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

Lori A. Beaman †

Abstract

This paper examines the dynamic implications of social networks for the labor market outcomes of refugees resettled in the U.S. A theoretical model of job information transmission shows that the relationship between social network size and labor market outcomes is heterogeneous and depends on the vintage of network members: an increase in network size can negatively impact some cohorts in a network while benefiting others. To test this prediction, I use new data on political refugees resettled in the U.S. and exploit the fact that these refugees are distributed across cities by a resettlement agency, precluding individuals from sorting. The results indicate that an increase in the number of social network members resettled in the same year or one year prior to a new arrival leads to a deterioration of outcomes, while a greater number of tenured network members improves the probability of employment and raises the hourly wage.

1 Introduction

Whether refugee or resident, social networks play an important role in the U.S. labor market. Studies from the 1930s onward report that between 30 and 60% of jobs are found through informal social network contacts (Bewley, 1999; Ioannides and Loury, 2004). A number of studies provide empirical evidence of network-based job referrals and informational spill-overs in the U.S. labor market (Bayer, Ross, and Topa, 2008; Munshi, 2003; Topa, 2001; Laschever, 2009). Economists have argued that networks are important in the labor market because of market imperfections, such as

*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 appreciate the valuable and significant input that many people have given me on this paper over the years, and I wish to thank in particular Imran Rasul, two anonymous referees, Chris Udry, Mark Rosenzweig, Rohini Pande, Frank Limbrock, and Jeremy Magruder. All errors are my own.

†Department of Economics, Northwestern University. Email: l-beaman@northwestern.edu.

firms' difficulties screening applicants (Montgomery, 1991). However, relatively little attention has been paid to labor market dynamics in the empirical literature and the possibility of within-network competition for job information. Competition within a network could mitigate the network's ability to overcome labor market imperfections.

The literature has mainly focused on how membership in a social network can be advantageous for labor market outcomes and has highlighted that larger social networks are better than small networks. Empirical studies by Munshi (2003) and Edin et al. (2003) both find a positive relationship between the number of network members and successful labor market outcomes in the U.S. and Sweden, respectively.[1] However, Calvo-Armengol (2004) argues theoretically that pairwise-equilibrium stable networks are inefficient because within-network competition creates a negative externality, and increases in network size can actually increase unemployment. Here, I propose a dynamic model of a social network with multiple cohorts that shows how social networks both help transmit information and also heighten competition. I test the model empirically using data on refugees resettled in the U.S.

Refugees are an important and useful population for testing this theory. The United States has a long history of refugee resettlement, having accepted over 2.4 million refugees and asylum seekers since 1975. Hundreds of thousands of asylum seekers apply for protected status in Europe every year. There is an active international policy debate on the best way to resettle refugees with a particular focus on the question of how living in an enclave affects integration. The existing evidence is equivocal,[2] and Edin et al. (2003) argue that there is heterogeneity in the effect of living in an enclave on earnings according to ethnic group "quality." While this literature focuses on the broader question of the economic consequences of living in ethnic enclaves, I will argue

[1] An exception is Wabha and Zenou (2005), who find that the probability of finding a job through a social network is concave with respect to population density in a city, suggestive of competition.

[2] Edin et al. (2003) find a positive relationship between the stock of immigrants and labor market outcomes while Borjas (2000) finds a negative relationship between the proportion of individuals within a city from the same country and assimilation for refugees in the U.S.

that understanding dynamics and social network structure can help explain these results using a systematic framework.

The refugee context also facilitates an analysis of social networks in a dynamic context since the relocation process creates an identification strategy with plausibly exogenous variation in network structure. The number of refugees who are resettled to different cities and from different countries of origin varies every year. I focus on refugees resettled by one resettlement agency, and refugees resettled by this agency without family already in the U.S. cannot choose where to reside within the U.S. I argue that this institutional feature, since the data contain all of the individual characteristics of the refugees known by the agency at the time of placement, allows me to exploit the variation in the size of a refugee's social network over time to test the predictions of the theoretical model.

The model is one in which individuals share job information with social network members within an overlapping generations framework. The overlapping generations structure highlights the importance of the short-run negative correlation in employment status across certain network members discussed in Calvo-Armengol and Jackson (2004). Depending on the vintage of other network members, having access to a larger network may actually lead to a deterioration of individuals' labor market outcomes due to competition among unemployed members for job information.[3] However, the relationship between network size and outcomes changes over time: an increase in the size of a given cohort will first decrease the employment rate of cohorts who arrive close in time to the large cohort, but will improve outcomes for those cohorts that arrive sufficiently later. The relationship between the size of a social network and labor market outcomes is therefore heterogeneous and dynamic.

The main challenge in identifying network effects is separating the causal impact of the

[3] The competition effect arises not because of an increase in labor supply in the face of fixed demand: this effect occurs even when the probability of receiving job information is constant, irrespective of network size.

network from the role of unobservable characteristics shared by network members, especially as individuals potentially sort themselves into localities and networks based on these factors.[4] To address the identification problem and disentangle the competition and information effects 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. The IRC selects the geographic location for refugees without family already in the U.S. after receiving a document from the U.S. State Department with a limited amount of demographic information on each refugee. The individual characteristics known by the IRC at the time of a refugee's placement are also available in the data set. By defining a refugee's social network as refugees of the same nationality resettled in the same city, I ensure that network size is uncorrelated with unobserved individual characteristics.

