# Ethnic Networks and Employment Outcomes

Eleonora Patacchini* Yves Zenou†

October 22, 2009

**Abstract**

Using a panel of local authority-level data in England between 2003 and 2007 and spatial statistics techniques, we find that (*i*) the higher the percentage of a given ethnic group living nearby, the higher the employment rate of this ethnic group; (*ii*) this effect decays very rapidly with distance, losing significance beyond approximately 90 minutes travel time. These results are interpreted using the network model of Calvó-Armengol and Jackson (2004). Two results are put forward: (*i*) the individual probability of finding a job is increasing in the number of strong and weak ties; (*ii*) the longer the length of ties, the lower is this effect. Such predictions are applied to our analysis by approximating the social space by the geographical space. Ethnicity is the chosen dimension along which agents' social contacts develop and, as a result, we use ethnic population density to capture social interactions within the given ethnic group.

**Key words:** Ethnic minorities, population density, social interactions,weak and strong ties, panel data, spatial correlation.

**JEL Classification:** A14, C21, C33, J15, R12, R23.

*Università di Roma "La Sapienza", Facoltà di Scienze Statistiche, P.le Aldo Moro, 5, 00185 Rome, Italy. E-mail: Eleonora.Patacchini@uniroma1.it

†Stockholm University, Department of Economics, 106 91 Stockholm, Sweden, IFN, and GAINS. E-mail: yves.zenou@ne.su.se.

# 1 Introduction

The increase in the number of individuals from different ethnic groups is one of the most significant changes in Europe (but also in the United States) during the last decades. For example, according to the Census, Britain's population grew by 4 percent in the 1990s but 73 percent of this growth was due to ethnic minority groups, which grew by about 1.6 million individuals compared to 600,000 in the white population (Lupton and Power, 2004). While the diversity of social groups can be considered as a source of benefits, the persistence of ethnic minority' identities is, however, often perceived as a threat or source of frictions by natives. For example, in november 2007, Morrissey, the former Smiths singer, has sparked controversy by claiming British identity has disappeared because the country has been "flooded" by immigrants.

To better understand these issues, it is of paramount importance to analyze the socio-economic determinants and outcomes of the spatial distribution of ethnic minorities. The existing studies on such issues are, however, still scarce and unsatisfactory. In particular, the study of the relationship between a community's ethnic population density and its employment rate, which is in general found to be negative, is typically plagued by different econometric problems that renders the results inconclusive and difficult to interpret.

For the UK, the available evidence is very limited. Clark and Drinkwater (2000, 2002) basically document that poor areas where ethnic minorities live are associated with higher unemployment rates and lower self-employment rates. Their cross section regression analysis, however, includes only few area-level controls. It is thus well possible that their correlations capture the effects of *unobserved area characteristics*, such as tastes for discrimination, language difficulties, spatial mistmach. Beside the presence of unobserved and unobservable area characteristics, some *endogenous sorting* of the ethnic population into areas mainly populated by other ethnic minorities can also be responsible for a negative correlation between ethnic population density and ethnic employment.

There is also evidence, at least for other European countries, that ethnic enclaves have a *positive impact* on labor market outcomes of immigrants. Using a natural experiment (i.e. a spatial dispersal policy under which refugees were randomly dispersed across locations), Edin et al. (2003) and Damm (2009), for Sweden and Denmark, respectively, find strong evidence that the size of ethnic enclaves are positively correlated with earnings and job finding. These authors explain these results by the fact that *ethnic networks* disseminate job information, which increases the job-worker match quality and thereby the hourly wage rate.

The aim of this paper is twofold. First, with non-experimental data, we study the relationship between a community's ethnic population density and its employment rate and determine if this relationhip is localized or more global. Second, we propose a plausible behavioral mechanism that is able to reproduce the evidence.

To be more precise, the methodological contribution of this paper lies in the use of *panel data techniques* to identify the claimed effect and in the use of *spatial regression techniques* to assess the spatial scale of such an effect. Once the influence of observable and unobservable factors defining the local context has been taken into account, as in Edin et al. (2003) and Damm (2009), we find that a positive and significant relationship between ethnic employment and ethnic density. Such an effect, however, is quite localized. This evidence holds when we control for further endogeneity issues (using GMM estimators and instrumental variables).

We then provide an interpretation of the results in terms of *social interactions.* To be more precise, we use as a theoretical background the dynamic model of Calvó-Armengol and Jackson (2004) who explicitly model social networks as graphs. If workers are linked to each other, then they exchange information about jobs. Strong ties are direct friends while weak ties are friends of friends of any length (Granovetter, 1983; Calvó-Armengol et al., 2007; Patacchini and Zenou, 2008). The following results are found. The individual probability of finding a job increases with the number of strong ties and weak ties. However, the longer the length of weak ties, the lower the individual probability of finding a job.

Clearly, a precise test of this model requires detailed information on all social contacts between individuals over time, which is unfortunately not available. However, one can use this mechanism to guide the interpretation of our results by approximating the *social proximity* by the *geographical proximity.* Since ethnic communities tend to be more socially cohesive, our conjecture is that the density of people living in the same area is a good approximation for the number of direct friends one has, i.e. *strong ties*, especially if the areas are not too large and if people belong to the same ethnic group.[1] In the same spirit, the density of individuals living in neighboring areas will be a measure of friends of friends, i.e. *weak ties.* Ethnicity is thus the chosen dimension along which agents' social contacts develop. Consistently with the theoretical model, we find that the higher is the percentage of a given ethnic group living nearby, the higher is the employment rate of this ethnic group. We also find that this effect decays very rapidly with distance, losing significance beyond approximately 90 minutes travel time.

As mentioned above, however, our data have some limitations, mainly because we use aggregate data. As a result, our analysis should be taken with caution. Nevertheless, it rules

[1] A similar approximation of the social space (approximated by the physical space) is used in Wahba and Zenou (2005) for the case of Egypt.

out some traditional explanations of the relationship between ethnicity and employment at the local level, such as those based on unobservable area characteristics, and suggests that peer effects might be at work.

Little is known, in fact, about the impact of social networks on labor-market outcomes of ethnic minorities. Networks of personal contacts mediate employment opportunities, which flow through word-of-mouth and, in many cases, constitute a valid alternative source of employment information to more formal methods. Such methods have the advantage that they are relatively less costly and may provide more reliable information about jobs compared to other methods. The empirical evidence reveals that around 50 percent of individuals obtain or hear about jobs through friends and family (Granovetter, 1974; Corcoran et al., 1980; Holzer, 1988; Montgomery, 1991; Gregg and Wadsworth, 1996; Addison and Portugal, 2001; Wahba and Zenou, 2005; Ioannides and Loury, 2004; Goel and Lang, 2009). Do these (positive) effects extend to ethnic groups in the labor market?

Because usually ethnic minorities experience higher unemployment rates, one may think that ethnic enclaves may be harmful to labor-market outcomes of minorities. Indeed, having fewer connections to employed workers makes it more difficult to receiving information about jobs and therefore reduces the chance of obtaining a job.

On the other hand, the hiring of new workers via employee referrals is presumed to be important for understanding ethnic divisions of labor because it creates a built-in bias toward incumbents: members of a particular ethnic group concentrate in specific jobs and when new employment opportunities become available at their workplace, they pass this information along to social contacts, often of the same race and ethnic background.

Some evidence can be found, for example, in Conley and Topa (2002). They examine the spatial distribution of unemployment in Chicago using different social and economic distance metrics. Their results indicate a clear dominance of the racial/ethnic distance metric and of the racial/ethnic composition variables in explaining the spatial correlation of unemployment. More direct evidence can be found in Falcon (2007) and Falcon and Melendez (2001). They show that Latinos in Boston are more likely to use personal networks to gain employment relative to other job search methods. Elliott (2001) finds that Latinos, especially newly arrived immigrants, are more likely than native-born Whites to enter jobs through insider referrals. He also finds that the correlation between insider referrals and ethnically homogeneous jobs is positive and significant only for native-born Blacks. Mouw (2002), using longitudinal data, finds that Black workers who used personal contacts to find employment did no worse compared to where they used formal methods. Munshi (2003) attempts to identify network effects among Mexican migrants in the U.S. labor market and to test whether the network improves labor market outcomes for its members. He finds that the same individual is more likely to be employed and to hold a higher paying nonagricultural job when his network is exogenously larger.

There are very few papers for Europe. Exceptions include Frijters et. al (2005) and Battu et al. (2005), both for the UK. They find that, though personal networks are a popular method of finding a job among ethnic minorities, they are not necessarily the most effective method in finding a job.

