



DISCUSSION PAPER SERIES

IZA DP No. 4423

## Do as the Neighbors Do: The Impact of Social Networks on Immigrant Employment

Fredrik Andersson  
Simon Burgess  
Julia Lane

September 2009

Forschungsinstitut  
zur Zukunft der Arbeit  
Institute for the Study  
of Labor

# **Do as the Neighbors Do: The Impact of Social Networks on Immigrant Employment**

**Fredrik Andersson**

*Risk Analysis Division, OCC*

**Simon Burgess**

*CMPO, University of Bristol  
and IZA*

**Julia Lane**

*National Science Foundation  
and IZA*

Discussion Paper No. 4423  
September 2009

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **Do as the Neighbors Do: The Impact of Social Networks on Immigrant Employment\***

Substantial immigrant segregation in the United States, combined with the increase in the share of the U.S. foreign-born population, have led to great interest in the causes and consequences of immigrant concentration, including those related to the functioning of labor markets. This paper provides robust evidence that both the size and the quality of an immigrant enclave affects the labor market outcomes of new immigrants. We develop new measures of the quality, or information value, of immigrant networks by exploiting data based on worker earnings records matched to firm and Census information. We demonstrate the importance of immigrant employment links: network members are much more likely than other immigrants to be employed in the same firm as their geographic neighbors. Immigrants living with large numbers of employed neighbors are more likely to have jobs than immigrants in areas with fewer employed neighbors. The effects are quantitatively important and robust under alternative specifications. For example, in a high value network – one with an average employment rate in the 90th percentile – a one standard deviation increase in the log of the number of contacts in the network is associated with almost a 5% increase in the employment rate. Earnings, conditional on employment, increase by about 0.7%.

JEL Classification: J61, J20

Keywords: social networks, immigrant enclaves, labor market intermediaries

Corresponding author:

Julia Lane  
The National Science Foundation  
4201 Wilson Boulevard  
Arlington, Virginia 22230  
USA  
E-mail: [jlane@nsf.gov](mailto:jlane@nsf.gov)

---

\* Thanks to seminar participants at the University of Bristol, IZA, the Upjohn Institute, and the Society of Labor Economics meetings, for their helpful comments. This document reports the results of research and analysis undertaken by the United States Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This document is released to inform interested parties of research and to encourage discussion. This research is a part of the United States Census Bureau's Longitudinal Employer-Household Dynamics Program (LEHD), which is partially supported by the National Science Foundation Grants SES-9978093 and SES-0427889 to Cornell University (Cornell Institute for Social and Economic Research), the National Institute on Aging Grant R01~AG018854-02, and the Alfred P. Sloan Foundation. The views expressed on statistical issues are those of the author(s) and not necessarily those of the United States Census Bureau, its program sponsors or data providers. Some or all of the data used in this paper are confidential data from the LEHD Program. The United States Census Bureau supports external researchers' use of these data through the Research Data Centers (see [www.ces.census.gov](http://www.ces.census.gov)). For other questions regarding the data, contact the US Census Bureau.

## ***1. Introduction***

Seven in ten of all immigrants reside in five states: California, New York, Texas, Florida and Illinois, although only one in three of the native population does so. One in three of all immigrants reside in just three cities: Los Angeles, New York & Miami, compared with one in eight of the native born. The more detailed the geography, the more concentrated the story: the 2,500 most immigrant-dense block groups account for almost half of all immigrants and only one in eight of the native population.<sup>2</sup>

This substantial segregation, combined with the increase in the share of the U.S. foreign-born population, has led to an extensive and mixed literature examining the causes and consequences of immigrant concentration<sup>3</sup>, including those related to the functioning of labor markets. Some evidence suggests that immigrant enclaves contribute to the ghettoization of ethnic communities. Competing evidence suggests that enclaves act as labor market intermediaries providing pathways to jobs.

In this paper we examine multiple dimensions of enclaves – geography, ethnicity and quality – to tease out the effect of enclaves on immigrant labor market outcomes. We do this by developing and using new measures of the information value of immigrant social networks, namely both the number of contacts and the quality of those contacts. The results are consistent with the view that new immigrants' success or failure in the labor market is influenced by the characteristics of the social networks in their local neighborhood. Immigrants living with large numbers of employed neighbors are more likely to have jobs than immigrants in areas with fewer employed neighbors. The effects are quantitatively important and robust

---

<sup>2</sup> Statistics based on authors' own calculations using Decennial data.

<sup>3</sup> We discuss those in some detail in next section.

under alternative specifications. For example, in a high value network – one with an average employment rate in the 90<sup>th</sup> percentile – a one standard deviation increase in the log of the number of contacts in the network is associated with almost a 5% increase in the employment rate. Earnings, conditional on employment, increase by about 0.7%.

We exploit the power of our data to examine the way that such gains might arise. Social networks provide job contacts and we examine the likelihood of network members working together in the same form, relative to others in the same neighborhood. We find strong evidence supporting this idea: immigrants who are members of ethnic enclaves are much more likely to be employed with neighbors than other immigrants. This gives some credence to the idea that ethnic enclaves are sustained because they act successfully as local labor market intermediaries.

## ***2. Background***

### ***Social Networks***

The burgeoning literature on social networks as drivers of both behavior and outcomes sheds a useful light on why enclaves might form (Jackson 2007). Briefly, networks have productive value, and the costs and benefits of belonging to a network depend on the network structure.

A recent paper (Calvo-Armengol and Jackson 2004) sketched out how social networks could influence employment outcomes: the more employed contacts an individual has, the more likely it is that the individual will learn about new job openings. This framework is important because it suggests that both the quantity (the number of contacts) and the quality (the employment status of those contacts) of networks are important in terms of getting jobs.

The Calvo-Armengol/Jackson framework has been extended by recent work that shows analytically how workers who are better socially connected are less likely to be unemployed and more likely to have jobs with higher earnings (Ioannides and Datcher-Loury 2004; Zenou 2007)

The formal description is consistent with a plethora of empirical literature documenting that social networks provide information about job openings (see, for example, (Rees 1966) (Granovetter 1995). There is abundant evidence that individuals use friends and relatives to find jobs – indeed personal ties are the source of between one in three and two in three job matches for low-wage workers (Ioannides and Datcher-Loury 2004). The social network evidence suggests that the use of informal contacts varies by age, race, and ethnicity as well as by location: those in high poverty neighborhoods and in large cities are substantially more likely to use informal networks. Finally, job search aided by personal ties is particularly productive for less well-educated workers in high poverty neighborhoods (Elliott 1999), a finding we echo below.