I use variation in the relative size and structure of refugee social networks across cities and ethnic groups over time to examine the dynamic relationship between social networks and labor market outcomes. By focusing on dynamics, the empirical analysis isolates both the competition and positive information effects from an increase in network size. This variation also facilitates a flexible econometric specification. Though unobserved individual characteristics are not correlated with network size, there may be unobserved nationality group or city level factors which are correlated with the agency's placement decisions. The econometric analysis allows for nationality group, city, and time heterogeneity.

The empirical analysis shows that an increase in network size has heterogeneous effects across network members, creating both negative and positive ramifications for employment outcomes. A one standard deviation increase in the number of network members who arrive in the U.S. one year prior to a newly arrived refugee lowers his probability of being employed by 4.9

[4] This is one of the "Reflection" problems articulated by Manski (1993).

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 4.3 percentage points. Among employed refugees, an increase in the number of senior network members has a strong positive effect on wages, but there is no evidence that a change in the number of network members who arrived in the current or previous year affects wages. The results suggest that by providing job information, networks affect wages both through the employment rate and job quality.

While we often think about migrants exploiting their pre-existing ties to generate a network in the U.S., as is the case for Mexican migrants in Munshi (2003), refugees are unlikely to have pre-existing ties with one another prior to resettlement in the U.S. It is therefore striking that refugees from the same country of origin provide support to one another. Common nationality therefore appears to be a strong basis for the creation of new networks, as seen in Bandiera et al. (2008).[5]

From a methodological standpoint, this paper shows that accounting for dynamics is essential for accurately assessing the role of social networks in the labor market. A static analysis of network effects, using, for example, 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. Moreover, as demonstrated in this paper, a static analysis may miss the presence of network effects completely, due to the two offsetting effects.

The remainder of the paper is organized as follows. The theoretical framework of job information transmission within social networks is outlined in section 2. Details on the institutional background and data are provided in Section 3, and Section 4 discusses the empirical strategy. The

[5]I thank an anonymous referee for pointing out this argument.

empirical results for both employment and wages are presented in section 5. Section 6 addresses alternative explanations for the empirical findings and robustness, and the paper concludes in section 7.

2 Theoretical Framework

2.1 A Model of Employment Rates

The theoretical framework builds on the models developed by Calvo-Armengol and Jackson (2004) and Boorman (1975). In the model, agents randomly receive job offers. If the agent is unemployed, he accepts the offer. However, if he is employed, he passes along the job offer to an unemployed network member. 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.[6]

The basic structure and timing of the model is as follows. Each agent lives and works for S periods. An agent's cohort is defined by the period they enter the labor market; in the empirical context of this paper, it is the time the refugee arrives in the U.S.. Each cohort c has N_c agents. If agent i in cohort c is employed at the end of period t, then $s_{ic}^t = 1$, and $s_{ic}^t = 0$ if agent i is unemployed. All agents within a cohort are identical, and the employment rate within the cohort at time t is denoted as s_c^t. The employment rate for a cohort in its first period in the labor market, i.e. a newly-arrived refugee cohort in this context, is indicated by s_c^t when $c = t$ (or equivalently, s_c^c). There is also the probability that any employed agent will lose his job at the very beginning of the

[6]This paper therefore also abstracts from the distinction made by Granovetter (1973) on weak versus strong ties. In doing so, this framework provides an additional reason for a within-competition effect other than the one highlighted by Calvo-Armengol (2004), which was driven by direct versus two-links-away network members.

period, captured by the exogenous breakup rate b. Information about job openings then arrives: an agent hears about a job opening with probability a, and the job arrival process is assumed to be independent across agents. If an agent is unemployed and receives job information, he fills the position. However, if the agent is already employed, he passes along the information to a randomly selected network member who is unemployed. Once job information arrives and is, if applicable, referred to unemployed members, jobs are immediately accepted.

Since each individual receives information directly with probability a, the total number of jobs available is scaled up as the size of the network increases. The advantage of this approach is that it enables the model to isolate the network effect since it is not confounded with a change in the ratio of labor demand to labor supply. This assumption also reflects the empirical setting in which the predictions are tested, as the size of the network is small compared to the entire economy.[7]

In this simple model, an individual can become employed either by receiving job information directly or through a network member passing information, captured by the term r^t defined below. This structure is formalized below:

$$s_c^t = a + r^t \qquad \text{if } c = t \tag{1}$$

The probability of becoming employed for an individual entering the market for the first time, i.e. $c = t$, is the probability of randomly getting an offer, a, plus the probability of receiving a referral, r^t, defined below. For individuals already in the market for at least one period, the probability of being employed is:

$$s_c^t = (1-b)s_c^{t-1} + (1-(1-b)s_c^{t-1})(a + r^t) \qquad \text{if } c \leq t \leq c + (S-1) \tag{2}$$

[7] As shown in table 2, 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 Phoenix, a change in the number of refugees arriving in each city in a given year is unlikely to have a general equilibrium effect on the job arrival rate or the distribution of wages. This validity of this assumption will be discussed further in the empirical sections.