Using a panel of local authority-level data in England between 2003 and 2007, we contribute to this strand of literature by showing that the popular view about a negative relationship between ethnic population density and employment rate might be flawed. In particular, our analysis seems to reveal that ethnic networks social contacts might improve labor market outcomes for network ethnic members. We also find that social networks are very localized.[2]

The paper is structured as follows. The next section discusses the different empirical problems underlying the relationship between a community's ethnic population density and its employment rate, and the achievements of our econometric approach. In Section 3 we expose the model of Calvó-Armengol and Jackson (2004), while highlighting and explicitating the relevant results in our context. We then bridge such a mechanism to our empirical analysis in Section 4 and we describe our data and empirical analysis (Section 4.1). Section 5 presents our empirical model and the estimation results both with OLS and IV estimators. Finally, Section 6 concludes.

# 2 Assessing the relationship between a community's ethnic population density and its employment rate

The assessment of the existence and the extent of the causal effect of local ethnic population density on local employment is a very difficult exercise. There are, in fact, at least three different issues that affect the usual finding of a negative relationship between a community's ethnic population density and its employment rate:

(*i*) A negative correlation between ethnic population density and employment can sim-

[2] This has been found before, at least for the United States. Using Census Tract data for Chicago in 1980 and 1990, Topa (2001) finds a significantly positive amount of social interactions across neighboring tracts, especially for areas with a high proportion of less educated workers and/or minorities. Bayer et al. (2008) also document that people who live close to each other, defined as being in the same census block, tend to work together, that is, in the same census block.

ply be driven by the presence of *unobservable factors* or/and by an *endogenous sorting of individuals into areas,* i.e. areas that attract ethnic populations may also have unobserved characteristics that reduce employment opportunities. Indeed, the areas where ethnic minorities are mostly concentrated are in general quite disadvantaged and the characteristics of such areas are typically associated with low prospects of finding a job. For instance, biased policing practices, low informal social control, lack of educational or economic opportunities, etc., which are typically difficult-to-measure variables, might result in a spurious negative correlation between density of ethnic population and employment at the local level.

($ii$) There is a *simultaneity/reverse causality bias* problem. Indeed, it may well be that higher ethnically dense areas produce less employment but it is also possible that areas with lower employment rates, which are typically poor areas, attract ethnic minorities, possibly because they cannot afford richer areas.

($iii$) High (low) employment rate areas are usually surrounded by high (low) employment rate areas. This creates spatial correlation that needs to be accounted for. Traditional studies of the relationship between ethnic population and employment show an average effect, thereby ignoring possible spillover effects at the local level, i.e. the effect of the levels of the variables in neighboring areas.

The construction of panel data, i.e. the availability of information on the same areas at different points in time, enables us to:

($a$) include *area-fixed effects* that control for the presence of *unobservable area-characteristics* in the regression models. Indeed, by using a within-group panel estimator, we purge our estimates from the possible existence of area characteristics that are constant over time, possibly correlated with the regressor of interest, i.e. ethnic population density, whose effects might otherwise be captured in the estimated coefficient of our regressor of interest.

($b$) include *a dependent variable lagged in time* in the regression models that accounts for dynamic effects, which also arise from the presence of unobservable factors that are varying over time. Indeed, by assuming the employment level today to be a function of the employment level in the previous period, we account for all factors contributing to the realization of such a level of employment in the previous period. Examples include education, activity rate, industrial structure, etc..

($c$) find suitable *instrumental variables* for tackling a possible simultaneity bias and/or an endogenous sorting of individuals into areas. Indeed, panel data *for the contemporaneous level of a regressor* offer the values of this regressor appropriately lagged in time as natural instruments.

In addition, by explicitly taking into account the geographical location of the areas, which requires the use of *spatial data analysis techniques*, we are able to appreciate the range of action of the effects. Most importantly, by creating ethnic population density proximity bands, we are able to appreciate to what extent such variables explain the spatial association between local ethnic employment and ethnic employment in neighboring areas.

Our empirical strategy will be explained in detail in Section 4. In the following section we present the theoretical background that will guide the interpretation of our findings.

# 3 Theoretical analysis

The aim of our theoretical framework is to understand how *strong* and *weak* ties affect the labor-market outcomes of workers. Indeed, given a network structure, we would like to see how the size of strong and weak ties affects the individual probability to obtain a job and thus the employment rate in the economy.

In the model, the main problem for each worker is to obtain information about jobs. Each worker is embedded in a network of social relationships, and his/her *direct friends* are his/her *strong ties* while *the friends of his/her friends* of any length are his/her *weak ties*. This worker can hear about a job either *directly* (if by chance he/she sees the job advertisement) or *indirectly* because one of their friends who belongs to their social network is employed, knows about this job and transmits the information to the worker. Observe that it is assumed that the probability to hear directly about a job is the same for someone who is employed and for someone who is unemployed.

## 3.1 Some notations and definitions from graph theory[3]

Denote by $n$ the number of individuals in a given social network $\mathbf{g}$, with $n = U + E$ ($U$ and $E$ are respectively the unemployment and the employment levels in the network). Therefore $N = \{1, \ldots, n\}$ is a set of individuals connected in some network relationship. A network is thus a list of unordered pairs of players $\{i, j\}$. These links are represented by a graph $\mathbf{g}$, where $g_{ij} = 1$ if $i$ is friend with $j$ (denoted by $ij$) and $g_{ij} = 0$ otherwise (*unweighted* graphs/networks). In our framework, links are taken to be reciprocal, so that $g_{ij} = g_{ji}$ (*undirected* graphs/networks). By convention, $g_{ii} = 0$. The set of $i$'s direct contacts is: $N_i(\mathbf{g}) = \{j \neq i \mid g_{ij} = 1\}$, which is of size $n_i(\mathbf{g})$.

One of the key features of networks/graphs is that not only *direct* but also *indirect* links that matter.

[3] For a more complete overview of these definitions, see Wasserman and Faust (1994) and Jackson (2008).

**Definition 1** *A path of length $k$ from $i$ to $j$ in the network $\mathbf{g}$ is a sequence $\langle i_0, i_1, ..., i_k \rangle$ of players such that $i_0 = i$, $i_k = j$, $i_p \neq i_{p+1}$, and $g_{i_p i_{p+1}} = 1$, for all $0 \leq p \leq k-1$, that is, players $i_p$ and $i_{p+1}$ are directly linked in $\mathbf{g}$. If such a path exists, then individuals $i$ and $j$ are path-connected.*

In words, a *path* between two individuals $i$ and $j$ is an ordered set of agents $(i, i_1, ..., i_k, j)$ of $N$, where an agent can appear several times, such that $i \neq j$. We say that a path belongs to the network $\mathbf{g}$ if $g_{ii_1} g_{i_1 i_2} ... g_{i_k j} \neq 0$.

To summarize:

**Definition 2** *An individual $i$ holds a **strong tie** with an individual $j$ if $g_{ij} = 1$. An individual $i$ holds a **weak tie** with an individual $j$ if individuals $i$ and $j$ are path-connected. The length $k$ of this (weak) tie is defined by the length of the path between individuals $i$ and $j$.*

## 3.2 The model

We now described the model of Calvó-Armengol and Jackson (2004). Time evolves in discrete periods indexed by $t$. The vector $s_t$ describes the employment status of the workers at time $t$. If individual $i$ is employed at the end of period $t$, then $s_{it} = 1$ and if $i$ is unemployed then $s_{it} = 0$.

A period $t$ begins with some agents being employed and others not, as described by the status $s_{t-1}$ from the last period. Next, information about job openings arrives. In particular, any given individual hears about a job opening with a probability $a$ that is between 0 and 1. This job arrival process is independent across individuals. If the individual is unemployed, then he/she will take the job. However, if the individual is already employed then he/she will pass the information along to a friend, picked at random among his/her unemployed friends. As stated above, graph or network $\mathbf{g}$ summarizes the links of all agents, where $g_{ij} = 1$ indicates that $i$ and $j$ know each other (strong tie), and share their knowledge about job information, while $g_{ij} = 0$ indicates that they do not know each other.

Observe that if an employed worker hears about a job but all his/her friends (i.e. direct links) are already employed, then the job is lost. We focus here on a model where wages are exogenous and identical for all workers. So there is no room in this model for an employed worker to exploit a job offer to increase his/her current wage.