A better understanding of the geographic dimension of social networks has been fostered by the new availability of very fine grained micro-data. New evidence shows that individuals who reside on the same block are more than 33 percent more likely to work together than people living on nearby blocks. Intriguingly, the effect is stronger when the individuals have similar sociodemographic characteristics and when at least one individual is well attached to the labor market (Bayer, Ross, and Topa 2008)

### ***The Formation of Ethnic Enclaves and their Consequences***

The reasons for the formation of ethnic enclaves have been well studied. Some

researchers have focused on the role of communication and language ((Chiswick 2002; Chiswick and Miller 1995; Bauer, Epstein, and Gang 2002). Others, consistent with the notion that social networks can act as labor market intermediaries for immigrants, have noted that enclaves may help create networks that provide access to employment opportunities and transportation (Cutler, Glaeser, and Vigdor 2008; Patel and Vella 2007) An analysis of Danish data (Damm 2006) finds that refugees with unfavorable characteristics self-select into ethnic enclaves.

The evidence on the consequences of enclaves is mixed. In terms of social outcomes, a number of authors have noted the likelihood of slower rates of cultural and economic assimilation (Lazear 1999), the promotion of a culture of poverty (Borjas 2006) as well as isolation and lock-in effects (Cutler and Glaeser 1997). There is also recent evidence that social networks form an important way in which immigrants select occupations (Patel and Vella 2007). In terms of employment and earnings outcomes, the preponderance of studies suggest that networks have a positive effect (see (Patel and Vella 2007) for a good review), and that network size is particularly important (Damm 2006). However, another recent study argues that the impact varies case by case – that the effect of enclaves depends on the characteristics of the people in the enclave (Cutler, Glaeser, and Vigdor 2007), and that the gains from segregation are greater for highly skilled groups. This latter finding is consistent with earlier findings on Swedish immigration (Edin, Fredriksson, and Aslund 2004) and draws at least partially on the notion that enclaves enable immigrants to form social networks that effectively make them act as intermediaries in either getting jobs or getting welfare.

### ***3. Empirical Strategy***

Our major aim is to determine the impact of geographically and ethnically defined social networks on the labor market outcomes of immigrants. We also aim to shed light on their

formation by investigating the degree to which neighbors in networks work together. We begin by assuming that the model governing individual labor market outcomes,  $Y$ , can be described by an augmented Mincer equation of the following form:

$$Y_i = X_i\beta + Network_i\alpha + \varepsilon_i \quad (1)$$

where  $i$  indexes individuals,  $X$  is a set of standard controls for productivity related characteristics,  $Network$  measures the information value of the social network for individual  $i$ , and  $\varepsilon$  is an error term.

The empirical identification of  $\alpha$ , the primary parameter of interest, in equation (1) is complicated for a number of reasons. One is the lack of direct measures on actual contacts. There is also a set of problems associated with omitted variables. One is geographic in nature: geographic differences in labor demand will affect individual outcomes for all in a given neighborhood, independent of any network effects. Another is social: discrimination will similarly affect all individuals in a given group regardless of network effects. A third is the endogeneity of the mobility decisions: individuals are more likely to move to high-employment areas, and select neighbourhoods based on homophily. We address each of these in turn below.

### **Measuring the Information Value of Social Networks**

A social network is commonly defined as a web of interconnected people who directly or indirectly interact with each other. Because of the difficulty in obtaining direct measures on contacts, a key challenge in the literature is how to estimate an individual's social network. Much of the previous literature assumes networks are geographically based and estimates



versions of:

$$Y_i = X_i\beta + C_{ij}\bar{Y}_{j(i)}\alpha + \varepsilon_i \quad (2)$$

where  $j(i)$  denotes the neighborhood of individual  $i$ ,  $C$  is some function of the number of available contacts in the neighborhood, and  $\bar{Y}$  is the neighborhood mean of  $Y$ .<sup>4</sup> This approach implicitly assumes that contacts are randomly distributed within the neighborhood and that there are no unobservable neighborhood characteristics that are correlated with the network variable. Because individuals choose their contacts, conditional independence of contacts is a strong assumption, especially since “neighborhoods” often have been defined in terms of rather large and heterogeneous geographical areas, such as Metropolitan areas<sup>5</sup>. We use Census tracts as our preferred measure of locality as they have two key advantages<sup>6</sup>. First, they are much smaller geographical areas, typically a tract consists of between 1,500 and 8,000 individuals. Second, they are designed to be relatively homogeneous in terms of population characteristics, economic status, and living conditions. This makes the assumption of conditional independence of contacts more plausible.

Another strand of the literature emphasizes ethnicity in defining networks. New immigrants are not only more likely to have already established contacts with previous

---

<sup>4</sup> Jencks and Mayer (1990) present a thorough survey of this literature. Papers have estimated neighborhood effects for a variety of socioeconomic variables, including crime, drug use, sexual behavior and educational attainment. Most papers tend to find a strong correlation between individual and mean neighborhood outcomes.

<sup>5</sup> Data are typically not available at a high enough level of geographic precision to permit more granular analysis.

<sup>6</sup> This is very different from the block by block specification employed by Bayer et al., which is based on the assumption that interactions with neighbors are very local in nature. This is unlikely to be the case for immigrant clusters, who are bound by language, religion and shared interest.

immigrants from the same country, but there are also arguably strong incentives for new immigrant to make contacts with other expatriates. Previous immigrants are uniquely well-suited to transfer knowledge about the host country, since they share language and migration experience, as well as possess knowledge about institutional differences between the origin and host country. For this reason, in the immigration literature, networks have been determined by ethnic groups rather than neighborhoods (Borjas 1995).

