The Stockholm University Linnaeus Center
for Integration Studies (SULCIS)

Unpacking the Causes of Ethnic Segregation across Workplaces

Magnus Bygren

Working Paper 2010:2
ISSN 1654-1189

Unpacking the Causes of Ethnic Segregation across Workplaces

Magnus Bygren

Abstract

Using a large sample of employees-within-workplaces, the author investigates the relative role of random and systematic sorting for ethnic segregation across workplaces. If employees, in a counterfactual world, were randomly allocated to workplaces, the level of ethnic segregation across workplaces would just be halved. The remainder of segregation - systematic segregation - is upheld because employees that are recruited to workplaces tend to be similar to those already employed there, not because underrepresented groups within workplaces are systematically screened out of them. This homosocial inflow of employees appears largely to be sustained by employers' tendency to select new employees from a pool of workplaces where its employees have been employed previously.

JEL codes: C15, J1, J2

Key words: workplaces, segregation, ethnicity, simulation

Financial support from the Swedish Council for Working Life and Social Research (FAS Dnr 2001-2442) and the Swedish Research Council (VR Dnr 421-2005-1424) is gratefully acknowledged.

Introduction

Countless casual and systematic observations indicate that immigrants in developed countries cluster in certain segments of the host country labor market. Compared to the native population, immigrants usually have a higher probability to hold low-qualified jobs, and typical sectors with a higher-than-average proportion of immigrants are manufacturing, construction, hotels, restaurants, and health care (OECD 2006). Within industries, immigrants also have a tendency to concentrate in certain workplaces (Åslund and Nordström Skans 2009).

An understanding of how this segregation comes about is important for at least two reasons. First, segregation is a structural property of the labor market, and as such it is important because it restrains between-group social interaction in the society. Social integration, in the sense that ascribed characteristics is unrelated to actual social interactions taking place, is a socially valued goal because it is assumed to reduce prejudice, stereotypes and unequal access to resources. As a substantial share of the interaction between people take place at workplaces (at least for the adult population) a decrease in segregation across workplaces would imply a decrease in segregation in the society as a whole, probably more so than would a decrease in residential segregation (Ellis, Wright, and Parks 2004). In other words, if we want to achieve integration in the society as a whole, a reduction of labor market segregation seems to be the key. Second, this topic is important because it represents a void in the scholarly literature on segregation. During the past decades there has been a surge in research on organizational (or work unit, team, workplace, etc.) demography, and its relevance for a range of outcomes (DiTomaso, Post, and Parks-Yancy 2007). However, despite its alleged importance for a number of group processes, group outcomes, and individual outcomes, the study of the *determinants* of segregation across work units has not been a great concern in labor market research – a notion that repeatedly has been pointed out (Mittman 1992; Tsui and Gutek 1999; Williams and O'Reilly 1998). Consequences of segregation, diversity and minority status constitutes a minor industry in sociological research. However, our knowledge about the processes leading to segregation in the labor market is still much wanting.

This study provides evidence on this process behind ethnic segregation in the context of an urban labor market in Sweden during the period 2001 – 2003. The contribution of the study to research is twofold. First, I will assess whether the segregation across workplaces can be accounted for using readily observable characteristics of employees and workplaces. Second, as changes in segregation across workplaces depend on the past composition of the workplaces' inflow and outflow of employees, I will investigate whether segregation across workplaces is sustained by the inflow or outflow of employees to and from workplaces. In this way a set of general "causal markers" of ethnic segregation across workplaces can hopefully be identified, and guide future research in this area.

Sources of segregation across workplaces

The demographic composition of a workplace at a given point in time is the result of its history of inflow and outflow of employees with different demographic characteristics (Stinchcombe, McDill, and Walker 1968). Differential mobility of various groups may either reproduce the existing composition of a workplace or alter it. A useful framework for explaining how demographic distributions are generated is therefore to view them as a final result of individuals entering and leaving workplaces.

One potentially crucial determinant of segregation across workplaces is the demographic composition in the labor pools from which workplaces recruit. If there is a large proportion of immigrants in the pool of employees available for recruitment to a workplace, a nondiscriminating employer should have a higher propensity to hire immigrants. Through studies of sex segregation across workplaces, we know that because these do not have identical demand for workers, sex segregation in industries and occupations give rise to workplace sex segregation (Bygren and Kumlin 2005; Carrington and Troske 1998). There are also other sources of this so-called baseline homophily, as the relative ease with which an employer can come into contact with potential immigrant employees may be affected in other ways too. If a workplace is located in, or close to, an area with a higher-than-average proportion of immigrants, the likelihood that the workplace will have a high proportion of immigrants should also be higher-than-average (again, if employers are nondiscriminating). Closeness can also be social, where potential employees are socially linked to the focal workplace in different ways, one of which is ties with employees that are already employed there. As emphasized by Granovetter (1974), informal networks are often used in the job seeking process, and employers frequently make use their employees' networks for recruitment purposes, primarily because this recruitment strategy is perceived to be cost-efficient, yielding applicants of higher quality compared to other recruitment methods (Marsden and Gorman 2001; Mencken and Winfield 1998). When an employer use network recruiting, the expectation is however that the demographic composition of the employees reproduce, as those individuals who are part of the network tend to resemble those already working for the employer (Marsden 1996). Studies have consistently shown a substantial proportion of jobs to be appointed through informal contacts (Ekström 2001; Mouw 2003; Wegener 1991). As immigrants usually hold less qualified jobs, and networks are segregated by ethnicity, it is relatively safe to say that recruitment through informal contacts tend to sort immigrants into less qualified jobs.

