee(s) in question. Contract length typically varied from one to three years and most contracts had a stipulation citing how and when contracts could be terminated. Contracts were effective only after local municipal government ratification. Most contracts included a clause stating that both partners in the agreement were to follow up refugee settlements in order to secure that settlement was in accordance with the goals stated in the national policy. Contracts were negotiated and re-negotiated during the program period and some terminated during the contract period. The municipality of Ragunda, for example, broke their contract of April 1991 in November 1991. Our reading of the contracts strongly suggests that municipals had large degrees of freedom in deciding the terms and conditions of these written agreements.

Empirically, we also find that there is a very low correspondence between actual refugee settlement and contracted levels. The percentage difference between actual and contracted refugee migration is plotted in Figure 1. Interestingly there are large, and both negative and positive deviations, in percentage terms in actual refugee settlement from the contracted levels. A more formal test of this relation is to estimate a regression of the form:

\[
\text{actual_settlements}_{jt} = \alpha_j + \lambda_t + \pi \text{written_agreements}_{jt} + v_{jt}
\]
where $\alpha_j$ is a municipality fixed-effect and $\lambda_t$ is a time fixed-effect. If the written agreements completely determined actual refugee settlements we would expect that $\pi = 1$. On the other hand, if $\pi \neq 1$, this would imply either greater leeway among refugees in determining where to settle or greater self-determination among municipalities in whether or not to abide by written agreements concerning potential placement. Thus, with equation (2) it is possible to make a joint test of the two claims made by Dahlberg et al.; that is to say (i) that refugees arriving to Sweden were not allowed to decide themselves where to settle and (ii) that municipalities could not affect the terms and conditions in the written contracts. It is not, however, possible to separate between these two potential explanations.

Table 1 presents results for the years 1986-1991, the period during which the placement policy is argued to be most “exogenous”. Results show that there is no (!) statistical relationship between actual and contracted refugee immigration to Swedish municipalities during this period. The point estimate of 0.007 is very close to zero and not significant. A formal test of $\pi = 1$ is also strongly rejected. Our results, therefore, clearly indicate that actual refugee placement is not determined by the written agreements generated by the refugee placement policy. As such, serious doubt is cast on whether actual refugee placement provides an exogenous source of variation for the endogenous share of immigrants in Dahlberg et al.’s main equation of interest.

Another way of empirically assessing to what degree the placement policy can be considered to generate exogenous, with respect to municipality characteristics, variation in immigrant shares is to analyze to what extent public housing vacancies determined both actual and contracted placement. Several studies argue that housing availability is the key determinant behind both contracted and actual refugee settlement during the policy period (Dahlberg et al., 2012, Åslund et al., 2011; Edin et al., 2003). The idea is that if the

---

6 Dahlberg et al. (2012) argue that variation in immigrant shares induced by the refugee placement policy is more likely to be exogenous during the initial years of the program, acknowledging that over time negotiations between municipalities and SIV concerning the possibility of settling fewer refugees increased. Åslund et al. (2011) also state that the policy formally was in place between 1985 and 1994 but that implementation was strictest between 1987 and 1991.

7 The natural or quasi-experimental approach emphasizes the importance of understanding the source of variation used to estimate key parameters. In the words of Meyer (1995) “researcher should seek to find variation that is driven by factors that are clearly identified and understood. One can then make an informed decision about the exogeneity of the variation and rule out other explanations.” Thus, if the variation is not well understood then it is hard to make a compelling case for exogeneity.

8 Dahlberg et al. (2012) claim that “the program is quite likely to provide exogenous variation conditional on a set of municipality-specific covariates”. Those mentioned include vacant public housing and local unemployment rates. Edin et al. (2003) state that “Government authorities placed refugees in localities that were deemed suitable according to certain criteria”. In practice, the availability of housing was the all-important factor. Finally, Åslund et al. (2011) argue that “Available public housing essentially determined the placement.”
relationship turns out to be weak or negative (contrary to expectation) then this would undermine the credibility of any argument suggesting that the refugee placement policy is exogenous conditional on this key observable factor.

Data on annual public housing vacancies, available at two different dates: March 1st and September 1st, were therefore collected. Table 2 displays the simple means for actual refugee settlement and the formal contracts for two groups of municipalities: those with no housing vacancies and those with available housing vacancies.⁹ Panel A in Table 2 shows means for the September data while Panel B displays the means for the March data. The comparison in Table 2 reveals that around 40 percent (e.g., 628/1704 or 682/1704) of contracted agreements were made with municipalities with no public housing vacancies during the years 1986-1991. On average, slightly more than 50 refugees where contracted to municipalities with no housing vacancies while, on average, about 70 refugees where contracted to municipalities with available housing vacancies.¹⁰ Moreover, the corresponding numbers for actual placements is 53 and 76 in municipalities with zero housing vacancies and with positive vacancies, respectively. Thus, the results from Table 2 reveal that the difference in contracted agreements or actual refugee placements is not very different between municipalities with no housing vacancies and those with positive vacancies. Put differently, municipalities with zero housing vacancies received a total of about 35,000 refugees during the period 1986-1991 while those with positive housing vacancies received about 77,000 refugees.