We explicitly combine these approaches by defining networks by both ethnicity crossed with neighborhood. Thus we specify the pool of available contacts in terms of the number of individuals of the same origin (country of birth) within a Census Tract. This is similar in spirit to other work on networks and welfare effects (Bertrand, Luttmer, and Mullainathan 2000). In particular, we replace  $network_i$  in equation (1) with  $C_{jk} \bar{Y}_{jk(i)}$ , where  $j$  indexes Census tracts,  $k$  country of birth, and  $C_{jk}$  is the log of the pool of available contacts defined by individuals in the same Census track and of same origin. Finally, because we are interested in the impact of social networks on labor market outcomes, and the potential contacts with employers, we weight the number of contacts by whether or not they have jobs, namely,  $\bar{Y}_{jk(i)}$  which is the employment rate of that group. This yields an empirical specification of

$$Y_{ijk} = X_i \beta + C_{jk} \bar{Y}_{jk(i)} \alpha + \varepsilon_{ijk} \quad (3)$$

### **Extending the Specification**

We extend the specification to address the obvious concerns that would be raised with a direct estimation of (3). Two of the three types of omitted variables bias that could potentially affect the empirical approach – the geographic and ethnic biases mentioned above – can

be directly addressed by using fixed effects. In particular, we can exploit the variation along the geographic and ethnic dimensions of the enclave by controlling for both the census tract itself as well as the country of birth. The inclusion of census tract fixed effects enables us to control for areas with a strong or weak labor demand, since individual outcomes for all in that neighborhood, should be roughly equally affected. The inclusion of ethnic specific fixed effects enables us to control for the effect of discrimination against a certain immigrant group is subject to discrimination, since individual outcomes for all of that ethnicity should be roughly equally affected.

The third major potential bias is if omitted personal characteristics are correlated with the network measure particularly with the number of available contacts,  $C_{jk}$ . Lazear, in particular, has argued that the larger the size of the ethnic group, the less necessary it is for immigrants to assimilate (Lazear 1999). This suggests that immigrants who do not live in numerically large enclaves are likely to be more assimilated than those who do. We include the log of the pool of available contacts,  $C$ , directly as a regressor in the specification to control for this possibility. We return to the issue of endogeneity below.

The final empirical specification, then, is:

$$Y_{ijk} = X_i\beta + C_{jk}\bar{Y}_{jk(i)}\alpha + \gamma_j + \delta_k + C_{jk}\theta + \varepsilon_{ijk} \quad (4)$$

This specification addresses concerns about differences in labor demand by including area fixed effects ( $\gamma_j$ ), differences in employer treatment of ethnic groups by including ethnic fixed effects ( $\delta_k$ ), as well as correlation between unobserved personal characteristics and the network terms based on preferences to live among one's own group through  $C_{jk}$ .

One potential source of selection bias still remains. Including  $C_{jk}$  only controls for self-selection to the extent that correlation between the network variable and  $C_{jk}$  is fixed across immigrant groups. If that is not the case, self-selection that is differential across immigrant groups could still contaminate the results. We follow the technique used by Bertrand et al (2000) to address this type of selection bias by instrumenting  $C_{jk}$  with the number of individuals from country  $k$  who reside in the metropolitan area. It is chosen as an instrument based on the idea that it is correlated with the scope for co-location of members from the same immigrant groups. While we cannot exclude the possibility of self-selection across MSAs it seems plausible that the scope for self-selection is greater within than across MSAs. We discuss this further in section 5.

#### ***4. Data***

Our data provide a unique opportunity to examine the role of immigrant networks in the labor market. We briefly describe the general characteristics of the dataset, and the characteristics that enable us to explore very detailed levels of geography, to identify large numbers of recent immigrants, to identify the labor market outcomes of these immigrants and to measure the quantity and quality of the potential enclave network.

The new database that enables us to match workers with past and present employers has been assembled at the Longitudinal Employer-Household Dynamics Program at the U.S. Census Bureau (Abowd, Haltiwanger, and Lane 2004). This database consists of quarterly records of the employment and earnings of almost all individuals from the unemployment insurance (UI) systems of a number of US states in the 1990s – these provide the key link between workers and firms. This dataset has been extensively described elsewhere

(Haltiwanger, Lane, and Spletzer 2007), but it is worth noting a number of advantages. Earnings are typically accurately reported, since there are financial penalties for misreporting. The data are current, and the dataset is extremely large. The UI records have also been matched to internal administrative records at the Census Bureau that contain information on date of birth, place of birth, race, and sex for all workers, thus providing limited demographic information. One limitation of the data is that there are no direct data links between workers and establishments, but only between workers and firms. Thus, for about 30% of the workforce – whose employing firm consists of more than one establishment -- we cannot tell with certainty in which particular establishment a worker is employed, if the employing firm consists of more than one establishment. Thus, probabilistic links are used to impute a place of work for workers who work for multi-unit businesses.<sup>7</sup>

The information that permits the identification of the place of residence of each worker is derived from an extract from the Census Bureau’s Statistical Administrative Records System (StARS), which incorporates data from the Internal Revenue Service (IRS) 1040 and 1099, Housing and Urban Development (HUD), Medicare, Indian Health Service, Selective Service System to derive a “best” residential address. The LEHD program uses geography coding software to map addresses into latitude and longitude coordinates. With these coordinates, locations are assigned to higher-level geographies, such as block group and tract. Well over 90% of the records in the relevant years have geocodes of high quality at the block level or finer.

---

<sup>7</sup> The probabilistic links are derived from a statistical model that conditions on distance between candidate units and place of residence as well as longitudinal restrictions based on the demographics of the firm (Freedman, Lane, and Roemer 2008; Abowd, Stephens, and Vilhuber 2008).

Our ability to identify large groups of immigrants is driven by the existence of universal administrative records systems at the Census Bureau which are used to provide intercensal population estimates. These data are anonymized and then matched to information from the Social Security Administration, under strict privacy and security protocols, to provide country of birth information. Although estimates are that about one third of immigrants into the US are illegal (Passel and Suro 2005), many of those get a taxpayer identification number (or TIN), since the penalties for tax evasion are very severe, and the Internal Revenue Service does not share data with the Immigration and Naturalization Service.<sup>8</sup> The data are also matched to the long form of the 2000 Decennial Census to derive information on education. Labor market outcomes, as well as the information about the firms for which immigrants work can be determined by matching the individual information to the universe unemployment insurance data

From this very large scale dataset, we follow a subset of immigrants who arrived to the country between 1997 and 1998 and located in one of 12 large metropolitan areas (with a combined population of more than 43 million).<sup>9</sup> We choose immigrants from the 31 largest immigrant groups, with at least one thousand immigrants in their group (defined by country of birth). This gives us a sample of approximately 171,000 recent immigrants, of whom 67,000 are

---

<sup>8</sup> While an illegal immigrant with a TIN will generally be included in our data we are not able to provide estimates of what fraction of illegal immigrants are included in our sample, because we have no independent information available on legal status nor do we have available reliable estimates of the number of immigrant workers who are illegal that are comparable to the estimates of immigrant workers in the LEHD database.