Processes that occur once employees have been recruited to a workplace may also contribute to segregation. In particular, processes of exclusion and isolation may, if these are driven by differences in ethnicity, enhance segregation through differences in turnover between different groups. Through social categorization and similarity attraction mechanisms individuals may define themselves and others using salient social categories such as sex, ethnicity/race and age (van Knippenberg and Schippers 2007).

In general, individuals seek to enhance the positive distinctiveness of their own group in relation to other groups, prefer the company of fellow in-group members and sometimes discriminate against out-groups (Abrams and Hogg 1990). Proceeding from these premises, Pfeffer (1985) formulated the following causal narrative on turnover: 1. turnover is produced, at least in part, by conflict and disagreement, as well as by a lack of group cohesion. 2. As cohesion, consensus and agreement are facilitated by demographic similarity, those who are most different in a workplace should be those who experience the lowest degree of integration in a group, and possibly have a higher degree of conflict and disagreement with others. 3. Therefore, they should also have the highest turnover rates, and the outflow of employees from workplaces should as a consequence contribute to ethnic segregation.

There are, in other words, reasons to expect both the inflow of employees to and outflow of employees from workplaces to contribute to segregation across workplaces. However, to the knowledge of the author, no study has addressed the empirical question of the relative importance of these flows for segregation. The present study represents a step towards filling this void in the literature on the determinants of ethnic segregation in the labor market.

Measuring segregation across workplaces

Carrington and Troske (1997), Winship (1977), and others have highlighted an important problem of upward bias in segregation indices when the units across which segregation is analyzed are small, or the minority share is small. To illustrate, say that all workplaces consist of two employees each, and that there is a 50/50 mix of men and women in the work force. With a purely random allocation of men and women to workplaces, only 50% of the workplaces are expected to have an even mix of men and women, and segregation, as measured by a standard segregation index, would be quite high. With a 10/90 mix, measured segregation would be extreme. The lesson is that the benchmark against which one should compare observed levels of segregation is not the zero level of perfect, and sometimes unattainable, integration. A more reasonable benchmark is that which would obtain had individuals been randomly allocated to existing units. For many real-world phenomena, observed segregation is probably a sum of random and systematic allocation rules, and in order to find out the extent of systematic segregation, we should net out segregation originating from random forces.

Following Carrington and Troske (1997), let D be the observed Duncan & Duncan dissimilarity index (D could without loss of generality be replaced by another segregation index with a range {0, 1}), and let D^* be $E(D)$ implied by the random allocation of a population to units. Their index of systematic dissimilarity, $\hat{D}$, is defined as

$$\hat{D} = \begin{cases} \dfrac{D - D^*}{1 - D^*} & if \quad D \geq D^* \\ \dfrac{D - D^*}{D^*} & if \quad D < D^* \end{cases} \qquad (1)$$

When $\hat{D} = 0$, the observed dissimilarity D is equal to the expected dissimilarity D^*, and no systematic segregation is at hand. When $\hat{D} = 1$, complete segregation is at hand, and as D^* is zero, all segregation is systematic. When $\hat{D} = -1$, complete integration is at hand, at the same time as $D^* > 0$, which implies that all desegregation is systematic. $\hat{D}$ is different from the standard measure of dissimilarity D in two ways. First, the baseline value zero corresponds to the segregation level under random allocation rather than perfect integration. Second, values of D that are greater than D^* are remapped into the systematic segregation interval {0, 1}, and values of D that are less than D^* are remapped into the systematic desegregation interval {-1, 0}. When the sizes of units to which individuals are allocated vary, there is no simple expression for D^*. Instead, this quantity can be estimated through simulation of a number of counterfactual random allocations of employees to workplaces. For each of these, D is computed and the average of these is an estimate of D^*.

One may ask whether differences in observable characteristics between natives and immigrants can explain, or hint at an explanation, of systematic segregation. Immigrants and natives come to the labor market with different kinds of assets relevant to employers. For example, irrespective of ethnicity, we would expect employees with similar education to cluster in the same workplaces. Thus, if immigrants and natives differ with regard to education we can expect this kind of heterogeneity to produce segregation. As argued previously, workplaces also differ in their geographical and social closeness to potential immigrant employees, and heterogeneity with regard to this kind of closeness can be expected to give rise to segregation. To analyze the influence of this heterogeneity on segregation, I used a method suggested by Åslund and Nordström Skans (2009), and reassigned immigrant status on the basis of probabilities of being an immigrant given a set of observed characteristics. I first assign for each individual i a value from a random variable X distributed uniformly between 0 and 1, and a propensity score $P(Z_i)$ predicting whether a person is an immigrant, using a vector Z, where $P(Z_i) = P(\text{Immigrant} = 1 \mid Z)$. Vector Z, here, yields probabilities of being an immigrant for employees in different categories of education, industry, and occupation, in workplaces varying in their geographical and social closeness to immigrants. $P(Z_i)$ is calculated nonparametrically using observed cell-frequencies of immigrants and natives for all combinations of Z. If $X_i \leq P(Z_i)$ for individual i s/he is assigned a counterfactual immigrant status, otherwise not. I call the segregation that results from this counterfactual D_Z. If $D > D_Z > D^*$, the fraction of systematic segregation attributable to observables is equal to $\frac{\frac{D_Z - D^*}{1-D^*}}{\frac{D-D^*}{1-D^*}}$, which simplifies to $\frac{D_Z - D^*}{D - D^*}$.