Finally, the last thing that happens in a period is that some agents lose their jobs. This happens randomly according to an exogenous breakup rate, $\delta$, between 0 and 1. We are able to write the probability $\mathbb{P}_{ij}$ of the joint event that individual $i$ learns about a job and this job ends up in individual $j$'s hands. It is equal to:

$$\mathbb{P}_{ij}(\mathbf{s}) = \begin{cases} a & \text{if} \quad s_i = 0 \text{ and } i = j \\ a/\sum_{k:s_k=0} g_{ik} & \text{if} \quad s_i = 1,\ s_j = 0, \text{ and } g_{ij} = 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where the vector $\mathbf{s}$ describes the employment status of all the individuals at the beginning of the period. In (1), $a$ is the probability to obtain a job information without using friends and relatives. Three cases may then arise. If individuals $i$ and $j$ are unemployed ($s_i = s_j = 0$), then the probability that $j$ will obtain a job is just $a$ since individual $i$ will never transmit any information to $j$. If individual $i$ is already employed and his/her friend $j$ is not ($s_i = 1$, $s_j = 0$), then individual $i$ transmits this job information to all his/her direct unemployed neighbors, who total number is $\sum_{k:s_k=0} g_{ik}$. We assume that all unemployed neighbors are treated on equal footing, meaning that the employed worker who has the job information does not favor any of his/her direct neighbors. As a result, the probability that an unemployed worker $j$ is selected among the $\sum_{k:s_k=0} g_{ik}$ unemployed direct neighbors of an employed worker $j$ is given by: $a/\sum_{k:s_k=0} g_{ik}$. Finally, if individual $j$ is employed, then he/she does not need any job information, at least in the current period.

### 3.3 Impact of strong ties on employment probabilities

The first result obtained by Calvó-Armengol and Jackson (2004) is not surprising and has also been showed in a static framework (see Calvó-Armengol, 2004, and Calvó-Armengol and Zenou, 2005).

**Proposition 1** *The higher $n_i(\mathbf{g})$, the number of strong ties individual $i$ has, the higher his/her individual probability of obtaining a job.*

Indeed, if an individual has more strong ties, then he/she is more likely to hear on average about more jobs through his/her friends and relatives but his/her chance to find a job directly does not increase since $a$ is not affected by the size of the network. This result is quite intuitive since when the number of direct connections increases, the source of information about jobs is larger and people find it more easy to obtain a job through their friends and relatives. This is the first prediction of our model, which implies that workers have a greater chance to find a job the higher the number of their strong ties. Observe that the individual probability to find a job through strong ties for individual $j$ is obviously not given by (1) since $\mathbb{P}_{ij}(\mathbf{s})$ is the probability that only one individual, $i$, who hold a strong tie with $j$, and who is aware of some job, will transmit this information to individual $j$. To determine the individual probability of obtaining a job for $j$, one has to do the calculation for all the direct friends of $i$.

## 3.4 Impact of weak ties on employment probabilities

We would like now to study the impact of weak ties (as defined by Definition 1) on the individual probability of finding a job. Calvó-Armengol and Jackson (2004) show that, in steady-state, there is a positive correlation in employment status between two path-connected workers. As we will see, this result is not at all easy to obtain since, in the short run, the correlation is negative. Indeed, in a static model, if an employed worker is directed linked to two unemployed workers, then if he/she is aware of a job, he/she will share this job information with his/her two unemployed friends (see (1)). These two persons, who are path-connected (path of length two) are thus in competition and one (randomly chosen) will obtain the job and be employed while the other will stayed unemployed. So their employment statuses will be negatively correlated (see Calvó-Armengol, 2004, and Calvó-Armengol and Zenou, 2005).

Let us now give the intuition why this negative correlation result does *not* hold in a dynamic labor-market model. Consider the star-shaped network described in Figure 1 with three individuals, i.e. $n = 3$ and $g_{12} = g_{23} = 1$. Suppose the employment from the end of the last period is $\mathbf{s_{t-1}} = (0, 1, 0)$. In the figure, a darkened node represents an employed worker (individual 2), while unemployed workers (1 and 3) are represented by lightly-colored nodes. Conditional on this state $\mathbf{s_{t-1}}$, the employment states $s_{1t}$ and $s_{3t}$ are negatively correlated. As stated above, this is due to the fact that individuals 1 and 3 are "competitors" for any job news that is first heard by individual 2.

$[Insert\ Figure\ 1\ here]$

Despite this negative (conditional) correlation in the shorter run, individual 1 can benefit from individual 3's presence in the longer run. Indeed, individual 3's presence helps improve individual 2's employment status. Also, when individual 3 is employed, individual 1 is more likely to hear about any job that individual 2 hears about. These aspects of the problem counter the local (conditional) negative correlation, and help induce a positive correlation between the employment status of individuals 1 and 3.

In what follows, we describe how we obtain this long-run positive correlation. Consider again the network described in Figure 1 but without imposing any employment status to workers. In that case, there are eight possible employment states: $000, 100, 010, 001, 110, 101,$

011, 111, where for example 000 means that all individuals 1, 2, and 3 are unemployed. As a result, the state of the economy $s_t$ evolves following a Markov process $\mathcal{M}(a, \delta)$ where $a$ is the job-arrival rate that takes place in the first half of each period, while $\delta$ is the job-destruction rate that takes place at the second half of each period. We gather the Markov transitions into a matrix $\mathbb{P}_{ij} = \Pr\{s_{t+1} = i \mid s_t = j\}$, where $i, j \in \{000, 100, 010, 001, 110, 101, 011, 111\}$, that is, rows correspond to $t+1$ while columns correspond to $t$ (the columns sum up to one as in all Markov matrices).

As highlighted above, an important issue in this case is the short-run negative correlation versus the long-run (possibly) strictly positive correlation. To sort out the short and longer run effects, we divide $a$ and $\delta$ both by some larger and larger factor, so that we are looking at arbitrarily short time periods. We call this the "sub-division" of periods. More precisely, instead of analyzing the Markov process $\mathcal{M}(a, \delta)$, we can analyze the associated Markov process $\mathcal{M}(a/T, \delta/T)$, that we name the $T-$*period subdivision* of $\mathcal{M}(a, \delta)$, with steady state distribution $\mu^T$. We show that there exists some $T'$ such that, for all $T \geq T'$, the employment statuses of any path-connected agents are positively correlated under $\mu^T$. Consider $\mathcal{M}(a/T, \delta/T)$. For this Markov process, at every period, every shock (be it a job arrival $a/T$ or a job breakdown $\delta/T$) is very unlikely when $T$ is high enough. Having two or more shocks in every such period is thus much less unlikely. Instead of analyzing $\mathcal{M}(a/T, \delta/T)$, we analyze an approximated Markov process $\mathcal{M}^*(a/T, \delta/T)$ where we only keep track of one-shock transitions, and disregard transitions involving two or more shocks. We denote by $\mu^{*T}$ the corresponding steady-state distribution. The higher $T$, the closer are the transitions of the approximated Markov process $\mathcal{M}^*(a/T, \delta/T)$ to that of the true Markov process $\mathcal{M}(a/T, \delta/T)$, and so the closer is $\mu^{*T}$ to $\mu^T$.

Calvó-Armengol and Jackson (2004) show that, with a high enough $T-$period subdivision, for $n$ individuals and any social network structure, we have:

**Proposition 2** *Under fine enough subdivisions of periods, the unique steady-state long-run distribution on employment is such that the employment statuses of any path-connected agents are positively correlated.*

The proposition shows that despite the short-run conditional negative correlation between the employment of competitors for jobs and information, in the longer run any interconnected workers' employment is positively correlated. This implies that there is a clustering of agents by employment status, and employed workers tend to be connected with employed workers, and vice versa. The intuition is clear: conditional on knowing that some set of agents are employed, it is more likely that their neighbors will end up receiving information about jobs, and so on. The benefits from having other agents in the network outweigh the local negative correlation effects, if we take a long-run perspective.

**Proposition 3** *The longer the length of two-path connected individuals (i.e weak ties), the lower is the correlation in employment statuses between these two individuals.*

Indeed, the correlation between two agents' employment is (weakly) decreasing in the number of links that each an agent has, and the correlation between agents' employment is higher for direct compared to indirect connections. The decrease as a function of the number of links is due to the decreased importance of any single link if an agent has many links. The difference between direct and indirect connections in terms of correlation is due to the fact that direct connections provide information, while indirect connections only help by indirect provision of information that keeps friends, friends of friends, etc., employed. In other words, the longer the path in the social network between two individuals, the weaker is the effect of job transmission.

# 4 Bridging the model to the empirics

Let us summarize our theoretical results. We have shown that:

(*i*) The individual probability of finding a job is increasing in the number of strong ties each individual has (Proposition 1);

(*ii*) The individual probability of finding a job is increasing in the number of weak ties each individual has (Proposition 2);

(*iii*) The longer the length of weak ties, the lower the individual probability of finding a job (Proposition 3).