<sup>9</sup> The MSAs included in the sample are *Austin, Chicago, Dallas, Daytona Beach, Houston, Los Angeles, Miami, Orange County, Philadelphia, Pittsburg, Riverside, and Ventura.*

Mexicans. Educational information from the 2000 Decennial Census is available for 18,500, little more than 1 out of 10 observations in this sample.<sup>10</sup>

Employment status and earnings reflect reported total annual UI earnings in 1999. We consider an individual to be employed if annual earnings exceed \$5,000, which avoids counting as employed individuals with very limited attachment to the labor market. In addition, in order to avoid spurious correlation between individual outcomes and the network measure, the calculation of the network is based on the stock of immigrants in the area who arrived prior to 1997 and their labor market status in 1999.

Table 1 shows the distribution of the 31 largest recent immigrant groups in our sample across country of origin. Not surprisingly, Mexican immigrants dominate the sample accounting for almost 40% of the total sample size of 171,000 immigrants in the sample, although the immigrants come from all over the world. The distribution of immigrants across country of origin in the sub sample of approximately 18,500 individuals that match the 2000 Decennial Census (not reported) is comparable to the distribution presented in Table 1.

Table 2a compares the basic characteristics of the recent immigrants with the native population. Recent immigrants are disproportionately young (58% are between the ages of 25-34 while only 28% of natives fall in the same age bracket), and have low education (42% do not have a high-school diploma as compared to 8% among natives). However, conditional on having at least high-school diplomas recent immigrants tend also to have college degrees to a larger

---

<sup>10</sup> The expected match rate to the Decennial is about 1 in 6. However, the match rate for certain subpopulations, including immigrants, is known to be lower.

extent than natives. Given that the sample consists of recent immigrants and given the varying reasons for emigrating to the U.S. these results are not surprising.

Table 2b shows the importance of immigrant concentration. As compared to the native population, very large fractions of the neighbors (defined by other individuals residing in the same Census Tracts) of recent immigrants are also immigrants (36% vs 19% for natives). Even more striking is the finding that so many immigrants in Census Tracts have neighbors from the same country of origin -- well beyond what would be predicted based on the overall composition in the MSA. Neighbors of recent immigrants are 555% more likely to be born in the same country as the recent immigrant as compared to if they were randomly dispersed across Census Tracts within the MSAs. By contrast, neighbors of native born residents are only 8% more likely to be born in the U.S. than if native born residents were randomly dispersed across Census Tracts within the MSAs.

Finally Table 2a provides basic information on labor market outcomes for recent immigrants as compared to natives. Recent immigrants are less likely to be employed (an employment rate of 63% vs. 74% for natives and, conditional on employment, tend to have much lower average earnings (\$24,000 as compared to \$44,000). The difference in earnings is greater further up in the wage distribution.

## ***5. Difference-in-difference Results***

Before turning to the results from estimating equation (4), it is illustrative to present simple difference-in-difference calculations on the importance of immigrant enclaves. The problem with a single difference is of course that we cannot distinguish enclave effects from other area-specific effects, i.e. the enclave population might perform better not because they are in an



enclave, but because the enclaves happen to be in areas where everyone performs better.

We classify each neighborhood (Census tract) as an immigrant enclave for a particular immigrant group if the actual concentration of immigrants from that country exceeds what would be expected based on the group's overall share in the metropolitan area by a factor of 2. Based on such classification we can then split the sample into 4 groups, depending on whether the individual is a member of the immigrant group that defines enclaves or not and whether the individual resides in an enclave or not. For example, assume the overall share of the population born in Mexico is 20% in a particular Metropolitan area then all Census Tracts with a share of Mexicans that exceeds 40% would be classified as Mexican enclaves. The first difference amounts to comparing outcomes for Mexicans who reside in Mexican Enclaves with those Mexicans who do not. To control for common area effects, we subtract from this difference a second difference comparing outcomes for non-Mexican Immigrants who do and do not reside in Mexican Enclaves.

Table 3 performs the difference-in-difference composition for Mexican Immigrants. The results suggest that there are positive returns associated with enclaves. Taken at their face value, the enclave effects are such that the employment rate is 0.39% higher and average earnings almost \$10,000. In particular, the earnings results suggest that recent Mexican immigrants residing in an enclave have slightly lower earnings than those that do not. Thus, the positive estimate is the result of a much higher earnings penalty for recent non-Mexican immigrants residing in Mexican enclaves.

There are several reasons why we need to interpret these basic difference-in-difference results on the labor market impact of enclaves with a great deal of caution. First, Table 3

estimates the effects for Mexican immigrants only and it is quite possible that the effects are different for other immigrants groups' respective enclaves (see Cutler, Glaeser, and Vigdor 2008). Indeed, when we estimate the impact of Chinese enclaves (not reported; available from the authors) we find fairly large negative earnings effect associated with enclaves. More importantly, the results do not capture differences in underlying productive characteristics between the four groups that define the composition. Table 4 makes it clear that there are notable differences in the basic demographic characteristics of the four groups. For instance with respect to education the level of education is lower among those Mexican immigrants that reside in an enclave, presumably contributing to why recent Mexican immigrants have lower average earnings in enclaves. The discrepancy in education with respect to enclave status is even large for non-Mexican immigrants, presumably contributing to why the earnings penalty is greater for this groups and to why there is an overall positive earnings premium associated with enclaves for their members. Thus we view these results as descriptive rather than causal in nature.

## ***6. Main Results***

Table 5 summarizes the estimation results of the specification described in equation (4) for our sample of individuals who immigrated between 1997 and 1998. The two outcome variables of interest are the individual's employment status in 1999, and, for the employed population only, the log of annual earnings in 1999. We use two basic estimation techniques: ordinary least squares and an instrumental variables approach to address the self-selection issue.

The outcomes are estimated as a linear function of the network variable, which is measured as the log of the number of individuals from the same country of origin in the Census tract who arrived to the country before 1997 (contacts in the network) interacted with their

employment rate in 1999 (mean labor market value of network); the log of the number of contacts in the network; fixed effects for country of birth and for Census Tract (not displayed); and demographic controls, including gender, age and education.

The results in Table 5 confirm the importance of network effects for recent immigrants: the coefficients associated with the network variable are positive and highly significant in all specifications, suggesting that recent immigrants who are members of networks that are well established in the labor market enjoy higher employment rates and, if employed, higher earnings. The results also confirm that the coefficients associated with the demographic controls display the expected signs.