The first term is the probability of being employed in period $t-1$ and retaining the job. All unemployed members can receive referred information and randomly receive offers. Therefore, the second term captures the probability of becoming employed as in equation (1), weighted by the probability of being unemployed. The probability of receiving job information through an employed network member is represented by:

$$r^t = \sum_{k=t-S+1}^{t-1} \frac{aN_k(1-b)s_k^{t-1}}{\sum_{k=t-S+1}^{t} N_k - (1-b)\sum_{k=t-S+1}^{t-1} N_k s_k^{t-1}} \tag{3}$$

The numerator is the total number of jobs which are available in the network to be passed, determined by the number of employed individuals who receive information in that period. The probability, r^t, is this numerator divided by the number of potential recipients. Potential recipients are those who are unemployed at the beginning of that period, after the exogenous breakup has occurred.

As an individual receives a random draw of a job offer in every period he is in the labor market, the probability of employment is weakly increasing with time spent in the market. Two testable predictions are derived from this model.

Proposition 1 *For all values* $0 < a < 1$ *and* $0 < b < 1$*, an increase in cohort size* N_j *decreases* s_c^j *for all* c*.*[8]

Proof: See appendix.

That is, the immediate effect of an increase in the size of the arriving cohort is a decline in the employment rate for all cohorts in the market in that period, including the one which is made exogenously larger. The intuition is that since s_c^{j-1} does not change, increasing N_j only increases the number of unemployed individuals seeking job information from network members while leaving the number of employed members unchanged. The striking part of the prediction

[8]This claim holds for all values of a and b such that $s_c^j \neq 1$ for all c and j.

is that the deleterious effect from an increase in network size is not the result of an increase in labor supply driving down wages or employment rates in equilibrium since the assumption that each individual faces a constant rate a of hearing about a job directly ensures that labor demand is held fixed. Instead, the negative effect comes from competition between network members for information provided by already employed individuals. However, this competition effect dampens over time as highlighted in Proposition 2.

Proposition 2 *The impact of an increase in N_j on s_k^k is monotonically increasing from period j to period $j + S - 1$.*

Proof: See appendix.

Propositions 1 and 2 show that despite an initial negative impact on all cohorts, the negative effect is mitigated for new cohorts that enter the labor market over time, as more members of cohort j become employed with new draws in each period. Numerical analysis of the model shows that $\frac{\partial s_k^k}{\partial N_j} > 0$ for at least one cohort (the cohort that enters in the period before the large cohort $k = j + S - 1$ exits) and usually earlier cohorts for all parameter values. While an increase in the size of cohort j negatively impacts cohorts who arrive close in time to period j, the increase in cohort size eventually increases the employment rate for cohorts who arrive sufficiently later. As cohort j's employment rate increases over time, its larger size becomes an asset to the entire network.

To illustrate the model's predictions, Figure 1 provides an example where $a = .35$, $b = .2$, and $S = 4$. The graph shows a comparison in the employment rates of a control network with constant cohort size and a treatment network that experiences a one-time shock, doubling the size of cohort j. Both the treated cohort j and the cohort $j + 1$ entering in the next period experience a lower employment rate in their first period in the market than they would have experienced in the absence of the cohort size shock. However, the following cohorts, $j + 2$ and $j + 3$, experience higher employment rates in all periods in which 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. Propositions 1 and 2 are robust to alternative assumptions regarding the arrival rate, such as allowing the job information arrival rate to depend on employment status.

2.2 Wages

Calvo-Armengol and Jackson (2007) analyze a general model which includes stochastic wages. They find that despite a short-run competition effect, there is a positive steady-state correlation across network members in terms of employment rates and wages.

I incorporate wages into the framework introduced in section 2.1 in the following way. With probability a, an individual i receives job information which has an attached wage offer, w^o_{ict}. As before, 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 $w^o_{ict} > w_{ict}$. Alternatively, if $w^o_{ict} < w_{ict}$, the offer is passed on to a randomly selected network member who is either unemployed or employed with a wage less than w^o_{ict}. Accordingly, an individual with a wage of x always has a positive probability of receiving an offer with a wage greater than x through the network. This ensures that the agent's optimal policy is to accept the highest offer received in any period if he is unemployed, even if that offer is low. This is a classic result from the search literature with on the job search, established by Burdett (1978) and used by Mortensen and Vishwanath (1994) in their model of search through a social network. The agent's dynamic optimization problem is presented in the appendix with a brief description of the standard assumptions about the agent's utility during unemployment necessary for this result. Since all offers are eventually used, the predictions for employment are identical to those described above.

Wages are *iid* draws from a distribution $F(w)$ with support $[\underline{w}, \bar{w}]$. $G^t(w)$ is the cumulative distribution function (CDF) of accepted wages; this distribution is determined endogenously in the model. $H(w)$ is the CDF of wages available to be passed within the network and is also endogenously determined. As in Burdett (1978), wages increase with tenure in the labor market. This result is not driven by an explicit return to experience; instead, wages improve as individuals receive more draws from the wage distribution the longer they are in the market as they discard low wage offers.

Individuals can receive job offers through two channels in the model: the direct channel, via a random draw from $F(w)$, and the indirect channel through the network. The distributions of wages available through the two channels differ systematically.