To conduct a more formal test of the correspondence, we run regressions of the following form:

\[ w_{jt} = \alpha_j + \lambda_t + \delta \text{housing\_vacancies}_{jt} + n_{jt} \]

where \( w_{jt} \) is either actual refugee placement or contracted levels through the written agreements, \( \alpha_j \) is a municipality fixed-effect and \( \lambda_t \) is a time fixed-effect. If housing vacancies completely determine the placement policy, we would expect that \( \delta = 1 \).

Results are reported in Table 3. Panel A shows results for the written agreements (contracts) while Panel B shows results for actual refugee settlement. The first thing to note is that all estimates are negative (significant for contracted levels) which is contrary to expectation given the presumption that housing vacancies should essentially determine the

---

⁹ The average number of available rental apartments is close to 40 in municipalities with positive housing vacancies.

¹⁰ The results that on average more than 50 refugees are placed in municipalities with zero vacancies are at odds with claims made by Åslund et al. (2011) that “Assigning a refugee to a municipality was conditional on having found a vacant apartment within that particular municipality.”
refugee placement policy (both contracted levels and actual placements of refugees). A negative correlation suggests that it is the urban areas with more limited public housing availability that continued to attract refugee migration.

In short, our empirical analysis of the Swedish refugee placement policy suggests that (i) there is little or no relationship between contracted agreements and actual refugee settlement and (ii) that there is a negative relationship between the refugee placement policy and housing vacancies, irrespective of whether the policy is defined by actual refugee settlement or contracted agreements. These results are at odds with arguments made in previous studies suggesting that the refugee placement policy can be considered exogenous conditional on observable factors such as housing availability although such claims have not earlier been empirically tested.11

We now turn to a re-analysis of Dahlberg et al. (2012) in order to determine if the results presented in this paper are sensitive to choice of policy instrument. Importantly, we also investigate how their analysis is affected by sample attrition bias (they have an attrition rate of 66%) as well as the theoretically best way to measure immigrants, i.e., according to citizenship, as is their preferred choice, or according to country of birth which is arguably a better measure of ethnicity (e.g. Borjas 1992, 1995).

3. Re-analysis of the Dahlberg et al. study

Dahlberg et al. (2012) study the relationship between preferences for redistribution and ethnic diversity where Swedish local governments are used as a testing ground. They exploit survey data from the Swedish National Election Studies to measure preferences for redistribution.12 The following causal relationship is estimated:

\[
\text{Preferences}_{ijt} = \alpha_j + \lambda_t + \beta\text{Immigrants}_{jt} + u_{ijt}
\]

where the index \(i\) denotes individuals, \(j\) municipalities and \(t\) time period. \text{Immigrants} is measured by the share of immigrants with non-OECD citizenship residing in municipality \(j\) at

---

11 Folke (2011) also investigates the determinants of the refugee placement policy and finds that actual settlement is correlated with the political make-up of the local government. This finding is, in turn, related to the more general question concerning the use of policies as a source of exogenous variation when policies are created by policy makers, who among other factors take into account the preferences of the electorate in their policy making, as discussed by Besley and Case (2000).

12 Note that the data used in Dahlberg et al. is not publicly available at the Journal of Political Economy (JPE) homepage but will eventually be posted at the Swedish National Data Service (SND) (as study SND 0906). In order to access this data, a formal request at SND is required, subject to the approval of the authors. To facilitate replication of our results, we have instead secured permission to post the data from the Swedish National Election Studies directly on the JPE Web site. Formal permission was granted by Henrik Oscarsson Ekegren (Henrik.Oscarsson@pol.gu.se) at the Swedish National Election Studies, Department of Political Science, University of Gothenburg, Box 711, SE 40530, Gothenburg, Sweden
time period $t$. Time periods correspond to election years: 1985, 1988, 1991 and 1994, when the survey data was collected.\textsuperscript{13} The parameter of interest is $\beta$ which measures the causal effect of immigration on preferences for redistribution under the assumption that $E[u|\text{Immigrants}]=0$.

### 3.1 Choice of Instrument

The major contribution of Dahlberg et al. is that they claim to avoid the endogeneity problem that immigrants self-select into municipalities based on, among other attributes, community attitudes towards redistribution, by using an instrument generated by the Swedish refugee placement policy. The instrument used by Dahlberg et al., however, is not the refugee placement policy, that is to say the number of contracted refugees via written agreements between the municipality and SIV, but rather the actual number of refugee settlements in respective municipality. Below, we show that their results of a negative causal impact of ethnic diversity on preferences for redistribution do not hold when the contracted number of refugees is used as the instrumental variable rather than the actual number of refugees.

### 3.2 Sample attrition bias

Another problem with the Dahlberg et al. study is the very large sample attrition due to unnecessary and unmotivated sample restrictions not discussed in the paper. The Swedish National Election Studies is a survey conducted each election year consisting of a representative sample of the Swedish population. The survey has a rotating panel design in which half of the sample has been interviewed in connection with the previous election and the other half in connection with the succeeding election.