The coefficient associated with the log of available contacts in the network enters the specifications either as insignificant or, in the case of the IV results for earnings in Column (4), with a negative sign at the 5% significance level. This finding can be interpreted two ways: it could simply confirm that the quality dimension of the network is important – e.g. a large number of contacts in the network could be detrimental if the average labor market information value the network carries is low. It could also be interpreted as supporting the hypothesis that any positive effects associated with a large number of contacts by itself, without adjusting for quality of contacts, is offset by the effects of self-selection away from one's peers among individuals with favorable unobservable productivity related characteristics.

Turning specifically to the IV results, the first stage results are strong. The coefficient on the log of the size of the immigrant group in the MSA is 0.46 and highly significant when regressed on the log size of the immigrant group in the tract. The variation across MSAs accounts for 63% of the variation across tracts. The small difference between the OLS estimates

of the network coefficients in columns (1) and (3), on the one hand, and the IV estimates in columns (2) and (4), on the other hand, suggests that self-selection does not account for the estimated network effects. Since the IV estimates are subject only to potential selection bias across MSAs while the OLS estimates are subject to potential selection biases within and across MSAs it is possible to calculate the relative importance of selection across and within MSAs under the null-hypothesis that selection accounts for the results. If endogenous location decisions account for the results, the true network effect would be zero,  $\alpha_{TRUE} = 0$ . As a result, the bias in our estimated results could be derived by comparing the coefficient from the regression of the MSA results ( $\rho_{MSA}$ ) and from the regression of the tract level results ( $\rho_{TRACT}$ ) to derive the bias from our OLS and IV results as follows:

$$\left. \begin{aligned} \hat{\alpha}_{OLS} &= \hat{\rho}_{MSA} + \hat{\rho}_{TRACT} \\ \hat{\alpha}_{IV} &= \left( \frac{1}{R_{IV}^2} \right) \hat{\rho}_{MSA} \end{aligned} \right\} \Rightarrow \frac{\hat{\rho}_{MSA}}{\hat{\rho}_{TRACT}} = \frac{\hat{\alpha}_{OLS} - \hat{\alpha}_{IV} R_{IV}^2}{\hat{\alpha}_{IV} R_{IV}^2}$$

where  $R_{IV}^2$  is the R-square from the first stage in the IV-estimation. Because it is presumably easier to move within than across MSAs, we would expect the ratio to exceed 1 if the hypothesis that  $\alpha_{TRUE} = 0$  is correct. However, the ratio is 0.796 for the employment equation and 0.810 for the earnings equation. This implies that it is unlikely that the results could be driven completely by endogenous location decisions.

The effects of networks are quantitatively important. Since the coefficient for the network variable is hard to interpret, in the last row of the table we compute the expected increase in the employment rate and earnings resulting from an expansion of the number of contacts in the network. In a high value network – one with an average employment rate in the

90<sup>th</sup> percentile -- a one standard deviation increase in the log of number of contacts in the network is associated with almost a 5% increase in the employment rate. Earnings, conditional on employment, increase by about 0.7%. By contrast, an individual who lives in a low value neighbourhood who increases his or her contacts by the same amount actually reduces employment probabilities and earnings levels<sup>11</sup>.

In summary, our results demonstrate the importance of network effects in the sense that recent immigrants enjoy a higher degree of success in the labor market if they are part of social networks that are well established in the labor market. Before we turn to analyse possible mechanisms underlying the network effects, we first confirm the robustness of these basic results by investigating the sensitivity of the findings to functional form and sample restrictions.

### **Robustness Checks**

We examine the sensitivity of the results to the geographic unit of analysis. In the first panel of Table 6 we report the network coefficients when we define neighborhoods using Census blocks, which are much smaller geographic entities than Census tracts. As noted earlier, these are likely to be too small to capture the full reach of social networks for immigrants. However, we still find a significant impact of networks on both employment probabilities and

---

<sup>11</sup> It is worth noting that these estimates are likely to be quite conservative and result in an understatement of network effects, because of the many controls included correcting for potential sources of omitted variable biases that may serve as proxy for network effects in their own right. For instance, the area and the country-of-birth specific fixed effects are likely to pick up some of the true network effects if the reach of networks goes beyond country of birth or neighborhoods defined by Census tracts.

earnings levels, although the estimated impact is about one third lower than using a Census tract unit of geography.

We also examined the sensitivity of the results to a different specification of the functional form. Our basic enclave measure is a continuous log linear function of the size of the ethnic population in the Census Tract or the block. We also defined enclaves as a dichotomous variable based on whether the enclave population is at least 200% larger than expected. The results are reported in the second panel of Table 6. Although the coefficients are not directly comparable to our basic results, the signs and significances support the importance of network effects.<sup>12</sup>

The importance of network effects could arguably differ along other dimensions. In particular, it has been argued in the literature that social networks as defined by neighborhoods are relatively more important for certain demographic groups. In Table 7, we report results excluding the numerically dominant Mexican immigrants; we also exclude individuals living in the very large Los Angeles MSA. It has also been argued that network effects are more important for women than for men, so we also estimate the same specification excluding women. Finally, since another strand of the literature suggests that networks are more important for less educated than more educated individuals, we estimate the regression excluding the highly educated.

In each case, the results confirm the importance of network effects and that network effects are broad, but we also note interesting patterns. Network effects seem to be particularly

---

<sup>12</sup> Again, we compute the expected increase in the employment rate and earnings resulting from a change in enclave status: for regions with employment rates in the 90<sup>th</sup> percentile enclave status is associated with a 4.79% increase in the employment rate and increases in earnings, conditional on employment, of 0.73%. The detrimental effects of enclave status for regions with employment rates in the 10<sup>th</sup> percentile is of the same order of magnitude.

important for both Mexicans and individuals in the Los Angeles MSA. That the network effects seem to be quantitatively more important for immigrants who are women and who have low education levels are findings that are consistent with the idea of more limited mobility opportunities among former group and the reliance of informal job search channels among the latter.

## ***7. The Link between Networks and Jobs***

The work of Bayer et al. has provided evidence that geographic residential proximity influences employment and earnings outcomes. Our focus is similar: we can directly examine whether recent immigrants get jobs in the same firms as their neighbors. We do this in two ways.