The simulation sketched out above only answer what would happen to segregation if a package of observables were accounted for, but the gross and net effect on segregation of each of these observables is unknown. To clarify the size of these effects, one may compare the values of D_Z estimated with different subsets of vector *Z* to the value of D_Z using the whole vector *Z*. The way these comparisons are used to compute gross and net effects is described in the Appendix.

Data and Variables

I estimate the influence of different kinds of factors on ethnic segregation across workplaces, using a dataset including every individual who worked or lived in the Stockholm's län during the period 2001 – 2003, and were between 18-64 years old in a given year. With these selection criteria, I effectively use (1) a sample of 80,151 workplaces that employed 913,795 individuals during the period, and (2) an overlapping sample of 1,286,959 individuals who lived in the focal area during the period. The second sample can be used to assess the relation between residential segregation and workplace segregation. All individuals and workplaces in Sweden have unique identification numbers that are used in various registers, which can be used to link individuals and workplaces across registers. The workplace affiliation of an employee is based on employers' (mandatory) reports of their employees' remuneration to the Swedish tax authorities each year. They then report all employees that have been employed in their workplaces during the preceding year.[1] The definition of a workplace is that used by Statistics Sweden; a physical place with one employer and one address. For around ten percent of the employees, work is carried out outside of this type of setting. Because no specific physical workplace could be determined for these cases, I excluded them from sample (1). Examples include home-helpers, street cleaners, salesmen, construction workers, substitutes, and employees who work at home. I define a person as employed in a workplace in a given year if s/he receives remuneration there during this year. I define a person to be an immigrant if s/he was not born in Sweden (i.e., s/he is a 'first generation immigrant').

At the individual level, I include the following variables in *Z*: educational field of highest education (using the Swedish Educational Terminology SUN-2000), occupation (based on Standard for Swedish Occupational Classification, SSYK-1996, a national adaptation of ISCO-88), and industry (based on the Swedish Industry Classification, SNI-92). At the workplace level, I include in *Z* the share of immigrants in the "geographical neighborhood" of the workplace, and the share of immigrants in the "social neighborhood" of workplace (both to be defined below).

For each individual, there is an associated probability of being an immigrant, given the vector of observed characteristics. For education, industry and occupation the procedure for constructing each Z_k was as follows: the share of immigrants in an as fine-

[1] Obviously, employers only report those employees that officially has been employed by them, excluding 'black' workers. The extent to which this is a problem for the present study is hard to tell. As long as employers pay taxes for *some* of the 'black' employees' wages, they are in the data, otherwise not. Non-observations as a consequence of this is likely to be larger for small service-oriented workplaces in the private sector.

grained category was computed (113 occupational, 342 educational and 773 industry categories). Thereafter, I constructed for each of these a set of five dummy variables, indicating which fifth in the cumulative distribution of the share of immigrants they belong. The reason for keeping the number of categories for each Z_k low at five is to minimize the risk of overfitting. Also, with each Z_k having an equal number of categories, they should have an equal "fighting chance" to explain observed segregation.[2]

As the boundaries of social as well as geographical neighborhoods are far from straightforward to define, the operationalization of these variables necessarily have an element of arbitrariness to them. As for physical neighborhoods, I use Statistics Sweden's detailed geographical SAMS-codes (Small Area Marketing Statistics). SAMS areas are defined by Statistics Sweden, in collaboration with Swedish municipalities, to designate relatively small and socially homogeneous neighborhoods. The number of SAMS areas in Stockholm's län is in this dataset 875, with an average population of 4,065. I define the geographical neighborhood of a workplace to be the total share of immigrants in the SAMS areas whose midpoints are within a circle with a radius of five kilometers from the workplace. From this quantity, I constructed five dummy variables indicating quintiles of the underlying share of immigrants in the neighbourhood of the workplaces. I compared this baseline definition to ones taking into account the population density and demographic composition of the focal area and neighboring

[2] As the number of combinations of Z increase, it gets possible to predict with a high degree of accuracy, or even determine, whether a person is immigrant. As with the case of overfitting, an overly complex regression model may fit perfectly to data, but the ability of the model to generalize beyond the data is thereby limited. The number of combinations of Z should therefore be kept at a reasonably low level.

areas as a decaying function of distance to the focal area. As there were no substantial differences in results using these definitions, I used the one described above.

The share of immigrants in a workplace's "social neighborhood" is operationalized giving equal weight to all employee-years that the employees currently employed there has spent at other workplaces during the years preceding the current one. In the data, it is possible to track individuals and their workplace affiliation retrospectively to 1990. To obtain a pure measure of the social surrounding, it should exclude the focal individual from the calculation of the share of immigrants in her previous workplaces. More formally, the share of immigrants in each employee-year, at another workplace, contributes to the workplace calculation of the mean of immigrants in the social neighbourhood of a workplace, as follows:

$$\frac{1}{it}\sum_i \sum_t \left[\frac{p_t n_t - imm_i}{n_t - 1} \right] \tag{2}$$

where i index individuals and t time (year in the period preceding the current year), p_t equals the share of immigrants, n_t equals the number of employees, and imm_i is an indicator with value 1 if individual i is an immigrant. As for the physical neighborhood variable, I constructed 5 dummy variables indicating quintiles of the underlying share of immigrants in the social neighbourhood of the workplaces. In total, there are 3125 (5^5) unique combinations of the variables included in vector Z.