Let us now approximate the *social proximity* by the *geographical proximity*, drawing a link between the social and geographical spaces.

Take an individual with a given set of characteristics (family, age, education, gender...), living in area $a$ and belonging to race $r$.

The individual probability of finding a job of individual of type$-r$ in local authority $a$ is measured using the ratio between the number of type$-r$ employed workers in local authority $a$ and the total working age population of type$-r$ individuals residing in local authority $a$ - rescaled to be between 0 and 1- (defined as the employment rate of ethnic minorities of type $r$

living in local authority $a$, $e_a^r$, hereafter). The group$-r$ population density in local authority $a$ ($d_a^r$, hereafter) is the ratio between the total working age population of group $r$ residing in local authority $a$ and the total working age population residing in local authority $a$. Our conjecture is that the density of people living in the same area is a good approximation of the number of direct friends one has, i.e. *strong ties*, especially if the areas are not too large and if people belong to the same ethnic group (Topa, 2001). In the same spirit, the density of individuals living in neighboring areas will be a measure of friends of friends, i.e. *weak ties*. Using this approach, other things being equal (i.e. fixing the characteristics of the area) the theoretical predictions are as follows.

($i$) The employment rate of type$-r$ individuals living in area $a$ is higher, the higher the percentage of type$-r$ individuals living in area $a$.

($ii$) The employment rate of type$-r$ individuals living in area $a$ is higher, the higher the percentage of type$-r$ individuals living in the neighboring areas of $a$.

($iii$) This effect should decrease with the distance between area $a$ and its neighboring areas.

## 4.1 Data

Our empirical analysis is based on a panel of local authority-level data in England from 2003 until 2007. The data source is the quarterly UK Labour Force Survey (LFS).[4] Because of small sample size per area of ethnic minority groups, we aggregate data from two subsequent quarters, so to end up with ten waves of the panel. The *local authority* is the finer level of spatial disaggregation of the English local government structure. In England, there is indeed a mix of single-tier and two-tier local government. Our definition of "local authority" considers single-tier (unitary) authorities together with the lower-tier authorities in areas of two-tier local government. We thus deal with very small spatial units. For example, in London there are 33 local authorities (London boroughs). Excluding areas with missing or incomplete information on our target variables (roughly the 10% of the total), we are left with a final sample of 301 local authorities in England.

We measure distance between areas by the average road journey time (in minutes) between the centres of the areas.[5] Indeed, driving times represents how agents' contacts develop

[4]The data are available through the Office of National Statistics (ONS). We acknowledge the original data creators, depositors or copyright holders, the funders of the Data Collections and the UK Data Archive. They bear no responsibility for their further analysis or interpretation.

[5]Distances in travel times and kilometers are estimated using Microsoft Autoroute 2002. The Microsoft

better than other measures of proximity, such as physical distance or contiguity (see, e.g. Conley and Topa, 2002). The estimated road journey time between areas in the sample varies between 9.7 minutes and 514 minutes, with a median journey time of approximately 198 minutes. This spatial approach is essential to implement the test of our point (iii) above. Let us be more precise. In the original Calvó-Armengol and Jackson (2004) model the length of ties is measured by the path connecting individuals (see definitions 1 and 2). Since we approximate the social space by the physical space, to measure ties of different lengths we create *proximity bands* based on driving time between areas and we measure the population density by ethnic group within each proximity band. To be specific, for each local authority $a$ in our sample, we create new variables containing the densities of ethnic population within 30 minutes driving time from local authority $a$; within 30-60 minutes, and so on. We assume that the population of each local authority is concentrated at the economic centre of the local authority, so that each time band (e.g., 30-60 minutes) contains the population densities of all areas whose centre is in the band (e.g., within 30-60 minutes from the centre of local authority $a$).[6] By comparing estimates across rings it is possible to assess the impact of the density of own-race individuals living nearby and how far this effect extends.

## 4.2 Empirical strategy

The conjecture underlying our empirical strategy is that the social and geographical spaces should be correlated for individuals within the same ethnicity, and more so given that our areas (local authorities) are quite small. In fact, beside the specific studies on labour market outcomes of ethnic minorities and ethnic networks such as the ones cited in the Introduction (i.e., Falcon, 2007; Falcon and Melendez, 2001; Elliott, 2001; Mouw, 2002; Munshi, 2003), there is a rich socio-economic literature on patterns of relations among individuals, documenting that social networks appear to be fairly homogeneous with regard to certain socio-demographic attributes. Indeed, individuals are likely to associate with people who are similar, i.e., *assortative matching* or *homophily*.[7] This tendency is particularly strong

---

Autoroute software computes the driving time between two locations on the basis of the most efficient route given the road network in 2002, and allowing for different average speeds of travel depending on the type of road.

[6] The analysis has also been performed assuming that the population is evenly distributed within the area (see, Rosenthal and Strange, 2008 and Rice and al., 2006, for details on this approach). The results remain qualitatively unchanged.

[7] For an overview on the homophily literature, see McPherson et al. (2001) and Jackson (2008). For some theoretical foundation, see Currarini et al. (2009).

among ethnic groups.[8,9] Therefore, it is reasonable to assume that the density of individuals of a given ethnic group is a good approximation for the size of social contacts that each individual of that ethnicity is exposed to. Although those relationships may neither be personal nor strong, those contact are channels of information. Such approximation, however, is not necessarily true for whites. Marsden (1988) shows, for instance, that the chance of observing a black-black tie is 4.2 times higher than that generated by pure random matching, whereas this value is only 2.6 for whites. Therefore, we run our empirical analysis for whites and nonwhites in order to provide valuable insights about the validity of our approach. If the percentage of white people living nearby (the majority) does not affect the white employment rate, whereas we find some effects for ethnic minorities, this will increase our confidence in the interpretation of the results in terms of "ethnic network effects" on ethnic employment rates. Such a comparison also rules out a possible interpretation of the results in terms of "agglomeration effects", i.e. about the fact that the population density is only capturing effects deriving from the degree of urbanization of the area rather than from social interactions. In that case, an alternative explanation would be that, in denser areas, there are relatively more jobs than in less populated areas so that the chance to find a job also increases. If we find a non significant effect for white and a significant one for nonwhite, it would then be difficult to explain why a positive relationship between employment rate and population density is only obtained for nonwhites and *not* for whites. Similar reasonings apply for all other factors that are supposed to increase labor market conditions regardless of the race.

Our data allows us to distinguish between different ethnic groups. However, for our approach to be consistent, we need to use narrowly defined ethnic groups. Clearly, for example, the density of Asian people living nearby cannot be a good approximation of the social contacts a Black individual has, and, likewise, the density of Chinese people nearby may not adequately capture the social ties of a Pakistani. This is due to important cultural differences between these groups. Therefore, our investigation has been performed separately for the different ethnic minorities that can be unambiguously identified in our data , namely "Black Caribbean", "Black African", "Indian", "Pakistani", "Bangladeshi" and "Chinese". We believe that within each of group, there is a relative cultural homogeneity. The ethnicity is thus captured by the index $r = W, BC, BA, I, P, B, C$.

[8] See e.g. Moody (2001), Marmaros and Sacerdote (2006), Bayer et al. (2007).

[9] It has also been shown that investments in public goods, tastes for redistribution, and other forms of civic behavior are more common in racially or ethnically homogenous communities (see Alesina and La Ferrara , 2005, for an overview of this literature).

# 5 Empirical model and estimation results

For each ethnic group $r = W, BC, BA, I, P, B, C$, we estimate the following panel regression model:

$$e^r_{a,t} = \alpha e^r_{a,t-1} + \sum_{\sigma} \gamma^r_\sigma d^r_{\sigma,a,t} + \eta_a + \varepsilon^r_{a,t}, \qquad |\alpha| < 1, \qquad a = 1, ..., N; \; t = 2, ..., T, \tag{2}$$

where $e^r_{a,t}$ is the employment rate of ethnic group $r$ in local authority $a$ at time $t$, $e^r_{a,t-1}$, the same variable at time $t-1$, and $d^r_{\sigma,a,t}$ denotes the population density of ethnic group $r$ in local authority $a$ within the proximity bands $\sigma$. The error term is composed of an area-specific fixed effect, $\eta_a$, controlling for cross-area differences constant across time and by a white noise error component, $\varepsilon^r_{a,t}$. Observe that the empirical model does neither include any measure of the average human capital characteristics of the different areas, nor other features of the local structure of the economy (such as the unemployment rate for example). Indeed, dealing with sub-annual data, we assume that the impact of these characteristics on the employment rate in each area is captured through the inclusion of (time) lagged values of the employment rate $e^r_{a,t-1}$. In other words, we use area fixed effects to purge our estimates from the effects of area characteristics that are constant over time and we assume that the impact of time-varying variables on the employment rate in each location is captured through the inclusion of (time) lagged values of the employment rate.[10] Observe, however, that the inclusion of the lagged dependent variable is not only a purely modelling device to approximate the effects of other area characteristics on local employment. It is also consistent with the *dynamic* theoretical model presented in Section 3.