Lemma 1 *The distribution of wages through the direct channel, $F(x)$, first order stochastically dominates the distribution of wages available through the network, $H(x)$.*

Proof: See appendix.

The intuition is that jobs which are passed through the network are those which are rejected by the original recipient, and therefore the network offers jobs with on average lower wages. If all network members were employed at $\bar{w}$, however, then $F(x) = H(x)$.

The model in section 2.1 predicts that an increase in the size of an arriving cohort j decreases s_j^j because r^j declines. An increase in N_j also affects the average wages of those employed.

Proposition 3 *For all values $0 < a < 1$ and $0 < b < 1$, an increase in cohort size N_j (weakly) increases the average wage for those employed in cohort j in period j, denoted as w_j^j.*

Proof: See appendix.

This result is driven by the fact that the fraction of individuals who gained their jobs through the direct channel relative to the indirect channel increases in period j after an increase in N_j. Since $F(x)$ dominates $H(x)$, average wages (weakly) increase for that cohort.

After an initial increase in average wages for cohort j, average wages for cohorts who enter thereafter decline. Proposition 2 showed that the fraction of individuals employed through the indirect channel increases with each subsequent period. This means that average wages then decline for cohorts $j+1$, $j+2$, etc., relative to cohort j as more individuals become employed with lower wages through the network channel. Simulation of the model shows that for all values of a and b, there is a monotonic decline in wages for each cohort entering after cohort j.[9] Panel A of Figure 2 demonstrates the model for two sets of parameters, both of which generate average employment levels similar to the data. The figure presents the average employment rates and wages conditional on employment for six cohorts in their first period in the market, after break-up occurs. Each cohort is in the labor market for five periods in this case, i.e. $S = 5$, and cohort j is tripled in size. For example, the graph shows the average employment rate and average wages among the employed for cohort $j-1$, which represents cohorts who just entered the market when all cohorts are the same size. When $a = 300$ and $b = 150$, 66% of cohort members are employed at the end of that cohort's first period in the market. The figure on the left in Panel A shows that the employment pattern is the same as described in the previous section: the larger cohort, j, drives down employment rates for itself and the cohorts who enter right afterwards. However, cohorts $j+3$ and $j+4$ experience higher employment rates. The figure on the right shows that wages follow the opposite pattern. Cohort j has higher than average wages, since fewer of the employed took network jobs. In contrast, cohort $j+4$ has lower than average wages at the end of its first period in the labor market as a higher fraction of the employed in that cohort becomes employed through the indirect channel.

[9] For the simulations I use the uniform distribution. The uniform distribution used in the figures has an average of 5 plus 5.15 to reflect the minimum wage. The search literature has argued that either Pareto or exponential distributions are appropriate as wage distributions (Lancaster and Chesher, 1983; Lynch, 1983). An exponential distribution shows results consistent with those for the uniform distribution. I provide an example with an exponential distribution in the next section, with a mean of 11.

2.3 Information Passing Technology

In the sections above, I modelled the network as perfectly efficient: job offers are passed on until all information is used. However, this is a strong assumption and may not be realistic. If at least some of the job information arrives in the network as employee referrals, then only individuals who know the employee may be eligible for the position. Jobs may also fill quickly, creating a constraint on the amount of time available to pass along job information. I model an inefficient network by assuming there is a limited number of rounds of job information passing and no coordination across network members. When information can only be passed once, for example, individuals who receive more than one offer become employed at the highest wage, discarding the lower wage offers. This implies that some of the offers available in the network are lost. The percentage of available information in the network in period j that is lost due to constraining passing to one round is expressed by the following term:[10]

$$L^j = \frac{\bar{N}^j}{a\hat{E}^{j-1}} \sum_{z=2}^{a\hat{E}^{j-1}} \binom{a\hat{E}^{j-1}}{z} (\frac{1}{\bar{N}^j})^z (\frac{\bar{N}^j - 1}{\bar{N}^j})^{a\hat{E}^{j-1}-z} (z-1) \tag{4}$$

where $\bar{N}^j$ is the total size of the network in period j. The employment rate for a given cohort entering the market is accordingly:

$$s_j^j = a + \sum_{k=j-S+1}^{j-1} \frac{aN_k(1-b)s_k^{j-1}}{\bar{N}^j - (1-b)\sum_{k=j-S+1}^{j-1} N_k s_j^{j-1}} (1 - L^j) \tag{5}$$

This implies that the employment level declines due to the inefficiency. It is straightforward to show that the average wages of those employed decrease, as low wage draws from $H(w)$ are now discarded. Panel B of Figure 2 demonstrates the inefficient model using the same parameter values

[10] The binomial distribution describes the fraction of individuals in the network who get multiple offers, assuming the number of network offers available is sufficiently large to disregard replacement. For example, the percentage who get z offers is $\binom{aE}{z}(\frac{1}{N})^z(\frac{N-1}{N})^{aE-z}$.

as in Panel A. The average employment rate falls from 66% to 32% for $a = 300$ and $b = 150$ and average wages rise by about \$.75.

