Social Networks and the Dynamics of Labor Market Outcomes: Evidence from Refugees Resettled in the U.S. *

Lori A. Beaman †

November 2007

Abstract

This paper examines the dynamic implications of social networks for the labor market outcomes of political refugees resettled in the U.S. Using a theoretical model of job information transmission within social networks, the paper shows that the relationship between the size of a social network, the vintage of network members and labor market outcomes is non-monotonic. To test this prediction, I use an empirical strategy which exploits the fact that resettlement agencies distribute refugees across cities, precluding individuals from sorting into locations. The results indicate that an increase in the number of social network members resettled in the same year or one year prior leads to a deterioration of labor market outcomes, while a greater number of long-tenured network members improves the probability of employment and raises the hourly wage for newly arrived refugees.

*I am indebted to the resettlement group of the International Rescue Committee for providing access to the data for this paper and for teaching me about the resettlement process. I also thank Joe Altonji, Pat Bayer, Ramona Bruhns, Anne Case, A.V. Chari, Dean Karlan, Fabian Lange, Frank Limbrock, Rohini Pande, Mark Rosenzweig, Petia Topalova, Chris Udry, and seminar participants at several workshops and seminars for valuable feedback. All errors are my own.

†Department of Economics, Northwestern University and RWJ Scholars in Health Policy, UC Berkeley. Email: l-beaman@northwestern.edu.
1 Introduction

Social networks are generally viewed by economists as a partial solution to information problems or other market frictions. However, their effect on labor market outcomes is more difficult to assess. This paper examines one mechanism through which social networks affect labor market outcomes - job information transmission - and shows that changes in network size have heterogeneous effects on labor market outcomes across network members. Using new data on refugees recently resettled in the U.S., the paper provides empirical evidence on the dynamics between changes in the size of a network, the tenure of its members, and labor market outcomes.

The empirical literature has shown that informal search methods play a major role in the labor market: between 30 to 60% of jobs in the U.S. are reported to be found through informal social network contacts (Ioannides and Loury, 2004). Bayer, Ross, and Topa (2005) and Topa (2001) find spatial evidence of network-based job referrals and informational spill-overs, respectively. The existing evidence suggests that information transmission and job referrals within networks improve outcomes. For example, Munshi (2003) shows that among Mexican migrants to the U.S., an exogenous increase in the size of a social network significantly increases its members’ probability of employment and the probability of employment in a higher-wage industry. Edin et al. (2003) also find a positive relationship between the stock of immigrants in a municipality and earnings among refugees in Sweden.¹

Relatively little attention has, however, been paid to within-network competition for job information.² Calvo-Armengol and Jackson (2004) develop a model which suggests that the structure

¹In a related paper, Laschever (2005) provides evidence that social networks affect the post-war employment outcomes of American World War I veterans, using the war draft as an exogenous source of variation in an individual’s social network.

²Wabha and Zenou (2005) provide evidence which suggests competition by showing that among the employed, the probability of finding a job through a social network is concave with respect to population density. Using a model of employer referrals similar to that of Montgomery (1991) as motivation, Munshi (2003) also finds that the positive effect is driven by the number of senior network members and an insignificant, but positive, effect for those who arrived in the previous 3 years. This serves as further motivation for investigating the empirical importance of within-network competition.
of a social network affects persistence in unemployment levels within the network, and that initial differences in employment levels across networks can lead to long-run inequality between groups. They also show that in the short-run, there is a negative correlation in employment outcomes between some individuals within a network since they compete for a finite number of known jobs.

In this paper, I extend their approach to analyze the short-run labor market implications of this competition effect when social network size changes over time. As in Boorman (1975) and Calvo-Armengol and Jackson (2004), I assume that individuals have a random probability of receiving job information, and this information is either used to obtain a job or passed on to a member in the individual’s social network. This basic structure is embedded into an overlapping generations framework.

The model predicts that, depending on the vintage of other network members, having access to a larger network may lead to a deterioration of individuals’ labor market outcomes due to competition among unemployed members for job information. The competition effect arises not because of an increase in labor supply in the face of fixed demand but occurs even when the probability of receiving job information is constant irrespective of network size. This implies a non-monotonic relationship between the size of a social network and labor market outcomes, as a function of network members’ tenure in the market. Changes in social network size will differentially influence labor market outcomes over time: an increase in the size of a given cohort will first decrease the employment rate and average hourly wage of cohorts who arrive close in time to the large cohort, but will improve outcomes for those cohorts that arrive sufficiently later. Competition over job information within the network can therefore, in the short run, mitigate the network’s ability to overcome labor market imperfections.

While the literature has generally found a positive relationship between networks and employment outcomes, the theoretical and empirical results are mixed for wages (Ioannides and Loury,
2004; Mortensen and Vishwanath, 1994; Bentolila et al., 2004). The job information transmission model presented in this paper predicts that wage offers follow the same pattern as the probability of employment. However, a change in network size affects both the number of offers a person receives but also the average wage received since referred wages are lower on average than random wage draws. As a result, the relationship between wages conditional on employment and network size is theoretically ambiguous since these two mechanisms act to offset one another. The empirical analysis therefore evaluates both the impact of an increase in network size on wage offers, including wages of zero for the unemployed, as well as the employed sample.

In order to test these predictions empirically, I compiled a data-set on refugees resettled in the U.S. between 2001 and 2005 using administrative records from the International Rescue Committee (IRC), a large resettlement agency. A unique aspect of the resettlement process is that refugees without family in the U.S. do not choose their destination city: the sample of refugees analyzed in this study comprises of those whose geographic location was selected by the IRC. The empirical strategy exploits this institutional feature. The resettlement process precludes individuals from sorting into localities based on unobservable individual characteristics. Furthermore, all individual characteristics used by the IRC when placing refugees into particular cities are available in the data. This addresses the classic identification problem of sorting: when individuals select their place of residence, it is difficult to distinguish the role of networks from other common unobservable characteristics. In the primary analysis of the paper, a social network is defined as the number of refugees without family already in the U.S. from the same country of origin who are resettled by the IRC in the same city. Network size is therefore uncorrelated with unobserved individual characteristics; variation in the relative size and structure of refugee social networks across cities and ethnic groups over time is then used to examine how networks affect labor market outcomes.

---

3This is one of the "Reflection" problems isolated by Manski (1993).
To address the possibility that the resettlement agency makes placement decisions based on unobserved ethnic group and city level factors, I include in the empirical analysis city-year, ethnic group-year and city-ethnic group fixed effects.

The empirical analysis using refugees resettled in the U.S. shows that an increase in network size has heterogeneous effects across network members, creating both negative and positive ramifications for employment outcomes. I find that a one standard deviation increase in the number of network members who arrive in the U.S. one year prior lowers the probability of employment for a new arrival by 4.9 percentage points. Conversely, as predicted by the model, an increase in the number of tenured network members improves the labor market outcomes for recently arrived refugees. An analogous increase in the number of network members who have two years tenure in the U.S. increases the employment probability by 4.6 percentage points. The analysis also shows that average hourly wages follow the model’s prediction. Among those who are employed, there is a strong positive effect from an increase in the number of senior network members on wages but an insignificant impact from changes in the number of network members who arrived in the current or previous year. This suggests that by providing job information, networks affect wages both through the employment rate but also by affecting job quality once jobs are found.

This paper argues from a methodological standpoint that evaluating the composition of a social network within a dynamic context is necessary to accurately assess the role of social networks in the labor market. A static analysis of network effects, as in using the stock of immigrants as the relevant network measure, is likely to miss important heterogeneity in the way network-based job information flows influence outcomes. In some cases, as demonstrated in this paper, estimating the static effect of an increase in the total size of the network, inclusive of all cohorts, will mask the presence of network effects completely.

In the following section, I present a theoretical framework of job information transmission
within social networks. Details on the institutional background and data are provided in Section 4, and Section 5 discusses the empirical strategy. The results of the empirical analysis on both the probability of employment and wages are presented in section 6. Section 7 addresses alternative explanations for the empirical findings and will exploit the natural experiment of September 11, 2001 to evaluate the response of social networks to an exogenous negative labor market shock. The paper concludes in section 8 with a brief discussion on policy implications for the refugee resettlement process.

2 Theoretical Framework

2.1 A Model of Employment Rates

The theoretical framework is an extension of the model developed by Calvo-Armengol and Jackson (2004) and Boorman (1975), incorporated into an overlapping generations setting. The objective of Calvo-Armengol and Jackson’s work is to show that in the steady-state, there is positive correlation of employment outcomes across time and across all agents within a network. By embedding this model into an overlapping generations framework and analyzing the short-run dynamics from changes in cohort size, I generate concrete predictions which can be tested empirically. To do this, I make the simplifying assumption that all individuals within a network are connected, which eliminates the distinction made by Calvo-Armengol and Jackson between direct and indirect connections.

The basic structure and timing of the model is as follows: each agent lives and works for \( S \) periods, and each cohort \( c \) has \( N_c \) agents. If agent \( i \) in cohort \( c \) is employed at the end of period \( t \), then \( s_{tc}^i = 1 \) and accordingly \( s_{tc}^i = 0 \) if agent \( i \) is unemployed. Since all agents within a cohort are identical, it is preferable to work with the employment rate within the cohort at time \( t \), denoted as
There is a positive probability that any employed agent will lose his job at the very beginning of the period at the exogenous breakup rate $b$. Information about job openings then arrive: any agent hears about a job opening with probability $a$, and the job arrival process is assumed to be independent across agents.

Since each individual receives information directly with probability $a$, the total number of jobs available in the economy is scaled up as the size of the network increases. I therefore assume that the size of the network is small compared to the entire economy. The advantage of this approach is that it enables the model to isolate the network effect directly. This assumption also reflects the empirical setting in which the predictions will be tested. As shown in table 1, the average cohort size in the sample of refugees used in this paper is less than 30. Since the resettlement locations are medium-sized cities, including cities such as Atlanta, Phoenix, and Salt Lake City, a change in the number of refugees arriving in each city in a given year is unlikely to have any general equilibrium effect on the job arrival rate or the distribution of wages.

If an agent is unemployed and receives job information, he will fill the position. However, if the agent is already employed, he will pass along the information to a randomly selected network member who is unemployed. Job information is shared with equal probability to any unemployed member in the network, regardless of which cohort he belongs to. Once job information arrives and is referred to unemployed members where suitable, jobs are immediately accepted.

This structure can be formalized in the following way:

$$s^t_c = a + r^t$$ \hspace{1cm} \text{if } c = t \hspace{1cm} \text{(1)}$$

$$s^t_c = (1 - b)s^{t-1}_c + (1 - (1 - b)s^{t-1}_c)(a + r^t) \hspace{1cm} \text{if } c \leq t \leq c + (S - 1) \hspace{1cm} \text{(2)}$$

$$r^t = (1 - b) \sum_{k=t-S+1}^{t-1} N_k s^{t-1}_k \frac{a}{\sum_{k=t-S+1}^{t-1} N_k - (1 - b) \sum_{k=t-S+1}^{t-1} N_k s^{t-1}_k} \hspace{1cm} \text{(3)}$$
where $r^t$ represents the probability of receiving job information through an employed network member. The probability of being employed for an individual entering the market is simply the probability of receiving job information directly, $a$, plus the probability an already employed network member passes him information, captured by the term $r^t$. The probability of receiving job information from the network, $r^t$, is the total number of jobs which are available in the network to be passed, divided by the number of potential recipients. The number of available jobs is the number of employed individuals in the network multiplied by the probability that each receives job information randomly. The number of potential recipients the number of individuals who are unemployed at the beginning of that period, after the exogenous break-up has occurred. For cohorts who have previously been in the market, the probability of being employed is the probability of having a job in the previous period and keeping it, at rate $1 - b$, plus $s^t_c$ weighted by the probability of being unemployed. This simple model can be used to show a couple of predictions which can be tested empirically.

**Proposition 1** For all values $0 < a < 1$ and $0 < b < 1$, an increase in cohort size $N^i_j$ decreases $s^j_c$ for all $c$.\(^4\)

*Proof:* See appendix.

The intuition is that since $s^t_c^{-1}$ does not change, increasing $N^i_j$ only increases the number of unemployed individuals seeking job information from network members while leaving the number of employed members unchanged. The result is a decline in the employment rate for both the cohort which is made exogenously larger and all other cohorts in the market in that period. The effect is present despite the fact that any individual in the market has the same probability of hearing about a job directly as before. This highlights the fact that this competition effect arises from the dynamics within the network since the larger market conditions remain constant.

\(^4\)This claim holds for all values of $a$ and $b$ such that $s^t_c \neq 1$ for all $c$ and $i$. 

8
**Proposition 2** The impact of an increase in $N_j$ on $s^k_j$ is monotonically increasing between $k = j$ and $j + S - 1$.

**Proof:** See appendix.

Proposition 1 shows that the initial effect of increasing cohort $j$ is to decrease the employment rate for that cohort. Proposition 2 shows that the impact on the subsequent cohort, $j + 1$ is less negative than on $j$ itself. Likewise, \( \frac{\partial s^{j+2}_{j+2}}{\partial N_j} > \frac{\partial s^{j+1}_{j+1}}{\partial N_j} > \frac{\partial s^j_j}{\partial N_j} \) for the case when $S = 3$. The idea is that as cohort $j$ gains experience in the labor market, its employment rate rises. Its negative impact on employment is then mitigated over time.

Numerical analysis of the model shows \( \frac{\partial s^k_k}{\partial N_j} > 0 \) for at least $k = j + S - 1$ and usually earlier cohorts. That is, an increase in the size of cohort $j$ first negatively impacts cohorts who arrive close in time to period $j$ but then increases the employment rate for cohorts who arrive sufficiently later.\(^5\) As cohort $j$’s employment rate increases over time, its larger size becomes an asset to the entire network. To illustrate this and the predictions outlined in Propositions 1 and 2, Figure 1 provides an example where $a = .35; b = .2$. The graph shows a comparison in the employment rates of a control network with constant cohort size and that of a treatment network in which the size of cohort $c$ is doubled. All subsequent cohorts after $j$ have the same size as the control cohort.