Table 1 about here

Descriptives are reported in Table 1. Immigrants and natives are unequally distributed across almost all categories indicating occupation, industry, field of education. Further, the likelihood that an immigrant is employed in a workplace increase with the density of immigrants in the geographical and social neighborhood of a workplace. Note in particular the high degree of separation of natives and immigrants with regard to education and the social neighbourhoods of workplaces.

Findings

In Figure 1, I report the size distribution of the workplaces in the sample in two ways: one displaying the share of workplaces of varying size, and one displaying the share of employees in workplaces of varying size. The size distribution of workplaces is extremely skewed. One third of all workplaces just employ a single person, and the number of workplaces with more than 3,000 employees (≈ ln 8) is too small to be discernible in the figure. The share of employees distributed over different workplaces is more even, with a center of gravity on workplaces with about 50 employees (≈ ln 4), but also with substantial shares of employees in the ends of the distribution as well. As argued previously, when the size of units to which individuals is small, standard measures of segregation are upwardly biased. In the present case, a third of the workplaces in the sample have just one employee, so adjustment for this bias is motivated.

Figure 1 about here

As mentioned, when the size of units to which individuals are allocated vary, there is no simple expression for random segregation (D^*), which has to be simulated. I did this using the records of existing workplaces in a given year, among which existing employees were hypothetically redistributed randomly. Thereafter, I computed *D* for the observed distribution. I iterated this procedure 100 times for each year, and took the mean value of *D* for these simulations to equal D^*. The standard deviation for D^* in the simulations were at a maximum .001, suggesting the efficiency of the estimated values to be high. In Table 2, Actual segregation, random segregation, systematic segregation, and segregation conditional on *Z* (*D*, D^*, $\hat{D}$, and D_Z, respectively) for the years 2001 to 2003 are displayed.

Table 2 about here

That there is a high degree of stability over time is clear. Actual segregation is within the span .427 – .432 over the three observation years. The expected segregation under random allocation is even more stable, staying within the span .201 – .203 throughout the period. The size relation between actual segregation and random segregation tells us that a substantial part of the segregation across workplaces, about half of it, can be attributed to purely random allocation forces. Still, the level of systematic segregation, $\hat{D}$, indicates that the observed level of segregation is substantially higher than that expected under random allocation, and that about two thirds of actual segregation . Immigrants and native-borns are, in other words, systematically sorted, or systematically sort themselves, into different workplaces.

Segregation increases dramatically conditioning on observable characteristics, implying that most of systematic segregation can be explained using observable characteristics of individuals and workplaces. In the last row I report how large a proportion of systematic segregation that is "explained" with observables, which turns out to be very high around 95 percent for all years. That is, almost all of the *systematic* sorting by birthplace across workplaces can be explained by observable differences between individuals and workplaces.

The simulations reported above only answer what would happen to segregation if a package of observables were accounted for, but the gross and net effect on segregation of each of these observables is unknown. To clarify the size of these effects, I compared the values of D_Z simulated with different subsets of vector *Z* to the value of D_Z using the whole vector *Z*. In Table 3, I report gross and net effects of the observables on segregation (cf. the background section for definitions of these). Because there is a very stable pattern of effects over time, I restricted the analysis to segregation in 2002. Looking first at gross effects, we can see that the social neighborhood of a workplace appears to exert a very large effect on segregation. Including this factor alone in a simulation generating a counterfactual distribution of immigrants and natives over workplaces account for as much as 84 percent of systematic segregation across workplaces. The occupation and industry of an employee also have large gross effects on segregation, accounting for 48 percent and 63 percent of systematic segregation, respectively, whereas the effects of field of employee education and the geographical neighborhood of the workplace appear to be smaller. Turning to net effects, we can see that most of these turn out to be very close to zero, with one exception: the social neighborhood of the workplace. The predictive value of adding this factor to the equation outperforms that of all other factors. Adding this factor increase the explained proportion of systematic segregation by 18 percentage points, net of the other factors. All other factors each have net effects around 1 percent of systematic segregation. In other words, the social neighborhood measure outperforms each one of the other measures by a factor of 18. If we compare the combined effect of all other factors to that of the social neighborhood, it still outperforms them by a factor of three.

Table 3 about here

In a (fixed) set of workplaces there are *a priori* two routes by which systematic segregation across workplaces can come about. Employees may, either through the use of ethnicity directly or indirectly through variables correlated with ethnicity, be sorted by place of birth into different workplaces and employees may, through the same kind of sorting principles, be sorted by place of birth out of workplaces. To investigate which of these flows that seems to be responsible for producing segregation, I simulated a number of counterfactual flows of employees, and compared these to the actual flows that took place during the period of study.

There are two kinds of (potentially overlapping) employees in workplace j in 2002 that are relevant here. First, employees that were not employed in the workplace in the preceding year (2001), denoted by e_j (e for entrants). Second, employees that were not employed in the workplace in the year following the current one (2003), denoted by l_j (l for leavers).