The inefficiency also affects how a change in network size alters the employment rate and average wages. Consider first the effect of an increase in N_j on s_j^j. Differentiating L with respect to N shows that the fraction of network jobs lost declines with an increase in N, holding E fixed. This corresponds to the change in L^j from an increase in N_j, since $\hat{E}^{j-1}$ is not affected by the change in N_j. As there are more recipients for the same number of job offer draws, there are fewer individuals who get more than one offer. Therefore, the initial negative competition effect will be attenuated due to the inefficiency. Differentiating L with respect to the number of network referral offers available, holding $\bar{N}$ constant, shows that more jobs are lost due to the inefficiency when there are more jobs available in the network. This corresponds to the impact of the change in N_j on cohorts that enter after cohort j. The total network size is then constant but E increases with each period as the large cohort increases its employment rate. This implies that once the impact of increasing cohort j increases the employment rate in, for example, period $j + 3$, the inefficiency will also reduce the gains for those cohorts. Simulation of the model shows that the inefficiency attenuates the impact of a change in cohort size, but the same pattern shown in section 2.1 holds. Panel B of figure 2 shows that, compared to Panel A, the effect of increasing the size of cohort j on the employment rates of entering cohorts is smaller for each entering cohort.

Wages in an Inefficient Network

There are two ways in which a change in network size affects wages in the inefficient model. First, as before, a change in N_j affects the proportion of the employed in the network who receive job information directly versus indirectly, which alters average wages (the “composition effect”). Second, a change in N_j affects the number of job offers an individual receives (the “inefficiency” effect). An increase in N_j will decrease the number of wasted offers for cohort j in period j as discussed above. Since the low wage network offers are less likely to be discarded, this reduces average wages. In contrast, an increase in N_j that increases s^{j+4}_{j+4} will increase average wages since employed individuals are more likely to get multiple offers. The inefficiency effect is strongest for high levels of a: more job offers generated create a larger scope for individuals to receive multiple offers. Figure 3 demonstrates how, for a given value of b, the prediction for average wages among the employed flips as a increases.

At intermediate levels of a, it is the composition effect which dominates for cohort j, increasing w^j_j. However, the inefficiency effect dominates for cohorts entering later, in this case for cohorts $j+3$ and $j+4$. Therefore, w^{j+4}_{j+4} increases due to the shock in N_j. Why is the composition effect the strongest in period j? The intuition is that the negative effect from a decline in r^j is larger, in absolute value terms, than the positive effect from r^{j+3}. The result is a non-monotonic U shape at intermediate levels of a. Note the average employment level for the parameter values generating the U shape are often similar to the employment rates observed among refugees in the sample. Figure 4 demonstrates the same U shape pattern in a simulation with an exponential distribution of wages and where agents are allowed to pass on job information in two rounds. This demonstrates that the U shape pattern is robust to alternative formulations of the inefficient model.

The next section introduces the data and context where the theory will be tested. To summarize the model predictions, Propositions 1 and 2 hold for all versions of the model, while Proposition 3 only holds if the network is perfectly efficient. If the network does not achieve efficiency, we may observe one of three patterns in the data, as demonstrated in Figure 3.

3 Institutional Environment and Data

3.1 Refugee Resettlement Process

As of 2007, there were over 11.4 million refugees under the Office of the UN High Commissioner for Refugees (UNHCR) mandate. Resettlement of refugees to North America and Europe is a key strategy used by the UNHCR to find solutions to long-term, persistent refugee crises. The United States has a long history of refugee resettlement. While refugees are not the largest immigrant class in the U.S., they constitute a sizeable component. For example, 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 networks in the labor market. Research on the economic performance of refugees in the U.S. is limited, largely due to data constraints.[11] This paper focuses exclusively on refugees, who are distinct from asylum seekers. Refugees are a well-defined group: according to Immigration and Nationality Act 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.

How does one become a refugee in the U.S.? INS (now USCIS) officers adjudicate individual cases in refugee processing centers around the world.[12] 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.

[11] 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.

[12] These centers are most often within refugee camps, but individuals can also apply for refugee status in local U.S. embassies.

The PRM's final role in the resettlement process is to allocate all accepted cases to one of ten contracted voluntary resettlement agencies who provide social services to refugees. This study focuses on refugees who did not attain admittance via family reunification (non-family reunification refugees) and who were resettled by one voluntary resettlement agency, the IRC. For these individuals, the IRC has sole discretion in determining where the refugee will be resettled among its 16 regional offices.[13] The IRC receives information from the State Department about each refugee's characteristics, such as country of origin and demographic information. 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 employee meets the refugee or his family 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.

The first responsibility of the IRC is to provide core services to refugees during the first 30 days in the U.S. Core services include finding housing, providing furniture and basic supplies, facilitating access to English as a Second Language (ESL) courses and employment services, assisting with administrative processes such as enrolling in food stamps or Temporary Assistance for Needy Families (TANF) if applicable, and facilitating access to health care if needed. The IRC can also enroll qualified refugees into the government Matching Grant (MG) program, providing financial and employment services support for the first 120 days. However, the refugee must be willing to accept the first suitable job offered. The objective is to promote refugees' ability to become self sufficient within 120 days. The belief that refugees should begin to work as quickly as possible, irrespective of their language or educational backgrounds, is widely held by voluntary agencies and the PRM. They argue that refugees are better off working immediately and pursuing ESL and other

[13]The process for determining where a refugee will be resettled within the U.S. may differ significantly across resettlement agencies; accordingly I describe here only on IRC's procedures.

training courses in the evening. Some regional offices provide additional long-term support, similar to what is provided by other non-profit organizations, including ESL courses and financial literacy training. The IRC is required by the PRM to track labor market outcomes for refugees at 90 days after arrival and 120 days for participants in the MG program as a way to monitor success.