Each agent is in the market for 4 periods. The treated cohort, $j$, experiences a lower employment rate in their first period in the market, but by period 4, the larger cohort size leads to a slightly higher employment rate. $s^{j+1}_{j+1}$ is represented as “Cohort $j+1$” in time period 1 in Figure 1. Similar to the pattern displayed by cohort $j$, the initial employment rate is lower than it would have been in the absence of the cohort size shock, but this effect is largely gone by cohort $j + 1$’s second period in the market. In fact by time period 3, the cohort reaches a higher employment rate than the counterfactual cohort. The following cohorts, $j + 2$ and $c + 3$, both receive gains in the employment

\(^5\)The positive effect on later cohorts, according to the numeric analysis, holds for all parameter values even though the analytic results only show a monotonically increasing effect.
rates for all 4 periods these cohorts are in the market.

In the simple model described above, job information arrives to all agents in the economy at the same rate regardless of their current labor market status. The model can be made more general to allow for the job information arrival rate to depend on employment status. I examined this more general model both when the arrival rate is higher for the already employed and when the unemployed are more likely to exogenously receive job information. Propositions 1 and 2 are robust to these alternative specifications.\footnote{The alternative specification can be formalized in the following way:}

### 2.2 A Model of Employment Rates and Wages

Subsequent work by Calvo-Armengol and Jackson (2007) analyzes a more general model which includes stochastic wages. In this model, job information that arrives exogenously also includes a wage. This leads to different behavior than in the above model. An employed individual will now switch jobs if he receives job information with an offer wage higher than his current wage. The implications of this more general framework is that in the steady-state, information passing leads to positive correlation between the employment and wage status of agents who are connected by a social network. There is again, though, the possibility of a negative correlation in wages across certain agents due to within network competition.

I incorporate wages into the overlapping generations framework used above in the following

\[
s_c^t = a_1 + r^t \quad \text{if } c = t
\]

\[
r^t = (1 - b) \sum_{k=-S+1}^{t-1} N_k s_k^{t-1} - \frac{a_2}{\sum_{k=-S+1}^{t-1} N_k} \sum_{k=-S+1}^{t-1} N_k s_k^{t-1}
\]

where \(a_1\) represents the arrival rate for unemployed individuals while \(a_2\) is the arrival rate for the employed.

The two cases of interest are when \(a_1 > a_2\) and when \(a_2 > a_1\). The former represents the situation when search intensity is higher among the unemployed. By contrast, \(a_2 > a_1\) may occur is job information is generated primarily through employee referrals. Numerical analysis shows that an increase in \(N_j\) continues to result in a decline in \(s_j^t\), with the impact on subsequent cohorts monotonically increasing until it is positive. An increase in \(a_2\) has a larger impact on increasing employment rates than an analogous change in \(a_1\). Whether \(\frac{\partial s_j^k}{\partial N_j}\) is larger, in absolute terms, when \(a_2 > a_1\) or \(a_1 > a_2\) depends on the parameter values examined.
way: with probability $a$, an individual receives job information which now also contains a wage. If the individual who receives the job information is unemployed, he takes the job. However, if the individual is employed, he accepts the job if $w_{ict}^o > w_{ict}$, where $w_{ict}^o$ denotes the offer wage from the new job information received by employed individual $i$. Alternatively if $w_{ict}^o < w_{ict}$, the offer is passed to a randomly selected unemployed network member.\(^7\) Wages are iid draws from the uniform distribution $w \sim U[w, \bar{w}]$. $w_c^e$ denotes the average wage for employed network members in cohort $c$ in period $c$.\(^8\)

The analysis of the model is done by simulation. I present here one numerical example. Figure 2 reflects the results of simulating the model with $a = .40$, $b = .05$ and agents working in the market for 5 periods, i.e. $S = 5$. Wages are distributed $w \sim U[5.15, 45.15]$, where $\bar{w} = 5.15$ reflects the minimum wage law. The thought experiment here is to triple cohort size $N_j$ and evaluate the effect on employment rates and wages of cohorts $j$, $j + 1$, $j + 2$, and $j + 3$. As in Figure 1, all cohorts except $j$ are the same size. Both the employment rates and the average wages of cohorts $j$ and $j + 1$ are lower in the first period than the levels that would have been achieved under the counterfactual. The effect on cohort $j + 2$ in its first period in the market, however, is close to zero while cohorts $j + 2$ and $j + 3$ show initial gains from the increase in cohort $j$.

In this model the effect on wages and employment is more subtle than in the simpler model with constant wages. For a given employment rate, the job information available in the network for unemployed members is diminished, since employed network members with low wages are unlikely

---

\(^7\)This could produce an incentive for network members to behave strategically. It may be preferable for an unemployed network member to refuse the job offer and wait another period to be eligible for job information from other network members. This assumption can, however, be relaxed so that information is passed randomly to anyone in the network. In this case, the magnitude of the impact from a change in network size decreases rather sharply but still demonstrates the pattern of first increasing competition then improving outcomes for later entering cohorts.

\(^8\)The search literature has argued that either pareto or exponential distributions are appropriate as wage distributions (Lancaster and Cheshire, 1983; Lynch, 1983). I therefore tested the sensitivity of the model predictions by assuming wages are exponentially distributed with $w \sim E(\lambda = 3) + 5.15$. Analysis of all other parameter values show consistent results as with the uniform distribution: the prediction holds for the employment for all parameter values, and wages conditional on employment continue to be ambiguous for certain values, despite the support of the wage distribution being much smaller.
to provide information to the unemployed. The only jobs which are passed are those with sufficiently low wages that the employed network member who initially receives the job information rejects the offer. This also implies that individuals who become employed through passed job information will have wages that are lower on average than those who become employed by receiving job information directly.

There are therefore two effects working in opposite directions on the average wages of the employed. The primary effect of an increase in network size is to change the number of job offers an individual receives, thereby affecting the wage. However, the offsetting effect is due to changes in the proportion of individuals who receive their jobs directly versus indirectly, i.e. through the network. Since these jobs have wages which are on average lower, this can counteract the primary effect.

There is thus the possibility that wages do not follow the same pattern as in the above example for all parameter values. An increase in $N_j$ may lead to a lower proportion of jobs being attained through the network. Since these jobs have lower wages on average, the average wage of those who are employed may actually increase. It is possible to have within-period increases in the wage due to an increase in $N_j$.

There is therefore no general prediction with regards to wages of the employed. An analogous Claim to the Propositions shown for the model with constant wages does hold, however, for employment rates and wages in the entire network.

**Claim 1** For some values of $(a,b)$ and an increase in $N_j$, there exists $\tilde{k}$ such that $\forall k \leq \tilde{k}$, $\frac{\partial w_k}{\partial N_j} < 0$ and $\frac{\partial s_k}{\partial N_j} < 0$. For $p > \tilde{k}$, $\frac{\partial w_p}{\partial N_j} > 0$ and $\frac{\partial s_p}{\partial N_j} > 0$.

where here $w^j_j$ represents the average hourly wage of the entire network, including a wage of 0 for those who are employed. The offsetting effect from changes in the composition of network members obtaining jobs through direct versus indirect channels is not strong enough to change the prediction
regarding wages of the entire network.

The model therefore predicts that employment and wage rates will be inversely correlated with the number of recently arrived refugees, but positively correlated with the number of senior network members.\textsuperscript{9} The striking part of the prediction is that the deleterious effect from an increase in network size does not come from an increase in labor supply with a fixed labor demand. Instead, the negative effect comes from competition between network members for information provided by already employed individuals. It is a within-network information competition effect and not a result of an increase in labor supply driving down wages or employment rates in equilibrium. The assumption that each individual faces a constant rate \( a \) of hearing about a job directly ensures that labor demand is held fixed, so that the latter effect is not driving the model prediction.

3 Institutional Environment and Data

3.1 Refugee Resettlement Process

The United States has a long history of refugee resettlement, having accepted around 2.4 million refugees and asylees since 1975. In 2005, 70,000 refugees were authorized for admission to the U.S. compared with the 55,000 immigrants who were permitted entry in 2005 through the diversity lottery system. Refugees come from a wide variety of countries and flee their homes for different reasons, from war-related violence to religious persecution to retribution for political views. The process through which refugees gain access to the U.S. creates a unique opportunity to look at the role of ethnic networks. Limited research has looked at the economic performance of refugees

\textsuperscript{9}A key difference between the theoretical model presented in section 2.1 and this model is how efficient the network is in passing along job information. The main prediction for employment rates holds in the two extreme cases where: information is passed until it arrives at an unemployed individual (section 2.1) and when information is only passed once and discarded immediately if unused (as described here). The model predictions also hold in a model with a moderate level of efficiency, where individuals who are already employed and receive an offer with a higher wage will pass on their discarded jobs to unemployed network members. The key difference with this formulation is an increase in the average employment rate.
in the U.S., largely due to data constraints. Refugees are a well-defined group: according to the Immigration and Nationality Act (INA) Section 101, they are individuals living abroad who have a well-founded fear of persecution based on race, religion, nationality, social group status or political opinion in their home country. This study does not look at asylees, i.e. those who travel by their own means to the United States and then apply for protected status upon arrival.

How does one become a refugee? Total refugee admissions levels and processing priorities are set each year by the president, in consultation with Congress. INS (currently USCIS) officers adjudicate individual cases in refugee processing centers around the world; most often these centers are within refugee camps, but individuals can also apply for refugee status in local U.S. embassies. Once the INS designates an individual as having refugee status, the Bureau of Population, Refugees, and Migration (PRM) is responsible for overseas processing and transportation to the U.S.

The PRM’s final role in the resettlement process is to allocate all accepted cases to one of ten contracted voluntary resettlement agencies. The resettlement agencies are responsible for acquiring housing, providing initial benefits including cash assistance and in-kind support, as well as providing access to resources such as ESL training and job assistance. I use data from one voluntary resettlement agency: the International Rescue Committee (IRC). In this paper I look specifically at individuals who are granted refugee status directly, excluding both asylum seekers and refugees who attained admittance via family reunification. For these individuals, the IRC has the sole discretion in determining where the refugee will be resettled among its 16 regional offices. The IRC receives information from the State Department about each refugee’s characteristics, including basic information such as country of citizenship plus demographic information including age, gender, marital status and education. With this information, the IRC decides to send each refugee or refugee family to one of its 16 regional offices. It is important to note that no IRC

\(^{10}\) Two exceptions are Cortes (2004) and Borjas (2000). Cortes (2004) argues that refugees perform worse relative to other immigrant groups in the short-run but eventually surpass the other groups. Gould, Lavy, and Paserman (2004) use flows of Ethiopian refugees into Israel to estimate the impact of school quality on educational outcomes.
employee meets the refugee or his family members until the allocation process has been completed, which is generally within one week of the State Department contacting the agency. The refugee travels directly from his place of residence overseas to the chosen IRC regional office within the U.S.

3.2 Placement Policy

The IRC does not have an explicit placement rule when distributing refugees across regional offices, although they do follow a few general guidelines. First, the IRC seeks to place refugees in locations where there is the presence of a pre-existing ethnic or nationality-based community. They also attempt to choose a regional office based on language competencies. The goal is to send each refugee to an office which has either a staff member of a volunteer who speaks the same language as the refugee. Individual refugees or refugee families who have special medical problems, such as HIV, are only sent to particular offices which specialize in such cases.

In addition to policies oriented towards achieving a good match between an individual refugee and a city, the IRC also budgets for the total number of refugees expected to arrive in each regional office. To do this, each regional office is budgeted a total number of people per year plus a target for refugees who do not have family already in the U.S. at the time of arrival, designated as non-family reunification refugees. These numbers are estimated using projected numbers from the State Department on how many refugees are expected to be admitted to the U.S. from each region of the world. Often the actual numbers can vary substantially from those anticipated, however, since the actual number of refugees who arrive from a region can be volatile. There is also a great deal of uncertainty about the number of family reunification cases arriving each year. Since family reunification cases are predestined for particular offices, this shifts the allocation of non-family reunification cases and often the total number of refugees sent to each city away from budgeted
numbers. Finally, the overall number of refugees sent to a particular office is also a function of employment statistics at the regional office level.

As for the remaining information provided to the IRC by the PRM, the IRC reports using a limited amount of this information in the allocation process. Given that this is difficult to verify, the data set used in this analysis fortunately includes all information given to the IRC prior to each refugee’s arrival. In fact, the data was compiled from the very forms provided to the IRC from the PRM. I can therefore control for individual characteristics which the IRC uses in the allocation process. This is important since it removes the problem of sorting based on unobserved characteristics which exists in other studies estimating social network effects.\footnote{Bertrand et al. (2000), for example, evaluate the role of networks in welfare participation using similar empirical strategy with neighborhood and language group fixed effects, but there remains the possibility of differential selection into MSAs based on unobserved preferences.}

\section*{3.3 Data}

The data from the IRC comprises of over 1,700 male adults who arrived in the U.S. between 2001 and 2005.\footnote{There are three groups whose placement do not follow the above guidelines due to special circumstances, the Meskhetian Turks, the Somali Bantu and the Kakuma Youth, and are therefore excluded from the sample.} All sample respondents did not have family members already in the U.S. to assist in their resettlement, and the IRC therefore placed all of these individuals using the placement policy described above. There are three components to this data. A fairly rich set of demographic variables which were compiled by the INS and the PRM prior to the refugee’s arrival in the U.S. is available, including ethnicity, date of birth, country of first asylum, the size of the family being resettled, initial English language level and education received in the home country. This data is comprehensive of all individual characteristics known by the IRC at the time of placement and retrieved from the forms the IRC received from PRM. Labor market outcomes, in particular employment status and hourly wage, were collected by the IRC at 90 days after each refugee’s arrival. Finally, data on the total number of individuals (inclusive of all ages) placed in each
IRC regional office by nationality from 1997 through 2004 were retrieved from archived aggregate reports. Unfortunately, individual-level data prior to 2001 are currently unavailable.

