

ETHNIC ENCLAVES AND THE ECONOMIC SUCCESS OF IMMIGRANTS—EVIDENCE FROM A NATURAL EXPERIMENT*

PER-ANDERS EDIN
PETER FREDRIKSSON
OLOF ÅSLUND

Recent immigrants tend to locate in ethnic “enclaves” within metropolitan areas. The economic consequence of living in such enclaves is still an unresolved issue. We use data from an immigrant policy initiative in Sweden, when government authorities distributed refugee immigrants across locales in a way that we argue is exogenous. This policy initiative provides a unique natural experiment, which allows us to estimate the causal effect on labor market outcomes of living in enclaves. We find substantive evidence of sorting across locations. When sorting is taken into account, living in enclaves improves labor market outcomes for less skilled immigrants: the earnings gain associated with a standard deviation increase in ethnic concentration is 13 percent. Furthermore, the quality of the enclave seems to matter. Members of high-income ethnic groups gain more from living in an enclave than members of low-income ethnic groups.

I. INTRODUCTION

In most countries, immigrants tend to be spatially concentrated [Stark 1991]. This is particularly true for the United States, where almost three-quarters of immigrants resided in only six states in 1998; see Borjas [1999]. It is also true for Sweden, where the share of the foreign-born population living in the three largest metropolitan areas outstripped the share of the native population by 18 percentage points in 1997. In this paper we are concerned with the labor market consequences of this

* We have benefited from the useful comments of two anonymous referees, Joshua Angrist, George Borjas, Marianne Bertrand, Kenneth Carling, Jan Ekberg, Edward Glaeser, Claudia Goldin, Peter Hedström, Lawrence Katz, Magnus Lofström, Regine Riphahn, Dan-Olof Rooth, seminar participants at European University Institute, Göteborg University, IIES, SOFI, Stockholm School of Economics, Tilburg University, Tinbergen Institute, University College London, Uppsala University, the CEPR conference on “Marginal Labour Markets in Metropolitan Areas,” the 2000 meeting of EALE/SOLE, the NBER Summer Institute 2000, and the CEPR European Summer Symposium in Labour Economics 2000. We thank Lisa Fredriksson for expert data assistance. We also thank Sven Hjelmsskog, Stig Kattilakoski, Christina Lindblom, Anders Nilsson, Kristina Sterne, and Lena Axelsson of the Immigration Board and Anna Gralberg of the Ministry of Culture, who generously found time to answer our questions. This research has been partly financed through a grant from the Swedish Council for Work Life Research (RALF).

© 2003 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology.

The Quarterly Journal of Economics, February 2003

location pattern. In particular, we ask the question: how does living in ethnic enclaves affect labor market outcomes?

There are several competing hypotheses about the effects of living in enclaves. One hypothesis is that residing in an enclave is associated with a lower rate of acquisition of host country skills, e.g., language. Another hypothesis is that the enclave represents a network that increases the opportunities for gainful trade in the labor market, e.g., by disseminating information on job opportunities. So, the causal effect of living in an enclave is *a priori* ambiguous in sign.

Estimating the causal effect of enclaves is no easy task. Because of sorting, treating the ethnic composition of an area within a city as exogenous is problematic. Residential location is a household choice variable, and this choice is very likely influenced by unobserved attributes also affecting labor market outcomes. Recent research has made progress in addressing this issue. One approach is to use the variation across metropolitan areas arguing that sorting is less problematic in this case; see Bertrand, Luttmer, and Mullainathan [2000], Cutler and Glaeser [1997], Dustmann and Preston [1998], and Gabriel and Rosenthal [1999]. Another approach is to use parental choices of neighborhoods, where the assumption is that this choice is exogenous with respect to the outcome of the offspring; Borjas [1995] is an example. In general, these studies suggest that disadvantaged groups such as blacks and immigrants are hurt by segregation.

Although the identification strategies briefly reviewed above are less susceptible to sorting problems, we think that the validity of the implicit exogeneity assumptions is an open question. For instance, differences in unobserved individual attributes across metropolitan areas still remain a concern.

In this paper we take a different approach. Our exogenous source of variation comes from a Swedish government policy concerning the initial location of refugee immigrants.¹ This policy was viable between 1985 and 1991. 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. Our maintained hypothesis is that, because of the policy, the initial location of immigrants is independent of

1. The Katz, Kling, and Liebman [2001] analysis of Moving to Opportunity in Boston is another example where a policy experiment is used to identify the causal effect of interest.

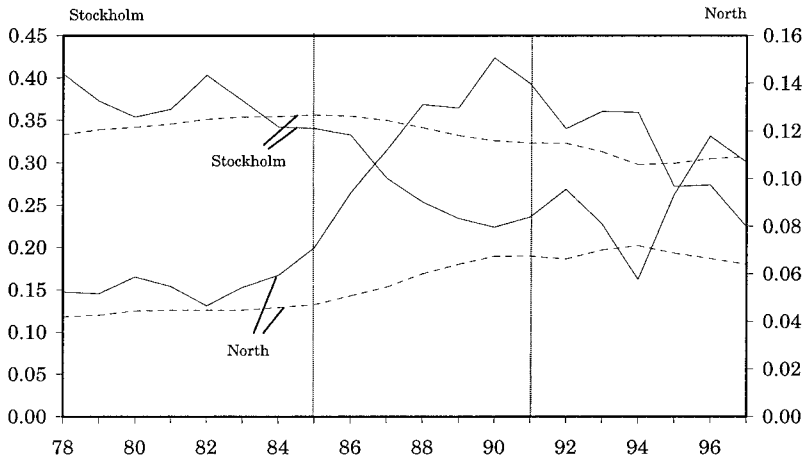


FIGURE I

Share of Non-OECD Immigrant Inflow (solid) and Stock (dashed) Located in Stockholm and in the North of Sweden, Respectively, 1978–1997

“Stockholm” refers to the county of Stockholm, and “North” to the six northernmost counties of Sweden. Authors’ calculations using the LINDA immigrant sample.

unobservable individual characteristics. Hence, this “natural experiment” enables us to reexamine the question of the economic consequences of living in enclaves.

The government settlement policy had real consequences for immigrant location. This is illustrated in Figure I, where we plot the share of the immigrant inflow and the immigrant stock that resides in Stockholm and the north of Sweden, respectively. Prior to 1985, refugees were allowed to settle in a neighborhood of their own liking. In 1985 the immigrant shares in Stockholm and the north of Sweden stood at 36 and 5 percent, respectively. By 1991 the share living in Stockholm had been reduced by more than 3 percentage points, while the share residing in the north increased by 2 percentages. Thus, the policy initiative clearly increased the dispersion of immigrants across Sweden.

Our results can be briefly summarized as follows. We find pervasive evidence of sorting across local labor markets. In cases where the size of the enclave matters for outcomes, the estimates that account for sorting differ drastically from those that do not. For low skilled immigrants, estimates that suffer from sorting bias associate an earnings *loss* of 4 percent with a standard deviation increase in ethnic concentration. Estimates that do not

suffer from this problem suggest an earnings *increase* of almost 13 percent. Furthermore, the earnings gains are increasing in the quality of the enclave. Members of high-income ethnic groups gain more from living in an enclave than members of low-income ethnic groups.

The remainder of the paper is outlined as follows. By way of background, Section II gives a description of the institutional setting and discusses whether the policy introduced in 1985 provides as a natural experiment. In Section III we outline a simple framework that we use as a guide to specification and interpretation. We then turn to the empirical analyses. Section IV describes the data. In Section V we report estimates of the effects of ethnic concentration on earnings. Section VI concludes.