3.2 Placement Policy

The IRC does not have an explicit placement rule for 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 a pre-existing ethnic or nationality-based community. They also attempt to send each refugee to an office which has either a staff member or a volunteer who speaks the same language as the refugee, though they hire translators as needed. 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. 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 (non-family reunification refugees). The estimates are generated using projections from the State Department on the number of admitted refugees expected from each region of the world. Often, the actual numbers differ substantially from those anticipated, as refugees flows 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, the allocation of non-family reunification cases across regional offices has to be adjusted to accommodate the predestined cases. Finally, the total number of refugees sent to a particular office is also a function of average employment statistics at the regional office level. Overall, the IRC employee who is solely in charge of placement states that the effectiveness of strategic decision-making is limited since she never knows when a refugee who is assigned to the IRC by the State Department will actually be allowed to travel. To highlight the stochastic component, consider 2005: there were cases that were given refugee status in 2001 but who arrived in 2005 due to delays associated with heightened September 11, 2001 security requirements.

Based on interviews with IRC staff, the organization uses a limited amount of the remaining information provided by the PRM in the allocation process. Given that this is difficult to verify, the econometric strategy discussed in section 4 will allow for the possibility that additional information is used.

The placement policy means that, on average, each nationality group has approximately five main cities they are resettled in. Once a city is established as a site for a particular group, table 1 shows that it regularly receives refugees from that country. The table presents the correlation matrix for the number of people the IRC allocated to each nationality/regional office pair, i.e. the size of each cohort across four years from 1997-2005. The strongest correlation is between periods t and $t-1$, and the correlation monotonically decreases thereafter in the time elapsed between cohorts.

3.3 Data

The data from the IRC is comprised of over 1,700 male adults who arrived in the U.S. between 2001 and 2005.[14] Sample respondents did not have family members already in the U.S. to assist in their resettlement and were subject to the placement policy described above. There are three components to these data. First, a set of demographic variables were compiled by the INS and the PRM prior to the refugee's arrival in the U.S. and given to the IRC, including ethnicity, date

[14]There are three groups whose placement does 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.

of birth, country of first asylum, the size of the family being resettled, initial English language level for some cases, and education received in the home country. The econometrician observes all individual characteristics which the IRC knew during the allocation process.

Labor market outcomes, in particular employment status and hourly wage, were collected by the IRC at 90 days after each refugee's arrival. Summary statistics for the key variables included in the empirical analysis are found in Table 2. The average employment rate is 66%, after only 90 days in the U.S. While only providing information on short-term labor market outcomes is a limitation of the data, the period of time is not too short for refugees to have found jobs. In fact, as an intense job search period for refugees, this time period is particularly interesting for examining how social networks affect labor market access. Moreover, the short-term nature of the data makes it easier to isolate the effect of the network since other factors, such as differential investment in U.S.-specific human capital or returns to experience, are limited or non-existent.

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 unavailable.

A wide variety of nationalities is represented in the data. The largest groups are from Afghanistan, Bosnia, Liberia, Somalia, and the Sudan, although there are, in total, 38 different nationality groups. The IRC has 16 offices where they resettle non-family reunification cases.[15] The sample excludes those refugees who are HIV positive, who comprise 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 nationality group's network in a given geographic space, I define the social network as non-family reunification refugees of the same nationality who were resettled in the same regional office. Since the aggregate data are available from 1997

[15] 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.

onwards, this measure of network size for an individual includes fellow refugees resettled in the four years prior to that individual's arrival. The reason the network is restricted to refugees without family already in the U.S. 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 having to 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 can still be "resettled" by the IRC. In contrast, since the IRC rents an apartment for each non-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 are used nevertheless in the robustness analysis in section 6.2.2.

Two additional data sources are used. The 2000 census available through IPUMS is used to construct an alternative network measure as discussed in section 6.2.1. In order to provide intuition on the magnitude of the results, I use a survey of refugees and asylum seekers collected by the Department of Health and Human Services' Office of Refugee Resettlement (ORR) between 1993 and 2004. The data does not identify refugees who entered the U.S. through family reunification, however, so the ORR sample is therefore not 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. Using labor market outcomes as of 90 days after arrival and the aggregate data on IRC placements from 1997-2005, the model predictions are tested using the following econometric specification:

$$Y_{ijkt} = \alpha + \gamma_1 N_{ijk(t)} + \gamma_2 N_{jk(t-1)} + \gamma_3 N_{jk(t-2)} + \gamma_4 N_{jk(t-3)} + \gamma_5 N_{jk(t-4)} + X_{ijkt}\beta + \delta_{jt} + \phi_k + \epsilon_{ijkt} \quad (6)$$

for individual i from nationality group j in city k who arrived at time t. I will refer to k as city but in practice it is the metropolitan area that the IRC regional office covers; for example, Atlanta includes the greater Atlanta metro area including suburbs. Y_{ijkt} represents either employment status or wages for individual i. $N_{jk(t-1)}$ and $N_{jk(t-2)}$ are the total number of refugees who arrived during the fiscal year one year and two years prior to refugee i's arrival, respectively. $N_{jk(t-3)}$ and $N_{jk(t-4)}$ are analogously defined. N_{ijkt} is the number of refugees from nationality group j resettled by the IRC in city k who arrived in fiscal year t up to i's specific date of arrival. Note that refugees who arrived after i are excluded from N_{ijkt} since they would be not be acting as competitors nor providers of job information to individual i. The error term is clustered at the nationality group-city level.