There is a wide variety of ethnic groups and nationalities in the data. The largest groups are from Afghanistan, Bosnia, Liberia, Somalia, and the Sudan, although there are in total 38 different ethnic groups represented. The IRC has 16 offices where they resettle non-family reunification cases.\textsuperscript{13} The sample excludes those refugees who are HIV positive, which are less than 1% of the sample, since these refugees spend a substantial portion of their initial 90 days under medical supervision.

In order to get an estimate of the size of each ethnic group’s network in a given geographic space, I will be using two different measures. The primary analysis will define the social network as non-family reunification refugees from the same nationality who were resettled in the same regional office. Since the aggregate data is available from 1997 onwards, this measure of network size for an individual will include fellow refugees resettled in the four years prior to that individual’s arrival. The relevant network is defined to include only those individuals without family in the U.S. prior to their arrival. The reason for this restriction is twofold. First, while not modelled explicitly, an incentive for participation in the network is insurance: even if an individual is employed now, there is a positive probability of becoming unemployed in future periods and may then rely on the network to gain a job. Refugees with family members who are already established in the U.S. would need to depend less on the social network formed by refugees who largely have not known each other for more than 90 days. The second reason is that the resettlement experience is different across these two groups. Family reunification refugees can be located far away from the regional office but still “resettled” by the IRC. By contrast, since the IRC rents an apartment for each non-

\textsuperscript{13}The offices are: Abilene, TX, Atlanta, Baltimore, Boston, Charlottesville, Dallas, New York, New Jersey, Phoenix, Salt Lake City, San Diego, Seattle, Tucson, Washington DC, and Worcester, MA. Atlanta, Baltimore, Dallas, Phoenix, and Salt Lake City are the largest.
family reunification refugee, they tend to be clustered together spatially. Moreover, the two types of refugees are less likely to interact since family reunification refugees receive less resettlement services from the IRC and are accordingly less likely to meet fellow refugees in the office or at IRC-sponsored events. The data on the number of family reunification refugees resettled during this time period will also be used in the econometric analysis, as discussed in section 6.

The second measure of the size of the social network comes from the 2000 Census data available through IPUMS. I calculate the size of the network at level of the metropolitan statistical area (MSA). I define a social network by either nationality or ethnicity, depending on the availability of the relevant code in IPUMS.\textsuperscript{14} IPUMS data also provides the age of each network member as well as the year of arrival in the U.S. Therefore I can create a network size variable which is specific to the year of arrival of the network members. The information on age also allows me to restrict the network to only prime age adults. Since the Census does not obtain information on the foreign born’s visa type or residency status/citizenship, this measure will include all immigrant types, ranging from illegal immigrants to permanent residents and naturalized citizens.

Supplemental information is also available from a survey of refugees and asylees collected by the Department of Health and Human Services’ Office of Refugee Resettlement (ORR) between 1993 and 2004. There is no information available in the data indicating which refugees were family reunification cases, however, so this sample may therefore not be precisely comparable to the IRC sample.

4 Empirical Strategy

The primary objective of this paper is to empirically test the predictions of a simple model of job-related information flows in social networks. The model corresponds nicely to the empirical setting.

\textsuperscript{14}For example, I can identify some ethnic groups which cover multiple countries, such as the Kurds.
Propositions 1 and 2 predict that having a larger number of network members who arrived in the same year, corresponding to the \( N_c \) cohort, will decrease the probability of a new refugee obtaining employment. An increase in the number of refugees who arrived the year prior, analogous to a change in \( N_{c-1} \), will impact employment outcomes more positively than the effect from an increase in \( N_c \). In particular, a negative impact from an increase in \( N_c, N_{c-1} \) and a positive effect from \( N_{c-2} \) and \( N_{c-3} \) would be consistent with the model. Wages should exhibit the same differential pattern across network cohorts.

Using labor market outcomes as of 90 days after arrival and the aggregate data on IRC placements from 1997-2005, the model predictions will be tested using the following econometric specification:

\[
Y_{ijkt} = \alpha + \gamma_1 N_{ijk(t)} + \gamma_2 N_{jk(t-1)} + \gamma_3 N_{jk(t-2)} + \gamma_4 N_{jk(t-3/t-4)} + X_{ijkt} \beta + \delta_{jt} + \phi_k + \epsilon_{ijkt}
\]

for individual \( i \) in country of origin \( j \) in city \( k \) who arrived at time \( t \). \( Y_{ijkt} \) represents either employment status or wages for individual \( i \). \( N_{jk(t-1)}, N_{jk(t-2)} \) are the number of refugees who arrived during the fiscal year one year, and two years prior to refugee \( i \)'s arrival. \( N_{jk(t-3/t-4)} \) is analogously the number who arrived three and four years prior. Therefore the network variables are the same for all refugees who arrive in the same fiscal year, are resettled in the same regional office and share the same country of origin/ethnicity. \(^{15} \) \( N_{ijkt} \) is the number of refugees from country of origin \( j \) resettled by the IRC in regional office \( k \) who arrived in fiscal year \( t \) up to \( i \)'s specific date of arrival. Those individuals who arrived after \( i \) are excluded from \( N_{ijkt} \) since they would be not be acting as competitors nor providing job information to individual \( i \). \(^{16} \) Negative point estimates

---

\(^{15} \) Since network size comes from aggregate data, this measures is the total number of refugees by nationality including children. Unfortunately, without individual records prior to 2001, this is as precise measure as available.

\(^{16} \) For example, an individual who arrived in December could not influence the 90 day labor market outcomes of a refugee who arrived in January. An alternative measure would be the number of refugees who arrived during same year of arrival up until \( i \)'s date of arrival plus 90 days.
of $\gamma_1$ and $\gamma_2$ and positive estimates of $\gamma_3$ and $\gamma_4$ would be consistent with the model.

The difficulty in consistently estimating $\gamma_1$, $\gamma_2$, $\gamma_3$ and $\gamma_4$ is separating the network effect from common unobservable attributes which are shared by members of a group. Networks which are defined by group identity and geography are particularly susceptible to bias from sorting. If individuals choose their locations based on factors which are not observable to the econometrician, and these factors are common among group members, then it is difficult to separate the effect of having a larger number of network members in a city on labor market outcomes from the impact of common characteristics network members share. In the case of refugee resettlement, the institutional environment provides a strategy to mitigate this problem of correlated unobservables.

There are two main threats to identification in this environment. The first originates from sorting along unobservable individual characteristics. The second class of correlated unobservables is omitted city and ethnic group characteristics. The first is addressed by including a flexible functional form to span the information set available to the IRC at the time of placement. In this case, $X_{ijkt}$ captures individual characteristics which are correlated with network size. The remaining individual attributes in $\epsilon_{ijkt}$ can not be a source of bias since they are not known by the IRC at the time of placement and accordingly are uncorrelated with $N_{jk}$.

Since the IRC resettles multiple ethnic groups across multiple cities, there is variation in social network size across cities, ethnic groups and over time. This facilitates a fixed effects strategy to minimize the unobservable factors common among ethnic groups or within cities which are correlated with network size. In my preferred specification, I include only $\delta_{jt}$ and $\phi_k$ controls. Time variant heterogeneity at the nationality/ethnic group level is captured by $\delta_{jt}$. Thus if one particular ethnicity has lower human capital on average or if the types of people who become refugees vary across sending countries, these common factors within a group are captured. By varying across years of arrival, this term also allows there to be unobservable differences within a
group across cohorts. Comparing labor market outcomes across cities is a challenge since there are numerous differences, such as macroeconomic variation, which are difficult to quantify completely. If the IRC uses this information when allocating refugees across cities, then these factors would be correlated with network size. Therefore unobservable factors at the city level, and affect all ethnic groups equally, are controlled for using metropolitan-area fixed effects, $\phi_k$. There are, however, three additional sources of bias. There may be shocks at the city level, denoted as $\xi_{kt}$, a match quality between ethnic groups and cities, $\tau_{jk}$, and finally shocks to match quality, $\nu_{jkt}$. The first two can be addressed through fixed effects and will be discussed further in section 5.1.1.

Could systematic variation in factors which affect particular ethnic groups in certain cities generate a relationship between network size and labor market outcomes as predicted by the model? In order to do so, there must be a very special sequence of shocks which is identical in all comparative advantage industries in all cities. Furthermore, these shocks must be known and anticipated by the IRC in order to be correlated with the time variation in network size. This is unlikely to occur. During extensive discussions with the IRC, they stated that they believe that all refugees can succeed in each of their resettlement locations, and therefore do not make allocation decisions to maximize match quality. There is also a stochastic time lag between when a refugee is assigned to the IRC by the State Department, and accordingly allocated to a regional office, and the actual arrival date. In some cases, there can be a full year between these two events. This would make it extremely difficult for the IRC to exploit time variant shocks of this nature. More concretely, if this was the strategy used by the IRC, we would observe the flows of refugees from a particular group to exhibit an oscillating pattern. Table 2 shows that the number of refugees from a nationality group who are resettled in the same city are positively correlated across all four year time periods. The table presents the correlation matrix of the number of people the IRC allocated to each nationality/regional office pair, i.e. the size of each cohort across 4 year periods from 1997-2005.
The strongest correlated is between time $t$ and $t-1$ and is thereafter monotonically decreasing in the time elapsed between cohorts. A correlation between $\nu_{jkt}$ and network size therefore unlikely to generate estimates of $\gamma_1$, and $\gamma_2$ which are negative and positive estimates of $\gamma_3$ and $\gamma_4$ under the null hypothesis of no network effects.

The employment analysis will be done using a linear probability model. There are a large number of categorical control variables to capture the unobserved heterogeneity across ethnic groups and cities, and an LPM is therefore easy to estimate. Furthermore, since the mean employment rate is 64%, the LPM should perform well. The error term is clustered at the nationality group/regional office level.

5 Empirical Results

5.1 Probability of Employment

5.1.1 Main Results

To begin the analysis of the effect of networks on labor market outcomes among refugees resettled in the U.S., I follow much of the existing empirical literature and estimate the effect of the stock of network members on the probability of employment. Columns 1 and 2 of table 3 show that this analysis leads to puzzling results. In columns 1 and 2, an increase in the number of refugees from country $j$ resettled in city $k$ from years $t$ though $t-4$ increases the probability of employment for a new arrival. This specification includes nationality-year, and city controls. However, once city-nationality and city-year controls are included, the effect becomes insignificant and the point estimates are negative. In this case, the analysis would be inconclusive regarding the existence of social networks providing job information to newly arrived refugees.

By contrast, the results from analyzing employment in a dynamic context consistently
confirm the predictions of the information transmission model. Columns 3 through 6 of table 3 show that a larger number of network members who arrived in the current and prior year strongly decreases the probability of employment for a new entrant. A one standard deviation increase in $t - 1$ network size decreases the probability of employment by 4.9 percentage points. Given that the mean level of employment in the sample is 64%, this constitutes a decline of over 7%. Analysis done with the ORR data shows that an additional year in the U.S. lead to an increase in the employment rate of 3.4%. Therefore this negative network effect is an economically significant factor in determining refugee labor market unemployment. In contrast to the first two columns in the table, this component is lost in an analysis of the stock of the network, instead of looking for the dynamic component of social network behavior. The approach used in this paper therefore sheds new light on how networks function and affect on the labor market outcomes of network members.

As is consistent with the model, however, a larger number of refugees with two to four years of experience living in the U.S. prior to a new refugee’s arrival has a positive and statistically significant effect on employment. The number of refugees resettled in year $t - 2$ has the largest effect on the probability of employment. In this case a one standard deviation increase in $t - 2$ network size raises the probability of employment by 4.6 percentage points. In this specification, the number of refugees who arrived in the prior 3 and 4 years are combined, and the estimates of $\gamma_3$ and $\gamma_4$ are jointly significant at the 5% level. One-sided tests indicate that while $\hat{\gamma}_1$ is not statistically more negative than $\hat{\gamma}_2$, $\hat{\gamma}_2 < \hat{\gamma}_3$ as predicted by the model. It is fairly surprising that the coefficient on the number of refugees who arrived in years $t - 3$ and $t - 4$ is smaller than that of the $t - 2$ network, although it is still positive and statistically significant. One reason for this is that out migration is likely to be higher for refugees who had been resettled 3 or more years prior to the new arrival.\(^{17}\) Out migration within the first 90 days is 6.96%, and therefore it is quite

\(^{17}\)The data used to measure network size is the total number of refugees who were placed in a given city in a given year, and I do not observe whether those individuals continue to live in their initial location.
plausible that the smaller coefficient reflects the fact that this variable has more measurement error in representing the true number of network members currently available to the new arrival. Attenuation bias would then push down the size of the coefficient compared to that of the $t - 2$ cohort. An alternative explanation, although less likely, is that if refugees remain in the social network for only 3 years, then the effect from an increase in the number of refugees who arrived 4 years prior would have a positive but smaller effect on 90 day outcomes.

The coefficients on the control variables are as expected, although the interpretation is unclear given that the coefficients are a mixture of the causal relationship and the selection rule used by the IRC. Age displays a concave relationship with the employment rate, increasing at a decreasing rate. Household size is negative, reflecting that a larger household may also contain more potential workers, thereby diminishing the incentive to work or providing the opportunity to invest in human capital for any given individual.