II. THE INSTITUTIONAL BACKGROUND

Sweden and the United States are sometimes portrayed as being opposite extremes in terms of the scope of the welfare state and the functioning of the labor market. Some aspects of the immigration experience are, however, remarkably similar. The share of the foreign born in the total population is about the same. In 1997, 11 percent of the Swedish population was foreign born. By comparison, the foreign born population in the United States amounted to 10 percent in 1998. Moreover, the probability of residing in an "ethnic neighborhood" is of a similar magnitude in the two countries.² Further, in both countries, immigrants from developing countries are more likely to live in an ethnic neighborhood. Of course, there are also differences between the two countries. For instance, immigrant source countries are obviously different. Also, the share of refugees in the total inflow of immigrants over the past 30 years has been substantially larger in Sweden.

Our analysis concerns refugee immigrants, and we now turn to presenting the Swedish refugee policy in more detail. Unfortunately, there is very little documentation about the practical implementation of the placement policy. Therefore, part of the

2. An ethnic neighborhood is defined as a neighborhood where the share of the ethnic group residing in the neighborhood is at least twice as large as the share of the ethnic group in the population. According to Borjas [1999], 48 percent of individuals with non-U. S. ancestry resided in an enclave in 1979. Doing an analogous calculation for Sweden, we find that 42 percent of the average first-generation immigrant lives in an ethnic neighborhood in 1997; see Edin, Fredriksson, and Åslund [2001] for further details about this calculation.

information is based on interviews with placement officers and other officials of the Immigration Board.

*II.A. The Swedish Refugee Placement Policy*³

The placement policy was introduced in 1985. The Immigration Board was then given the responsibility of assigning refugee immigrants to an initial municipality of residence. The Board was to place all political immigrants, except those who arrived for family reunification reasons.

The introduction of the placement policy was a reaction to immigrant concentration in large cities. The idea was to distribute asylum seekers over a larger number of municipalities that had suitable characteristics for reception, such as educational and labor market opportunities. At first, the intention was to sign contracts with about 60 municipalities, but due to the increasing number of asylum seekers in the late 1980s, a larger number became involved; in 1989, 277 out of Sweden's 284 municipalities participated. The factors that initially were supposed to govern placement were more or less abandoned. Instead, the availability of housing became the deciding factor.

Formally, the policy of assigning refugees to municipalities was in place from 1985 to 1994. The strictest application of the assignment policy was between 1987 and 1991. For our purposes, this is the most attractive time period, since it contained very small degrees of freedom for the individual immigrant to choose the initial place of residence. During 1987–1991 the placement rate, i.e., the fraction of refugee immigrants assigned an initial municipality of residence by the Immigration Board, was close to 90 percent.

II.B. From Port of Entry to Placement

As a way of describing the placement policy during 1987–1991, we describe the handling of a typical asylum seeker from the border to the final placement. An asylum seeker was placed in a refugee center while waiting for a decision from the immigration authorities. Refugee centers were distributed all over Sweden, and there was no correlation between the port of entry and the location of the center. However, immigrants were sorted by native language when placed in centers.

3. This section draws on The Committee on Immigration Policy [1996] and The Immigration Board [1997].

There was a long wait for a residence permit. The mean duration between entry into Sweden and the receipt of a permit (conditional on receipt) varied between three and twelve months during 1987–1991; see Rooth [1999].⁴ There was a much shorter wait for a municipal placement after receiving the permit, partly because there were explicit goals in terms of the duration of this spell.

When it came to the municipal placement, weight was given to immigrant preferences. Most immigrants, of course, applied for residence in the traditional immigrant cities of Stockholm, Gothenburg, and Malmö. However, there were very few housing vacancies in these locations, in particular during the second half of the 1980s when the housing market was booming. When the number of applicants exceeded the number of available slots, municipal officers may have selected the “best” immigrants. There was no interaction between municipal officers and refugees, so the selection was purely in terms of observed characteristics; language, formal qualifications, and family size seem to have been the governing criteria. Preferences were given to highly educated individuals and individuals who spoke the same language as some members of the resident immigrant stock. Single individuals were particularly difficult to place, since small apartments were extremely scarce.

After having been assigned to an apartment, immigrants’ main source of income was welfare (i.e., social assistance). They could live on welfare while participating in introductory Swedish courses. Receipt of welfare was not conditional on residing in the assigned municipality, and the central government reimbursed the local governments for their welfare expenditures. So there was little incentive for the immigrants to stay on in an assigned municipality, if they could realize their preferred choice. The main cost, apart from moving costs, consisted of delayed enrollment in language courses.

On the basis of the above description, we think that it is realistic to treat the municipal assignment as exogenous with respect to the random components of the outcomes of interest, *conditional on observed characteristics*.⁵ The strictest application

4. Whether individuals were subjected to the placement policy or not depended solely on when they received their residence permits.

5. Potentially, selection by municipal officers based on observed skills suggests that the high skilled may have been more likely to influence their preferred option. On the whole, rates of postplacement mobility do not suggest that the

of the assignment policy was between 1987 and 1991. We have chosen to base our empirical work on placements during 1987–1989. The last year is an obvious choice, since the probability of being “exogenously” placed is increasing in the total inflow of residence permits and the tightness of the housing market. There was a hike in the number of new residence permits in 1989, and the housing market peaked in that year. Given the choice of 1989, we can follow individuals for a maximum of eight years. To increase the size of the sample we added two additional years. We chose 1987 and 1988 since we wanted to follow individuals over time for as long as possible.⁶

III. THEORETICAL BACKGROUND AND EMPIRICAL FRAMEWORK

This section begins by offering a brief account of the literature on why enclaves could affect outcomes. Then we adapt the schooling model of Card [1999] to our setting in order to discuss the bias of the OLS estimator and outline our empirical approach.

III.A. *Why Does Living in Enclaves Affect Outcomes?*

There are at least four types of explanations for why enclaves influence the outcomes of individuals living there: (i) slower rate of acquisition of host country skills; (ii) “network effects”; (iii) “spatial mismatch”; and (iv) human capital externalities.⁷ Although we present each explanation separately, they are not mutually exclusive.

The hypothesis that the enclave *decreases the rate of host country skill acquisition* seems to have been among the prime motives for the reform that we are utilizing. According to this view, the ethnic enclave provides less interaction with natives and reduces the incentives for acquiring, e.g., language skills. The

highly qualified were more likely to exercise their preferred option when being assigned to a municipality.

6. As a guide to the selection of years, we calculated the ratio of the inflow of residence permits and the stock of vacant public rental apartments. This ratio stood at 10 in 1989, at 4 in 1988 and 1990, and at 2 in 1987 and 1991.

7. To this list one could potentially add relative factor supplies and compensating differentials. This story would go as follows. If the typical immigrant has preferences for living with members of his own ethnic group, then he is willing to pay a price for living in that area. The price corresponds to the movement along the labor demand curve as labor supply increases. The equilibrium sorting of individuals will feature a negative correlation between wages and ethnic concentration. The correlation is simply due to preferences and does not have a causal interpretation. For this reason, this story is not included above.

enclave, hence, hinders the move to better jobs and reduces earnings in the longer run.

More of a positive view is contained in stories that emphasize *network effects*. The enclave represents a network that increases the opportunities for gainful trade in the labor market, e.g., Portes [1987] and Lazear [1999]. Further, the network disseminates valuable information on, e.g., job opportunities, and constitutes an environment where the immigrant is less exposed to the discrimination encountered elsewhere in the labor market. The enclave would thus improve labor market outcomes, in particular for recent immigrants and for individuals who have difficulty integrating into the labor market. Of course, the enclave may also provide information on matters that are not conducive to success in the labor market, such as welfare eligibility; e.g., Bertrand, Luttmer, and Mullainathan [2000].