Dahlberg et al. base their estimation on data from respondents in the rotating panel only. Specifically, they exploit the rotating panel feature of the survey, i.e., that the same individual is surveyed at two consecutive points in time, e.g., $t$ and $t-3$. The following difference transformation of equation (4) is estimated:

\begin{equation}
(5) \quad (\text{Preferences}_{ijt} - \text{Preferences}_{ijt-3}) = \beta(\text{Immigrants}_{jt} - \text{Immigrants}_{jt-3}) + (u_{ijt} - u_{ijt-3})
\end{equation}

where the difference is taken over a three year-period (which corresponds to the date of election surveys). Dahlberg et al. then instrument the variable, $(\text{Immigrants}_{jt} - \text{Immigrants}_{jt-3})$, by actual refugee inflows between $t-3$ and $t$.\textsuperscript{14}

\textsuperscript{13} There were 284 municipalities in Sweden from 1985 to 1991 and 286 municipalities from 1992 to 1994. Note that Dahlberg et al. (2012) incorrectly state that they have data from 288 municipalities (p. 42).

\textsuperscript{14} Note that they do not transform all of their control variables, such as the political variables and housing vacancies. As a result, any changes in these variables are not controlled for and can result in omitted variable bias.
From an identification point of view, access to panel data, or in this case, a rotating panel is not necessarily useful in a difference-indifference set-up such as that of Equations 4 and 5. As the placement policy generates variation at the municipal level, what matters for identification is that group (municipal) means at the group level are consistently estimated (Angrist and Pischke, 2008 and Blundell and MaCurdy, 1999). Thus, repeated-cross sections or panel data are equally good as long as these data are representative of the underlying population (via random sampling) since otherwise the sample mean will be not be a consistent estimator of the population mean. However, repeated cross-sections typically have fewer problems with attrition than panel or rotating panel data. Thus, if attrition rates are high in panel data, repeated cross-sectional data are preferred since representativeness is then maintained.

To illustrate the sample attrition problem in the Dahlberg et al. study, Table 4 provides information about the number of survey respondents in the 1985, 1988, 1991 and 1994 elections. Panel A shows information for the repeated cross-section while Panel B shows information for the rotating panel. Due to non-response on the question of interest, only 9,620 individuals out of total of 14,297 can potentially be used in the analysis. Thus, the attrition rate is 33 percent in the repeated cross-section. In the rotating panel, the maximum sample size is 5,571 while the available number of respondents is 2,703 leading to an attrition rate of 51 percent. Estimation in Dahlberg et al, however, is based on even fewer observations at 1,917 individuals. They therefore have an attrition rate of 66 percent from the original rotating panel which is likely to exacerbate any sample attrition bias.

The average number of individuals per municipality and year in the Dahlberg et al. sample is less than three. There are also about 70 municipalities out of 284 in each panel that have no observations meaning that, on average, 25% of all municipalities are dropped from their analysis. In contrast, in the repeated-cross section, the average cell size (municipal-year)

---

15 Dahlberg et al. seem to confuse panel data with rotating panel data with regards to random sampling. Panel data without attrition only requires random sampling at a single point in time to maintain its representativeness with the underlying parent population. However, a rotating panel requires random sampling across rotation periods in order to maintain representativeness. For example, the set of individuals in the rotating panel of 1985/1988 are a completely different set of individuals than those in the rotating panel of 1988/1991. To pool these two rotations together therefore requires random sampling at each rotation. In other words, the rotating panel does not solve any identification problems that cannot be solved by a pure repeated cross-section analysis. Indeed, given a large attrition rate across time periods within a rotation, repeated cross-section is preferable to a rotating panel since it maintains representativeness over time thus providing an unbiased estimate of the population treatment effect.

16 Deaton (1997, p.19) writes in his book, *Analysis of Household Surveys*, that “panel data have a number of specific problems. One of the most serious is attrition”. Another problem with rotating or panel data is that survey respondents may change their behavior after being surveyed, as discussed by Peterson et al. (2001) and Bartels (2000). In sharp contrast, independent cross-sections are much less affected by such problems.
is almost nine and, on average, only eight municipalities have no observations. As a result, treating the survey data as a repeated cross-section rather than a rotating panel is superior with regards to sample attrition bias. In addition, since repeated cross-section data provides us with 9,620 useable observations, it is likely to be preferred on efficiency grounds as well.\footnote{The type of data (panel, rotating panel, or independent cross-section) that will be most statistically efficient depends on the degree of temporal autocorrelation in the quantity being estimated. The formulas are given in Hansen et al. (1953, p. 268-72) and are discussed in the context of developing countries by Ashenfelter et al. (1986).}

The reason Dahlberg et al. have only 1,917 observations instead of the available 2,703 obeservations, is partly due to the fact that a number of sample restrictions are made which are not reported in the paper.\footnote{The authors kindly provided us with their STATA do-files documenting the sample restrictions made.} First and foremost, many sample restrictions were made based on responses to survey questions not used in the final analysis of the study. In the 1985 and 1988 election surveys, only individuals that answered all of the following questions and answered according to the first five options on a six point scale, (i) very good proposal, (ii) fairly good proposal, (iii) neither good nor bad proposal, (iv) fairly bad proposal, and (v) very bad proposal, were included in the analysis.\footnote{We have used the original English translation of the survey questions as provided by the Swedish National Election Studies.}

1. Retain nuclear power, also after 2010
2. Reduce third world aid
3. Increase the proportion of health care run by private interest
4. Reduce defense spending
5. Reduce the public sector
6. Increase economic support to immigrant so they can maintain their own culture

In the 1991 and 1994 surveys, two additional restrictions were made based on the following questions:

7. Reduce income differences in society
8. Accept fewer refugees into Sweden

Individuals that responded to any of these survey questions according to the six option (6)“do not know/do not want to answer” are dropped from the analysis as well as all responses coded as (i) a missing answer or (ii) a missing observation.\footnote{When the authors use the 1982 survey for the placebo analysis, they use four survey questions (i.e., questions 2, 4, 5 and 6) to make their sample restrictions, implying that sample restrictions are inconsistent over time.} These unnecessary sample restrictions reduce the number of individuals used in estimation by 581.