Columns 3 and 4 of table 3 includes city, and nationality group-year indicator variables. If individuals are able to influence the way the INS adjudicates individual cases or the timing between when an individual is granted refugee status and when he is allowed to travel to the U.S., there could be differences in the quality of cohorts over time. If the IRC is aware of these differences and changes its placement policy accordingly, then network size would be correlated with the error term. For example, individuals who enter the U.S. after a large cohort may differ in an unobservable way from those who enter at the same time as the large cohort. In that case, $\tau_{jt}$ would be correlated with $N_{jk(t)}$ for all $t$. To address this concern, all estimates include fixed effects for nationality-year.

The specification estimated in column 3 contains a limited number of demographic covariates. There are accordingly individual characteristics which may have been used by the IRC when choosing an individual’s location, thereby being in $\epsilon_{ijkt}$ and correlated with network size. Column 4 addresses this potential concern by including a wider range of individual characteristics. The
estimates of $N_{jk}(t)$ for all $t$ are robust to the inclusion of these variables as can be seen in column 4 of table 3. The coefficients are largely unchanged and continue to be significant.\textsuperscript{18}

To return to the correlation matrix presented in table 2, is the covariance structure between the network variables themselves causing the observed pattern? The positive correlation across all periods indicates that this is not the case. Further confirmation that the correlation between the network variables is not creating a spurious pattern will be shown in section 6.3 where an alternative measure of network size is used.

An alternative hypothesis and a common problem in identifying network effects based on geographic variation is that there may be city-ethnic group specific matches which are both unobserved and correlated with network size. This would arise if, for example, there are characteristics or skills which are common to all individuals in ethnic group $j$ which receive a higher return in particular cities $k$. Thus if particular immigrant groups receive differentially higher wages in a particular city and the IRC uses this information while making decisions on how to distribute refugees, network size would be endogenous. There are two reasons this is unlikely to be the case. First, the IRC themselves state that they take no position on whether certain cities are preferable for particular ethnic groups. Not unlike the canonical pool player subconsciously calculating angles while playing, though, this alone does not definitively rule out the possibility that there are not other unobservable characteristics of a city-ethnic group pair which are being used to determine placement. The second argument is that a comparative advantage story alone would not generate a negative $\hat{\gamma}_1$ and $\hat{\gamma}_2$ and positive $\hat{\gamma}_3$ and $\hat{\gamma}_4$. It would create a uniformly upward bias, not a differential effect between recently arrived refugees and tenured refugees.

The structure of the data also allows a richer set of fixed effects than used in columns 3

\textsuperscript{18} Another placement criteria used by the IRC is to send each refugee to an office where someone can speak the same language as the refugee. Including two discrete variables indicating whether the refugee was placed in an office with either a staff member or a volunteer who speaks at least one of the languages spoken by the refugee does not result in significant changes to the coefficients of interest. Having a volunteer who speaks the same language positively impacts employment while there is no impact from staff members.
and 4 of table 3. Specifically I can include nationality-city, nationality-year and city-year controls since the network variables varies at the nationality-city-year level. The results of this specification are shown in columns 5 and 6 of table 3. Despite the large number of additional controls this requires, column 5 shows that the results are robust to this richer specification.\textsuperscript{19} The estimates are in fact larger than in the specification used in columns 3 and 4, although only the coefficient on $N_{jk(t)}$ is statistically different. One reason for this change is that controlling for city-year and nationality-city year better capture group and city-specific resources available to refugees from the IRC during their first 6 months in the U.S. Column 6 includes the full set of control dummies plus a wider set of individual characteristics. In this case, $\hat{\gamma}_4$ is statistically insignificant but qualitatively similar in size. $\hat{\gamma}_3$, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ remain strongly significant and of the same sign and magnitude.

This specification removes the possible bias originating from time invariant unobserved city-ethnic group match quality. The city-year dummy variables also remove the possibility that city-level employment shocks are influencing the estimates. These additional controls help to identify the causal effect of network size on employment as long as there are not year-specific shocks which vary at the city-ethnicity level and are used by the IRC to determine placement.

In contrast to the analysis of the stock of network size shown in columns 1 and 2 in table 3, the results are not sensitive to the inclusion of additional fixed effects. This highlights the problem in that estimation. By properly structuring the network variables to reflect the dynamic relationship between network size and labor market outcomes, the presence of network-based job information transmission is easily detected and not sensitive to the specification used.

\section*{5.1.2 Robustness Analysis}

\textbf{Falsification Test}

\textsuperscript{19}There are 198 nationality-city, 72 city-year and 115 nationality-year pairs.
The structure of the data also facilitates a falsification test using the 2001-2004 sample.\textsuperscript{20} I test whether the number of refugees who arrive in year $t + 1$ impact the probability of employment. Since there is no possible interaction between sample refugees and $t + 1$ refugees, there should be no significant relationship. This exercise therefore serves to test whether there is a cyclical pattern to IRC placement of refugees which may be confounded with the social network mechanism of interest.

Columns 1 through 4 in table 4 are consistent with the maintained identification assumption. Columns 1 and 3 show the baseline results for the sub-sample of refugees who arrived between 2001 and 2004 using nationality-year and city and then the full set of fixed effects, respectively. The results are quite similar to those in columns 3 through 6 in table 3. Columns 2 and 4 show the relationship between $y_{ijkt}$ and $N_{jk(t+1)}$. In neither specification is there a significant relationship and the most flexible specification in Column 4 shows a very small, positive, point estimate. This evidence is inconsistent with the idea of time-variant match quality shocks which are correlated with network size driving the results. It also casts doubt on the concern that there is a dynamic relationship between IRC staffing resources and the number of refugee arrivals.\textsuperscript{21}

Heterogeneous Effects by Initial English Level

Social networks play an important role in many aspects of a refugee’s life. In particular, network members may be helpful in providing translation services either during the job search process or on-the-job. Does the network effect estimated in this paper reflect job information transmission or simply the ability of network members to translate for network members at the work site? A larger number of network members available to translate at the workplace is likely to increase the probability recently arrived refugees find employment at those sites, and particularly so for those

\textsuperscript{20}Refugees who arrived in 2005 must be excluded since I do not observe the number of refugees who arrived in 2006.

\textsuperscript{21}The number of staff members and volunteers at the office and an indicator for whether the case workers speak a common language with the refugee is additional information available to address this concern. The relationship between network size and outcomes are robust to the inclusion of these variables as well.
individuals who arrive in the U.S. with no or low English knowledge. I therefore test to see if there are heterogeneous network effects depending on refugees’ initial English levels. Columns 5 and 6 of table 4 shows results of including \( N_{jk(t)} \) for all \( t \) and \( N_{jk(t)} \) for all \( t \) interacted with initial English ability. The level effects are consistent with the previous analysis and the interaction terms are insignificant. This indicates that the coefficients on \( N_{jk(t)} \) for all \( t \) are unlikely to only reflect the ability of social networks to provide English language services.

### 5.2 Job Quality: Wages

Both measures of network size provide complementary evidence on the importance of job information flows for employment within social networks. I now turn to the role of networks in determining hourly wages. Columns 1 and 2 of table 5 show the effect of network size on wages for the employed sample. Recall that the theoretical predictions regarding these effects were ambiguous. There are two offsetting factors: on one hand, an increase in \( N_{jk(t)} \) will decrease wages since an individual will receive less job offers, thereby reducing the ability to choose the highest paying offer. However, the proportion of individuals who receive job information indirectly, through other employed network members, will decline. Since these wages are lower on average, the average wages of those who are employed rises. Columns 1 and 2 of table 5 are broadly consistent with the model’s prediction. The size of the network in periods \( t - 2 \) and \( t - 3/t - 4 \) are positive and statistically significant. There is no evidence that junior network members, those who arrived in years \( t \) and \( t - 1 \), impact average hourly wages. These results constitute weak evidence of the information transmission model. The more senior network members are having a strong, positive effect on hourly wages while those network members who arrived more recently have no discernable effect. This is also suggestive that the effect of the network in changing the number of wage offers an individual receives is stronger than the compositional effect.
The inclusion of additional demographic information on individual refugees in column 2 does not have a large effect on the estimates. The initial English level has a large impact on the average hourly wage. Again, this must be interpreted with caution since the estimate may not reflect the causal relationship between English level and wages since the IRC may use this information when making geographic placement decisions.

In order to test Claim 1, I estimate equation 4 with the full male sample and impute wage offers for the unemployed as zero. Claim 1 argues that variation in network size will have heterogeneous effects on hourly wages (unconditional on employment) due to the dynamic relationship between network size and wages. An increase in the number of network members who arrive in the same period will have a negative impact on wages. The impact from an increase in the number of more senior members on a new arrival $i$’s wages will be monotonically increasing in the time elapsed between the additional network member’s arrival in the U.S. and individual $i$. The estimates are consistent with the model’s prediction. As shown in column 3 of table 5, one standard deviation increase in the number of network members who arrived in time $t - 1$ decreases the wage by $.70. An increase of one standard deviation in $N_{jk(t-2)}$ increases hourly wage by $.50$. Using this measure of wages, these results reflect both the effect of the network on employment and the direct effect on wages. These results support the intuitive notion that network members become increasingly valuable to new arrivals as their exposure in the labor market in the U.S. increases. Including a wider range of demographic and other control variables as in column 4 of table 5 leads to little change in the network coefficients.

However, there remains a potential problem in estimating the wage equation with the full sample. The model predictions imply that the effect of network size should have an effect on offer wages. However, the data provided by the IRC only provides wages for those individuals who are employed, and as such offer wages for those who are unemployed are unknown. According to the
model, individuals who are unemployed have received no job offers. Therefore, a wage of zero for these individuals is the correct offer wage. However, there may be a censuring problem if some individuals reject an offer because their reservation wage is higher than the wage offer. While this is not in the model, I address this concern in the empirical analysis.

The classic solution to this problem is to estimate a structural model of wage offers and labor market participation. Without a suitable exclusion restriction, however, classic selection models are not necessarily identified (Heckman, 1974). One alternative solution, as in Neal and Johnson (1996) and Johnson, Kitamura, and Neal (2000), is to impute unobserved wages as zero and estimate the wage equation using least absolute deviations (LAD). Under this assumption that all unemployed individuals receive wage offers below the median offer made to employed workers with comparable skills, LAD estimation is unaffected by imputing unobserved wage offers as zero.

The analysis of wages using LAD estimation shows results consistent with OLS results. As shown in columns 5 and 6 of table 5, a larger number of refugees who arrived in years $t$ and $t-1$ negatively impact the average hourly wage of a new arrival. The point estimates are, however, smaller than in the OLS specification and have larger standard errors. The analysis does not include the full set of control variables for city-time, nationality-time and city-nationality since LAD estimation is difficult with large numbers of dummy variables. This additional specification does suggest that the main results in columns 1 and 2 are not driven by wage censuring.\footnote{The results of the LAD estimation in table 5 depend on the assumption that unemployed refugees receive offer wages which are below the median wage of those employed with similar observable characteristics. While it is difficult to provide direct evidence on the validity of the assumption, I will note that the refugees in sample generally arrive in the U.S. with low English ability and find employment in low skilled service positions, such as housekeepers, in low-skilled industries. It is therefore likely that those who are unable to gain employment in the initial 90 days after arrival are those who would otherwise have low wage offers.}

The OLS results in columns 1 and 2 of table 5 show that tenured network members positively influence hourly wages of recently arrived refugees. Taken in conjunction with the LAD estimates, the analysis provides consistent evidence supporting the job information transmission model.
5.3 Alternative Specifications

5.3.1 Evidence using Fraction Employed

Given that the model’s predictions are driven by the distinction between employed and unemployed network members, ideally the size of the network would be broken down along those lines. Unfortunately, since there are no individual records prior to 2001, and the IRC only collects employment and wage data as of 90 days after arrival, this is not possible. However, restricting the sample to those refugees who arrived between 2003 and 2005 allows for an analysis which is at least suggestive of this preferable specification. By making the assumption that there is persistence in employment outcomes over time, i.e. that the probability of employment at 90 days is a good predictor of whether that individual will be employed later, I can construct the number of network members who arrived during the previous two years who are likely to be employed at time $t$. Indeed, table 7 confirms that an increase in the number of individuals who were unemployed as of 90 days after their arrival in the U.S. is negatively associated with employment rates of refugees who arrive in time $t$, up to two years after the arrival of the network members. Conversely the number of refugees who were employed as of 90 days after arrival is positively correlated with employment outcomes. This same pattern is found for wages for the sample who are employed.

5.3.2 Alternative Network Measure: Census Data

To further test the model, I also use Census data to construct a measure of network size which includes all individuals from a country of origin group in a given metropolitan area.\textsuperscript{23} This measure will include all immigrants groups, not only refugees. Since the data structure differs from the network measure used above, the empirical specification varies as well. In order to test the

\textsuperscript{23}Most networks are defined at the level of the MSA, however some include multiple MSAs. For example, refugees resettled in the New York office can be resettled in either New York-Northeastern NJ MSA or the Nassau Co., NY MSA. Thus the network size includes both MSAs since there is likely to be contact between individuals across this geographical space.
hypothesis using the 2000 Census data, the size of the network is restricted to those who arrived most recently in the U.S., specifically those who arrived in 1999. I then look for a differential effect of this network for refugees who arrived in 2001 and 2002.

\[ Y_{ijkt} = \alpha + \phi_1 N_{jk(t=1999)} + \phi_2 N_{jk(t=1999)} * \lambda_{2001} + X_{ijkt} \beta + \delta_j + \phi_k + \lambda_{2001} + \epsilon_{ijkt} \]  

(5)

\( Y_{ijkt} \), \( X_{ijkt} \), \( \delta_j \), \( \phi_k \), and \( \epsilon_{ijkt} \) are defined as above. As described above, \( N_{jk(t=1999)} \) is the size of the network for those immigrants who arrived in 1999 according to the Census, and \( \lambda_{2001} \) is an indicator for those refugees who arrived in 2001. Estimates showing \( \phi_1 \) to be positive and \( \phi_2 \) to be negative would be consistent with the model. The differential network effect across the two cohorts is therefore captured by \( \phi_2 \): an increase in the number of network members who arrived in 1999 would have a smaller or negative impact on labor market outcomes for those who arrived in 2001 than for those who arrived in 2002. By 2002, network members who arrived in 1999 would have acquired additional job information, becoming employed themselves, and be better able to provide referrals to newly resettled refugees. These network members would, however, be more likely to be competitors for job information with those who arrived more closely to them in time, namely refugees in the 2001 cohort. In this specification, the error term is again corrected for clustering at the nationality group/regional office level.