Propositions 1 and 2 of the model predict that having a larger number of network members who arrived in the same year will decrease the probability of a new refugee obtaining employment and his wage.[16] The competition effect diminishes with the length of time elapsed between the arrival of the two cohorts, and the positive information effect eventually dominates. The model

[16] In the model, increases in the size of an arriving cohort affects employment rates in a nonlinear fashion: larger increases in the cohort size increases the effect size but at a decreasing rate. Unfortunately due to the sample size, I am not able to estimate this nonlinear relationship.

does not predict how long the competition effect will dominate, so there is a set of coefficients which are consistent with the model. One example is negative point estimates of γ_1 and γ_2 and positive estimates of γ_3, γ_4 and γ_5.[17]

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, it is difficult to separate the effect of having a larger network from the correlation between common characteristics network members share and their labor market outcomes. 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; the first originates from sorting along unobservable individual characteristics, and the second is omitted city and nationality group characteristics correlated with network size as a result of the IRC's placement strategy. The former is addressed by including a flexible functional form of covariates to span the information set available to the IRC at the time of placement. X_{ijkt} then captures the individual characteristics which are correlated with network size. The remaining individual attributes in ϵ_{ijkt} are uncorrelated with N_{jk} since they are not known by the IRC at the time of placement.

Since the IRC resettles multiple nationality groups across multiple cities during many years, there is variation in social network size across cities, nationality groups and over time. This variation facilitates a fixed effects strategy to minimize the second concern of unobservable factors at the city and nationality group level.[18] In the preferred specification, I include only δ_{jt} and ϕ_k controls. Time variant heterogeneity at the nationality group level is captured by δ_{jt}. The specification allows groups to differ, for example, if one nationality group has lower human capital on average or if the

[17]A cohort and time period are defined as one year due to data limitations. One year may correspond to multiple periods and cohorts in the model. Therefore, we could have observed all γ parameters to be positive or negative.

[18]The empirical strategy is close to that used by Bertrand et al. (2000) to look at how social networks affect welfare use in the U.S.

types of people who become refugees vary across sending countries. Since this term varies across year of arrival, there can be unobservable differences in quality within a group across cohorts. Unobservable factors at the city level, such as variation in labor demand, may also affect labor market outcomes and may influence IRC's placement decisions. Metropolitan-area fixed effects, ϕ_k, are therefore included. There are three additional potential sources of bias. Shocks at the city level (ξ_{kt}), match quality between nationalities and cities (τ_{jk}), and finally shocks to match quality (ν_{jkt}). The first two can also be addressed through fixed effects since the econometric strategy exploits time variation to test the predictions of the dynamic model. The main results will show both the preferred specification and one including ξ_{kt}, δ_{jt} and τ_{jk}. Finally, I will argue in Section 6.1 that time varying shocks to a nationality-city pair are unlikely to explain the pattern observed.

5 Empirical Results

5.1 Probability of Employment

5.1.1 Main Results

Table 3 shows the results of analyzing employment in a dynamic context. The results from estimating the preferred specification, including city and nationality group-year fixed effects, are in columns 1 and 2. Column 1 shows that a larger number of network members who arrived in the current and prior years strongly decreases the probability of employment for a new entrant.[19] This is in contrast to the finding in Munshi (2003) that network members always have a weakly positive effect on employment. 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 66%, this constitutes a decline of over 7%. Another way to frame the magnitude of the

[19]A linear probability model is used due to the number of indicator variables, particularly in column 3, though a probit framework provides very similar results for both specifications. In particular, the coefficient on $N_{jk(t-3)}$ is significant at the 10% and slightly larger in magnitude than the LPM estimates.

effect is to compare it to an approximation of the returns to experience. Analysis done with the ORR data shows that each additional year spent in the U.S. is associated with an increase in the employment rate of 3.4%. The negative competition effect from a one standard deviation increase in the number of network members arriving a year prior more than offsets the positive effects of an additional year of residence. The negative network effect is therefore an economically significant factor in determining short-run refugee labor market unemployment rates.