The *spatial mismatch* hypothesis emphasizes discrimination in the housing market; see Ihlanfeldt and Sjoquist [1998]. Since immigrants face restrictions in the housing market, they are *forced* to segregate in an enclave. The enclave, in turn, may be distant from employment opportunities. Therefore, individuals living in the enclave will fare worse than otherwise similar immigrants who have escaped housing market discrimination. In this view, it is not the enclave as such that hampers success in the labor market, but rather that the enclave is distant from employment opportunities.

Models featuring *human capital externalities* are also based on residential segregation. In this instance, however, segregation is not necessarily bad—it all depends on the quality of the enclave, e.g., the stock of human capital; see Cutler and Glaeser [1997] and Borjas [1998]. If residential segregation implies that skilled members of an ethnic group live in the enclave, and individuals primarily interact with members of their own ethnic group, then disadvantaged members such as recent immigrants gain from living in the enclave.

The conclusion from this brief review is that the causal effect of living in an enclave is ambiguous in sign. The net effect on outcomes is thus an empirical question. To determine the net effect, we estimate what must be interpreted as reduced-form relationships between measures of labor market outcomes and, among other things, the size of the enclave. We will thus not be able to test any of the above hypotheses. However, we argue that our estimates have a causal interpretation.

III.B. On the Estimation of the Outcome Equations

Here we present a simple model to discuss the sources of bias of OLS. To illustrate our main points, we adapt the schooling model of Card [1999] to our setting. Consider an immigrant who derives utility from living with members of his own ethnic group and the consumption of goods. The immigrant maximizes utility by making a location choice, where each location is characterized by some measure of ethnic concentration (e). For simplicity, we assume that there is a continuum of location choice such that $e \in [0,1]$. Utility is given by

$$(1) \quad U_i = u_i(e_i) + \ln y_i.$$

The objective of the individual is to maximize (1) subject to a market opportunity locus:

$$(2) \quad \ln y_i = \alpha_i + \beta_i e_i,$$

where α_i reflects general aptitudes in the labor market and β_i is the marginal return to living in an enclave. The first-order condition is $u'_i + \beta_i = 0$. To proceed, assume that $u'_i = \mu_i - ke$, where k is a positive constant. Then the optimal e (if it is interior) satisfies

$$(3) \quad e_i^* = (\mu_i + \beta_i)/k.$$

Let us revert to the earnings equation. Rewrite this relationship as

$$(4) \quad \ln y_i = \alpha + \beta e_i + (\alpha_i - \alpha) + (\beta_i - \beta)e_i = \alpha + \beta e_i + \eta_i.$$

Our interest concerns the parameter β —the average return to living in an enclave. Consider the OLS estimate of β . The probability limit of the OLS estimate, b_{OLS} , is

$$(5) \quad \text{plim} b_{OLS} = \beta + \lambda_0 + \lambda_1 e.$$

The parameters λ_j are theoretical regression coefficients: $\lambda_0 = \text{cov}(\alpha, e)/\text{var}(e)$; and $\lambda_1 = \text{cov}(\beta, e)/\text{var}(e)$, where e is given by (3). The coefficient λ_0 measures ability bias, while λ_1 is related to bias because of self-selection.

In general, we cannot say much about the sign of the bias, but it is easy to think of examples producing either positive or negative bias. If the preferences of high ability immigrants are such that they would not like to live in the enclave, i.e., $\text{cov}(\alpha, \mu) < 0$,

and the return to living in the enclave is homogeneous, i.e., $\beta_i = \beta$, then OLS is biased downwards. On the other hand, if immigrants respond to β_i and $\text{cov}(\alpha, \beta) > 0$, OLS produces upward biased estimates.

What kinds of assumptions are necessary to estimate the average return to living in an enclave consistently? In practice, we have two estimation alternatives—an IV or a control function approach. For our instrument to be of any use, it must be independent of the random coefficients in (4). In the major part of the analysis, we also impose an exclusion restriction in the sense that only variables associated with the current location have an effect on earnings.

Given the above assumptions, two questions arise. What do (linear) instrumental variables estimate? Under what conditions does IV estimate the average return to living in an enclave consistently? In our application, IV estimates the average treatment effect (ATE) if the return to enclaving is homogeneous in the population (i.e., $\beta_i = \beta$), or when there is *ex ante* homogeneity but *ex post* heterogeneity in β (provided that the individual return is uncorrelated with the individual marginal utility of living in an enclave; i.e., $\text{cov}(\beta, \mu) = 0$). However, as soon as there is self-selection by comparative advantage, IV estimation requires special assumptions in order to be consistent for ATE.

In this paper we rely on linear IV in the empirical analysis. So, in practice, we assume either that the return is homogeneous or that individuals do not act on their individual specific return.⁸ We do, however, let β vary by observed characteristics. We have also experimented with an alternative estimation strategy. If we impose some additional structure, we can apply the selection correction method proposed by Garen [1984] to estimate ATE. The Garen approach yields estimates that are very similar to those presented below.

IV. AN INITIAL LOOK AT THE DATA

This section provides an initial look at the data. We begin by describing the data and sample selections (subsection IV.A). Then

8. In the heterogeneous returns case we would like to think of the IV estimate as giving a weighted average of individual returns among those who were induced to change "treatment dose" because of the value of the instrument. Notice, though, that this is a heuristic interpretation since our particular estimation problem has not been considered in the literature.

we focus on two issues that may compromise our identification strategy (subsection IV.B). The first issue is whether some individuals were able to influence their initial placement. In particular, does the initial size of the enclave vary with individual characteristics in a way that may cause concern? The second issue concerns the predictive power of the instruments and for whom we identify the effects. The main source of identification comes from individuals who stayed on in the assigned residences. So we are interested in the extent of subsequent mobility and the characteristics of those who stayed relative to those who moved.

IV.A. Data and Sample Selection

The empirical analysis is based on data from the LINDA database. Among other things, LINDA contains a panel of around 20 percent of the foreign-born population. Moreover, the data are cross-sectionally representative. Data for immigrants are available from 1970 and onwards. The data set is based on a combination of income tax registers, population censuses and other sources; for more details, see Edin and Fredriksson [2000].

All individual variables in our analysis are based on register information. Earnings and some of the demographic characteristics (gender, age, education, family composition) are obtained from the income tax registers, which also contain information on country of birth and year of immigration from the population registers. Throughout we use ethnicity as shorthand for country of origin, although it is the latter information that the data contain.

Local characteristics are measured at the municipality level.⁹ At the time there were 284 municipalities, and the median municipality had close to 16,000 inhabitants.¹⁰ We use public population counted data to represent the characteristics of the municipalities. The sizes of various population groups are obtained from Statistics Sweden (SCB). The unemployment rates are based on unemployment registers at the Labor Market Board (AMS). Importantly, we use population counts of the number of

9. Refugee immigrants were assigned to an apartment so, in principle, there is scope for defining smaller "neighborhoods." However, calculating the number of immigrants for smaller geographical units requires micro data covering the entire population, something that is unavailable to us at present.

10. The municipality is a smaller entity than the Swedish definition of a local labor market. For instance, the Stockholm metropolitan area includes 26 municipalities. In the aggregate, there are roughly 2.5 times as many municipalities as there are local labor markets.

foreign citizens from a particular source country to represent the size of the ethnic group in a municipality. The number of foreign citizens is used rather than the number of foreign born, since all measures available for the latter will be plagued by sampling error. However, the correlation between the number of foreign citizens and the number of foreign born across municipalities appears to be high.¹¹ Descriptive statistics, along with the definition of some of the key variables, are reported in the Appendix.