**Probability of Employment**

Using the 2000 Census to create a second network measure allows for flexibility in the definition of the social network. This measure expands the potential network members to those who come from the same country of origin or ethnic group but who may have different immigration status. In this case, individual network members have self-selected into their preferred location based on a number of unobserved factors. So while this measure of the network is more susceptible to
selection bias due to comparative advantage, it does test the generality of the job information sharing effect across two independently constructed network measures. Columns 1 and 2 in table 6 show that the estimates are as expected from the model. The effect of a larger number of network members from 1999 increases the probability of employment for those refugees who arrived in 2002. More specifically, increasing network size by one standard deviation increases the probability of employment for the 2002 cohort by 6.7%. The interaction term between network size and the indicator for arrival in 2001 is negative. This shows that relative to those refugees who arrived in 2002, an increase in the network size has a smaller effect on the probability of employment. The sum of the two coefficients is negative but small and statistically insignificant. This is consistent with the information transmission model: those refugees who arrive less than 2 years after the network members do not gain from an increase in network size while those who arrived sufficiently later do experience the positive influence of the network in terms of job information. While not shown in table 6, these results are also robust to the inclusion of a richer set of demographic variables.

Wages

Changing the estimation approach to use Census data as in equation 5 provides qualitatively similar results. Column 3 of table 6 indicates that the OLS estimates show no significant effect of network size on wages of those employed although the signs of the point estimates are as expected from the model. The coefficients in columns 4 and 5 from using the full sample with LAD estimation are more informative. The network effect for refugees who arrived in 2002 is positive and statistically strong. A one standard deviation in network size wages by $0.56. The interaction of network size with the dummy indicating arrival in 2001 is negative and statistically significant as in the employment regression. The sum of the two network coefficients is negative but statistically insignificant. This closely parallels the results found in the employment results as well as those predictions of the theoretical model.
6 Alternative Explanations

The analysis in the previous section interprets the empirical findings in terms of the theoretical framework presented in section 2. This section considers three alternative hypotheses to the model of network-based job information transmission: labor market competition with fixed labor demand (6.1), firm learning (6.2), and measurement error in relation to secondary migration (6.3). I conclude that there is little evidence that any of these alternative interpretations can be a primary explanation for the empirical findings in section 5.

6.1 Labor Market Competition and September 11th

The theoretical framework and empirical evidence sections have both argued that an increase in the number of network members who arrived recently in the U.S. exacerbates competition within the social network. Under this framework, the negative coefficient on $N_{jkt}$ and $N_{jk(t-1)}$ reflects this negative within-network competition. This section considers whether an increase in labor supply in the face of fixed labor demand is an alternative explanation for these estimates.

First, the number of new arrivals in a given ethnic network is small in each year, around 30 people on average. I argue that it is therefore unlikely that such a small addition to the total labor market in cities such as Atlanta and Salt Lake City will influence the unemployment rate or equilibrium wage level. However, segmented labor markets may make these numbers non-trivial.

A more formal way to discern between these two different “competition” effects is a test using September 11, 2001. Anecdotal evidence from the International Rescue Committee suggests that the terrorist attacks on September 11, 2001 had a large, negative impact on refugees’ labor market outcomes. The number and variety of opportunities available to refugees diminished significantly. The effect of this shock, however, differs in the two models. The network model will predict, as will be explained below, a diminished competition effect after 9/11. Conversely, the fixed labor demand
model would lead to exacerbated competition after a negative shock in labor demand due to the terrorist attacks. Two potential channels for the negative effect on employment outcomes post-9/11: an increase in xenophobia that decreased the ability of refugees to gain employment and a negative shock to the tourism industry, where many refugees were employed.\textsuperscript{24} This natural experiment also tests how refugee social networks respond to an exogenous negative shock to employment opportunities and allows us to evaluate whether the response is consistent with the job information transmission model.

In terms of the model presented in section 2, the 9/11 shock can most simply be analyzed in the model as a shock to the arrival rate, $a$.\textsuperscript{25} Figure 3 shows how the treatment effect from an increase in the size of cohort $c$, more specifically $\frac{\partial s_k}{\partial N_c}$ for some $k$, varies with $a$. The treatment effect is the impact from a change in the size of one specific cohort, $c$, on the probability of employment of an arriving cohort $k$. The figure represents the specific case where $b = .20$, each individual lives for 4 periods, and the treated cohort is doubled in size. The first panel demonstrates how the treatment effect of doubling cohort $c$ varies with $a$: for values of $a$ between 0 and .40, the treatment effect becomes stronger, i.e. more negative, as $a$ increases. However, the relationship between $a$ and $\frac{\partial s_k}{\partial N_c}$ is nonlinear. In particular, for values of $a$ such that the network is close to full employment, in this case approximately $a = .40$, the treatment effect quickly converges to zero.\textsuperscript{26} Since the average employment rate in the sample is 66%, however, it is unlikely that the portion of the model in which an increase in $a$ lowers competition is relevant empirically. Therefore, the conclusion from the model is that for a given change in network size, an increase in $a$ will exacerbate the competition effect, making $\frac{\partial s_k}{\partial N_c}$ more negative.

\textsuperscript{24}Table with the distribution of industries from 2001-2003 is available from the author upon request.
\textsuperscript{25}Alternatively, the negative 9/11 shock could also have increased the exogenous break-up rate $b$. The prediction is in fact the same.
\textsuperscript{26}The intuition for this nonlinearity is that for low levels of $a$, there are few referrals that can be made by the network; as $a$ rises, there is more available job information and therefore more scope for competition. However, for high levels of $a$ such that employment is almost 100%, there is little competition since all are likely to receive job information and become employed directly.
The second panel in Figure 3 graphically depicts the model’s prediction for the corresponding positive scenario: where an increase in $N_c$ raises the employment rate of the cohort which enters two periods afterwards, $c + 2$. In the example presented here, the impact on the employment rate of cohort $c + 2$ is positive for all values of $a, b$ and $N_c$. An increase in $a$ strengthens this effect for values of $a$ up to approximately .40 until it again displays the type of nonlinearity discussed above.

By interpreting the effect of 9/11 as an exogenous shock which decreased $a$, the model would predict that an increase in the size of the network would have a more dampened effect on labor market outcomes after September 11, 2001. The corresponding econometric specification is:

$$
Y_{ijkl} = \alpha + \sum_{j=1}^{4}(\gamma_j N_{ijk(t-2)} + \pi_{j} N_{ijk(t-3)} * Post9/11) + X_{ijkl}\beta + \delta_{jt} + \phi_{jk} + \epsilon_{ijkl}
$$

The hypothesis is that $\pi_p$ should be of the opposite sign of $\gamma_p$ for all $p$. Table 8 shows the results of estimating equation 6 using two different functional forms. Columns 1 and 2 estimate the effect of an increase in network size captured by two variables: the number of network members who arrived in years $t$ and $t - 1$ and the number of members who arrived in years $t - 2, t - 3$ and $t - 4$. The interaction term between the number of refugees who arrived in years $t$ and $t - 1$ and the post 9/11 indicator is positive as predicted by the model but statistically insignificant. An increase in the number of senior network members, captured by an increase in $N_{jkt}$, has a negative impact as in the previous analysis, but the interaction with the post 9/11 indicator is positive. Column 2 includes additional demographic characteristics, and in this specification the effect is marginally significant at the 10% level.

Columns 3 and 4 test for a differential network effect after 9/11 from an increase in the number of network members who arrived in the same period, $N_{jkt}$, and two or more years prior,

\[\text{There are a limited number of observations, 472, prior to September 11, 2001. I therefore combine the variables of interest to maximize statistical power.}\]
In this specification, the interaction term between the number of network members resettled in the same city in the same year and the post 9/11 indicator is precisely estimated and positive. This implies that while still present, the competition effect is weaker after the exogenous shock of 9/11 than before. Furthermore, the effect from an increase in $N_{jk(t-2/t-3/t-4)}$ is more muted after September 11, 2001. The interaction term is estimated with a p-value of .11 in column 4 but is consistently estimated across all specifications to be around .0004 in size.

The analysis therefore suggests that after September 11, 2001, the influence of social networks on newly arrived refugees’ employment outcomes diminished. We could alternatively imagine that the effect of September 11 would be to draw ethnic communities together more tightly, causing the interaction between $N_{jk(t-2/t-3/t-4)}$ and post 9/11 to be positive. This would likely bias the analysis against finding the result consistent with the model’s prediction. The exogenous shock of September 11, 2001 therefore lends additional support to the job information transmission model: while the specific pattern of network effects as a function of network members’ tenure in the U.S. is present both before and after the incident, the network’s impact on employment outcomes of new arrivals is dampened afterwards, consistent with a decline in the arrival rate of jobs. The result is furthermore inconsistent with the fixed labor demand model.

6.2 Are Firms Learning over Time?

What roles do firms play in recruiting and hiring within these networks? If firms learn about the unobservable ability of a group through experience, then a statistical discrimination model could generate a positive correlation between the number of individuals from a group in the labor market.

---

28 The previous analysis showed that an increase in $N_{jk}$ and $N_{jk(t-1)}$ lead to a negative effect on employment rates. However, I specifically estimate the impact from an increase in $N_{jk}$ in columns 3 and 4 since an increase in $N_{jk}$ is unambiguously negative in the model. An increase in $N_{jk(t-1)}$, however, can be of either sign in the model. Decreasing $a$ or increasing $b$ could shift the effect from being negative to positive. For this reason the interaction between post 9/11 and network size in year $t$ is the easiest to interpret.

29 The model presented here does not capture whether a network is densely or close-knit. The work by Calvo-Armengol and Jackson (2004), however, does show that these dimensions can have an impact on how effective networks are in providing job information to members.
and the performance of new arrivals. As firms observe a new group in the labor market, either through direct contact or by information from other firms, a prior belief on that group’s ability could rise over time. Therefore the number of co-ethnics who previously arrived would be positively correlated with the labor market outcomes of new arrivals, not due to within-network information sharing but due solely to firms’ behavior. This does not alone explain the results presented in table 3 since there is also a negative correlation between the number of recently arrived co-ethnics the new arrivals’ performance. A model of firm learning coupled with a model of fixed labor demand, as discussed above, or a short-term xenophobic backlash could conceivably be combined to generate both of the relationships observed in the data.

Section 6.1 uses the natural experiment of 9/11 to argue against the model of classic labor market competition. This test also contradicts the hypothesis of a xenophobic backlash as we would expect xenophobia to have increased after 9/11.

There are two additional empirical tests which can be used to distinguish the firm learning hypothesis from the network job information transmission model. The first tests whether family reunification refugees from the same country of origin show the same relationship with labor market outcomes of new arrivals. The second test uses variation in the data on the length of tenure of nationality groups in the U.S.

6.2.1 Family Reunification Refugees

In the primary analysis in section 5, I define the relevant network as being the number of non-family reunification refugees resettled in a city from the same country of origin. I argue that family reunification networks do not participate in the same social networks as those who arrive without family. However, do family reunification refugees also influence labor market outcomes?

The objective of this section is twofold: testing for a relationship between family reunification and
sample refugees can both validate the assumption that these two groups do not constitute one larger network and investigate the role of firm learning. The test is therefore to include both the number of network members and the number of family reunification refugees of group $j$ in city $k$ in periods $t$ through $t - 4$. The network model would predict that there is no correlation between the number of family reunification refugees and sample refugees’ labor market outcomes.

If firms are learning over time about a group, it is unlikely that firms update their beliefs and screening behavior only when exposed to non-family reunification refugees from a particular country.\footnote{It would be useful to show that family reunification refugees are employed in the same industries and occupations as sample refugees. This would undermine the argument that these two groups of refugees are simply in different labor markets and accordingly firms are only exposed to one group or another. Unfortunately the data is not currently available to do this.} Firms furthermore are unlikely to even observe which refugees arrived in the U.S. through family or non-family reunification channels. We would also expect that both types of refugees would be competing in the labor market; in that case, the negative relationship between the number of arrivals in the same year or 1 year prior and performance of new arrivals would be observed for both groups.

Family reunification refugees are immediately reunited with their family already in the U.S. and these refugees’ family members may have self-selected into their localities. Therefore, the exercise in this section tests the joint hypotheses that family reunification refugees are not the relevant social network group for non-family reunification refugees, that firm learning is limited, and that unobserved factors at the $jkt$ are not correlated with network size. Sorting among family reunification refugees may bias upwards the coefficients.

Table 9 shows no evidence that the number of family reunification refugees resettled in a city impacts the labor market outcomes of non-family reunification refugees of the same nationality. All specifications used in table 9 show that these coefficients are statistically insignificant both independently and jointly. Moreover, the sign of the coefficients are not consistent with sorting,
social network participation of family and non-family reunification refugees, nor with firm learning.\footnote{Edin et al. (2003), however, find evidence of refugees sorting across locations in their study using data from Sweden. While the analysis here does not show sorting along factors which are common to an entire nationality group-city pair, there is still the possibility of sorting based on characteristics which vary within an ethnic group, such as education level. Since I do not currently have individual data with labor market outcomes for family reunification refugees, I cannot evaluate whether there is in fact sorting along these other dimensions.} The coefficients of interest are also largely unchanged from their counterparts in table 3. In fact, the coefficients in column 1 of each table are not statistically different from one another.