Another questionable sample restriction made by Dahlberg et al. is to exclude all individuals that moved to a different municipality between survey periods. In total, there are
205 individuals out of 2,703 that change municipality of residence. Such attrition naturally induces sample selection problems. This type of migration problem can, however, be solved by using information on previous municipality of residence (as discussed in Angrist and Pischke, 2008). In other words, the problem of survey respondents who migrate can be eliminated by defining comparison groups based on prior (to survey response) place of residence as this cannot be affected by the treatment.

In total, these unwarranted sample restrictions reduce the sample to 1,927 observations (excluding one observation dropped due to an extreme value on their instrument). The Dahlberg et al. sample, however, is even smaller at only 1,917 individuals. It is unclear in the paper or the do-files what other sample restrictions are made to reduce the sample size further. One possibility is, of course, that there are missing values on other control variables of interest. The data we compiled for this replication do not reveal any such missing values in other control variables.

3.3 Definition of immigrants
Dahlberg et al. measure the municipal share of immigrants as the proportion with non-OECD citizenship (according to OECD membership prior to 1994). Given the question of interest, the impact of ethnic heterogeneity on preferences for redistribution, it seems more natural to define immigrants according to country of birth. A definition based on citizenship is likely to mask part of the heterogeneity Dahlberg et al. aim to measure as a proportion of individuals with non-OECD origin are likely to have changed citizenship and will not be counted as immigrants. Indeed, in a European perspective, Sweden is characterized by a relatively high naturalization rate, 65 percent of the foreign born are Swedish nationals (OECD, 2006). As citizenship is not an observable characteristics, the correlation between native preferences for redistribution and ethnic diversity is more likely to be based on the number and composition of the foreign born in any given municipality. At the very least, both series should be shown in order to determine to what degree results are sensitive to choice of definition.

4. Results
In this section, we provide results from our re-analysis of Dahlberg et al. (2012) regarding (i) choice of instrument, (ii) sample attrition bias and (iii) definition of immigrants.

We start by analyzing the reduced form relationship between preferences for redistribution and the choice of instrument as we do not want to rely on there being a first stage relationship between the instrument and the endogenous variable, share of immigrants (regardless of definition). Estimation of the reduced form equation is, however, still
informative since in a just-identified instrument variable model, the $P$-value for the reduced form effect of the instrument is approximately the $P$-value from the second-stage (e.g., Chernozhukov and Hansen, 2008). To use the words of Angrist (2012), “if you can’t see the relationship you’re after in the reduced form, it ain’t there!”.

Results are estimated for three samples: (i) the selected sample used by Dahlberg et al, (ii) the full sample available in the rotating panel of the survey (without unnecessary sample restrictions) and (iii) the cross-sectional survey data. In addition, two model specifications are estimated and reported; (i) No controls except for panel (time) effects and municipality-fixed effects, (ii) all the control variables used in Dahlberg et al. We report two sets of standard errors: cluster-robust and homoscedasticity-only standard errors since the robust standard errors may be more biased than the homoscedasticity-only standard errors (Angrist and Pischke, 2008)

The results reported in Column 1 and 2 of Table 5 using the Dahlberg et al. instrument, actual refugee settlement, illustrate a number of important points. Notice, again, that our replication results are based on a slightly larger sample than the Dahlberg et al. sample (1,927 observations rather than 1,917) as discussed earlier. We do, however, come close to replicating their reduced form estimate of -0.172, as can be seen in Panel A (Columns 1 and 2).21

A comparison of the different samples used in estimation indicates that there is a large bias due to attrition. The reduced form estimate is most biased in the Dahlberg et al. sample where the attrition rate is the largest at 65 percent (Panel A). Coefficient estimates are in the range [-0.15, -0.19]. The bias gets smaller when the number of observations is increased in the rotating panel (Panel B). Here the attrition rate is 51 percent and the corresponding coefficient estimates are in the range [-0.10, -0.14]. The bias gets smaller still when estimation is based on the repeated cross-section where the attrition rate is the smallest at 33 percent. Now the estimates are in the range [-0.05, -0.07] and no longer significant. Our results therefore reveal a clear pattern showing that the estimated reduced form effects are reduced by two-thirds when sample attrition bias is corrected for. In the specification without any control variables (Column 1), the point estimate is reduced from -0.15 to -0.05. Note that the bias also seems to depend on the control variables. The (absolute) size of the coefficient estimate tends to be 25 to 40 percent higher when control variables are added to the specifications in Column 2.