The inflow simulation sets the probability of a particular employee, in the pool of all entrants between 2001 and 2002, to be entering a vacant position in workplace *j*, to be equal to $\frac{e_j}{\sum_j e_j}$, where the summation is over all workplaces. The outflow simulation sets the probability of any employee to leave workplace *j* between 2002 and 2003 to be equal to $\frac{l_j}{n_j}$, where n_j is the number of employees in the workplace in 2002. I compare the counterfactual segregation that results from the inflow simulation to that of actual segregation in 2002, and I compare the counterfactual segregation that results from the outflow simulation to that of actual segregation for the same year, but with actual leavers between 2002 and 2003 removed. The result of these hypothetical scenarios answers the question of which one of the inflow to, and outflow from workplaces, which tends to uphold segregation across workplaces. I calculated the unadjusted segregation as well as the systematic segregation that would obtain using these sorting rules, and report the results in Table 4.

Table 4 about here

The main message of these simulations is that segregation is generated by bias in the inflow of employees to workplaces, not the outflow of employees from workplaces. The scenario of incoming employees being randomly distributed over workplaces would decrease systematic segregation by about a third, from 0.285 to 0.191. In the scenario of employees leaving workplaces randomly, systematic segregation among the remaining employees would remain more or less unaltered, with just a marginal decrease from 0.282 to 0.279. That is, there is basically no systematic pattern in that underrepresented (or overrepresented) groups within workplaces disproportionately leave these, but there is a clear pattern of immigrants and natives being recruited to different workplaces. A conclusion that may be drawn with reasonable certitude is therefore that the inflow of employees to workplaces is *the* flow that generates ethnic segregation across workplaces.

A question that may be raised with regard to this finding is whether we can account for this pattern with the observed characteristics of individuals and workplaces, measured through the vector *Z*. The relevant simulation to answer this question would, for the case of employees entering workplaces, within cells defined by *Z*, randomly re-allocate entering employees to vacant positions in existing workplaces. That way, the distribution across workplaces of observed individual and workplace characteristics would be the same as that actually observed, but "identical" immigrant and native employees would counterfactually be randomly reshuffled between workplaces. The question this simulation answers is "How would segregation be affected if employers were blind to ethnicity, but chose their new employees according to *Z*?". Gross and net effects may be computed following the same logic as for the cross-sectional case. A gross effect on systematic segregation is the difference in systematic segregation between a random allocation of employees to vacant positions and an allocation using only $Z_{k=m}$. A net effect on systematic segregation is the difference in systematic segregation between an allocation using the whole vector *Z* to an allocation using $Z_{k \neq m}$. The maximum amount of explainable systematic segregation is in this case equal to actual systematic segregation in 2002 (.285), subtracted with the counterfactual segregation after random allocation of entrants (.191). In Table 5, I report the fractions of this explainable part that can be attributed to the observables included in *Z*. The results are straightforward to interpret. The social neighborhood of a workplace, as defined previously, appears to be the most important driving force behind ethnic bias in the recruitment of employees to jobs in workplaces. Although the gross effects of the other dimensions of segregation are large and positive, they are much lower than that of the social neighborhood. The net effect of the social neighborhood dwarfs those of the other variables by a factor of around 40.

Table 5 about here

Discussion

For the empirical case at hand, I have shown ethnic segregation across workplaces to be upheld through both random and systematic forces. If employees, in a counterfactual world, were randomly allocated to workplaces, about half the level of actual ethnic segregation across workplaces would still obtain. The remainder of segregation is systematically generated, because immigrants and natives are sorted *into* different workplaces, largely because immigrants and natives, through their work related networks, appear to be socially linked to different workplaces.

An important caveat is that the variables used to explain segregation may be endogenous to the outcome. Occupational and residential segregation may be caused by workplace segregation. Immigrants and natives may move to the area where they have a job, which means that workplace segregation may give rise to residential segregation. Importantly, including the share of immigrants in the employees' previous workplaces introduces workplace segregation at an earlier point in time as a statistical explanation of workplace segregation at the present time. The share of immigrants in the employees' previous workplaces may further correlate with unobserved industry or branch characteristics.

A fairly certain conclusion is however that ethnic segregation across these workplaces is generated through bias in the inflow of employees to workplaces, not through bias in the outflow of employees from workplaces. This is highly relevant to theories of homosociality applied to the labor market (e.g., Pfeffer 1985; Williams and O'Reilly 1998). If homosocial processes affect segregation, they appear to do it in the process of hiring employees to workplaces, but not through processes leading to between-group differences in quits and lay-offs. Once a person is employed in a workplace, these kinds of processes appear irrelevant, at least as far as the average ethnic bias in employees leaving workplaces is concerned. If they operate, counteracting processes apparently outweigh them.

Conclusion

The main conclusion of this study is that systematic ethnic segregation across workplaces is upheld not because immigrants and natives differ in the formal assets they bring to the labor market, but primarily because work-related networks appear to be segregated by country of birth. There appears to be nonnegligible degree of path dependency in the proportion of immigrants in the workplaces individual employees move between (for a similar result, see Hedström and Collett 2009) Thus, the process by which immigrants are matched to jobs, and – perhaps more importantly – the process by which immigrants are matched to their *first* jobs in the host country, appears to be important in order to understand how ethnic segregation across workplaces is generated and upheld.

Tables

Table 1. Descriptives of the explanatory variables (pooled 2001 – 2003).