We cannot identify refugee immigrants directly from our data. Instead, we identify them by country of origin. As a general rule, we include immigrants from countries outside Western Europe that were not members of the OECD as of 1985. The only exception to this rule is Turkey, which is included since it was the origin of a substantial inflow of refugee immigrants during the period.

We exclude persons belonging to a household with an adult already residing in Sweden, since they were likely to have immigrated as family members and, consequently, were not placed. We base our analysis on individuals aged 18–55 at the time of entry into Sweden. Last, we focus on the immigration waves during 1987–1989 for reasons outlined above.

After imposition of all these restrictions, we are left with a sample of 9883 individuals. If the sample is restricted to those with positive annual earnings, 6418 individuals remain. There are 270 municipalities of assignment and 98 nationalities represented in the sample with positive earnings. The biggest source countries (ordered by size) are Iran, Chile, Ethiopia, Romania, and Poland.

Another feature of the data that is relevant for our analysis is that we observe the region of residence at the end of the year. Thus, the observed initial location may differ from the actual initial placement if individuals move during their first year. This introduces a measurement error in initial placement, an issue we will address in the empirical analysis.

IV.B. Initial Placement and Subsequent Mobility

To shed some light on the question of whether some individuals were more likely to influence their placement than others, we divide individuals into two groups: those who were placed in

11. For the 42 largest source countries, the correlation between these two measures is 0.96.

TABLE I
INDIVIDUAL CHARACTERISTICS BY INITIAL PLACEMENT

	Initial placement	
	Enclave	No enclave
Female	.44	.45
Age	37.3 (7.7)	37.6 (7.4)
Years of schooling	11.3 (3.0)	11.7 (2.9)
Married	.63	.62
Kid ≤ 15 years of age	.55	.57
No. of individuals	3094	3324

Standard deviations are in parentheses. An enclave is defined as described in the main text. Years of schooling are imputed from highest degree attained. Individuals with missing information on education were given the same number of years of schooling as those with less than nine years of schooling. All characteristics are measured eight years after immigration. The sample is restricted to those with positive earnings at that point in time.

an enclave and those who were not. An “enclave” is defined for each ethnic group as a municipality where ethnic concentration (the size of the ethnic group relative to the population in each municipality) was at least twice as large as the share of the ethnic group in the entire population.

Table I compares the characteristics of refugee immigrants who were placed in an enclave with those who were not. Around 48 percent were placed in a municipality that we define as an enclave. There are practically no differences between individuals who were placed in an enclave and those who were not. The only significant difference is in (imputed) years of schooling: those who were placed outside an enclave have 0.4 years more schooling. Although small, the difference in education may be worrying. Therefore, we probed more deeply into this issue by running a regression where we related the size of the ethnic group in the assigned municipality to a wider set of individual characteristics including a set of source country dummies and year of entry dummies. The source country dummies are bound to be significant because some countries make up a larger share of the immigrant population than others. Apart from these nationality dummies, however, age was the only characteristic that entered significantly in the regression. In particular, there were no differences across educational groups. Therefore, we conclude that

TABLE II
INDIVIDUAL AND LOCAL CHARACTERISTICS BY MOBILITY STATUS

	Mobility status	
	Stayer	Mover
Individual characteristics		
Female	.48	.41
Age	38.1 (7.9)	36.6 (7.0)
Years of schooling	11.2 (2.9)	11.8 (3.0)
Married	.66	.59
Kid ≤ 15 years of age	.58	.53
Eastern Europe	.24	.15
Africa	.10	.13
Middle East	.33	.49
Asia	.10	.08
South America	.23	.14
Local characteristics		
Ethnic concentration (percent)	.32	.34
Immigrant density (percent)	7.85	8.38
Population size (1000)	217.0	234.8
Unemployment rate (percent)	5.87	5.76
No. of individuals	3492	2926

Standard deviations are in parentheses. Years of schooling are imputed from highest degree attained. Individuals with missing information on education were given the same number of years of schooling as those with less than nine years of schooling. All characteristics are measured eight years after immigration. The sample is restricted to those with positive earnings at that point in time.

sorting on observed characteristics does not seem to be a problem.

What is the extent of subsequent mobility, and who moves on to another municipality? Table II addresses these questions by reporting individual and local characteristics by mobility status.

We first note that 46 percent of the sample has moved to another municipality within eight years after entering Sweden. Thus, since the majority stayed on in the assigned municipality, our instruments—i.e., the characteristics of the assigned municipality—will have predictive power in the first-stage regressions. The rate of mobility may seem high; however, it is a generic feature of the Swedish immigration experience that there is substantial mobility out of the initial location. In other work we have compared the rate of mobility among those who were placed with the rate of mobility for refugee immigrants who entered prior to

the placement policy; see Åslund [2000]. Much of the raw difference in mobility rates between these two cohorts (around 5 percentage points) disappears when mobility rates are standardized with respect to individual characteristics.

The differences in individual characteristics are broadly in line with what one would expect in any analysis of mobility. Movers are more likely to be young and single, and have more education than stayers. There are some differences across source regions. Refugee immigrants from the Middle East and Africa tend to move to a greater extent than immigrants from other regions.

Our account of the workings of the assignment policy suggests that education and family size may have influenced the probability of realizing the preferred option. To examine whether this was the case, we look at migration propensities by education and family status. As Table II has little to offer in these respects—the highly educated and singles are more likely to move in any situation—we calculate before and after (and difference-in-differences) estimates for immigrants arriving before and after the reform. With respect to education we find no evidence suggesting that the high skilled were able to realize their preferred choice to a greater extent than the less skilled. Relative relocation rates of individuals with a university degree appear to be higher after the introduction of the placement policy. Moreover, we found no differences by family or marital status.

In terms of the local characteristics eight years after immigration the mover and stayer groups are fairly similar. Those who moved entered larger municipalities with a greater share of immigrants and slightly lower rates of unemployment relative to those who stayed on.¹² A comparison of characteristics of the assigned and destination municipalities for movers reveals that they escaped smaller municipalities, where immigrant density was low and unemployment high.

To conclude this section, we wish to emphasize two points. First, there is little evidence of a correlation between the characteristics of the assigned region and observed individual attributes. Second, since the majority of the individuals stay on in

12. The big city municipalities of Stockholm (715,000 inhabitants), Gothenburg (452,000 inhabitants), and Malmö (247,000 inhabitants) are very influential in the averages, which are based on individual data.

the assigned municipality, the instruments will have strong predictive power.

V. THE EFFECTS OF LIVING IN ENCLAVES: EMPIRICAL ESTIMATES

In this section we estimate the earnings effects of living in ethnic enclaves. We first present a set of baseline estimates for the full sample and explore to what extent these estimates differ across skill (education) groups. We then turn to analyzing to what extent the quality of the enclave matters. The final section examines whether the enclave effects vary with time in Sweden.