## 5.1 OLS estimates with diagnostics for spatial effects

Table 1 reports the results obtained with five proximity bands: up to 30 min, 30 to 60 min, 60 to 90 min, 90 to 120 min, 120 to 150 min. The table contains the within groups estimates, i.e. OLS where all variables are expressed in deviations from their area-specific means (taken over time)[11].

[10] However, the inclusion of a list of contemporary area-specific controls, such as proportion of high-skilled population, proportion of car owners, local industrial structure, white and nonwhite population mass, etc., does not change qualitatively the results on our target variables.

[11] As $T$ becomes large, the within groups estimator is consistent, even in the presence of lagged dependent variables (or other endogenous regressors). Thus, with our panel of 10 points time-length, any bias from using within groups is likely to be minimal (see Nickell, 1981, for example).

Let us begin by documenting to what extent the spatial population bands explain the spatial association between employment in a local area and its neighboring areas. The first column of each block of the table shows the results when the population density proximity bands (i.e., the term $\sum_{\sigma} \gamma_{\sigma}^{r} d_{\sigma,a,t}^{r}$) are not included in the specification of model (2). Under this specification, the spatial dependence in employment rate may result in an omitted spatially lagged dependent variable (i.e. the average employment rate in neighboring areas). We report the Lagrange Multiplier test for an omitted spatially lagged dependent variable (LM test) in the last row of the table. The null hypothesis is the absence of spatial dependence (i.e., that the effect of the spatially lagged dependent variable is zero). A significant value of the test provides evidence of the existence of spatial dependence in the data that is not fully captured by the model specification.

Looking at the results for the model specification without population density proximity bands (first column of each block), the test provides clear evidence of unexplained spatial dependence both for Whites and all ethnic groups. This indicates that the model does not incorporate all channels of interdependence between areas.

Our revisitation of the theoretical mechanism presented in Section 3 postulates that such a contagion/spillover effect is explained by the diffusion of information between adjacent areas, which is captured by population density. Our intuition, however, is that this is true only for ethnic minorities. Indeed, when we add the population density proximity bands to the model (second column of each block), for Whites the LM test remains statistically significant, whereas for all the ethnic minority groups the LM test now does not provide evidence of an omitted spatial lag, thus revealing that the ethnic population density bands entirely capture the spatial interactions in ethnic employment rate.[12] In other words, our results reveals that there should be something specific for ethnic minorities that is captured by the population density that explains spatial dependence and increase the ethnic employment rate. Once the effects of area observable and unobservable factors is taken into accont by our panel model estimation, we are left to think that we are really capturing "network effects". Indeed, if the ethnic population density is taken as a measure of the strength of social contacts, these results are consistent with our theoretical mechanisms. This is the more plausible interpretation of

[12]In the model specification with population bands, the null hypothesis of no spatial dependence is tested against an alternative of spatial dependence within a specified proximity. The tests are computed with spatial weight matrices $W = \{w_{ij}\}$, where $w_{ij} = 1$ if the estimated driving time between area $i$ and area $j$ is less than $d$ minutes and $w_{ij} = 0$ otherwise, for values of $d = \{30, 60, 90, 120, 150\}$. The highest values are reported in the table in each case. When the bands are not considered, a simple first-order contiguity matrix (i.e. a spatial weight matrix $W = \{w_{ij}\}$, where $w_{ij} = 1$ if area $i$ and area $j$ share a common border and $w_{ij} = 0$ otherwise) is used.

our results. For whites, on the contrary, the population density is not capturing any effects over and above area-fixed effects. The determinants of whites employment rates remains implicity captured by the spatial and time lagged dependent variables. Observe that they might also include network effects, for which however we have not a good proxy in this case.

[*Insert Table* 1 *here*]

Focusing now on the estimates of the spatial decay for the different ethnic minorities, we find positive and statistically significant effects, which are greatest within 30 minutes driving time. They then decrease quite sharply with travel time and have no effect beyond approximately 90 minutes. This pattern remains unchanged across the different ethnic minority groups. If we go back to our theoretical model, then this means that strong ties (i.e. population density of the same ethnic group within 30 minutes driving time) have a greater positive impact on the employment rate than weak ties of length 2 (i.e. population density of the same ethnic group within 30 to 60 minutes driving time), which, in turn, has a higher impact than weak ties of length 3 (i.e. population density of the same ethnic group within 60 to 90 minutes driving time), etc.

The magnitude of the effects, however, is different across ethnic groups, in particular between the Black and the Asian groups. Asians display higher values as well as higher rates of attenuation than Blacks. In particular, a one point percentage increase in the density of the Chinese population within 30 minutes driving time increases Chinese employment rate by roughly 0.15 percentage points. It has more than three times the impact of the density of the Chinese population 60 minutes away, and more than 10 times that of the density of the Chinese population at 90 minutes.

Because a substantial proportion of ethnic minorities concentrates in London, we have also performed our analysis excluding the London area (33 areas out of 301). The results remain qualitatively unchanged.[13]

## 5.2 Robustness check: Instrumental variable estimates

If areas that attract ethnic population have also exogenously determined characteristics (not directly observable) that affect employment, then the population density variables will be correlated with the error term. In other words, there may be some sorting of the population where ethnic minorities choose to live in areas mainly populated by individuals belonging to the same ethnic group. In that case, instead of social networks, we may capture some

[13] They are not reported here for brevity, they remain available upon request.

unobserved characteristics of workers if, for example, the higher the (unobserved) ability of ethnic minorities, the more likely they live in areas with the highest population density of the same ethnic group.

In our analysis, the adoption of a panel data estimator with area fixed effects removes any unobserved area characteristic constant over time, which could be responsible for such a sorting behavior. It does not account, however, for the effects of possible unobserved troubling factors that vary in the short-run. As it is standard, we attempt to address this problem by using an instrumental variable approach. We employ the Arellano and Bond (1991) instrumental variable estimator for dynamic panel data. This method consists of taking deviations from the area-specific time means to get rid of the unit-specific error term and combining valid instruments for the lagged dependent variable and the other endogenous variables in a GMM framework. Given the first-order autoregressive specification of our model, valid instruments for the (time) lagged dependent variable are variables that are lagged two-time periods or more. The use of population density variables lagged in time as instruments for the contemporaneous values does not solve all the problems, but at least mitigates endogeneity issues. In addition, following Ciccone (2002)'s approach, we also instrument the population density bands with total land area. The Sargan test of overidentifying restrictions (Sargan, 1958) is used to choose the appropriate set of instruments in the case study (in particular, the lags of the employment rate and of the population density to be included in the instrumental set). In our case, this test identifies as valid instruments variables lagged more than four periods (i.e. more than two years).

Our instrumental variable results for the model that includes all the population bands are reported in Table 2. The Sargan test does not reject the null of instruments' validity. In the last rows, we also report the tests for first-order and second-order serial correlation in the first-differenced residuals (M_1 and M_2 ). The consistency of the GMM estimators requires the absence of serial correlation in the original error term. In turn, this requires negative first-order, but no second-order correlation in the differenced error term. Table 2 reveals no evidence of misspecification. The results confirm the main findings from Table 1, thus revealing that possible endogeneity issues are properly accounted for by the inclusion of area-fixed effects. The estimated effects are only slightly higher in magnitude.

[*Insert Table* 2 *here*]

## 5.3 More intuition of our results

In order to better understand our mechanism, we would like now to investigate further our relationship between ethnic population density (capturing social networks and transmission of job information) and ethnic employment probabilities by looking at certain aspects of it. First, because ethnic minorities tend to be more self-employed than whites, we will see if our relationship is stronger or weaker for self-employed workers only. Second, because ethnic minorities tend to concentrate in specific jobs (or industries), we will investigate if our relationship is stronger in certain activities, i.e. information transmission is better in certain jobs. Finally, because ethnic minorities have different cultures and traditions concerning female work participation, we will also examine the relationship between ethnic population density and ethnic employment probabilities for female only.

Table 3 reports the estimation results for model (2) for self-employment only. The results remain qualitatively unchanged but the estimated effects are smaller in magnitude for all ethnic groups and at all distance bands. This could be explained by the fact that finding a job requires social networks, especially for ethnic minorities, while for self-employment this is still true but to a lesser extent.