	Immigrants		*Natives*	
	n	%	*n*	%
Occupation				
0.0-8.2 % immigrants	33,816	8.63	427,230	21.02
8.7-11.8 % immigrants	48,548	12.39	430,500	21.19
12.0-13.4 % immigrants	62,198	15.87	423,345	20.83
13.6-22.8 % immigrants	84,735	21.62	414,159	20.38
24.3-66.3 % immigrants	162,542	41.48	336,841	16.58
Industry				
0-9.2 % immigrants	35,179	8.98	444,993	21.90
9.2-11.7 % immigrants	47,726	12.18	438,933	21.60
11.7-16.4 % immigrants	63,493	16.20	405,955	19.98
16.4-22.1 % immigrants	89,541	22.85	394,753	19.43
22.1-100.0 % immigrants	155,900	39.79	347,441	17.10
Field of education				
0-8.2 % immigrants	28,390	7.25	446,158	21.96
8.2-10.8 % immigrants	43,737	11.16	446,402	21.97
10.8-14.7 % immigrants	58,878	15.03	433,076	21.31
15.1-23.4 % immigrants	70,173	17.91	310,984	15.30
23.4-100.0 % immigrants	190,661	48.66	395,455	19.46
Geographical neighborhood (workplace)				
3.5-14.8 % immigrants	54,205	13.83	400,067	19.69
14.8-15.7 % immigrants	70,775	18.06	414,722	20.41
15.7-18.2 % immigrants	80,348	20.51	428,911	21.11
18.2-25.9 % immigrants	88,736	22.65	399,810	19.67
25.9-60.4 % immigrants	97,775	24.95	388,565	19.12
Social neighborhood (workplace)				
0-9.3 % immigrants	28,711	7.33	443,331	21.82
9.3-12.3 % immigrants	42,216	10.77	450,911	22.19
12.3-15.9 % immigrants	58,472	14.92	431,939	21.26
15.9-25.5 % immigrants	85,230	21.75	397,905	19.58
25.5-100.0 % immigrants	177,210	45.23	307,989	15.16

Table 2. Observed, random, systematic and explained segregation across workplaces. Standard deviations of estimated values in brackets.

	2001	*2002*	*2003*
Actual segregation	0.427	0.429	0.432
Random segregation	0.203	0.201	0.203
Systematic segregation	0.280	0.285	0.287
Conditional segregation	0.405	0.410	0.410
"Explained" part of systematic segregation (%)	95.1	95.6	94.9

Table 3. Gross and net effects (fractions of the maximum positive effect = 1) of observable characteristics on systematic segregation, based on counterfactual cross-sectional employee allocations to workplaces in 2002.

	Gross effect	Net effect
1. Occupation of individual	0.481	0.010
2. Industry of individual	0.627	0.012
3. Field of education of individual	0.176	0.011
4. Geographical neighborhood of workplace	0.121	0.006
5. Social neighborhood of workplace	0.842	0.185
1 + 2 + 3	0.712	0.063
4 + 5	0.852	0.203

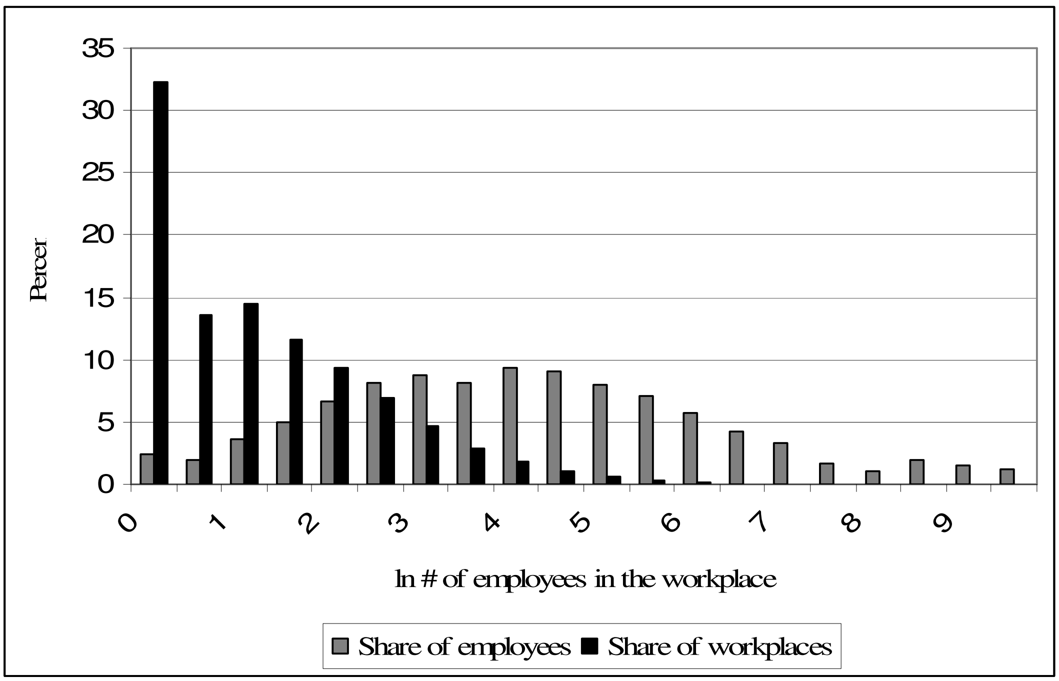
Table 4. Actual segregation (unadjusted and systematic), and counterfactual segregation (unadjusted and systematic) when random allocation rules/random removal rules have been used.