V.A. *The Baseline Estimates*

The basic strategy to free the estimates of simultaneity bias between earnings and region of residence is to use the settlement policy introduced in 1985 to obtain instruments for local variables. In effect, we use variables pertaining to the initial (assigned) municipality as instruments for municipal variables eight years later. Our maintained assumption is that the placement policy is independent of unobserved individual characteristics. Moreover, we assume that (initial) location does not have permanent effects on outcomes.¹³

We employ the following baseline specification:

(6)

$$\ln(\text{earnings})_{ijk(t+8)} = \alpha' \mathbf{X}_{ik(t+8)} + \beta \ln e_{jk(t+8)} + \delta_{j^*} + \delta_k + \delta_t + \epsilon_{ijk(t+8)},$$

where i indexes individuals, j municipalities, k countries of origin, and t years of immigration. We focus on the log of annual earnings and standardize for a set of individual characteristics \mathbf{X} , containing gender, age, age squared, marital status, education, and a dummy for the presence of young kids in the household. The earnings of individual i of ethnic group k is also related to the size of the ethnic group in municipality j ($e_{jk(t+8)}$). This measure of ethnic concentration is our key variable of interest. In order to control for omitted variables, the regressions include a full set of

13. An issue is whether this exclusion restriction is valid. To examine this issue, we have performed Sargan tests of overidentifying restrictions on slight variations of our basic model. Since our baseline specification is exactly identified, we cannot apply the Sargan procedure directly. Therefore, we amended the instrument set by including the lag of the size of the ethnic group in the assigned municipality. For the full sample, the parameter of interest was not affected by altering the set of instruments. Furthermore, the Sargan-statistics do not reject the estimated models.

country of origin fixed effects (δ_k), immigration year fixed effects (δ_t), and fixed effects for the assigned municipality (δ_{j^*}). We only control for the assigned municipality, as the coefficients on the current municipality may be endogenous because of sorting and instrumenting the full set of municipality dummies is simply too taxing on the data. Notice that the size of the ethnic group is identified since it varies across ethnic groups within municipalities. Summary statistics for the variables included in equation (6) are reported in the Appendix.

All results that we present are based on a weighting procedure. The rationale for applying a weighting scheme is that we do not actually observe refugees who were placed and where they were placed. There are three problems in the data related to this fact. First of all, around 10 percent of refugees were not subjected to the placement policy. Second, if refugees move within the year of receiving residence permits, we do not observe their municipality of placement (since we only observe residence at the end of the year). Third, a comparison of the number of individuals in our sample with the total number of allotted residence permits 1987–1989 suggests that we sample too many individuals. To alleviate these three problems, we use external information on the number of received refugee immigrants in each municipality. We reweigh our data such that the distribution of the initial locations of the individuals in our sample corresponds to the distribution of received refugees over municipalities. The weighting procedure is based on the number of refugees covered by grants from the Immigration Board.

We estimate equation (6) for the full sample, but also for different education groups. The first-stage equations in the IV (2SLS) procedure amount to regressing $\ln e_{jk(t+8)}$ on $\ln e_{jkt}$ and the remaining explanatory variables in (6). All first-stage regressions suggest that the instrument has strong predictive power: for the entire sample, the size of the ethnic group in the assigned municipality enters the equation with a coefficient of 0.41 and a t -ratio of 8.35; for low educated, the estimate is 0.47, and the t -ratio 9.58; and for high-educated the estimate is 0.38 with a t -ratio of 6.68.

Table III reports the results of the basic specification where the outcome of interest is the log of earnings conditional on

TABLE III
BASELINE ESTIMATES—DEPENDENT VARIABLE: $\ln(\text{EARNINGS})$

	Full sample		Low education (10 years or less)		High education (more than 10 years)	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
$\ln(\text{ethnic group})$	-.056 (.022)	.012 (.050)	-.053 (.024)	.174 (.088)	-.050 (.030)	-.057 (.080)
Female	-.071 (.081)	-.069 (.082)	-.087 (.128)	-.050 (.132)	-.004 (.098)	-.004 (.098)
Age	.066 (.023)	.068 (.022)	.079 (.038)	.099 (.040)	.054 (.030)	.054 (.031)
Age squared ($\times 10^{-2}$)	-.074 (.028)	-.075 (.027)	-.090 (.049)	-.112 (.052)	-.062 (.036)	-.062 (.037)
Married	.210 (.084)	.210 (.084)	.289 (.162)	.278 (.166)	.168 (.072)	.167 (.073)
Kid	-.027 (.075)	-.004 (.082)	-.115 (.132)	-.050 (.138)	.083 (.086)	.081 (.102)
Married*female	-.049 (.100)	-.032 (.100)	-.226 (.153)	-.207 (.162)	.012 (.106)	.011 (.106)
Kid*female	-.262 (.125)	-.278 (.124)	-.144 (.222)	-.223 (.229)	-.391 (.137)	-.389 (.139)
Education missing and <9 years	Ref. .078 (.060)	Ref. .077 (.059)	Ref. .097 (.069)	Ref. .084 (.070)		
High school ≤ 2 years	.204 (.088)	.209 (.087)			Ref.	Ref.
High school >2 years	.196 (.070)	.204 (.069)			-.013 (.081)	-.013 (.081)
University <3 years	.181 (.071)	.180 (.070)			.006 (.072)	.007 (.072)
University ≥ 3 years	.525 (.081)	.526 (.082)			.341 (.076)	.341 (.076)
Immigration year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Country of origin dummies	Yes	Yes	Yes	Yes	Yes	Yes
Municipality dummies	Yes	Yes	Yes	Yes	Yes	Yes
No. of individuals	6393	6393	2205	2205	4188	4188
Standard error of regression	1.44	.145	1.44	1.48	1.42	1.42

Standard errors are in parentheses. IV estimation is by 2SLS using the size of the ethnic group in the assigned municipalities as an instrument for the size of the ethnic group in the municipality eight years later. Estimates are weighted using the number of refugees by municipality covered by grants from the Immigration Board. Robust variance estimates, allowing for correlation across individuals residing in the same municipality.

having positive earnings.¹⁴ Since we are using annual earnings, we are capturing wage as well as employment variations. Columns (1), (3), and (5) report OLS estimates where we treat ethnic concentration as exogenous. Columns (2), (4), and (6) give the results of the IV procedure, which uses the ethnic concentration for each immigrant's initial placement (in year t) as an instrument for current (year $t + 8$) ethnic concentration. In order to interpret the estimates on the size of the ethnic group, we calculate the standard deviation relative to the mean in ethnic concentration within countries of origin. The relative standard deviation is 73 percent (the absolute size of the standard deviation is 0.24 percentage points).

According to the OLS estimates in column (1), ethnic concentration has a negative earnings effect. A standard deviation increase in ethnic concentration is associated with a significant earnings loss of around 4 percent. The corresponding IV estimates, however, suggest that this conclusion may be premature, as the estimate on ethnic concentration switches sign and becomes insignificant.

In the first two columns the effect of enclaves is restricted to be constant across groups. There are several reasons why this may not be true. For instance, less skilled immigrants are more prone to live in enclaves [Lalonde and Topel 1997], suggesting perhaps that they have the greatest benefit of living there or that they face restrictions in the housing market. Independent of the underlying cause of this location pattern, it is reasonable to expect that the return to living in the enclave varies by, e.g., educational groups.

To investigate this issue further, we have performed separate analyses for subgroups defined by gender and individual education. We found no differences in the returns to enclaving between males and females. There were some interesting differences across educational groups, however.

In the last four columns of Table III, we present estimates for two educational groups. In columns (3) and (4), respectively, we report OLS and IV estimates for the subsample with at most ten years of schooling. The corresponding estimates for immigrants with more than ten years of schooling are reported in columns (5)

14. This raises the issue of selection into the group with positive income. Notice, though, that the probability of having positive earnings is unrelated to the size of the enclave.

and (6). The effect of ethnic concentration exhibits an interesting pattern. There is a large positive effect of living in an enclave for the least skilled (i.e., least educated). A standard deviation increase in ethnic concentration causes a 13 percent increase in earnings in this group. The presence of strong sorting bias is evident:¹⁵ the estimated effect goes from significantly negative in OLS to large and positive using IV.¹⁶ For immigrants with (short) high school or more education, there is no evidence of an enclave effect. These results clearly indicate that the least skilled are the ones who gain most from living in ethnic enclaves.¹⁷ This is consistent with a story that emphasizes that enclaves are associated with ethnic networks that primarily benefit the least skilled.