[*Insert Table* 3 *here*]

Table 4 displays the results obtained from the estimation of model (2) when employment is disaggregated by industry. To save place, we report the results only for the Black population as a whole and the Asian population as a whole.[14] The results are not qualitatively different between Blacks and Asians. We always find that the coefficients are larger in magnitudes for the "distribution, hotel and restaurant" and "banking, finance and insurance" sectors. This means that the transmission of job information and thus ethnic social networks have a larger effect in these sectors than in others. For both Blacks and Asians, if one compares the distance decay across industries, we find that for the "distribution, hotel and restaurant" sector, the effect fades away after 90 minutes, it is non significant already after 60 minutes for the sectors "agriculture & fishing", "manufacturing & construction" and "transport & communication", whereas for the "banking, finance and insurance" sector, it is still significant at distances within 90-120 minutes. This suggests that social networks are less localized for

[14] As before, the results are significant for each ethnic minority group and not for Whites. When the analysis is performed for each Black and Asian minority group separately, we find essentially the same qualitative evidence of the aggregate groups .

the latter sector than for the other sectors.

[*Insert Table* 4 *here*]

We investigate finally the importance of social networks for ethnic female employment. In other words, we estimate model (2) for female employment only. Table 5 collects the evidence. It is striking to see that the estimated effects are not statistically significant for all groups (whites and nonwhites). This may be not surprising if we consider that there are fewer females employed in ethnic minorities and the channel through which job information is transmitted is usually between men. The LM test provides clear evidence of unexplained spatial dependence for all groups. This indicates that the spatial correlation in female employment rates is not explained by the diffusion of information between adjacent areas, as captured by our population density bands.

[*Insert Table* 5 *here*]

# 6 Conclusion

Because of different econometric problems the relationship between local ethnic population density and local employment rate is extremely difficult to assess and to interpret with non-experimental data. The methodological achievements of our analysis allow us to rule out some possible interpretations and to highlight a mechanism that is consistent with the evidence. Using panel data estimation techniques and spatial statistics methods, we find that the higher the percentage of a given ethnic group living nearby, the higher the employment rate of this ethnic group. However, this effect decays very rapidly with distance, losing significance beyond approximately 90 minutes travel time. For whites, the effects are not statistically significant. The puzzling question we challenge is the following one. Given that we control for observable and unobservable area characteristics and other endogeneity issues, how can we explain a positive relationship between population density and employment rate and its spatial decay pattern that holds only for ethnic minorities?

Our answer is that local social interactions between people of the same ethnicity can explain this positive relationship and its spatial trend. Conjecturing that the social space is highly correlated to the physical space for ethnic minorities in small areas, we present a theoretical framework based on Calvó-Armengol and Jackson (2004), which shows that the individual probability of finding a job increases with the number of strong ties and weak ties, and the longer the length of weak ties, the lower this probability. Our data are, however, limited for delivering conclusive results about the mechanisms at the basis of the complex relationship between employment and ethnicity. Our purpose here is to highlight that peer effects might be an important part of the story. If ethnic population density is interpreted as a proxy for the strength of social interactions, then our analysis suggests that they are quite localized and are relevant in explaining the spatial distribution of ethnic eployment.

# References


[1] Addison J.T. and P. Portugal (2002), "Job search methods and outcomes," *Oxford Economic Papers* 54, 505-533.

[2] Alesina, A. and E. La Ferrara (2005), "Ethnic Diversity and Economic Performance," *Journal of Economic Literature* 43, 762-800.

[3] Anselin, L. (1995), *Spacestat Version 1.80 User's Guide.* Morgantown, WV: Regional Research Institute, West Virginia University.

[4] Arellano, M. and S. Bond (1991), "Some test of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies* 58, 277-297.

[5] Battu, H., Seaman, P.T., and Y. Zenou (2005), "Job contact networks and the ethnic minorities," CEPR Discussion Paper No. 5225.

[6] Bayer, P., F. Ferreira, and R. McMillan (2007), "A unified framework for measuring preferences for schools and neighborhoods," *Journal of Political Economy* 115, 588-638.

[7] Bayer, P., Ross, S.L. and G. Topa (2008), "Place of work and place of residence: Informal hiring networks and labor market outcomes," *Journal of Political Economy* 116, 1150-1196.

[8] Calvó-Armengol, A. (2004), "Job contact networks," *Journal of Economic Theory* 115, 191-206.

[9] Calvó-Armengol, A. and M.O. Jackson (2004), "The effects of social networks on employment and inequality," *American Economic Review* 94, 426-454.

[10] Calvó-Armengol, A., Verdier, T. and Y. Zenou (2007), "Strong and weak ties in employment and crime," *Journal of Public Economics* 91, 203-233.

[11] Calvó-Armengol, A. and Y. Zenou (2005), "Job matching, social network and word-of-mouth communication," *Journal of Urban Economics* 57, 500-522.

[12] Ciccone, A. (2002), "Agglomeration effects in Europe," *European Economic Review* 46, 213-227.

[13] Clark K. and S.J. Drinkwater (2000), "Pushed out or pulled in? Self-employment amongst Britain's ethnic minorities", *Labour Economics* 7, 603-628.

[14] Clark K. and S.J. Drinkwater (2002), "Enclaves, neighbourhood effects and employment outcomes: Ethnic minorities in England and Wales," *Journal of Population Economics* 13, 5-30.

[15] Conley, T.G. and G. Topa (2002), "Socio-economic distance and spatial patterns in unemployment," *Journal of Applied Econometrics* 17, 303-327.

[16] Corcoran, M., L. Datcher, G.J. Duncan (1980), "Most workers find jobs through word of mouth," *Monthly Labor Review* 103, 33-35.

[17] Currarini, S., M.O. Jackson and P. Pin (2009), "An economic model of friendship: Homophily, minorities, and segregation," *Econometrica* 77, 1003-1045.

[18] Damm, A.P. (2009), "Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence," *Journal of Labor Economics* 27, 281-314.

[19] Doornik, J.A. (2001), *Ox: An Object-Oriented Matrix Language*, London: Timberlake Consultants Press.

[20] Edin, P.-A., Fredriksson, P. and O. Åslund (2003), "Ethnic enclaves and the economic success of immigrants. Evidence from a natural experiment," *Quarterly Journal of Economics* 118, 329-357.

[21] Elliott, J.R. (2001), "Referral hiring and ethnically homogeneous jobs: How prevalent is the connection and for Whom?" *Social Science Research* 30, 401-425.

[22] Falcon, L.M. (2007), "Social networks and latino immigrants in the labor market: A review of the literature and evidence," In: M. Montero-Sieburth and E. Melendez (Eds.), *Latinos in a Changing Society*, Westport, CT: Praeger.

[23] Falcon, L.M. and E. Melendez (2001), "The social context of job searching for racial groups in urban centers," In: A. O'Connor, C. Tilly and L. Bobo (Eds.), *Urban Inequality: Evidence from Four Cities*, New York: Rusell Sage Foundation, pp. 341-371.

[24] Frijters, P. Shields, M.A. and S. Wheatley-Price (2005), "Immigrant job search in the UK: Evidence from panel data," *Economic Journal* 115, F359-F376.

[25] Goel, D. and K. Lang (2009), "Social ties and the job search of recent immigrants," NBER Working Paper No. 15186.

[26] Granovetter, M. (1974), *Getting a Job*, Chicago: University of Chicago Press.

[27] Granovetter, M.S. (1983), "The strength of weak ties: A network theory revisited", *Sociological Theory* 1, 201-233.

[28] Gregg, P. and J. Wadsworth (1996), "How effective are state employment agencies? Jobcentre use and job matching in Britain," *Oxford Bulletin of Economics and Statistics* 58, 443-457.

[29] Hagan, J.M. (1994), *Deciding to be Legal: A Maya Community in Houston*, Philadelphia: Temple University Press.

[30] Holzer, H.J. (1988), "Search method use by the unemployed youth," *Journal of Labor Economics* 6, 1-20.

[31] Ioannides, Y. and L.D. Loury (2004), "Job information networks, neighborhood effects, and inequality," *Journal of Economic Literature* 42, 1056-1093.

[32] Jackson, M.O. (2008), *Social and Economic Networks*, Princeton: Princeton University Press.

[33] Lupton, R. and A. Power (2004), "Minority ethnic groups in Britain," CASE-Brookings Census Briefs 2, London: CASE.

[34] Marmaros, D. and B. Sacerdote (2006), "How do friendships form," *Quarterly Journal of Economics* 121, 79-119.

[35] Marsden, P.V. (1988), "Homogeneity in confiding relations," *Social Networks* 10, 57-76.