	Unadjusted segregation	Systematic segregation	Change in syst. seg. after the use of random rule
1. Segregation in 2002	0.429	0.285	
2. Segregation after removal of leavers (2002 – 2003)	0.437	0.282	
3. Counterfactual segregation after random allocation of entrants 2001 – 2002 (compare with 1)	0.375	0.191	-0.094
4. Counterfactual segregation after random removal of leavers 2002 – 2003 (compare with 2)	0.435	0.279	-0.003

Table 5. Gross and net effects of observable characteristics on segregation in the inflow of new employees to workplaces (fractions of the maximum positive effect = 1).

	Inflow 2001 – 2002	
	Gross effect	*Net effect*
1. Occupation of individual	.374	.004
2. Industry of individual	.456	.005
3. Field of education of individual	.194	.005
4. Geographical neighborhood of workplace	.083	.002
5. Social neighborhood of workplace	.749	.211
1 + 2 + 3	.573	.040
4 + 5	.752	.218

Figure 1. The share of employees across the size of workplaces. The share of workplaces across the size of workplaces.



Appendix: The computation of gross and net effects on segregation

Let $P(Z_{i,k=m})$ be the propensity score of individual i to be an immigrant given the value of $Z_{k=m}$, and let $P(Z_{i,k\neq m})$ be the propensity score of an individual to be an immigrant using all Z_k but $Z_{k=m}$ to predict it. Then the gross effect of $Z_{k=m}$ on systematic segregation may be defined to be the fraction of systematic segregation accounted for by allocating individuals conditional on $Z_{k=m}$ only: $\frac{E(D|P(Z_{k=m}))-D^*}{D-D^*}$ (if $D > E(D|P(Z_{k=m})) > D^*$). The net effect of $Z_{k=m}$ on systematic segregation may be defined to be the fraction of systematic segregation accounted for by including $Z_{k=m}$ in Z: $\frac{D_Z - E(D|P(Z_{k\neq m}))}{D-D^*}$ (if $D > E(D|P(Z_{k\neq m})) > D^*$).

References

Abrams, Dominic and Michael A. Hogg. 1990. "Social Identification, Self-Categorization and Social Influence." *European Review of Social Psychology* 1:195 - 228.

Bygren, Magnus and Johanna Kumlin. 2005. "Mechanisms of Organizational Sex Segregation: Organizational Characteristics and the Sex of Newly Recruited Employees." *Work and Occupations* 32:39-65.

Carrington, William J and Kenneth R Troske. 1997. "On Measuring Segregation in Samples with Small Units." *Journal of Business & Economic Statistics* 15:402-409.

—. 1998. "Sex Segregation in U.S. Manufacturing." *Industrial and Labor Relations Review* 51:445-464.

DiTomaso, Nancy, Corinne Post, and Rochelle Parks-Yancy. 2007. "Workforce Diversity and Inequality: Power, Status, and Numbers." *Annual Review of Sociology* 33:473-501.

Ekström, Erika. 2001. "Arbetsgivarnas rekryteringsbeteende." in *IFAU rapport 2001:3*: Institutet för arbetsmarknadspolitisk utvärdering (IFAU).

Ellis, Mark, Richard Wright, and Virginia Parks. 2004. "Work Together, Live Apart? Geographies of Racial and Ethnic Segregation at Home and at Work." *Annals of the Association of American Geographers* 94:620-637.

Granovetter, Mark S. 1974. *Getting a Job. A Study of Contacts and Careers*. Cambridge: Harvard University Press.

Hedström, Peter and Francois Collett. 2009. "Path-dependent paths: Tie-formation mechanisms and the evolution of interorganizational networks." Unpublished manuscript.

Marsden, Peter V. 1996. "The Staffing Process. Recruitment and Selection Methods." Pp. 133-156 in *Organizations in America*, edited by A. L. Kalleberg, D. Knoke, P. V. Marsden, and J. L. Spaeth. Thousand Oaks: Sage Publications.

Marsden, Peter V. and Elizabeth H. Gorman. 2001. "Social Networks, Job Changes, and Recruitment." Pp. 467-502 in *Sourcebook on Labor Markets. Evolving Structures and Processes*, edited by I. Berg and A. L. Kalleberg. New York: Kluwer Academic/Plenum Publishers.

Mencken, F. Carson and Idee Winfield. 1998. "In Search of the "Right Stuff:" The Advantages and Disadvantages of Informal and Formal Recruiting Practices in External Labor Markets." *American Journal of Economics and Sociology* 57:135-153.

Mittman, Brian S. 1992. "Theoretical and Methodological Issues in the Study of Organizational Demography and Demographic Change " Pp. 3-53 in *Research in the Sociology of Organizations*, vol. 10, edited by P. Tolbert and S. Bacharach. Greenwich, CT: JAI Press.

Mouw, Ted. 2003. "Social Capital and Finding a Job: Do Contacts Matter?" *American Sociological Review* 68:868-898.