First, we examine whether people in the neighborhood work for the same employer. We create a binary variable that indicates whether or not an individual works for the same employer as someone else in the enclave. We also compute the proportion of other enclave members working for the same employer. The first column of Table 8 provides evidence that new immigrants who locate in an enclave are much more likely than those not in the enclave to work in a firm that has at least one other enclave member in it. This holds true for the full sample, but is more important for women than for men, and more important for the less educated than the educated. The second column is consistent with the first: not only are new immigrants in an enclave more likely to work in a firm with other enclave members in it, but they work in firms with a high proportion of other enclave members. The same relative pattern for women and men, and for low and high educated workers, is evident as in the first column.

Second, we examine whether people who work for the same employer come from the same

neighborhood. The two dependent variable is an indicator variable equal to one if the individual has a coworker who resides in the same tract. Coworkers include everyone in the same firm and the sample is as before, recent immigrants with a minimum sample size of 1000. The first column in Table 9 reports the results controlling for enclave status and enclave quality. Enclave status takes the value 1 if the actual exposure to same country of birth individuals in the tract exceeds the expected exposure (based on MSA composition) by 200%. This is interacted with indicator variables for the labor market quality (employment rate) of the network, split into the lower quartile, middle half and upper quartile. In the regressions no enclave is the omitted category. The additional columns add controls: Column 2 adds controls for log firm size and log tract population, Column 3 adds demographic controls.

The results strongly support our hypothesis, i.e. enclave members have a higher probability of working with neighbors even after including controls. The strength of the enclave variable is increasing in its quality: networks with more employed members are more likely to find jobs with them for newly arrived immigrants. Interestingly, the effects are larger for low-educated workers, though there is no additional effect in the female sample.

## ***8. Conclusion***

This paper examines the effect of social networks on the labor market outcomes of recent immigrants. We use new measures of the quality of immigrant networks, combining administrative data, survey data and spatial data. The network measure combines both the size of the network in terms of the number of neighborhood\*ethnic group contacts, and the quality of the network in terms of the contacts' average employment status. We show that it is the quality as well as the size of the social network that determines the impact on labor market outcomes. The estimated effects are quantitatively significant and are robust to a variety



of possible misspecifications. The results are particularly strong for Mexicans, women and less educated workers.

We take the analysis step further and exploit the very detailed spatial data and the matched worker and firm information available in the LEHD data. We show that one channel through which network membership improves employment outcomes is a very direct one. Enclave members are disproportionately more likely to work with other enclave members than are other neighbors, suggesting a fairly direct role for the network as a labor market intermediary.

These results contribute to a number of related literatures. First, the results provide new evidence on the role of ethnic enclaves, showing that they can have a positive role improving employment and earnings for new immigrants who are enclave members. Of course, further research is warranted. Our analysis quantifies the way in which enclaves provide immigrants with pathways into jobs and higher earnings, at least in the relatively near term. Since other literature has suggested that enclaves delay eventual assimilation into the host society, future research using the same data should examine the labor market outcomes of these immigrants in a number of years time.

Secondly, the results contribute to the economic analysis of social networks (Jackson, 2007; Ioannides and Datcher-Lourey, 2004)). We have shown that network membership is productive for new immigrants. Thinking of these immigrant networks as labor market intermediaries opens lines of enquiry about the network itself, how it evolves over time, whether there are different 'business models' for networks. For example, some networks may help initially place new arrivals, but with success they move on both economically and

geographically; others may retain the individuals in the same area as they prosper. Of course, some individuals do not prosper in the network, though we have shown that on average they do, and again different network models may accommodate these individuals differently. In turn, the nature and evolution of different immigrant networks have important further implications for policy attitudes to enclaves. The data we have assembled for the present paper, plus the elapse of some time, will allow us to address these important questions.

## References

- Abowd, John, John Haltiwanger, and Julia Lane. 2004. Integrated Longitudinal Employee-Employer Data for the United States. *American Economic Review* 94:224-229.
- Abowd, John, Bryce Stephens, and Lars Vilhuber. 2008. The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators. In *Producer Dynamics: New Evidence from Micro Data*, edited by T. Dunne, J. B. Jensen and M. Roberts: University of Chicago Press.
- Bauer, Thomas, Gil Epstein, and Ira Gang. 2002. Enclaves, Language and the Location Choice of Migrants. In *IZA Discussion Paper 2002*. Bonn, Germany.
- Bayer, Patrick, Stephen L. Ross, and Giorgio Topa. 2008. Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy* 116 (6):1150-1196.
- Bertrand, Marianne, Erzo F. P. Luttmer, and Sendhil Mullainathan. 2000. Network Effects and Welfare Cultures. *Quarterly Journal of Economics* 115 (3):1019-1055.
- Borjas, George. 2006. Making it in America: Social Mobility in the Immigrant Population. In *NBER Working Paper 12088*.
- Borjas, George J. 1995. Ethnicity, Neighborhoods, and Human-Capital Externalities. *American Economic Review* 85 (3):365-90.
- Calvo-Armengol, Antoni, and Matthew O. Jackson. 2004. The Effects of Social Networks on Employment and Inequality. *American Economic Review* 94 (3):426-454.
- Chiswick, Barry. 2002. Do Enclaves Matter in Immigrant Adjustment?\*
- Chiswick, Barry, and Paul Miller. 1995. The Endogeneity between Language and Earnings: International Analyses. *Journal of Labor Economics* 13 (2):246-288.
- Cutler, David, and Edward Glaeser. 1997. Are Ghettos Good or Bad? *Quarterly Journal of Economics* 112:827-872.
- Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor. 2008. When are ghettos bad? Lessons from immigrant segregation in the United States. *Journal of Urban Economics* 63 (3):759-774.
- Damm, Anna Pil. 2006. Ethnic Enclaves and Immigrant Labour Market Outcomes: Quasi Experimental Evidence. In *University of Aarhus Department of Economics Working Paper*. Aarhus, Denmark.
- Edin, Per-Anders, Peter Fredriksson, and Olaf Aslund. 2004. Ethnic Enclaves and the Economic Success of Immigrants—Evidence From a Natural Experiment. *Quarterly Journal of Economics* 118 (1):329-357.
- Elliott, John. 1999. Social Isolation and Labor Market Isolation: Network and Neighborhood Effects on Less Educated Urban Workers,". *Sociological Quarterly*. 40 (2):199-216.
- Freedman, Matthew, Julia Lane, and Marc Roemer. 2008. New Approaches to Creating Data for Economic Geographers. *Journal of Official Statistics* 24 (1):133-156.