### 6.2.2 Established Networks

In order to see a positive relationship between the number of refugees from a country of origin and new arrivals within two years, firms must update their beliefs over a group’s ability quickly. The magnitude of the relationship between senior network members and new arrivals should decline in situations where firms have already had ample time to adjust their beliefs and hiring strategies. An additional test is therefore to compare the size of the coefficients on $N_{jk(t)}$ through $N_{jk(t-3/t-4)}$ for established and new networks. During the period 1997-2005, the evolution of refugee crises around the world and changing State Department policies led to large changes in which groups were given access to the U.S. as refugees. If firms are learning over time, we expect the effect for established networks to be less strong. I therefore estimate equation 6 with an indicator for established networks, in lieu of the 9/11 variable.\footnote{Vietnamese, Bosnians, Somalis and Iraqis are treated as established groups, or networks, since they were the largest groups within sample in 1997.} There is no significant differential effect among refugees in established networks.\footnote{The table is available from the author upon request.} This finding is more consistent with the network model, which would not predict a differential effect.

### 6.3 Measurement Error and Secondary Migration

A limitation of the data is that each refugee is observed at only one point in time. The empirical analysis relies on the assumption that refugees remain in their placement city for the next four
years. As already discussed in section 5, the imprecision in the estimates of $N_{jk(t-3/t-4)}$ may be due to measurement error induced by secondary migration within the U.S. If the measurement error is classical, it will create attenuation bias and therefore bias the coefficient downward. However, the level of measurement error is likely to grow over time as more refugees will have out-migrated after 4 years than after 1 year. This section therefore assesses whether non-random out-migration can generate a pattern as observed in section 5.

Consider first a theoretical framework where there are no social networks, but refugees can choose to leave their initial resettlement city at any time. The number of jobs available to refugees from a particular country is fixed. There is no passing of information across refugees: the probability of becoming employed is simply the number of jobs available divided by the total number of refugees in the city in that year. If a fraction of refugees out-migrate each year, a nonlinear pattern emerges. Using the notation from the empirical analysis, the effect of an increase in $N_{jkt}$ is negative due to labor market competition, irrespective of the network, and the negative effect will be mitigated over time as individuals leave the city. This implies that the effect of an increase in $N_{jk(t-2)}$ will be less negative than an increase in $N_{jk(t-1)}$ or $N_{jkt}$. The model can not, however, explain how the impact of $N_{jk(t-2)}$ is not only less negative than that of $N_{jkt}$ but also positive.

The above thought experiment is one where out-migration is randomly distributed across refugees. I also consider models where $p\%$ of either the employed or unemployed network out-migrate each period. Simulation of such models show that the same pattern emerges as when randomly selected network members move out of the labor market. Recall $S$ is the number of periods an individual is in the market. During the $S$ periods following the entry of the exogenously larger cohort, the average employment rate for the entering cohort is less than or equal to the rate
when cohort size is constant.\textsuperscript{34}

There is also a larger concern that the measurement error is correlated with other unobserved network characteristics. Analyzing the decision to out-migrate as a function of network size and composition during the first 90 days after arrival shows no significant relationship between $N_{jk(t)}$ through $N_{jk(t-3/t-4)}$ and binary variable indicating out-migration. While this is not perfect, since it requires the assumption that the decision to migrate is the same during the first 90 days as it is in subsequent years, it provides the best insight possible into migration decisions given the data. Due to data limitations, however, a rigorous analysis of refugees’ secondary migration decisions in the U.S. must be left to future research.

7 Conclusion

This paper presents evidence on the importance of ethnic networks in influencing access to local labor markets for refugees recently resettled in the U.S. The empirical results support a model of job information transmission within a social network. Both the size and the structure of the network, as measured by length of tenure of network members in the U.S., influence the labor market outcomes of newly arrived refugees. This provides an insight into the functioning of social networks and provides empirical evidence that within-network competition over job information can lead to an economically sizable negative impact on labor market outcomes. This result tempers the previous findings in the empirical literature on social networks which show that networks play a beneficial

\textsuperscript{34}When the unemployed are more likely to out-migrate, there are some conditions in which an increase in the size of an arriving cohort leads to a higher employment rate for a cohort entering the market $S + 1$ periods after the exogenously large cohort. The intuition is that the larger cohort leads to a reduction in the employment rate in that period, which induces more out-migration than occurs in the steady-state, resulting in fewer individuals in the market in period $c + S + 1$. The new entrants in period $c + S + 1$ may have therefore have a higher employment rate than those in the market prior to cohort $c$’s entrance. It is crucial to note that this is strikingly different than the network model, which generates a positive impact on employment for cohorts who are in the market simultaneously with cohort $c$. Therefore, in order for the out-migration model to generate the pattern observed empirically in section 5, refugees must only work for 2 periods since the coefficient on $N_{jk(t-2)}$ is positive. This is not only extremely unlikely based on intuition alone, but the ORR data shows that refugees continue to work after 2 years in the U.S.
role in overcoming market frictions. The existence of costs to having a larger social network is difficult to identify unless the dynamic relationship between employment, wages and social network structure is taken into account. Using network variables which capture the tenure composition of social networks is therefore crucial to accurately assess the full network effect.

A static analysis of the effect of total network size on labor market outcomes conflates these two opposite effects. This paper shows that ignoring the dynamic relationship can lead to erroneously fail to identify the presence of social networks in the labor market and sensitivity to the empirical approach used. This may also lead to contradictory conclusions, as highlighted by the results of Edin et al. (2003), who find a positive relationship between the stock of immigrants and labor market outcomes, in contrast to Borjas (2000), who finds a negative relationship between proportion of individuals within a city from the same country and assimilation for refugees in the U.S.

Evidence of social networks providing labor market information suggests that there are spill-overs from policy interventions. The negative estimated effect from an increase in social network size is driven by an increase in the proportion of unemployed network members. Therefore, a job training program which increases the employment rate of some individuals in a social network will generate a positive externality to other network members. According to the model, such an intervention would create a net increase in the information available to unemployed network members who were not directly exposed by the program. This makes the returns to programs providing employment and training services to refugees even higher than otherwise measured by looking at program participants alone. The model’s prediction is general to any social network providing job information to its members. Accordingly, we would expect there to be spill-overs in policy interventions which improve labor market performance more generally, particularly for other immigrant groups. Interventions to improve the employment rate of some immigrants within
a community would have a positive impact throughout the network, thereby reducing the need for social services from local communities.

The evidence in this paper sheds some light onto the debate over the optimal resettlement of refugees. Large numbers of refugees and asylum seekers are permanently resettled in Europe and North America due to prolonged and protracted conflicts around the world. During 2004, for example, 676,400 people applied for asylum and over 83,000 refugees were permanently resettled to third countries through UNHCR resettlement programs (UNHCR, 2005). However, there is no consensus on the optimal method of resettlement within the new destination country. Policies vary widely from the dispersal policies in some European countries to the clustering method used by at least some American resettlement agencies (Edin et al., 2004).

By showing empirical evidence that refugee social networks provide labor market information to its members, this paper suggests a drawback to immigrant dispersal policies. Sending refugees to areas with a community of tenured network members, who have achieved relatively high employment rates, could improve short-run labor market outcomes. Improving the short-run labor market outcomes of refugees would ease the fiscal burden refugees put on local municipalities, one motivation for dispersal policies. However, this analysis only looks at very short-run outcomes and therefore can not provide an estimate of the total costs of dispersal policies. In the long-run, there are additional considerations such as the way the network affects individuals’ incentives to invest in learning the host country language or other types of human capital. Similarly, it is difficult to conclude whether other immigrant groups in the U.S. should be encouraged to cluster together. There is not only the above concern, but an analysis of larger immigrant groups in the U.S. would also have to address the general equilibrium consequences of an increase in social network size. Identifying one specific mechanism through which networks affect labor market outcomes, job information transmission, does however provide one piece of the puzzle. However, the long run role
of social networks in creating incentives or disincentives for integration and investments in host
country-specific human capital remains an open question and an area of future research.

References


**A Appendix**

**A.1 Proposition 1**

**Proposition 1** For all values $0 < a < 1$ and $0 < b < 1$, an increase in cohort size $N_j$ decreases $s^j_c$ for all $c$.\(^{35}\)

Proof:

For cohort $j$: If $N_j$ increases, $s^j_j$ decreases. This is simple since the previous periods' employment rate, $s^{j-1}_c$, will be unchanged for all $c$. Since $s^{j-1}_j = 0$, $s^j_j$ can be written as:

$$s^j_j = \frac{a(N_j + \sum_{c' \neq j} N_{c'})}{N_j + \sum_{c' \neq j} N_{c'}(1 - (1 - b)s^{j-1}_{c'})}$$

Differentiating with respect to $N_j$ gives:

$$\frac{\partial s^j_j}{\partial N_j} = \frac{a}{N_j + \sum_{c' \neq j} N_{c'}(1 - (1 - b)s^{j-1}_{c'})^2} = \frac{a(N_j + \sum_{c' \neq j} N_{c'})}{[N_j + \sum_{c' \neq j} N_{c'}(1 - (1 - b)s^{j-1}_{c'})]^{2} < 0}$$

For cohorts $c < j$: Similarly, if $N_j$ changes, the employment rate for all other cohorts in time period $j$, $s^j_c$, decreases as well. Consider cohort $j - 1$, although this holds for all other cohorts in the market at time $j$:

$$s^{j-1}_{j-1} = (1 - b)s^{j-1}_j + (1 - b)s^{j-1}_{j-1}$$

$$\frac{\partial s^{j-1}_{j-1}}{\partial N_j} = \frac{a(N_j + \sum_{c' \neq j} N_{c'})}{N_j + \sum_{c' \neq j} N_{c'}(1 - (1 - b)s^{j-1}_{c'})^{2}}$$

Since $s^{j-1}_c$ is unaffected by change in $N_j$ for all $c$,

$$\frac{\partial s^{j-1}_{j-1}}{\partial N_j} = \frac{(1 - (1 - b)s^{j-1}_{j-1})a}{N_j + \sum_{c' \neq j} N_{c'}(1 - (1 - b)s^{j-1}_{c'})^{2}} = \frac{a(1 - (1 - b)s^{j-1}_{j-1})(N_j + \sum_{c' \neq j} N_{c'})}{[N_j + \sum_{c' \neq j} N_{c'}(1 - (1 - b)s^{j-1}_{c'})]^{2}}$$

\(^{35}\)This claim holds for all values of $a$ and $b$ such that $s^j_c \neq 1$ for all $c$ and $j$. 

47
\[
= -a(1 - (1-b)s_{j-1}^j)(1-b)\sum_{c'\neq j} s_{c'}^{j-1} \left[N_j + \sum_{c'\neq j} N_{c'}(1 - (1-b)s_{c'}^{j-1})\right] < 0
\]

since \((1 - (1-b)s_{j-1}^j) > 0\).

A.2 Proposition 2

**Proposition 2** The impact of an increase in \(N_j\) on \(s_k^j\) is monotonically increasing between \(k = j\) and \(j + S - 1\).

**Proof:**

Assume \(S = 3\) and \(N_k = 1 \forall k \neq j\).

**Step 1:** \(s_{j+1}^j(N_j) > s_j^j(N_j)\)

\[
s_j^j(N_j) = \frac{a(2 + N_j)}{2 + N_j - (1-b)(s_{j-1}^{j-1} + s_{j-2}^{j-1})}
\]

\[
s_{j+1}^j(N_j) = \frac{a(2 + N_j)}{2 + N_j - (1-b)(N_js_j^j + s_{j-1}^j)}
\]

Therefore, need to show: \(N_js_j^j + s_{j-1}^j > s_{j-1}^{j-1} + s_{j-2}^{j-1}\)

Using the steady-state properties of the economy,\(^{36}\) equation (7) will hold with equality if \(N_j = 1\).

Since (7) holds with equality when \(N_j = 1\), inequality holds if expression is increasing in \(N_j\).

\[
\frac{a(2 + N_j)(1 + N_j - (1-b)s_{j-1}^{j-1})}{2 + N_j - (1-b)(s_{j-1}^{j-1} + s_{j-2}^{j-1})} + (1-b)s_{j-1}^{j-1} > s_{j-1}^{j-1} + s_{j-2}^{j-1}
\]

(7)

Differentiating equation (8) with respect to \(N_j\) gives:

\[
\frac{\partial}{\partial N_j} = \frac{a(3 + 2N_j - (1-b)s_{j-1}^{j-1})}{2 + N_j - (1-b)(s_{j-1}^{j-1} + s_{j-2}^{j-1})} - \frac{a(2 + N_j)(1 + N_j - (1-b)s_{j-1}^{j-1})}{[2 + N_j - (1-b)(s_{j-1}^{j-1} + s_{j-2}^{j-1})]^2}
\]

\[
= \frac{a(3 + 2N_j - (1-b)s_{j-1}^{j-1})(2 + N_j - (1-b)(s_{j-1}^{j-1} + s_{j-2}^{j-1}) - a(2 + N_j)(1 + N_j - (1-b)s_{j-1}^{j-1})}{[2 + N_j - (1-b)(s_{j-1}^{j-1} + s_{j-2}^{j-1})]^2}
\]

\[
= a[(2 + N_j)(1 - (1-b)s_{j-2}^{j-1}) + (1+N_j)(2 + N_j - (1-b)(s_{j-1}^{j-1} + s_{j-2}^{j-1})) - (1-b)s_{j-1}^{j-1}(2 + N_j - (1-b)(s_{j-1}^{j-1} + s_{j-2}^{j-1})] > 0
\]

since \(1 + N_j - (1-b)s_{j-1}^{j-1} > 0\)

Since denominator is greater than zero, if numerator is greater than zero, then equation (7) holds.