[36] McPherson, M., L. Smith-Lovin, and J. Cook (2001), "Birds of a feather: Homophily in social networks," *Annual Review of Sociology* 27, 415-44.

[37] Montgomery, J. (1991), "Social networks and labor-market outcomes: Toward an economic analysis," *American Economic Review* 81, 1408-1418.

[38] Moody, J. (2001), "Race, school integration, and friendship segregation in America," *American Journal of Sociology* 107, 679-716.

[39] Mouw, T. (2002), "Racial differences in the effects of job contacts: Conflicting evidence from cross-sectional and longitudinal data," *Social Science Quarterly* 31, 511-538.

[40] Munshi, K. (2003), "Networks in the modern economy: Mexican migrantsin the U.S. labor market," *Quarterly Journal of Economics* 118, 549-599.

[41] Nickell, S. (1981), "Biases in dynamic models with fixed effects," *Econometrica*, 49, 1417-26.

[42] Patacchini, E. and Y. Zenou (2008), "The strength of weak ties in crime," *European Economic Review* 52, 209-236.

[43] Rice, P., Venables A.J. and E. Patacchini (2006), "Spatial determinants of productivity: Analysis for the regions of Great Britain," *Regional Science and Urban Economics* 36, 727-752.

[44] Rosenthal, S.S. and W. C. Strange (2008), "The attenuation of human capital externalities," *Journal of Urban Economics,* 64 (2), 373-389.

[45] Sargan, J. D. (1958), "The estimation of economic relationships using instrumental variables," *Econometrica* 26, 393-415.

[46] Topa, G. (2001), "Social interactions, local spillovers and unemployment," *Review of Economic Studies* 68, 261-295.

[47] Wahba, J. and Y. Zenou (2005), "Density, social networks and job search methods: Theory and applications to Egypt," *Journal of Development Economics* 78, 443-473.

[48] Wasserman, S. and K. Faust (1994), *Social Network Analysis: Methods and Applications*, Cambridge: Cambridge University Press.

Table 1. Employment and population density
- OLS estimates -

| | White | | Black Caribbean | | Black African | | Indian | | Pakistani | | Bangladeshi | | Chinese | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Population density | | | | | | | | | | | | | | |
| …within 30 min | - | 0.1961 | - | 0.0808 | - | 0.0787 | - | 0.1201 | - | 0.1101 | - | 0.1086 | - | 0.1509 |
| | | (0.44) | | (3.80) | | (4.03) | | (4.16) | | (3.56) | | (4.05) | | (4.14) |
| … within 30-60 min | - | 0.0920 | - | 0.0351 | - | 0.0397 | - | 0.0545 | - | 0.0399 | - | 0.0403 | - | 0.0425 |
| | | (0.83) | | (3.36) | | (3.85) | | (4.50) | | (3.36) | | (3.78) | | (3.45) |
| … within 60-90 min | - | 0.0311 | - | 0.0135 | - | 0.0175 | - | 0.0322 | - | 0.0130 | - | 0.0212 | - | 0.0129 |
| | | (1.01) | | (2.04) | | (2.50) | | (3.89) | | (3.02) | | (2.45) | | (2.77) |
| … within 90-120 min | - | 0.0024 | - | 0.0059 | - | 0.0080 | - | 0.0111 | - | 0.0078 | - | 0.0075 | - | 0.0099 |
| | | (0.25) | | (0.49) | | (1.10) | | (1.41) | | (0.89) | | (1.09) | | (0.79) |
| … within 120-150 min | - | 0.0019 | - | 0.0022 | - | 0.0084 | - | 0.0059 | - | 0.0025 | - | 0.0079 | - | 0.0012 |
| | | (0.16) | | (0.27) | | (0.40) | | (0.55) | | (0.39) | | (0.44) | | (0.54) |
| Time lag of dependent variable | 0.4510 | 0.4071 | 0.3115 | 0.2630 | 0.3350 | 0.3134 | 0.4305 | 0.3745 | 0.3849 | 0.3502 | 0.3459 | 0.2933 | 0.3377 | 0.2466 |
| | (6.16) | (5.10) | (4.33) | (4.29) | (5.25) | (5.02) | (5.18) | (5.22) | (7.18) | (6.82) | (5.33) | (5.27) | (3.96) | (3.35) |
| R-squared | 0.41 | 0.50 | 0.39 | 0.46 | 0.47 | 0.59 | 0.60 | 0.66 | 0.55 | 0.63 | 0.65 | 0.72 | 0.70 | 0.78 |
| LM test | 7.97 | 6.96 | 6.95 | 0.42 | 5.50 | 0.95 | 5.06 | 0.46 | 5.29 | 0.55 | 4.95 | 0.33 | 7.55 | 0.73 |
| | [0.00] | [0.01] | [0.01] | [0.52] | [0.02] | [0.33] | [0.02] | [0.50] | [0.02] | [0.46] | [0.03] | [0.57] | [0.01] | [0.39] |

Notes:
Dependent variable: employment rate by race group. The number of observations is 3,010 in all cases. Regional dummies are included. Within-group parameter estimates and *t*-ratios in parentheses are reported. LM test: Lagrange multiplier test for a spatial lagged dependent variable, distributed as a chi-squared with 1 degree of freedom with associated probability level in squared brackets. Estimation using SpaceStat v1.93 (Anselin, 1995).

Table 2. Employment and population density
- IV estimates -

| | White | Black Caribbean | Black African | Indian | Pakistani | Bangladeshi | Chinese |
|---|---|---|---|---|---|---|---|
| Own ethnic group population density | | | | | | | |
| …within 30 min | 0.1412 | 0.0890 | 0.0858 | 0.1299 | 0.1355 | 0.1150 | 0.1711 |
| | (1.06) | (4.16) | (4.10) | (4.53) | (3.67) | (4.16) | (4.29) |
| … within 30-60 min | 0.0810 | 0.0345 | 0.0439 | 0.0564 | 0.0410 | 0.0432 | 0.0505 |
| | (0.85) | (3.50) | (3.90) | (4.35) | (3.43) | (3.60) | (3.76) |
| … within 60-90 min | 0.0181 | 0.0201 | 0.0199 | 0.0331 | 0.0151 | 0.0219 | 0.0210 |
| | (0.97) | (3.05) | (2.75) | (3.90) | (2.57) | (2.76) | (2.89) |
| … within 90-120 min | 0.0028 | 0.0055 | 0.0087 | 0.0117 | 0.0059 | 0.0099 | 0.0137 |
| | (0.39) | (1.02) | (1.10) | (1.34) | (0.99) | (0.89) | (0.85) |
| … within 120-150 min | 0.0015 | 0.0049 | 0.0070 | 0.0066 | 0.0010 | 0.0061 | 0.0015 |
| | (0.17) | (0.43) | (0.38) | (0.58) | (0.42) | (0.69) | (0.19) |
| Time lag of dependent variable | 0.5030 | 0.2920 | 0.3333 | 0.3888 | 0.3944 | 0.3101 | 0.3055 |
| | (5.82) | (5.08) | (4.93) | (5.90) | (6.12) | (5.35) | (4.08) |
| Sargan test | 300.55 | 251.88 | 255.76 | 270.60 | 265.65 | 260.07 | 244.61 |
| | [0.40] | [0.97] | [0.95] | [0.84] | [0.89] | [0.93] | (0.98) |
| M_1 | 6.5054 | -11.2306 | -16.422 | 7.9892 | 7.8488 | 7.5004 | -10.0190 |
| | [0.00] | [0.00] | [0.00] | [0.00] | [0.00] | [0.00] | [0.00] |
| M_2 | 0.3468 | -0.4068 | 1.2524 | 0.6666 | -0.5642 | 0.9564 | -1.0416 |
| | [0.7288] | [0.4796] | [0.2104] | [0.5050] | [0.5726] | [0.3388] | [0.2976] |

Notes:
Dependent variable: employment rate by race group . The number of observations is 3,010 in all cases. Arellano and Bond (1991) parameter estimates and *t*-ratios in parentheses are reported. Instruments: lagged values of population density in the area within 30 kilometers; within 600 kilometers; within 90 kilometers; within 120 kilometers; within 150 kilometers; and of the dependent variable up to (t-4); corresponding total land area and regional dummies. Sargan test: Sargan test of overidentifying restrictions, distributed as a chi-squared with degrees of freedom given by the number of overidentifying restrictions, which are 295. M_1 and M_2: tests for first-order and second-order serial correlation in the first-differenced residuals, distributed as N(0,1) under the null of no serial correlation. The associated probability levels of all tests are in squared brackets. Estimation using Ox version 3.0 (Doornik, 2001).