21 Results in Table 2 of Dahlberg et al. (2012) indicate that the first stage coefficient is 0.497 and the second stage coefficient -0.347 implying a reduced form estimate equal to -0.172 (0.497*-0.347).
In Columns 3 and 4 of Table 5, we present results from estimation of the reduced form equation when the placement policy, the number of contracted refugees, is used as the instrument instead of actual refugee settlements. The first thing to note is that all the reduced form estimates are much smaller than the corresponding estimates in Columns 1 and 2. For example, the reduced form estimates in the limited rotating sample (Panel A) in Columns 3 and 4 are in the range \([-0.097, -0.106]\) while the instrument used by Dahlberg et al. produces results in the range \([-0.15, -0.19]\). The reduced form effects in both the extended rotating panel and the cross-section sample are smaller still and not significant at conventional levels. Moreover, the control variables no longer seem to introduce any bias in the reduced form estimates since estimates are quite insensitive to inclusion of covariates. This suggests that the placement policy (contracted refugees) is orthogonal to municipality characteristics whereas actual refugee migration is not. A comparison of results in Table 5 indicates that the significant result reported in Dahlberg et al. can only be reproduced by a combination of their (endogenous) instrument and the limited rotating sample. Results also suggest that a causal relationship between preferences for redistribution and the share of immigrants does not exist as we find no reduced-form relationship between the refugee placement policy (contracted refugees) and preferences.

Turning to the corresponding instrumental variable estimates of ethnic diversity on preferences for redistribution, results (displayed in Table 6) show an even stronger case for no causal relationship. This is especially noticeable in a comparison between the preferred specification of Dahlberg et al. using actual refugee inflows as the IV and the limited rotating sample, which produces a coefficient estimate of \(-0.385\) (Column 2, Panel A), and our preferred specification using contracted refugees as the IV and the largest possible sample, which produces an insignificant and small (and positive!) point estimate of 0.065 (Column 4, Panel C).

Next we analyze to what degree the definition of immigrants matters for the results presented in Dahlberg et al. As noted above, they define their independent variable of interest, the (municipal) immigrant share, as the share of people with non-OECD citizenship. Table 7 shows the first-stage estimates for the two definitions of immigrant share: citizenship (Columns 1 and 2) and foreign birth (Columns 3 and 4) and the two definitions of the instrument: actual refugees (Columns 1 and 3) and contracted refugees (Columns 2 and 4). The first thing to note from Table 7 is that there is large variation in the first-stage estimates. For example, comparing the first-stage estimates based on citizenship with those based on foreign birth reveals that that the first-stage estimate increases by almost 40 percent in the
limited rotating sample from 0.49 to 0.68 (in the specification with the full set of control variables).\textsuperscript{22}

This implies that Dahlberg et al. overstate the economic magnitude of their instrumental variable estimate with 40 percent, since the IV estimate is the ratio between the reduced-form result and the first-stage relationship. Moreover, such a large change in their IV estimate due to a re-definition of their key endogenous variable of interest also raises concerns about a causal interpretation of their findings. A perhaps even more disturbing fact is that there is\textit{ no} statistical first-stage relationship between the share of immigrants and the instrument when immigrants are defined according to country of birth and the instrument is defined as contracted refugees (Column 4, Panel C), i.e., the specification we argue to be most correct.

Finally, regarding statistical efficiency, the repeated cross-section analysis is more efficient than the rotating panel data approach as can be seen by comparing the standard errors in Panel A and Panel B of Tables 5 and 6 with those in Panel C of Tables 5 and 6. The standard error of the rotating panel is at least 40 percent larger than the corresponding standard errors of the repeated cross-section. Another noteworthy feature is that the cluster-robust standard errors (within parenthesis) are likely to be biased since they are typically smaller than the homoskedasticity-only standard errors (within brackets). This issue is discussed in Angrist and Pischke (2008) who suggest that in such cases, one should base inference on the largest standard errors. Dahlberg et al. base inference on the (smaller) clustered standard errors.

In summary, our re-analysis of Dahlberg et al. reveals that their results are non-robust as their choice of (i) instrument, (ii) sample, (iii) definition of immigrants and (iv) standard errors are all biased towards finding an economically significant relationship between the share of immigrants and preferences for redistribution when no such relationship exists.

5. Conclusions

In this paper, we have re-analyzed the results presented in Dahlberg et al. (2012) concerning the causal relationship between ethnic diversity and preferences for redistribution. Our results show that the previous characterization of the placement policy is inaccurate. First, there is no relationship between the formal agreements concerning refugee placement in municipalities and actual refugee settlement in municipalities. Second, we find that observable characteristics such as housing vacancies are\textit{ not} the key determinant of refugee placement, as previously argued in the literature. These new results put into question the extent to which the

\textsuperscript{22} The first-stage estimate reported in Dahlberg et al. is 0.497.
Swedish refugee placement policy can be considered as an exogenous source of variation in immigrant placement given key observable characteristics.

A re-analysis of the results reported in Dahlberg et al. (2012) indicates that their results are not robust to a correct characterization of the placement policy. In addition, we find that their study is plagued by severe sample attrition bias. Correcting for either of these two problems reveals that there is no relationship between ethnic diversity and preferences for redistribution.