\(^{36}\)Employment status reflect a finite-state irreducible and aperiodic Markov process as in Calvo-Armengol and Jackson (2004). Then by Freidlin and Wentzell (1984) and Young (1993), there exists a unique steady-state distribution associated with this process.
Step 2: \( s_{j+2}^{j+2}(N_j) > s_{j+1}^{j+1}(N_j) \)

\[
\begin{align*}
\quad & s_{j+2}^{j+2}(N_j) = \frac{a(2 + N_j)}{2 + N_j + (1 - b)(s_{j+1}^{j+1} + N_j s_j^{j+1})} \\
\quad & s_{j+1}^{j+1}(N_j) = (1 - b)s_j^j + (1 - (1 - b)s_j^j) \frac{a(2 + N_j)}{2 + N_j + (1 - b)(s_{j+1}^{j+1} + N_j s_j^{j+1})}
\end{align*}
\]

Need to show:

\[
\quad s_{j+1}^{j+1} + N_j s_j^{j+1} > N_j^j s_j^j + s_{j-1}^j
\]

Let \( N_j = 1 + x \). Rearranging equation (8) gives:

\[
(x(s_{j+1}^{j+1} - s_j^j) + (s_{j+1}^{j+1} - s_{j-1}^j) + x(s_j^{j+1} - s_j^j) > 0 \quad (9)
\]

We can write \( s_{j+1}^{j+1} \), \( s_{j-1}^j \) and \( s_{j-2}^{j-1} \) in the following way:

\[
\begin{align*}
\quad & s_{j+1}^{j+1} = (1 - b)s_j^j + [1 - (1 - b)s_j^j]s_{j+1}^{j+1} \\
\quad & s_{j-1}^j = (1 - b)s_{j-1}^{j-1} + [1 - (1 - b)s_{j-1}^{j-1}]s_j^j \\
\quad & s_{j-2}^{j-1} = (1 - b)s_{j-2}^{j-2} + [1 - (1 - b)s_{j-2}^{j-2}]s_{j-1}^{j-1}
\end{align*}
\]

The left hand side (LHS) of equation (9) is then:

\[
LHS = (s_{j+1}^{j+1} - s_j^j)(2 - (1 - b)s_j^j) + x(s_j^{j+1} - s_j^j) + (1 - b)(s_j^j - s_{j-1}^{j-1})(1 - s_j^j)
\]

Using the fact that \( N_j s_j^j + s_{j-1}^j > s_{j-1}^{j-1} + s_{j-2}^{j-2} \) as shown above, this implies

\[
\quad s_{j+1}^{j+1} - s_{j-1}^{j-1} > s_{j-2}^{j-2} - s_{j-1}^{j-1} - x s_j^j > (s_{j-1}^{j-1} - s_j^j)(1 - (1 - b)s_{j-1}^{j-1}) - x s_j^j
\]

Substituting the above gives:

\[
LHS > (s_{j+1}^{j+1} - s_j^j)(2 - (1 - b)s_j^j) + x(s_j^{j+1} - s_j^j) + (1 - b)(1 - s_j^j)[(1 - (1 - b)s_{j-1}^{j-1})(s_{j-1}^{j-1} - s_j^j) - x s_j^j]
\]

\[
= (s_{j+1}^{j+1} - s_j^j)(2 - (1 - b)s_j^j) + (1 - b)(1 - s_j^j)(1 - (1 - b)s_{j-1}^{j-1})(s_{j-1}^{j-1} - s_j^j) + x(1 - (1 - b)s_j^j)(s_{j+1}^{j+1} - s_j^j) > 0
\]

since \( s_{j+1}^{j+1} > s_j^j \) as shown above and \( s_{j-1}^{j-1} > s_j^j \) as in Claim 1.
### Table 1: Summary Statistics

<table>
<thead>
<tr>
<th>IRC Data:</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>No. Obs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>33.99</td>
<td>11.05</td>
<td>1720</td>
</tr>
<tr>
<td>HH Size</td>
<td>2.76</td>
<td>2.04</td>
<td>1720</td>
</tr>
<tr>
<td>Employment rate</td>
<td>0.66</td>
<td></td>
<td>1720</td>
</tr>
<tr>
<td>Wage (conditional on employment)</td>
<td>7.48</td>
<td>1.36</td>
<td>1125</td>
</tr>
<tr>
<td>Spoke No English Upon Arrival</td>
<td>0.466</td>
<td></td>
<td>1453</td>
</tr>
<tr>
<td>Primary School</td>
<td>0.180</td>
<td></td>
<td>1720</td>
</tr>
<tr>
<td>Secondary School</td>
<td>0.464</td>
<td></td>
<td>1720</td>
</tr>
<tr>
<td>University or Above</td>
<td>0.202</td>
<td></td>
<td>1720</td>
</tr>
<tr>
<td>None, vocational or adult education</td>
<td>0.153</td>
<td></td>
<td>1720</td>
</tr>
<tr>
<td>Muslim</td>
<td>0.251</td>
<td></td>
<td>1720</td>
</tr>
<tr>
<td>IRC Exemption from Employment</td>
<td>0.059</td>
<td></td>
<td>1720</td>
</tr>
<tr>
<td># Refugees Resettled in Year t</td>
<td>10.32</td>
<td>13.47</td>
<td>1720</td>
</tr>
<tr>
<td># Refugees Resettled in Year t – 1</td>
<td>29.47</td>
<td>34.13</td>
<td>1720</td>
</tr>
<tr>
<td># Refugees Resettled Year t – 2</td>
<td>25.16</td>
<td>43.82</td>
<td>1720</td>
</tr>
<tr>
<td># Refugees Resettled Years t – 3 and t – 4</td>
<td>65.80</td>
<td>102.63</td>
<td>1720</td>
</tr>
<tr>
<td># Family Reunification Refugees Resettled in Year t</td>
<td>15.45</td>
<td>27.24</td>
<td>1720</td>
</tr>
<tr>
<td># Family Reunification Refugees Resettled in Year t – 1</td>
<td>18.77</td>
<td>48.25</td>
<td>1720</td>
</tr>
<tr>
<td># Family Reunification Refugees Resettled in Year t – 2</td>
<td>19.50</td>
<td>58.55</td>
<td>1720</td>
</tr>
<tr>
<td># Family Reunification Refugees Resettled in Years t – 3 &amp; t – 4</td>
<td>48.89</td>
<td>149.99</td>
<td>1720</td>
</tr>
</tbody>
</table>

#### 2000 Census Data:

| Network Members who Arrived in 1999 | 150.83 | 267.84 | 753 |

### Table 2: Correlation Coefficients of Refugee Cohort Sizes: 1997-2005

<table>
<thead>
<tr>
<th></th>
<th>Current Year</th>
<th>Prior Year</th>
<th>2 Years</th>
<th>3 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Prior</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Num Refugees Resettled in Current Year</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Num Refugees Resettled in Prior Year</td>
<td>0.5394</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Num Refugees Resettled in 2 Years before</td>
<td>0.2859</td>
<td>0.4744</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Num Refugees Resettled in 3 Years before</td>
<td>0.2711</td>
<td>0.3794</td>
<td>0.5892</td>
<td>1</td>
</tr>
<tr>
<td>Num Refugees Resettled in 4 Years before</td>
<td>0.2399</td>
<td>0.3473</td>
<td>0.3971</td>
<td>0.5815</td>
</tr>
</tbody>
</table>
Table 3: Employment Probability on Network Size

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td># Refugees Resettled in Year $t$</td>
<td>-0.236 **</td>
<td>-0.245 **</td>
<td>-0.340 **</td>
<td>-0.345 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.124)</td>
<td>(0.164)</td>
<td>(0.173)</td>
<td></td>
<td></td>
</tr>
<tr>
<td># Refugees Resettled in Year $t - 1$</td>
<td>-0.140 **</td>
<td>-0.117 *</td>
<td>-0.257 ***</td>
<td>-0.232 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.069)</td>
<td>(0.070)</td>
<td>(0.100)</td>
<td>(0.100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td># Refugees Resettled in Year $t - 2$</td>
<td>0.104 ***</td>
<td>0.100 ***</td>
<td>0.113 **</td>
<td>0.105 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.039)</td>
<td>(0.054)</td>
<td>(0.054)</td>
<td></td>
<td></td>
</tr>
<tr>
<td># Refugees Resettled in Years $t - 3$ and $t - 4$</td>
<td>0.037 **</td>
<td>0.034 **</td>
<td>0.056</td>
<td>0.044</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.043)</td>
<td>(0.042)</td>
<td></td>
<td></td>
</tr>
<tr>
<td># Refugees Resettled Years $t$ to $t - 4$</td>
<td>0.028 ***</td>
<td>-0.035</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.047)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.022 ***</td>
<td>0.025 ***</td>
<td>0.023 ***</td>
<td>0.022 ***</td>
<td>0.026 ***</td>
<td>0.025 ***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Age Sq</td>
<td>0.000 ***</td>
<td>0.000 ***</td>
<td>-0.0003 ***</td>
<td>-0.0003 ***</td>
<td>-0.0004 ***</td>
<td>-0.0004 ***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>HH Size</td>
<td>-0.015 **</td>
<td>-0.020 ***</td>
<td>-0.015 **</td>
<td>-0.013 **</td>
<td>-0.021 ***</td>
<td>-0.019 ***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>IRC Exemption from Employment</td>
<td>-0.540 ***</td>
<td>-0.552 ***</td>
<td>-0.540 ***</td>
<td>-0.543 ***</td>
<td>-0.557 ***</td>
<td>-0.557 ***</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.057)</td>
<td>(0.051)</td>
<td>(0.050)</td>
<td>(0.056)</td>
<td>(0.054)</td>
</tr>
</tbody>
</table>

P-value of education variables 0.449 0.250
P-value of initial English level variables 0.002 0.000
P-value of religion variable 0.318 0.581

No obs 1720 1720 1720 1720 1720 1720
Adjusted R squared 0.224 0.273 0.231 0.236 0.282 0.288

a SE are in parentheses & clustered by city-ethnicity.
b Columns 1, 3 and 4 include fixed effects for nationality-year and regional office.
c Columns 2, 5 and 6 include fixed effects for nationality-year, regional office-year and nationality-city.
d Columns 2 and 4 include additional individual covariates including: education, initial English level, religion.
e Rows 1-4 are multiplied by 100.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td># Refugees Resettled in Year $t + 1$</td>
<td>0.096</td>
<td>(0.099)</td>
<td>0.030</td>
<td>(0.179)</td>
<td></td>
<td></td>
</tr>
<tr>
<td># Refugees Resettled in Year $t$</td>
<td>-0.279 **</td>
<td>-0.306 **</td>
<td>-0.551 ***</td>
<td>-0.551 ***</td>
<td>-0.203</td>
<td>-0.253</td>
</tr>
<tr>
<td></td>
<td>(0.125)</td>
<td>(0.128)</td>
<td>(0.156)</td>
<td>(0.156)</td>
<td>(0.164)</td>
<td>(0.172)</td>
</tr>
<tr>
<td># Refugees Resettled in Year $t - 1$</td>
<td>-0.163 **</td>
<td>-0.158 **</td>
<td>-0.253 **</td>
<td>-0.248 **</td>
<td>-0.219 **</td>
<td>-0.192 **</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
<td>(0.078)</td>
<td>(0.126)</td>
<td>(0.124)</td>
<td>(0.094)</td>
<td>(0.096)</td>
</tr>
<tr>
<td># Refugees Resettled in Year $t - 2$</td>
<td>0.072 *</td>
<td>0.074 **</td>
<td>0.147 **</td>
<td>0.148 **</td>
<td>0.111</td>
<td>0.104</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.037)</td>
<td>(0.060)</td>
<td>(0.061)</td>
<td>(0.075)</td>
<td>(0.076)</td>
</tr>
<tr>
<td># Refugees Resettled in Years $t - 3$ &amp; $t - 4$</td>
<td>0.037 **</td>
<td>0.038 **</td>
<td>0.085</td>
<td>0.087</td>
<td>0.077 **</td>
<td>0.072 **</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.015)</td>
<td>(0.058)</td>
<td>(0.061)</td>
<td>(0.033)</td>
<td>(0.033)</td>
</tr>
<tr>
<td># Refugees Resettled in Year $t$ * No English</td>
<td>-0.005</td>
<td>(0.235)</td>
<td>(0.066)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.230)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Refugees Resettled in Year $t - 1$ * No English</td>
<td>0.135</td>
<td>0.107</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.087)</td>
<td>(0.095)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Refugees Resettled in Year $t - 2$ * No English</td>
<td>-0.067</td>
<td>-0.059</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.096)</td>
<td>(0.098)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Refugees Resettled in Years $t - 3$ &amp; $t - 4$ * No English</td>
<td>-0.027</td>
<td>-0.028</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.032)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No English</td>
<td>-0.077</td>
<td>(0.048)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P-value of F test for No English Interactions: 0.320 0.320