OECD. 2006. "International Migration Outlook: Annual Report 2006 Edition.". Paris: OECD Publishing.
Pfeffer, Jeffrey. 1985. "Organizational Demography: Implications for Management." *California Management Review* 28:67-81.
Stinchcombe, Arthur L., Mary S. McDill, and Dollie R. Walker. 1968. "Demography of Organizations." *The American Journal of Sociology* 74:221-229.
Tsui, Anne S. and Barbara A. Gutek. 1999. *Demographic Differences in Organizations. Current Research and Future Directions.* . Lanham, MD: Lexington Books.
van Knippenberg, Daan and Michaela C. Schippers. 2007. "Work Group Diversity." *Annual Review of Psychology* 58:515-541.
Wegener, Bernd. 1991. "Job Mobility and Social Ties: Social Resources, Prior Job, and Status Attainment." *American Sociological Review* 56:60-71.
Williams, Katherine Y. and Charles A. O'Reilly. 1998. "Demography and Diversity in Organizations. A Review of 40 Years of Research." *Research in Organizational Behavior* 20:77-140.
Winship, Christopher. 1977. "A Revaluation of Indexes of Residential Segregation." *Social Forces* 55:1058-1066.
Åslund, Olof and Oskar Nordström Skans. 2009. "How to Measure Segregation Conditional on the Distribution of Covariates." *Journal of Population Economics* 22:971-981.

The Stockholm University Linnaeus Center for Integration Studies (SULCIS)

SULCIS is a multi-disciplinary research center focusing on migration and integration funded by a Linnaeus Grant from the Swedish Research Council (VR). SULCIS consists of affiliated researchers at the Department of Criminology, the Department of Economics, the Department of Human Geography, the Department of Sociology and the Swedish Institute for Social Research (SOFI). For more information, see our website: www.su.se/sulcis

SULCIS Working Paper Series

2007:1 Arai, M & Skogman Thoursie, P., "Giving Up Foreign Names: An empirical Examination of Surname Change and Earnings"

2007:2 Szulkin, R. & Jonsson, J.O., "Immigration, Ethnic Segregation and Educational Outcomes: A Multilevel Analysis of Swedish Comprehensive Schools"

2007:3 Nekby, L. & Özcan, G., "Do Domestic Educations Even Out the Playing Field? Ethnic Labor Market Gaps in Sweden"

2007:4 Nekby, L. & Rödin, M., "Acculturation Identity and Labor Market Outcomes"

2007:5 Lundborg, P., "Assimilation in Sweden: Wages, Employment and Work Income"

2007:6 Nekby, L., Rödin, M. & Özcan, G., "Acculturation Identity and Educational Attainmnet"

2007:7 Bursell, M., "What's in a name? A field experiment test for the existence of ethnic discrimination in the hiring process"

2007:8 Bygren, M. & Szulkin, R., "Ethnic Environment during Childhood and the Educational Attainment of Immigrant Children in Sweden"

2008:1 Hedberg, C., "Entrance, Exit and Exclusion: Labour Market Flows of Foreign Born Adults in Swedish "Divided Cities"

2008:2 Arai, M, Bursell, M. & Nekby, L. "Between Meritocracy and Ethnic Discrimination: The Gender Difference"

2008:3 Bunar, N., "Urban Schools in Sweden. Between Social Predicaments, the Power of Stigma and Relational Dilemmas"

2008:4 Larsen, B. and Waisman G., "Who is Hurt by Discrimination?"

2008:5 Waisman, G. and Larsen, B., "Do Attitudes Towards Immigrants Matter?"

2009:1 Arai, M., Karlsson, J. and Lundholm, M. "On Fragile Grounds: A replication of "Are Muslim immigrants different in terms of cultural integration?"

2009:2 Arai, M., Karlsson, J. and Lundholm, M. "On Fragile Grounds: A replication of "Are Muslim immigrants different in terms of cultural integration?" Technical documentation.

2009:3 Bunar, N. "Can Multicultural Urban Schools in Sweden Survive Freedom of Choice Policy?"

2009:4 Andersson Joona, P and Nekby, L. "TIPping the Scales towards Greater Employment Chances? Evaluation of a Trial Introduction Program (TIP) for Newly-Arrived Immigrants based on Random Program Assignment – Mid Program Results."

2009:5 Andersson Joona, P and Nekby, L. "TIPping the Scales towards Greater Employment Chances? Evaluation of a Trial Introduction Program (TIP) for Newly-Arrived Immigrants based on Random Program Assignment"

2009:6 Arai, M., Besancenot, D., Huynh, K. and Skalli, A., "Children's First Names and Immigration Background in France"

2009:7 Çelikaksoy, A., Nekby, L. and Rashid, S., "Assortative Mating by Ethnic Background and Education in Sweden: The Role of Parental Composition on Partner Choice"

2009:8 Hedberg, C., "Intersections of Immigrant Status and Gender in the Swedish Entrepreneurial Landscape"

2009:9 Hällsten, M and Szulkin, R., "Families, neighborhoods, and the future: The transition to adulthood of children of native and immigrant origin in Sweden.

2009:10 Cerna, L., "Changes in Swedish Labour Immigration Policy: A Slight Revolution?"

2009:11 Andersson Joona, P. and Wadensjö, E., "Being employed by a co-national:
A cul-de-sac or a short cut to the main road of the labour market?

2009:12 Bursell, M. "Surname change and destigmatization strategies among Middle Eastern immigrants in Sweden"

2010:1 Gerdes, C., "Does Immigration Induce 'Native Flight' from Public Schools?
Evidence from a large scale voucher program"

2010:2 Bygren, M., "Unpacking the Causes of Ethnic Segregation across Workplaces"