Our results on the practical implementation of the Swedish refugee placement policy also speaks to the literature using the refugee placement policy as an exogenous source of variation in refugee sorting across municipalities (e.g., Edin et al., 2003; Edin et al., 2004; Åslund, 2005; Åslund and Rooth, 2007; Åslund and Fredriksson, 2009; Åslund et al., 2010; Åslund et al., 2011 and Grönqvist et al., 2012). Specifically, our results question the key identifying assumption in these studies that housing vacancies essentially determine the placement of refugees.

References


Hammar, Thomas (1992), “A crisis in Swedish refugee policy” in A. Daun et al., (eds), To make the world safe for diversity, Ethnology Institute, Stockholm University, Stockholm


Similä, Matti (1992), Det lokala flyktingmottagandet : konflikter och roller, Stockholm, CEIFO, Centre for Research in International Migration and Ethnic Relations, Stockholm University

Statens Invandrarverk (SIV) (1997), Individuell mångfald, Statens Invandrarverk, Norrköping

Soininen, Maritta (1992), Det kommunala flyktingmottagandet. Genomförande och organisation (Local refugee-care – implementation and organisation) Stockholm, CEIFO, Centre for Research in International Migration and Ethnic Relations, Stockholm University

Wickström, Eva (2008), ”Hela världen på vår tröskel:” Lokala reaktioner på en utlokaliserad Flyktingförläggning, Phd-disertation, Umeå University.

Figure 1: Percentage difference between actual and contracted immigration 1986-1991

Table 1. The relationship between actual refugee settlement and contracted refugees for the period 1986-1991

<table>
<thead>
<tr>
<th>Dependent variable = number of actual refugee settlements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of contracted refugees</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
</tr>
</tbody>
</table>

Note: A full set of municipality and time fixed effects are included in the specifications. Standard errors are reported in the parentheses and clustered at the municipality level.
Table 2. Available housing vacancies and the refugee policy: 1986-1991

<table>
<thead>
<tr>
<th></th>
<th>Housing vacancies = 0</th>
<th>Housing vacancies &gt; 0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Housing vacancies September 1&lt;sup&gt;st&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of contracted refugees</td>
<td>51</td>
<td>71</td>
</tr>
<tr>
<td>Number of actual refugee settlements</td>
<td>53</td>
<td>75</td>
</tr>
<tr>
<td>Number of observations</td>
<td>628</td>
<td>1,076</td>
</tr>
<tr>
<td><strong>Panel B: Housing vacancies March 1&lt;sup&gt;st&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of contracted refugees</td>
<td>52</td>
<td>71</td>
</tr>
<tr>
<td>Number of actual refugee settlements</td>
<td>53</td>
<td>76</td>
</tr>
<tr>
<td>Number of observations</td>
<td>682</td>
<td>1,022</td>
</tr>
</tbody>
</table>

Table 3. Test of whether housing vacancies determine refugee placement

<table>
<thead>
<tr>
<th></th>
<th>Housing vacancies: September 1&lt;sup&gt;st&lt;/sup&gt;</th>
<th>Housing vacancies: March 1&lt;sup&gt;st&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Panel A: Dependent variable: actual refugee settlements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacancies</td>
<td>-0.062</td>
<td>-0.037</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,704</td>
<td>1,704</td>
</tr>
<tr>
<td>Panel B: Dependent variable: Number of contracted refugees</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacancies</td>
<td>-0.084</td>
<td>-0.089</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,704</td>
<td>1,704</td>
</tr>
</tbody>
</table>

Note: A full set of municipality and time fixed effects are included in the specifications. Standard errors are reported in the parentheses and clustered at the municipality level.

Table 4. Sample definitions

<table>
<thead>
<tr>
<th></th>
<th>Sample sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Repeated cross-section data 1985-1994</td>
<td></td>
</tr>
<tr>
<td>Total repeated cross-section survey data</td>
<td>14,297</td>
</tr>
<tr>
<td></td>
<td>(85=3704, 88= 3694, 91= 3558, 94=3341)</td>
</tr>
<tr>
<td>Available sample</td>
<td>9,620</td>
</tr>
<tr>
<td></td>
<td>(85=2598, 88=2383 , 91= 2382, 94=2259)</td>
</tr>
<tr>
<td>Panel B: Rotating panel data 1985-1994</td>
<td></td>
</tr>
<tr>
<td>Total rotating panel survey data</td>
<td>5,571</td>
</tr>
<tr>
<td></td>
<td>(85/88=1901, 88/91=1956, 91/94=1714)</td>
</tr>
<tr>
<td>Available sample</td>
<td>2,703</td>
</tr>
</tbody>
</table>
|                                                | (85/88=898, 88/91=989, 91/94=815)
### Table 5. Reduced form relationship between preferences for redistribution and instruments