No obs: 1340 1340 1340 1340 1471 1471

---

a SE are in parentheses and clustered by city-ethnicity.
b Basic demographic controls include: age, age squared, HH size, IRC exemption from employment.
c Columns 1, 2, 5 and 6 include fixed effects for nationality-year and city.
d Columns 3 and 4 include fixed effects for nationality-year, city-year and nationality-city.
e Columns 1-4 only contain years 2001-2004.
f Column 6 includes additional individual covariates: education level, initial English level, religion.
g Coefficients are all multiplied by 100.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td># Refugees Resettled in Year $t^b$</td>
<td>0.161</td>
<td>0.230</td>
<td>-0.024**</td>
<td>-0.023*</td>
<td>-0.451</td>
<td>-0.644</td>
</tr>
<tr>
<td></td>
<td>(0.372)</td>
<td>(0.369)</td>
<td>(0.012)</td>
<td>(0.013)</td>
<td>(0.469)</td>
<td>(0.589)</td>
</tr>
<tr>
<td># Refugees Resettled in Year $t-1^b$</td>
<td>0.000</td>
<td>0.034</td>
<td>-0.021***</td>
<td>-0.018**</td>
<td>-0.453**</td>
<td>-0.396</td>
</tr>
<tr>
<td></td>
<td>(0.246)</td>
<td>(0.256)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.220)</td>
<td>(0.279)</td>
</tr>
<tr>
<td># Refugees Resettled Year $t-2^b$</td>
<td>0.576***</td>
<td>0.537***</td>
<td>0.011**</td>
<td>0.010**</td>
<td>0.488***</td>
<td>0.556**</td>
</tr>
<tr>
<td></td>
<td>(0.214)</td>
<td>(0.211)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.191)</td>
<td>(0.237)</td>
</tr>
<tr>
<td># Refugees Resettled Years $t-3$ &amp; $t-4^b$</td>
<td>0.440***</td>
<td>0.394***</td>
<td>0.006*</td>
<td>0.005</td>
<td>0.113</td>
<td>0.132</td>
</tr>
<tr>
<td></td>
<td>(0.137)</td>
<td>(0.134)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.087)</td>
<td>(0.109)</td>
</tr>
<tr>
<td>Age</td>
<td>0.083***</td>
<td>0.078***</td>
<td>0.245***</td>
<td>0.231***</td>
<td>0.133***</td>
<td>0.095***</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.021)</td>
<td>(0.039)</td>
<td>(0.046)</td>
<td>(0.029)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>Age Sq</td>
<td>-0.001***</td>
<td>-0.001***</td>
<td>-0.004***</td>
<td>-0.003***</td>
<td>-0.002***</td>
<td>-0.001***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Case Size</td>
<td>0.032</td>
<td>0.033</td>
<td>-0.130***</td>
<td>-0.114**</td>
<td>-0.045</td>
<td>-0.034</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.021)</td>
<td>(0.053)</td>
<td>(0.056)</td>
<td>(0.028)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>IRC Exemption from Employment</td>
<td>0.708</td>
<td>0.708</td>
<td>-4.11***</td>
<td>-4.09***</td>
<td>-6.15***</td>
<td>-6.04***</td>
</tr>
<tr>
<td></td>
<td>(0.645)</td>
<td>(0.578)</td>
<td>(0.423)</td>
<td>(0.407)</td>
<td>(0.229)</td>
<td>(0.294)</td>
</tr>
<tr>
<td>P-value of education variables</td>
<td>0.000</td>
<td>0.207</td>
<td></td>
<td></td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td>P-value of initial English level variables</td>
<td>0.097</td>
<td>0.000</td>
<td>0.000</td>
<td>0.030</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-value of religion variables</td>
<td>0.516</td>
<td>0.463</td>
<td></td>
<td></td>
<td></td>
<td>0.084</td>
</tr>
</tbody>
</table>

a SE are in parentheses and clustered by city-ethnicity.
b Rows 1, 2, 5 and 6 are multiplied by 100.
c Columns 1 and 2 are conditional on employment while Columns 3-6 use the full sample.
d Columns 1-4 include fixed effects for nationality-year, regional office-year and nationality-city.
e Columns 2, 4 and 6 include additional individual covariates including: education, initial English level, religion.
f Columns 3 and 4 are LAD estimates with FE for nationality group, year of arrival, and city.
<table>
<thead>
<tr>
<th></th>
<th>Employment</th>
<th></th>
<th>Wages</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Network size which arrived in 1999 (^g)</td>
<td>0.0284 **</td>
<td>0.0220 *</td>
<td>0.013</td>
<td>0.249 ***</td>
<td>0.172 *</td>
</tr>
<tr>
<td></td>
<td>(0.0121)</td>
<td>(0.0126)</td>
<td>(0.055)</td>
<td>(0.088)</td>
<td>(0.091)</td>
</tr>
<tr>
<td>Network size which arrived in 1999 * Refugee arrived in 2001 (^g)</td>
<td>-0.0275 **</td>
<td>-0.0253 **</td>
<td>-0.043</td>
<td>-0.238 ***</td>
<td>-0.201 **</td>
</tr>
<tr>
<td></td>
<td>(0.0114)</td>
<td>(0.0121)</td>
<td>(0.058)</td>
<td>(0.095)</td>
<td>(0.097)</td>
</tr>
<tr>
<td>Age</td>
<td>3.388 ***</td>
<td>3.020 ***</td>
<td>0.073 **</td>
<td>0.270 ***</td>
<td>0.229 ***</td>
</tr>
<tr>
<td></td>
<td>(0.912)</td>
<td>(1.002)</td>
<td>(0.034)</td>
<td>(0.068)</td>
<td>(0.072)</td>
</tr>
<tr>
<td>Age Sq</td>
<td>-0.0530 ***</td>
<td>-0.0482 ***</td>
<td>-0.0009 **</td>
<td>-0.0042 ***</td>
<td>-0.0037 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0122)</td>
<td>(0.0129)</td>
<td>(0.0005)</td>
<td>(0.0009)</td>
<td>(0.0009)</td>
</tr>
<tr>
<td>HH Size</td>
<td>-1.789 *</td>
<td>-1.474</td>
<td>-0.017</td>
<td>-0.162 **</td>
<td>-0.122</td>
</tr>
<tr>
<td></td>
<td>(1.093)</td>
<td>(1.046)</td>
<td>(0.028)</td>
<td>(0.078)</td>
<td>(0.076)</td>
</tr>
<tr>
<td>p-value of education variables</td>
<td>0.014</td>
<td></td>
<td>0.014</td>
<td></td>
<td>0.038</td>
</tr>
<tr>
<td>p-value of initial English level variables</td>
<td>0.066</td>
<td></td>
<td>0.066</td>
<td></td>
<td>0.003</td>
</tr>
<tr>
<td>p-value of religion variables</td>
<td>0.027</td>
<td></td>
<td>0.027</td>
<td></td>
<td>0.002</td>
</tr>
<tr>
<td>p-value of occupation variables</td>
<td>0.757</td>
<td></td>
<td>0.757</td>
<td></td>
<td>0.736</td>
</tr>
<tr>
<td>No obs</td>
<td>753</td>
<td>753</td>
<td>523</td>
<td>742</td>
<td>742</td>
</tr>
<tr>
<td>Adjusted R squared</td>
<td>0.187</td>
<td>0.199</td>
<td>0.300</td>
<td>0.183</td>
<td>0.207</td>
</tr>
</tbody>
</table>

\(^a\) Standard errors are in parentheses and clustered by city-ethnicity.
\(^b\) Sample restricted to refugees who arrived in 2001 and 2002.
\(^c\) All columns include fixed effects for nationality-year and regional office.
\(^d\) Column 3 uses only the employed sample. Columns 4 and 5 use the full sample.
\(^e\) Columns 2 and 5 also include: education, initial English level, religion and occupation variables.
\(^f\) Network Size is number of individuals in the 2000 Census who arrived in 1999 by place of birth/MSA.
\(^g\) Coefficients in row are multiplied by 100.
Table 7: Network Size Using Within Sample Employment Info

<table>
<thead>
<tr>
<th>Employment</th>
<th>Wage</th>
</tr>
</thead>
<tbody>
<tr>
<td># Refugees Unemployed during 2 Prior Years</td>
<td>-0.038 ***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
</tr>
<tr>
<td># Refugees Employed during 2 Prior Years</td>
<td>0.018 ***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>No obs</td>
<td>1487</td>
</tr>
</tbody>
</table>

a Network variables indicate employment status as of 90 days after the network member’s arrival.
b SE are clustered by city-ethnicity-arrival year pairs using only 2003-2005 sample.
c Also included are fixed effects for nationality, city and year of arrival, age, age squared and HH size.

Table 8: Employment: Shock of 9/11

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td># Network Members Resettled in Year t a</td>
<td>-0.708 ***</td>
<td>-0.720 ***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.183)</td>
<td>(0.183)</td>
</tr>
<tr>
<td># Network Members Resettled in Year t * Post 9/11 a</td>
<td>0.530 **</td>
<td>0.536 **</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.256)</td>
<td>(0.258)</td>
</tr>
<tr>
<td># Network Members Resettled in Year t and t − 1 a</td>
<td>-0.326 **</td>
<td>-0.315 **</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.136)</td>
<td>(0.139)</td>
</tr>
<tr>
<td># Network Members Resettled in Year t and t − 1 * Post 9/11 a</td>
<td>0.023</td>
<td>0.020</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.143)</td>
<td>(0.144)</td>
</tr>
<tr>
<td># Network Members Resettled in Year t − 1 a</td>
<td>-0.289 ***</td>
<td>-0.265 ***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.105)</td>
<td>(0.105)</td>
</tr>
<tr>
<td># Network Members Resettled in Years t − 2, t − 3 and t − 4 a</td>
<td>0.080 *</td>
<td>0.069</td>
<td>0.077 *</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.046)</td>
<td>(0.044)</td>
</tr>
<tr>
<td># Network Members Resettled in Years t − 2, t − 3 and t − 4 * Post 9/11 a</td>
<td>-0.043</td>
<td>-0.046 *</td>
<td>-0.039</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.028)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>No obs</td>
<td>1720</td>
<td>1720</td>
<td>1720</td>
</tr>
</tbody>
</table>

a SE are in parentheses and clustered by city-ethnicity. Coefficients in row are multiplied by 100.
b All columns include fixed effects for nationality-year, city-year and nationality-city.
c Columns 2 and 4 include additional individual covariates including: education, initial English level, religion.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td># Refugees Resettled in Year t $^e$</td>
<td>-0.245 **</td>
<td>-0.255 **</td>
<td>-0.348 **</td>
<td>-0.355 **</td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.125)</td>
<td>(0.164)</td>
<td>(0.172)</td>
</tr>
<tr>
<td># Refugees Resettled in Year t – 1 $^e$</td>
<td>-0.159 **</td>
<td>-0.135 *</td>
<td>-0.301 ***</td>
<td>-0.275 **</td>
</tr>
<tr>
<td></td>
<td>(0.074)</td>
<td>(0.075)</td>
<td>(0.114)</td>
<td>(0.114)</td>
</tr>
<tr>
<td># Refugees Resettled Year t – 2 $^e$</td>
<td>0.101 **</td>
<td>0.097 **</td>
<td>0.064</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.051)</td>
<td>(0.065)</td>
<td>(0.067)</td>
</tr>
<tr>
<td># Refugees Resettled Years t – 3 and t – 4 $^e$</td>
<td>0.053 **</td>
<td>0.053 **</td>
<td>0.049</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.023)</td>
<td>(0.053)</td>
<td>(0.055)</td>
</tr>
<tr>
<td>Age</td>
<td>0.023 ***</td>
<td>0.022 ***</td>
<td>0.026 ***</td>
<td>0.025 ***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Age Squared</td>
<td>-0.0003 ***</td>
<td>-0.0003 ***</td>
<td>-0.0004 ***</td>
<td>-0.0004 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>HH Size</td>
<td>-0.0152 **</td>
<td>-0.0137 **</td>
<td>-0.0210 ***</td>
<td>-0.0193 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0066)</td>
<td>(0.0071)</td>
<td>(0.0071)</td>
<td>(0.0076)</td>
</tr>
<tr>
<td>IRC Exemption from Employment</td>
<td>-0.543 ***</td>
<td>-0.547 ***</td>
<td>-0.559 ***</td>
<td>-0.560 ***</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.049)</td>
<td>(0.057)</td>
<td>(0.054)</td>
</tr>
<tr>
<td># Family Reunification Refugees Resettled in Year t $^e$</td>
<td>0.075</td>
<td>0.086</td>
<td>-0.055</td>
<td>-0.011</td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td>(0.108)</td>
<td>(0.158)</td>
<td>(0.147)</td>
</tr>
<tr>
<td># Family Reunification Refugees Resettled in Year t – 1 $^e$</td>
<td>0.001</td>
<td>-0.007</td>
<td>0.131</td>
<td>0.127</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.061)</td>
<td>(0.111)</td>
<td>(0.107)</td>
</tr>
<tr>
<td># Family Reunification Refugees Resettled in Year t – 2 $^e$</td>
<td>-0.040</td>
<td>-0.042</td>
<td>0.069</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td>(0.073)</td>
<td>(0.074)</td>
<td>(0.085)</td>
<td>(0.089)</td>
</tr>
<tr>
<td># Family Reunification Refugees Resettled in Years t – 3 and t – 4 $^e$</td>
<td>-0.008</td>
<td>-0.009</td>
<td>-0.081</td>
<td>-0.078</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.025)</td>
<td>(0.060)</td>
<td>(0.068)</td>
</tr>
</tbody>
</table>

p-value of F test for Family Variables: .664 .565 .250 .268
No obs: 1720 1720 1720 1720
Adjusted R squared: 0.230 0.235 0.281 0.287

---

a SE are in parentheses & clustered by city-ethnicity.
b Columns 1 and 2 include fixed effects for nationality-year and city.
c Columns 3 and 4 include fixed effects for nationality-year, city-year and nationality-city.
d Columns 2 and 4 include additional individual covariates including: education, initial English level, religion.
e Coefficients in row are multiplied by 100.
Figure 1: Graphical Example of Model with Constant Wages

Employment Rates from Simulated Model
Cohorts j and j+1

Employment Rates from Simulated Model
Cohorts j+2 and j+3

- Control Cohort
- Cohort j
- Cohort j+1
- Cohort j+2
- Cohort j+3
Figure 2: Graphical Example of Model with Wages

Employment Rates from Simulated Model
Cohorts j and j+1

Average Hourly Wages from Simulated Model
Cohorts j and j+1
Figure 3: Graphical Example of Model Varying $a$

![Graphical Example of Model Varying $a$](image-url)