Table 3. Self-Employment and population density by race
- OLS estimates -

| | White | Black Caribbean | Black African | Indian | Pakistani | Bangladeshi | Chinese |
|---|---|---|---|---|---|---|---|
| Own ethnic group population density | | | | | | | |
| …within 30 min | 0.0957 | 0.0539 | 0.0605 | 0.1049 | 0.0878 | 0.0699 | 0.1258 |
| | (1.39) | (2.49) | (3.22) | (3.99) | (3.04) | (3.28) | (4.04) |
| … within 30-60 min | 0.0444 | 0.0188 | 0.0305 | 0.0299 | 0.0204 | 0.0250 | 0.0305 |
| | (1.01) | (2.15) | (2.84) | (2.98) | (2.94) | (2.79) | (3.06) |
| … within 60-90 min | 0.0110 | 0.0101 | 0.0112 | 0.0209 | 0.0105 | 0.0168 | 0.0102 |
| | (0.79) | (1.99) | (2.43) | (2.39) | (2.65) | (2.04) | (3.02) |
| … within 90-120 min | 0.0054 | 0.0045 | 0.0055 | 0.0072 | 0.0023 | 0.0016 | 0.0049 |
| | (0.59) | (0.45) | (0.91) | (0.72) | (0.53) | (0.65) | (0.90) |
| … within 120-150 min | 0.0026 | 0.0017 | 0.0039 | 0.0030 | 0.0011 | 0.0012 | 0.0015 |
| | (0.33) | (0.12) | (0.30) | (0.28) | (0.45) | (0.60) | (0.28) |
| Time lag of dependent variable | 0.4369 | 0.2940 | 0.3003 | 0.3528 | 0.3199 | 0.3017 | 0.2530 |
| | (7.02) | (4.12) | (4.89) | (4.96) | (3.05) | (3.89) | (3.32) |
| R-squared | 0.38 | 0.40 | 0.47 | 0.56 | 0.51 | 0.58 | 0.61 |
| LM test | 6.57 | 0.86 | 0.94 | 0.62 | 0.50 | 0.95 | 0.42 |
| | [0.01] | [0.35] | [0.33] | [0.43] | [0.48] | [0.33] | [0.52] |

Notes:
Dependent variable: self-employment rate by race group. The number of observations is 3,010 in all cases. Regional dummies are included. Within-group parameter estimates and *t*-ratios in parentheses are reported. LM test: Lagrange multiplier test for a spatial lagged dependent variable, distributed as a chi-squared with 1 degree of freedom with associated probability level in squared brackets. Estimation using SpaceStat v1.93 (Anselin, 1995).

Table 4. Employment and population density by race and industry
- OLS estimates -

| | Agriculture & Fishing | | Manufacturing & Construction | | Distribution, hotels & restaurants | | Transport & communication | | Banking, finance & insurance | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Blacks | Asians | Blacks | Asians | Blacks | Asians | Blacks | Asians | Blacks | Asians |
| Own ethnic group population density | | | | | | | | | | |
| …within 30 min | 0.0235 | 0.0189 | 0.0228 | 0.0406 | 0.0597 | 0.0990 | 0.0165 | 0.0377 | 0.0767 | 0.0576 |
| | (2.25) | (2.17) | (2.42) | (3.56) | (4.48) | (3.67) | (3.69) | (2.98) | (3.34) | (3.15) |
| … within 30-60 min | 0.0154 | 0.0109 | 0.0125 | 0.0252 | 0.0321 | 0.0760 | 0.0135 | 0.0250 | 0.0352 | 0.0299 |
| | (2.10) | (2.06) | (2.05) | (2.74) | (3.36) | (2.86) | (2.58) | (2.67) | (3.19) | (2.76) |
| … within 60-90 min | 0.0117 | 0.0086 | 0.0096 | 0.0126 | 0.0219 | 0.0420 | 0.0102 | 0.0120 | 0.0210 | 0.0159 |
| | (1.00) | (0.98) | (0.71) | (1.05) | (2.16) | (2.18) | (1.20) | (1.03) | (3.00) | (2.49) |
| … within 90-120 min | 0.0084 | 0.0034 | 0.0045 | 0.0101 | 0.0150 | 0.0210 | 0.0069 | 0.0072 | 0.0128 | 0.0097 |
| | (0.51) | (0.40) | (0.16) | (0.87) | (1.05) | (1.01) | (0.98) | (0.60) | (2.51) | (2.15) |
| … within 120-150 min | 0.0044 | 0.0019 | 0.0015 | 0.0021 | 0.0060 | 0.0096 | 0.0049 | 0.0019 | 0.0078 | 0.0025 |
| | (0.43) | (0.21) | (0.13) | (0.32) | (0.10) | (0.43) | (0.45) | (0.27) | (1.09) | (0.88) |
| Time lag of dependent variable | 0.4505 | 0.4211 | 0.4230 | 0.3466 | 0.4552 | 0.3306 | 0.3419 | 0.3025 | 0.3109 | 0.2872 |
| | (4.05) | (3.15) | (4.68) | (4.90) | (5.02) | (5.36) | (4.65) | (3.29) | (3.90) | (2.85) |
| R-squared | 0.30 | 0.45 | 0.27 | 0.48 | 0.35 | 0.49 | 0.31 | 0.44 | 0.37 | 0.48 |
| LM test | 0.85 | 0.75 | 0.94 | 0.79 | 0.66 | 0.46 | 0.72 | 0.69 | 0.49 | 0.56 |
| | [0.35] | [0.39] | [0.33] | [0.37] | [0.42] | [0.50] | [0.40] | [0.41] | [0.48] | [0.45] |

Notes:
Dependent variable: employment rate by race group. The number of observations is 3,010 in all cases. Regional dummies are included. Within-group parameter estimates and *t*-ratios in parentheses are reported.
LM test: Lagrange multiplier test for a spatial lagged dependent variable, distributed as a chi-squared with 1 degree of freedom with associated probability level in squared brackets. Estimation using SpaceStat v1.93 (Anselin, 1995).

Table 5. Employment and population density by race: Female only
- OLS estimates -

| | White | Black Caribbean | Black African | Indian | Pakistani | Bangladeshi | Chinese |
|---|---|---|---|---|---|---|---|
| Own ethnic group population density | | | | | | | |
| …within 30 min | 0.0950 | 0.0643 | 0.0605 | 0.0434 | 0.0323 | 0.0352 | 0.0656 |
| | (1.25) | (1.10) | (1.35) | (1.18) | (1.02) | (1.40) | (1.33) |
| … within 30-60 min | 0.0645 | 0.0400 | 0.0299 | 0.0195 | 0.0210 | 0.0215 | 0.0250 |
| | (1.05) | (0.95) | (1.40) | (0.91) | (1.01) | (1.32) | (0.89) |
| … within 60-90 min | 0.0262 | 0.0233 | 0.0107 | 0.0095 | 0.0109 | 0.0089 | 0.0062 |
| | (0.97) | (0.65) | (0.70) | (0.88) | (1.06) | (1.15) | (0.70) |
| … within 90-120 min | 0.0114 | 0.0066 | 0.0038 | 0.0034 | 0.0052 | 0.0062 | 0.0045 |
| | (0.67) | (0.34) | (0.56) | (0.30) | (0.59) | (0.62) | (0.56) |
| … within 120-150 min | 0.0055 | 0.0036 | 0.0019 | 0.0009 | 0.0020 | 0.0008 | 0.0021 |
| | (0.41) | (0.23) | (0.14) | (0.28) | (0.38) | (0.40) | (0.42) |
| Time lag of dependent variable | 0.5454 | 0.4563 | 0.3850 | 0.3238 | 0.3099 | 0.3170 | 0.3940 |
| | (3.16) | (3.71) | (3.96) | (3.19) | (3.20) | (3.20) | (3.36) |
| R-squared | 0.37 | 0.42 | 0.43 | 0.36 | 0.34 | 0.38 | 0.41 |
| LM test | 7.87 | 7.50 | 7.69 | 7.90 | 7.66 | 7.88 | 7.92 |
| | [0.00] | [0.01] | [0.01] | [0.00] | [0.01] | [0.00] | [0.00] |

Notes:
Dependent variable: female employment rate by race group. The number of observations is 3,010 in all cases. Regional dummies are included. Within-group parameter estimates and *t*-ratios in parentheses are reported. LM test: Lagrange multiplier test for a spatial lagged dependent variable, distributed as a chi-squared with 1 degree of freedom with associated probability level in squared brackets. Estimation using SpaceStat v1.93 (Anselin, 1995).