- Granovetter, Mark. 1995. *Getting a job: A study of contacts and careers*. Chicago: University of Chicago Press.
- Haltiwanger, John, Julia Lane, and James Spletzer. 2007. Wage, Productivity and the Dynamic Interaction of Businesses and Workers. *Labour Economics* 14 (3):575-602.
- Ioannides, Yannis, and Linda Datcher-Loury. 2004. Job Information Networks, Neighborhood Effects, and Inequality,. *Journal of Economic Literature* 42:1056-1093.
- Jackson, Matthew O. 2007. The Study of Social Network in Economics. In *Missing Links: Formation and Decay of Economic Networks* edited by J. E. Rauch. New York: Russell Sage Foundation.
- Lazear, Edward P. 1999. Culture and Language. *Journal of Political Economy* 107 (s6):S95-S126.
- Passel, Jeffrey, and Roberto Suro. 2005. Pew Hispanic Center.
- Patel, Krishna, and Francis Vella. 2007. Immigrant Networks and Their Implications for Occupational Choice and Wages. In *IZA Discussion Papers*. Bonn, Germany.
- Rees, Albert. 1966. Information Networks in Labor Markets. *American Economic Review* 56 (2):559-66.
- Zenou, Yves. 2007. Social Interactions and Labor Market Outcomes in Cities. In *CEPR Discussion Papers*.

**Table 1: Distribution of sample across country of origin**

<b>Country of Origin</b>	<b>Fraction of all immigrants in sample</b>
Mexico	38.6%
India	8.6%
Cuba	7.5%
China	5.4%
Philippines	4.7%
El Salvador	2.8%
Vietnam	2.6%
Canada	2.4%
United Kingdom	2.3%
Republic of Korea	1.9%
Guatemala	1.8%
Japan	1.8%
Poland	1.8%
Columbia	1.8%
Nicaragua	1.7%
Honduras	1.6%
Bosnia and Herzegovina	1.2%
Peru	1.2%
Taiwan	1.1%
Pakistan	1.1%
Haiti	1.1%
Iran	0.9%
Jamaica	0.9%
Brazil	0.7%
France	0.7%
Venezuela	0.7%
Romania	0.7%
Union of Soviet Socialist	0.6%
Ukraine	0.6%
Dominican Republic	0.6%
Bulgaria	0.6%
Total observations	171,671

**Table 2: Sample characteristics of sample vs. native-born population**

	<b>Recent immigrant sample</b>	<b>Native-born population</b>
Gender		
- Male	50.50%	49.14%
- Female	49.50%	50.86%
Age		
- 25-34	57.64%	27.78%
- 35-54	37.68%	57.10%
- 55-64	4.69%	15.12%
Education		
- Less than high school	41.63%	8.53%
- High school diploma	17.32%	24.18%
- Some college (no degree)	9.79%	25.35%
- College degree	31.26%	41.94%
<b>Labor market outcomes in 1999</b>		
- Employment rate	63.3%	74.0%
- Average earnings, conditional on employment	\$24,421	\$44,440
- 25 <sup>th</sup> percentile earnings, conditional on employment	\$10,174	\$18,413
- Median earnings, conditional on employment	\$15,466	\$31,947
- 75 <sup>th</sup> percentile earnings, conditional on employment	\$25,290	\$51,067
Total observations	171,671	13,076,404

**Table 2b: Neighborhood Composition**

<b>Sample</b>	<b>Fraction of Neighbors in Census Tract that are immigrants (std. dev)</b>	<b>Actual fraction of neighbors in Census Tract that are from same origin as individual (relative to expected fraction)</b>
Native-born population	19.28% (17.00%)	9% (23%)
All Recent Immigrant	37.42% (23.65%)	428% (1,295%)
All Recent Immigrants by country of birth		
- Mexico	39.91% (20.94%)	176% (262%)
- India	21.42% (15.79%)	385% (597%)
- Cuba	72.04% (17.67%)	102% (124%)
- China	34.68% (20.96%)	1,471% (2,848%)
- Philippines	33.11% (21.05%)	200% (278%)
All Recent Immigrants by education level		
- Less than high school	43.33% (22.09%)	338% (1,135%)
- High school diploma	41.03% (23.64%)	509% (1,774%)
- Some college (no degree)	36.19% (23.44%)	368% (1,290%)
- College degree	28.82% (22.73%)	446% (1,042%)

**Table 3: Difference-in-Difference Estimates of Labor Market Returns to Mexican Enclaves**

Recent Immigrant from Mexico?	In Mexican enclave in 1999?		Difference
	Yes	No	
Employment in 1999			
Yes	80.99%	80.00%	0.99%
No	79.45%	78.85%	0.60%
Difference	1.54%	1.15%	0.39%
Average Earnings in 1999 (employed)			
Yes	\$13,127	\$14,555	(\$1,428)
No	\$13,177	\$24,262	(\$11,085)
Difference	(\$50)	(\$9,707)	\$9,657
25th Percentile Earnings in 1999 (employed)			
Yes	\$5,824	\$6,038	(\$214)
No	\$5,769	\$6,313	(\$544)
Difference	\$55	(\$275)	\$330
Median Earnings in 1999 (employed)			
Yes	\$11,853	\$12,114	(\$261)
No	\$11,540	\$13,201	(\$1,661)
Difference	\$313	(\$1,087)	\$1,400
75th Percentile Earnings in 1999 (employed)			
Yes	\$18,125	\$19,110	(\$985)
No	\$17,522	\$25,810	(\$8,288)
Difference	\$603	(\$6,700)	\$7,303



**Table 4: Demographic characteristics of the four groups contributing to the difference-in-difference composition**

	Mexican Immigrants in Mexican Enclave		Non-Mexican Immigrants in Mexican Enclave	
	Yes	No	Yes	No
<b>Gender</b>				
Female	56.03%	53.26%	47.53%	46.08%
Male	43.97%	46.74%	52.47%	53.92%
<b>Age</b>				
25-34	65.96%	68.31%	53.15%	52.67%
35-54	29.90%	28.37%	41.58%	42.49%
55-64	4.15%	3.33%	5.27%	4.84%
<b>Education</b>				
Less than high	71.84%	64.05%	44.84%	19.95%
High school	17.06%	18.15%	21.78%	18.06%
Some college (no	6.65%	9.18%	11.03%	12.86%
College degree	4.44%	8.63%	22.34%	49.13%
<b>MSA</b>				
Austin	1.28%	1.27%	2.38%	2.03%
Chicago	27.66%	10.83%	39.29%	19.15%
Dallas	7.94%	5.60%	10.61%	7.09%
Daytona Beach	0.00%	0.03%	0.00%	0.25%
Fort Lauderdale	0.01%	0.50%	0.24%	6.39%
Houston	11.03%	11.53%	12.69%	10.15%
Los Angeles	30.66%	46.07%	18.84%	25.56%
Miami	0.21%	0.86%	1.59%	18.55%
Orange County	14.10%	8.93%	11.82%	6.05%
Pittsburg	0.00%	0.11%	0.00%	1.72%
Riverside	4.00%	11.89%	1.54%	2.31%
Ventura	3.09%	2.38%	1.00%	0.75%