We have subjected the baseline specification to a number of robustness checks. Let us mention two of them. First, we have excluded education from the set of controls. The idea is that if observed and unobserved skills similarly affect the probability of being placed, then omitting education provides information on the omitted variable bias. Applying this procedure, the estimate on the size of the ethnic group drops in all specifications. Thus, there seems to be no upward bias resulting from the fact that those with high skills (observed and unobserved) have been able to influence placement. Second, we have also estimated equations based on alternative weighting strategies. In general, the results are very similar. In particular, note that the results for the least educated are not driven by the weighting procedure; unweighted estimates yield a slightly larger and more precise coefficient on the size of the ethnic group.

To sum up this subsection, we have two major findings. First, there is evidence that estimates of “neighborhood” effects that do not account for sorting may be severely biased. Our results indicate that high (unobserved) ability immigrants locate outside

15. A Hausman test for exogeneity of the size of the ethnic group rejects exogeneity for the low educated. The test statistic is $F(1,205) = 7.60$ (degrees of freedom within parentheses), with a p -value of 0.006. Exogeneity cannot be rejected in the full sample or in the group with higher education.

16. A comparison between the OLS and IV estimates does not yield a correct picture of the size of the sorting bias. This is so because the majority of our individuals remained in the locations in which they were originally placed. Sorting will clearly be present among individuals who moved to a new region. An OLS regression for low skilled movers indicates a larger sorting bias: the OLS estimate is -0.112 with a standard error of 0.033.

17. Interestingly, Berman, Lang, and Siniver [2000] find that this group gains the least from language acquisition.

ethnic enclaves to a greater extent. Second, and perhaps more importantly, less skilled immigrants derive a substantial positive return from living in ethnic enclaves. For this group, there is an earnings gain of 13 percent associated with a standard deviation increase in ethnic concentration.

V.B. The Quality of Enclaves

Several theoretical models suggest that the quality of the enclave is important for the effects of segregation. This is evident in the models that rely on human capital externalities, e.g., Cutler and Glaeser [1997] and Borjas [1998], where segregation may have a positive effect when skilled members of the ethnic group stay in the enclave. Similarly, it is reasonable to think that the quality of the enclave may be important in models that emphasize network effects.

Here we explore specifications along the lines of Bertrand, Luttmer, and Mullainathan [2000]. We rely on two different indicators to get at enclave quality. First, we use the average annual labor income of the ethnic group in 1986, i.e., the year prior to our first inflow year. Second, we use the share of self-employment in the ethnic group in 1986. Both these variables capture aspects of how the ethnic peers of our refugee immigrants have fared in the Swedish labor market. The main effect of our enclave quality variables are not identified separately from the country of origin fixed effects. The variables of interest are the interactions of quality and ethnic concentration.

In Table IV we report the estimates with measures of enclave quality. These are treated as endogenous variables, using initial placement as an instrument. As before, we report estimates for the full sample, as well as for the two education groups. The results indicate that the quality of enclaves matters for the earnings effect of ethnic concentration. In general, the main effect of ethnic concentration is much smaller now. Actually, the main effect is significantly negative for the full sample and for the high skilled sample, while it is zero for the less educated sample.

The negative main effect of ethnic concentration is counter-balanced by positive interaction effects with both our measures of enclave quality: ethnic income and ethnic self-employment rate. In the case of ethnic income, the estimates of the interaction effects are very similar across skill groups, even though the effect for the less skilled sample is less precisely estimated. The overall effect of a standard deviation increase in ethnic concentration on

TABLE IV
THE "QUALITY" OF ENCLAVES—DEPENDENT VARIABLE: $\ln(\text{EARNINGS})$

	Full sample		Low education (10 years or less)		High education (more than 10 years)	
	(1)	(2)	(3)	(4)	(5)	(6)
	IV	IV	IV	IV	IV	IV
$\ln(\text{ethnic group})$	-.221 (.109)	-.138 (.071)	-.031 (.220)	.027 (.114)	-.315 (.161)	-.267 (.113)
$\ln(\text{ethnic group}) * \text{ethnic inc}$.044 (.015)		.039 (.040)		.047 (.019)	
$\ln(\text{ethnic group}) * \text{ethnic self-employment rate}$		3.212 (.926)		2.964 (1.589)		4.592 (1.309)
Female	-.068 (.081)	-.071 (.080)	-.049 (.130)	-.039 (.133)	-.011 (.099)	-.018 (.098)
Age	.065 (.022)	.069 (.023)	.095 (.040)	.096 (.041)	.053 (.031)	.056 (.031)
Age squared ($*10^{-2}$)	-.072 (.027)	-.077 (.028)	-.108 (.052)	-.110 (.053)	-.061 (.037)	-.065 (.038)
Married	.204 (.088)	.203 (.084)	.287 (.168)	.290 (.162)	.147 (.076)	.143 (.073)
Kid	-.022 (.083)	-.025 (.082)	-.068 (.140)	-.080 (.134)	.057 (.106)	.059 (.102)
Married*female	-.044 (.101)	-.046 (.102)	-.216 (.163)	-.236 (.152)	.002 (.108)	.008 (.111)
Kid*female	-.264 (.126)	-.264 (.127)	-.213 (.228)	-.215 (.228)	-.367 (.141)	-.366 (.141)
Education missing and <9 years	Ref.	Ref.	Ref.	Ref.		
9–10 years	.077 (.059)	.071 (.059)	.086 (.069)	.079 (.069)		
High school ≤ 2 years	.202 (.086)	.200 (.087)			Ref.	Ref.
High school >2 years	.195 (.069)	.190 (.068)			-.013 (.083)	-.017 (.083)
University <3 years	.186 (.072)	.174 (.068)			.020 (.072)	.010 (.074)
University ≥ 3 years	.527 (.081)	.511 (.083)			.349 (.077)	.335 (.076)
Immigration year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Country of origin dummies	Yes	Yes	Yes	Yes	Yes	Yes
Municipality dummies	Yes	Yes	Yes	Yes	Yes	Yes
No. of individuals	6393	6393	2205	2205	4188	4188
Standard error of regression	1.45	1.45	1.47	1.47	1.43	1.43

Standard errors are in parentheses. IV estimation by 2SLS using values in the assigned municipalities as instruments for $(t + 8)$ local variables. Estimates are weighted using the number of refugees by municipality covered by grants from the Immigration Board. Robust variance estimates, allowing for correlation across individuals residing in the same municipality.

earnings turns out to be small and negative (-1.5 percent) in column (1) if we evaluate at sample means of ethnic income. If we evaluate the effect at one standard deviation above the mean, the effect is large and positive.¹⁸ In this case a standard deviation increase in ethnic concentration raises earnings by 4.6 percent.

The estimates where the ethnic rate of self-employment is used as a measure of enclave quality are similar to the ethnic income estimates. The interaction terms between ethnic concentration and ethnic self-employment are all positive and significant, though only at the 10 percent level for the less skilled sample. Evaluating the overall effect of ethnic concentration at sample means suggests an earnings loss of 1.9 percent from a standard deviation in ethnic concentration. With a doubling of the self-employment rate, the earnings gain is about 6.3 percent from the same variation in ethnic concentration.