<table>
<thead>
<tr>
<th></th>
<th>Instrument = actual refugees (Dahlberg et al)</th>
<th>Instrument = contracted refugees</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Panel A: Limited rotating panel sample of 1,927 individuals: Attrition rate = 65%</td>
<td></td>
</tr>
<tr>
<td>Reduced form effect</td>
<td>Panel B: Available rotating panel sample of 2,702 individuals: Attrition rate = 51%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Panel C: Available repeated cross-section sample of 9,620 individuals: Attrition rate = 33%</td>
<td></td>
</tr>
<tr>
<td>Control variables</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|                      | (1)                                      | (2)                                      | (3)                                      | (4)                                      |
| Reduced form effect  | -0.152 [0.071]                           | -0.189 (0.072)                          | -0.106 (0.084)                           | -0.097 (0.087)                           |
|                      | -0.069 [0.071]                           | -0.072 (0.079)                          | -0.079 (0.090)                           | -0.079 (0.096)                           |
|   \textit{P}-value: cluster-robust | 0.028 0.009                              | 0.209 0.257                              | 0.209 0.257                              | 0.209 0.257                              |
|   \textit{P}-value: homoscedasticity | 0.032 0.016                              | 0.239 0.300                              | 0.239 0.300                              | 0.239 0.300                              |

**Note.** The dependent variable is preferences for redistribution (“decrease social welfare spending”) measured on a five point scale. Standard errors clustered at the municipality level are within parentheses and homoskedasticity-only standard errors within brackets. All regressions controls for panel (time) fixed and municipality-fixed effects.
Table 6. The relationship between preferences for redistribution and the share of immigrants: instrumental variable estimates

<table>
<thead>
<tr>
<th>Instrument = actual refugees (Dahlberg et al)</th>
<th>Instrument = contracted refugees</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Panel A: Limited rotating panel sample of 1,927 individuals: Attrition rate = 65%</td>
<td></td>
</tr>
<tr>
<td>Share of immigrants</td>
<td></td>
</tr>
<tr>
<td>-0.294</td>
<td>-0.385</td>
</tr>
<tr>
<td>(0.141)</td>
<td>(0.156)</td>
</tr>
<tr>
<td>[0.138]</td>
<td>[0.161]</td>
</tr>
<tr>
<td>$P$-value: cluster-robust</td>
<td></td>
</tr>
<tr>
<td>0.038</td>
<td>0.014</td>
</tr>
<tr>
<td>$P$-value: homoscedasticity</td>
<td></td>
</tr>
<tr>
<td>0.033</td>
<td>0.017</td>
</tr>
</tbody>
</table>

| Panel B: Available rotating panel sample of 2,702 individuals: Attrition rate = 51% |
| Share of immigrants                         |                               |
| -0.202                                     | -0.291                        |
| (0.118)                                    | (0.140)                       |
| [0.121]                                    | [0.148]                       |
| $P$-value: cluster-robust                   |                               |
| 0.089                                      | 0.038                         |
| $P$-value: homoscedasticity                 |                               |
| 0.094                                      | 0.049                         |

| Panel C: Available repeated cross-section sample of 9,620 individuals: Attrition rate = 33% |
| Share of immigrants                         |                               |
| -0.067                                     | -0.099                        |
| (0.058)                                    | (0.069)                       |
| [0.060]                                    | [0.075]                       |
| $P$-value: cluster-robust                   |                               |
| 0.247                                      | 0.150                         |
| $P$-value: homoscedasticity                 |                               |
| 0.263                                      | 0.186                         |

Control variables | No | Yes | No | Yes

Note: The dependent variable is preferences for redistribution (“decrease social welfare spending”) measured on a five point scale. Standard errors clustered at the municipality level are within parentheses and homoskedasticity-only standard errors within brackets. All regressions controls for panel (time) fixed and municipality-fixed effects.
<table>
<thead>
<tr>
<th>Panel</th>
<th>Sample Size</th>
<th>Attrition Rate</th>
<th>First-stage effect Instrument=actual refugees</th>
<th>First-stage effect Instrument=contracted</th>
<th>First-stage effect Instrument=actual refugees</th>
<th>First-stage effect Instrument=contracted</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Limited rotating panel sample of 1,927 individuals</td>
<td>65%</td>
<td>0.492 (0.059)</td>
<td>0.424 (0.078)</td>
<td>0.677 (0.058)</td>
<td>0.396 (0.083)</td>
</tr>
<tr>
<td>B</td>
<td>Available rotating panel sample of 2,702 individuals</td>
<td>51%</td>
<td>0.472 (0.055)</td>
<td>0.406 (0.067)</td>
<td>0.652 (0.058)</td>
<td>0.371 (0.080)</td>
</tr>
<tr>
<td>C</td>
<td>Available repeated cross-section sample of 9,620 individuals</td>
<td>33%</td>
<td>0.724 (0.011)</td>
<td>0.450 (0.052)</td>
<td>0.668 (0.117)</td>
<td>0.162 (0.143)</td>
</tr>
</tbody>
</table>

Control variables: Yes, Yes, Yes, Yes

Note. Standard errors clustered at the municipality level are within parentheses. All regressions controls for panel (time) fixed and municipality-fixed effects.
Web appendix (Not to be published)

Replication issues

For our re-analysis, we collected data on the instrument used in Dahlberg et al.—actual refugee settlements in municipalities—directly from the published material kept in the Archives of the Swedish Immigration Board. When comparing our data with those provided by the authors, a number of inconsistencies were discovered. For example, six observations are coded as missing when the correct amount is zero and four observations are based on the contracted number of refugees rather than actual refugee settlement (as noted earlier).

The do-files provided by the authors also reveal that the endogenous variable of interest—the share of immigrants in the municipality (IM)—was missing for the year 1985 and replaced by a linear extrapolation between 1984 and 1986. This issue is not discussed in the paper.