**Table 5: Network effects**

Dependent Variable: Estimation Method:	Employed		ln(Earnings)	
	OLS	IV	OLS	IV
ln(POP <sub>k,j</sub> )*(EMP <sub>k,j</sub> -EMP)	0.064** (0.011)	0.054** (0.009)	0.135** (0.029)	0.113** (0.023)
ln(POP <sub>k,j</sub> )	0.002 (0.003)	-0.012 (0.007)	-0.010 (0.009)	-0.045* (0.017)
Age	0.017** (0.003)	0.017** (0.003)	0.065** (0.008)	0.064** (0.008)
age*age/100	- (0.004)	-0.024** (0.004)	- (0.010)	-0.090** (0.010)
Male	0.264** (0.007)	0.263** (0.007)	0.746** (0.018)	0.746** (0.018)
High-School Diploma	0.017 -0.010	0.016 -0.010	0.072** (0.026)	0.072** (0.026)
Some College	0.007 -0.013	0.006 -0.013	0.080* (0.033)	0.081* (0.033)
College Degree	0.047** (0.011)	0.046** (0.011)	0.418** (0.029)	0.419** (0.029)
Country of Birth Fixed Effects	Yes	Yes	Yes	Yes
Tract Fixed Effects	Yes	Yes	Yes	Yes
Observations	18,409	18,409	15,050	15,050
R-squared	0.10	0.10	0.19	0.19
Response to Shock   EMP <sub>k,j</sub> =P10	-5.09%	-4.29%	-0.73%	-0.61%
Response to Shock   EMP <sub>k,j</sub> =P90	4.84%	4.08%	0.70%	0.58%

Standard errors in parentheses; \* significant at 5%; \*\* significant at 1%

**Table 6: Network effects using alternative enclave metrics**

Dependent Variable: Estimation Method:	Employed		ln(Earnings)	
	OLS	IV	OLS	IV
ln(POP <sub>k,j</sub> )*(EMP <sub>k,j</sub> -EMP)	0.044** (0.011)	0.038** (0.008)	0.139** (0.028)	0.098** (0.021)
ENCLAVE <sub>jk</sub> *(EMP <sub>jk</sub> -EMP)	0.209** (0.051)	0.223** (.052)	0.498** (0.13)	0.339* (0.139)

Standard errors in parentheses; \* significant at 5%; \*\* significant at 1%

**Table 7: Network effects in alternative samples**

Observations deleted:				
None	Men	Highly educated	Mexicans	LA MSA
Employment equation				
0.064** (0.011)	0.054** (0.014)	0.050** (0.017)	0.040** (0.009)	0.050** (0.010)
Earnings equation				
0.135** (0.029)	0.138** (0.040)	0.132** (0.044)	0.068* (0.027)	0.109** (0.027)

Standard errors in parentheses; \* significant at 5%; \*\* significant at 1%

**Table 8: Do Enclave Members work Together?**

<b>Dependent Variable</b>	<b>Indicator of whether member of ethnic enclave works for same employer</b>	<b>Proportion of coworkers that are members of the enclave</b>
Sample		
Full	.083* (21.05)	.0095* (10.54)
Females	.096* (14.62)	.0099* (7.62)
Males	.071* (13.81)	.0091* (7.09)
High School and below	.099* (16.09)	.0136* (10.53)
More than High School	.076* (13.59)	.0069* (4.85)

Absolute value of t statistics in parentheses \* significant at 5%; \*\* significant at 1%

**Table 9: Do Co-Workers Live Near Each Other?**

**A: Sample = All**

<b>Dependent Variable</b>	<b>Indicator variable of whether any coworkers reside in same tract</b>	<b>Indicator variable of whether any coworkers reside in same tract</b>	<b>Indicator variable of whether any coworkers reside in same tract</b>
Enclave Q<p25	0.010* (2.22)	0.08** (20.81)	0.068** (4.95)
Enclave p25<=Q<p75	0.053** (15.35)	0.061** (20.42)	0.068** (7.27)
Enclave Q>=p75	0.160** (35.46)	0.115** (29.77)	0.119** (9.47)
Controls for log firm size and log tract population	No	Yes	Yes
Demographic controls	No	No	Yes
Observations	136,220	136,220	15,031
R-squared	0.01	0.27	0.28
Absolute value of t statistics in parentheses * significant at 5%; ** significant at 1%Demographic controls include gender, education, age, age squared and country of birth. The reduction in sample size when demographic controls are included is due to match to Decennial			

### B: Sample = Low Education Workers

Dependent Variable	Indicator variable of whether any coworkers reside in same tract	Indicator variable of whether any coworkers reside in same tract	Indicator variable of whether any coworkers reside in same tract
Enclave   Q < p25	0.011 (0.60)	0.077 (4.64)**	0.070** (3.57)
Enclave   p25 ≤ Q < p75	0.048** (3.69)	0.058** (4.99)	0.065** (5.48)
Enclave   Q ≥ p75	0.187** (9.41)	0.146** (8.39)	0.143** (7.90)
Controls for log firm size and log tract population	No	Yes	Yes
Demographic controls	No	No	Yes
Observations	8,827	8,827	8,827
R-squared	0.01	0.24	0.26

### C: Sample = Females

Dependent Variable	Indicator variable of whether any coworkers reside in same tract	Indicator variable of whether any coworkers reside in same tract	Indicator variable of whether any coworkers reside in same tract
Enclave   Q < p25	-0.003 (0.36)	0.081** (14.19)	0.063** (3.15)
Enclave   p25 ≤ Q < p75	0.058** (11.39)	0.068 (15.41)**	0.067** (4.91)
Enclave   Q ≥ p75	0.158** (23.28)	0.103** (17.67)	0.126** (6.85)
Controls for log firm size and log tract population	No	Yes	Yes
Demographic controls	No	No	Yes
Observations	61,607	61,607	6,959
R-squared	0.01	0.28	0.28