Overall, these results suggest that the characteristics of the enclave are important for the effects of living in an enclave. Enclaves with high ethnic income or high ethnic self-employment drive the positive effect of residing in an enclave. Residing in an enclave with less than average quality, according to our measures, may actually have a negative effect on earnings.

V.C. Do the Effects Vary with Time in the Enclave?

In the above analysis, we estimate the relationship between contemporaneous ethnic concentration and earnings. However, most models of the effect of living in an enclave suggest that time spent in the enclave is what matters for outcomes.¹⁹ Here, we provide some tentative evidence on whether the effect of living in enclaves is contemporaneous or cumulative.

Let us start by emphasizing that the main source of identification in our setting comes from individuals who stayed on in the assigned municipalities. Thus, those for whom we primarily identify the effect will have lived in the local area for eight years. So, in practice, the IV estimates will be more or less the same, independent of whether the effect is contemporaneous or cumulative.²⁰ However, if the effect is cumulative, we may have mis-

18. An increase in ethnic income of one standard deviation is equivalent to increasing income from 54 percent to 76 percent of mean native earnings.

19. This is also supported by the empirical results of Glaeser and Maré [2001] in the context of urban wage premiums.

20. If the exclusion restriction is right, the IV-estimate may be thought of as the treatment effect of increasing the size of the enclave cumulated over eight

TABLE V
REDUCED-FORM ESTIMATES OF THE "QUALITY" OF ENCLAVES AT DIFFERENT POINTS
IN TIME
DEPENDENT VARIABLE: $\ln(\text{EARNINGS})$

Year after immigration	$t + 2$	$t + 3$	$t + 4$	$t + 5$	$t + 6$	$t + 7$	$t + 8$
$\ln(\text{ethnic group})_t$	-.010 (.042)	-.065 (.048)	-.096 (.061)	.007 (.064)	-.045 (.051)	-.115 (.060)	-.101 (.050)
$\ln(\text{ethnic group})_t * \text{ethnic inc}$	-.004 (.006)	.006 (.007)	.009 (.009)	.002 (.011)	.011 (.008)	.021 (.010)	.022 (.009)
No. of individuals	5473	5455	5332	5200	5327	5637	6393
Standard error of regression	1.16	1.28	1.31	1.32	1.30	1.24	1.44

Standard errors are in parentheses. $\ln(\text{ethnic group})$ refers to the size of the ethnic group in the assigned municipality. Apart from that, the regressions include the same explanatory variables as column (1) in Table IV. Estimates are weighted using the number of refugees by municipality covered by grants from the Immigration Board. Robust variance estimates, allowing for correlation across individuals residing in the same municipality.

specified the equation for the part of the sample that moves since it should potentially include the characteristics of previous regions of residence.²¹

To shed some light on whether the effects are contemporaneous or cumulative, we estimate the reduced-form relationship corresponding to column (1) of Table IV at different points in time after immigration. We estimate such equations for each year from $t + 2$ to $t + 8$, where t is the year of immigration, for the individuals included in the original regression. The rationale for excluding the first two years is that they coincide with the introductory period when individuals could live off welfare.²² If the effects of enclave size are increasing over time in Sweden, then this suggests that the effect is cumulative in nature.

Table V gives the estimates on size of the ethnic group in the assigned municipality and the corresponding interaction term with ethnic income: the results of this exercise are summarized

years. Relative to the reduced form, the IV estimate corrects for the fact that not all individuals complied with the assigned treatment dose.

21. If the time spent in the enclave affects the accumulation of host country human capital, we have incorrectly excluded the size of the ethnic group in the assigned municipality. If it primarily is a cumulating network effect that operates, however, the issue of misspecification depends on whether the benefits of the network are kept after moving to a new region.

22. The probability of having positive earnings in year t is around 40 percent for individuals with positive earnings in $t + 8$. The probability of having positive earnings in $t + 2$, on the other hand, is 86 percent for these individuals.

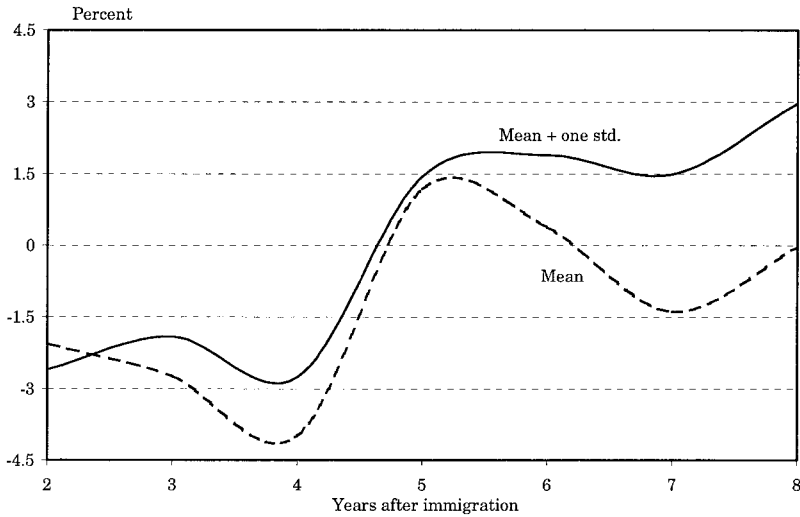


FIGURE II

The Reduced-Form Earnings Effect of Ethnic Concentration over Time

The figure is based on Table V. The dashed line shows the effect evaluated at mean ethnic income; the solid line shows the effect evaluated at one standard deviation above the mean.

in Figure II. The figure shows the earnings effect of increasing ethnic concentration by a standard deviation for two different values of ethnic income. The dashed line shows the effect evaluated at mean ethnic income, and the solid line gives the effect evaluated at a standard deviation above the mean. Notice that the interaction term is significant only in $t + 7$ and $t + 8$. There is some evidence suggesting that the effects are cumulative. According to the reduced-form coefficients, the effect of living in a high quality enclave turns positive after five years in the host country and becomes significant, at a slightly higher level, after eight years.

VI. CONCLUDING REMARKS

The main purpose of this paper has been to provide estimates of the causal effect on economic outcomes of living in an enclave. To this end, we have made use of an immigrant policy initiative in Sweden, when government authorities distributed refugee immigrants across locales in a way that may be considered exoge-

nous. This policy initiative provides a unique natural experiment, which allows us to handle the endogeneity problem due to the residential choice of individuals.

Our empirical analysis suggests three principal conclusions. First, we find evidence of sorting across locations. There is a downward bias in estimates that do not account for sorting. Second, when accounting for the endogeneity of residential choice, we find that earnings rise with increases in ethnic concentration for individuals at the lower end of the observed skill distribution. The earnings gain associated with a standard deviation increase in ethnic concentration is about 13 percent (and statistically significant) for less educated immigrants. Third, we find that the positive effects of enclaves are increasing with the "quality" of enclaves. Immigrants in ethnic groups with high earnings or high self-employment rates have higher returns to living in an enclave. Immigrants belonging to an ethnic group with less than average earnings may actually lose from residing in the enclave. This result is in line with previous research in Borjas [1995] where ethnic capital during childhood has a positive influence on adult outcomes.

We have also presented suggestive evidence implying that the effect of living in the enclave is cumulative in nature. Members of high-income ethnic groups gain more from ethnic concentration after eight years in Sweden than after, say, four years in the host country. The precise time pattern of the effects of ethnic concentration is an interesting topic worthy of further research.