Other measurement issues have to do with the control variables used in the Dahlberg et al. analysis. There are two available annual series on housing vacancies in March and September and it is unclear which series is used in Dahlberg et al. (2012). Here it is important to stress that the data on municipal housing vacancies also used by Dahlberg et al. covers public entities only, which constitute about 40% of the total aggregate housing supply while private housing and housing co-operatives (tenant ownership) are not covered.

Local unemployment rates can also be measured in two different ways, one which includes the change in the definition of unemployment that occurred in 1990 and another that uses the same (old definition) of unemployment for the entire time period, 1985-1994. Again, it is not clear in the paper which series is used in the analysis.

Likewise, there are two potential sources of information on social welfare spending; one published by the National Board of Health and Welfare and the other from local government budget records. The latter data contain seven missing values for the period 1985-1994 while the former have no missing values. Again, there is no discussion in the paper concerning which source is used in the analysis.

---

23 Dahlberg provided us with data on the instrument used in their study. There are 24 differences between the authors’ measures of actual refugee immigration and our measure based on information gathered directly from the SIV archives for the period 1986 to 1990.

24 Note that we do not have access to the full data set compiled by the authors but have rather put together the necessary information ourselves from publicly available sources.

25 It is also wrongly stated in the paper that information on housing vacancies is not available before 1985 (see footnote 25 in Dahlberg et al. (2012). Data on housing vacancies for the years 1983-1988 can, for example, be collected from the readily available Statistics Sweden’s publications “Statistiska meddelanden”:Bo 34 SM8801 or Bo 35 SM 8801.
With these issues in mind, the following choices are made in our replication of the Dahlberg et al. results: (i) we use corrected data (24 observations) in estimation based on their instrument (actual refugee settlement) from information gathered directly at the SIV Archives (ii) we use the correct, and not extrapolated, values for share of immigrants per municipality in 1985, (iii) we use the measures of social welfare spending with no missing values and housing vacancies from the September surveys, and (iv) we use the definition of unemployment which is consistent over time.

**Data on the outcome(s)**

The survey data from Swedish election studies are taken from:

1. **SND 0217 - Swedish election study 1985**
2. **SND 0227 - Swedish election study 1988**
3. **SND 0391 - Swedish election study 1991**
4. **SND 0570 - Swedish election study 1994**

This data were provided by SND (Swedish National Data Service)
http://snd.gu.se/en

**Data on the refugee placement program**

1. Yearly data on the number of flat-rate payments (“utbetalda schablonbidrag”) for the period 1985-1994
2. Yearly data on the number of contracted slots for refugees (“kommunplatser för flyktingar enligt avtal”) for the period 1985-1994

The data was collected from published material in the Archives of the Immigration Board

Contact:
Thina Carlsson
Migrationsverket
Förvaltningsarkivet
010-485 67 41
e-mail contact: Thina.carlsson@migrationsverket.se

**Data on immigration**

Data on foreign citizenship and country of birth for the years 1985, 1988, 1991 and 1994
The data was provided by Statistics Sweden (e-mail contact: befolkning@scb.se).

**Data on housing vacancies**

There are two available data series on housing vacancies in semi-public bodies
1. Yearly data on unlet dwellings in multi-dwelling buildings dwellings per March 1st
2. Yearly data on unlet dwellings in multi-dwelling buildings dwellings per September 1st

These data were collected from the publications “Statistiska meddelanden” : (i) Bo 34 SM8801, (ii) Bo 35 SM 8801, (iii) Bo 35 SM 9401, and (iv) Bo 34 SM 9401.
Population
Local government’s population per January 1st. The data was provided by Statistics Sweden (e-mail contact: befolkning@scb.se).

Welfare spending
There are two available data series

1. Social welfare spending 1 was downloaded from the following link:

The data is collected by The National Board of Health and Welfare (“Socialstyrelsen”) and is provided by Statistics Sweden (e-mail contact: bo.thyden@scb.se). There are 7 missing values for the period 1985-1994.

2. Social welfare spending 2 was taken from the local governments budget records (“Räkenskapssammandrag för kommuner”) (see http://www.scb.se/RSkommuner/). These data was provided by Statistics Sweden (e-mail contact: offentlig.ukonomi@scb.se).

Unemployment
There are two available data series

1. Unemployment series 1. These data was provided by the Swedish Public Employment Service (“arbetsförmedlingen”). e-mail contact: birgitta.i.andersson@arbetsformedlingen.se

2. Unemployment series 2. This data was provided by Anders Forslund at IFAU (Institute for Evaluation of Labor Market and Education Policy). e-mail contact: anders.forslund@ifau.uu.se

The difference between the two unemployment series for the period 1985-1994 is that there is a new definition of the unemployment after 1989. The second series has the same (old) definition of unemployment during the whole period 1985-1994.

Tax base
The local government income tax base in period t (i.e., taxable personal income as reported in t-2). The data was provided by Statistics Sweden.

Political characteristics
Data on Socialist majority status (the Social Democrats and the Left Party), the Green Party and the New Democrats was taken from the municipal elections1985, 1988, 1991 and 1994. The data was provided by Statistics Sweden. e-mail contact: valstatistik@scb.se