APPENDIX

Table VI reports the definitions and primary sources of data for variables used in the empirical analysis. Information on gender, age, and marital status comes from the Income tax registers. The definitions of these variables should be obvious and are not included.

Table VII presents means and standard deviations of the variables relevant for the empirical analysis.

TABLE VI
VARIABLE DEFINITIONS AND PRIMARY SOURCES OF DATA

Variable	Definition	Primary source of data
Individual characteristics		
Annual earnings	The sum of income from work and income from business. All transfers unrelated to the employment position are excluded.	Income tax registers, Statistics Sweden (SCB).
Education	Highest degree attained.	Education register, SCB.
Kid	Dummy for the presence of kids ≤ 15 years of age in the household.	Income tax registers, SCB.
Immigration year	Year of receipt of residence permit.	Population register, SCB.
Country of origin	Immigrant source country.	Population register, SCB.
Local/group characteristics		
Size of ethnic group (e_{jk})	Number of foreign citizens from source country k residing in municipality j .	Population statistics (population counted data), SCB.
Number of immigrants	Number of foreign citizens residing in municipality j .	Population statistics (population counted data), SCB.
Unemployment rate	Percent of the population aged 16–64 registered as unemployed in municipality j .	Unemployment register (population counted data), the Labor Market Board.
Mean income in ethnic group 1986	Average earnings (including zeros) among individuals aged 18–64 from source country k .	Income tax register, SCB. Authors' calculations based on 20 percent sample of immigrants.
Self-employment rate in ethnic group 1986	Percent of ethnic group k , aged 18–64, with income from business as their primary source of income.	Income tax register, SCB. Authors' calculations based on 20 percent sample of immigrants.

TABLE VII
SUMMARY STATISTICS: MEANS (STANDARD DEVIATIONS)

Variable	All	Weighted	Stayers	Movers
Local/group characteristics				
ln(ethnic group) (lne)	5.104 (2.111)	4.942 (2.137)	4.995 (2.057)	5.234 (2.167)
Ethnic concentration (percent) (standard deviation within country of origin)	.330 (.240)	.319 (.239)	.320 (.238)	.343 (.243)
Mean income in ethnic group 1986 (10,000 SEK)	4.564 (1.883)	4.532 (1.888)	4.892 (1.808)	4.173 (1.897)
Self-employment in ethnic group 1986	.035	.034	.036	.033
Individual characteristics				
ln(earnings)	10.955 (1.513)	10.927 (1.533)	11.075 (1.424)	10.812 (1.601)
Age	37.437 (7.506)	37.443 (7.460)	38.136 (7.872)	36.602 (6.955)
Female	.447	.459	.482	.406
Married	.624	.631	.657	.586
Kid	.560	.568	.583	.533
Education: missing	.023	.022	.021	.025
<9 years	.139	.137	.184	.112
9–10 years	.182	.182	.193	.170
High school ≤2 years	.190	.197	.198	.181
High school >2 years	.190	.189	.188	.193
University <3 years	.144	.143	.121	.171
University ≥3 years	.131	.130	.117	.149
Region of origin: Eastern Europe	.201	.202	.241	.154
Africa	.117	.113	.104	.133
Middle East	.403	.402	.326	.494
Asia	.089	.089	.101	.075
South America	.190	.194	.229	.144
Immigration year: 1987	.276	.273	.280	.271
1988	.357	.355	.364	.348
1989	.367	.372	.356	.381
No. of individuals	6418	6393	3492	2926

All variables refer to the situation in $t + 8$, where t is the year of immigration. The “standard deviation within country” is a weighted average of the standard deviations within each country of origin. The column headed “weighted” reports weighted means as used in the regression analysis. “Stayers” are those who stayed on in the assigned municipality between t and $t + 8$. “Movers” are those who moved between these two time points.

REFERENCES

- Åslund, Olof, "Immigrant Settlement Policies and Subsequent Migration," Working Paper No. 2000:23, Department of Economics, Uppsala University, 2000.
- Berman, Eli, Kevin Lang, and Erez Siniver, "Language-Skill Complementarity: Returns to Immigrant Language Acquisition," NBER Working Paper No. 6832, 2000.
- Bertrand, Marianne, Erzo Luttmer, and Sendhil Mullainathan, "Network Effects and Welfare Cultures," *Quarterly Journal of Economics*, CXV (2000), 1019–1055.
- Borjas, George, "Ethnicity, Neighborhoods, and Human-Capital Externalities," *American Economic Review*, LXXXV (1995), 365–390.
- , "To Ghetto or Not to Ghetto: Ethnicity and Residential Segregation," *Journal of Urban Economics*, XLIV (1998), 228–253.
- , *Heaven's Door—Immigration Policy and the American Economy* (Princeton, NJ: Princeton University Press, 1999).
- Card, David, "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, Volume 3, Orley Ashenfelter and David Card, eds. (Amsterdam: North-Holland, 1999).
- The Committee on Immigration Policy, *Sverige, framtiden och mångfalden. Slutbetänkande från Invandrarpolitiska kommittén*, SOU 1996:55 (Stockholm: Fritzes, 1996).
- Cutler, David, and Edward Glaeser, "Are Ghettos Good or Bad?" *Quarterly Journal of Economics*, CXII (1997), 827–872.
- Dustmann, Christian, and I. Preston, "Attitudes to Ethnic Minorities, Ethnic Context and Location Decision," CEPR Discussion Paper No. 1942, 1998.
- Edin, Per-Anders, and Peter Fredriksson, "LINDA—Longitudinal Individual Data for Sweden," Working Paper No. 2000:19, Department of Economics, Uppsala University, 2000.
- Edin, Per-Anders, Peter Fredriksson, and Olof Åslund, "Ethnic Enclaves and the Economic Success of Immigrants—Evidence from a Natural Experiment," CEPR Discussion Paper No. 2729, 2001.
- Gabriel, Stuart, and Stuart Rosenthal, "Location and the Effect of Demographic Traits on Earnings," *Regional Science and Urban Economics*, XXIX (1999), 445–461.
- Garen, John, "The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable," *Econometrica*, LII (1984), 1199–1218.
- Glaeser, Edward, and David Maré, "Cities and Skills," *Journal of Labor Economics*, XIX (2001), 316–342.
- Ihlanfeldt, Keith, and David Sjoquist, "The Spatial Mismatch Hypothesis: A Review of Recent Studies and Their Implications for Welfare Reform," *Housing Policy Debate*, IX (1998), 849–892.
- The Immigration Board, *Individuell mångfald: Invandrarverkets utvärdering och analys av det samordnade flyktingmottagandet 1991–1996* (Norrköping: Statens invandrarverk, 1997).
- Katz, Lawrence, Jeffrey Kling, and Jeffrey Liebman, "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *Quarterly Journal of Economics*, CXVI (2001), 607–654.
- LaLonde, Robert, and Robert Topel, "Economic Impact of International Migration and the Economic Performance of Migrants," in *Handbook of Population Economics*, Mark Rosenzweig and Oded Stark, eds. (Amsterdam: North-Holland, 1997).
- Lazear, Edward, "Culture and Language," *Journal of Political Economy*, CVII (1999), S95–S126.
- Portes, Alejandro, "The Social Origins of the Cuban Enclave Economy of Miami," *Sociological Perspectives*, XXX (1987), 340–372.
- Rooth, Dan-Olof, "Refugee Immigrants in Sweden—Educational Investments and Labour Market Integration," Ph.D. thesis, Lund Economic Studies No. 84, Department of Economics, University of Lund, 1999.
- Stark, Oded, *The Migration of Labor* (Oxford: Blackwell, 1991).