Consistent with the model, 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 4.3 percentage points. The numbers of refugees who arrived in the prior three and four years are not statistically significant individually but are jointly significant at the 5% level. One-sided tests indicate that while $\hat{\gamma}_1$ is not statistically more negative than $\hat{\gamma}_2$, $\hat{\gamma}_2 < \hat{\gamma}_3$. Though the coefficients on the numbers of refugees who arrived in years $t-3$ and $t-4$ are positive and jointly significant, they are not larger than that of the $t-2$ network as the model would predict. This may be the result of a limitation of the data: I only observe refugees right after their arrival in the U.S.. Secondary migration, both in and out of city k, is likely to be higher for refugees who were resettled three or more years prior to the new arrival.[20] The out-migration rate within the first 90 days was 6.96%, and in-migration is unobserved. In the model, systematic in-migration into networks which received a large cohort two or more years ago would lead to an attenuation of the effect of $N_{jk(t-3)}$ and $N_{jk(t-4)}$.[21] Finally, the point estimates on $N_{jk(t-2)}$ are

[20]The measure of 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.

[21]A model that has no job information sharing within networks but in which the total number of job offers per network is fixed would also generate a negative competition effect. It would not, however, generate both the negative and positive effects. I explored whether a model with fixed job offers where individuals either leave the resettlement city over time or others arrive can generate the same pattern as the network model. As long as more people do not out migrate than the additional people who were sent to the city, neither random nor systematic secondary migration

not different from $N_{jk(t-3)}$ or $N_{jk(t-4)}$ in column 1 of table 3. As discussed in section 2.3, The inefficiency in the network leads to attenuation of the effect of a change in network size. For some parameter values, an increase in cohort j increases the employment rate for cohorts $j+2$ through $j+4$, but the effect on those three cohorts is very similar to one another. Given the limited sample size, this is another hypothesis for why we do not observe $\hat{\gamma}_3$ and $\hat{\gamma}_4$ greater than $\hat{\gamma}_2$. Unfortunately it is difficult to disentangle between these possible explanations.

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. The specification estimated in column 1 of table 3 contains a limited number of demographic covariates. To ensure that the individual characteristics known by the IRC at the time of placement are sufficiently controlled for, column 2 includes a wider range of control variables than column 1. The coefficients of interest are largely unchanged and continue to be significant. The education variables, whose coefficients are suppressed for brevity, are not jointly significant; a higher measured English level is, though, positively correlated with employment status.

Could systematic variation in unobserved factors, which affect particular nationalities in certain cities, generate a spurious relationship between network size and labor market outcomes as predicted by the model? A "comparative advantage" would arise if, for example, there are characteristics or skills common to all individuals in nationality group j which receive a higher return in particular cities k. If the IRC uses information on how the returns to skills vary across cities in placement decisions, network size would be endogenous. I provide two pieces of evidence against this concern. First, if the IRC used information on match quality, there would be a positive relationship between network size in time t and outcomes. That is, the organization would be able to exploit this information and send more refugees to a place which has a high match quality,

generates the positive relationship between large cohorts that arrived two years prior and new arrivals' labor market outcomes.

generating a positive relationship between N and outcomes even from year t. This is in contrast to table 3. Second, I can also rule out this explanation by including a richer set of fixed effects than is used in columns 1 and 2 of table 3. Nationality-city fixed effects can be included, in addition to nationality-year and city-year, since the network variables vary at the nationality group-city-year level. Despite the large number of additional controls this requires, column 3 shows that the results are robust to the inclusion of the above listed fixed effects.[22] The estimates on $N_{jk(t)}$ and $N_{jk(t-1)}$ are more negative than in the specification in columns 1 and 2, with $N_{jk(t-1)}$ significantly more negative at the 5% level. The patterns on $N_{jk(t-2)}$ and $N_{jk(t-3)}$ differ slightly but both are consistent with the model's predictions. However, the effect of $N_{jk(t-4)}$ is the opposite sign than expected though the estimate is very imprecisely estimated. The effect of having a network member with four years' experience in the U.S. may be highly co-linear with having a permanently larger social network, which is captured by the nationality-city fixed effect, resulting in an insignificant effect of $N_{jk(t-4)}$. Unobserved time invariant match quality is therefore not driving the results. The city-year fixed effects remove the possibility that time variant city employment shocks are influencing the estimates. These additional controls help to identify the causal effect of network size on employment under the maintained assumption that year-specific shocks do not vary at the city-nationality level and are not correlated with network size. Column 4 shows that the results are robust to the inclusion of a wider set of individual characteristics.

If the resources of a field office are constrained, the IRC may reduce services to refugees in a large cohort. This is an alternative explanation for the negative relationship between outcomes and own cohort size. Since the results in columns 3 and 4 in table 3 include city-year dummies, an overall increase in the number of refugees going to a given office is not driving the negative effect. The only potential omitted variable bias stems from variation in IRC resources across cities which

[22]There are 198 nationality-city, 72 city-year and 115 nationality-year pairs.

are nationality group-specific. The IRC's main activities which are group specific relate to whether the refugee has someone in the office, either staff, volunteer or available translator, who speaks his language. This may affect the refugee's ability to be prepared more generally for the labor force. It may also affect the type of employment assistance the refugee receives: if one case worker handles all refugees from a given country, and contacts with employers are case worker-specific, an increase in the size of an arriving cohort could hinder employment outcomes because the case worker's contacts are stretched too thin. To investigate both of these possibilities, I re-estimated the main specifications including the number of staff members and volunteers who speak one of the refugee's languages. In results available upon request, the coefficients of interest are not significantly altered.[23] The number of volunteers who speak the same language positively impacts employment while there is no impact from the number of staff members speaking the refugee's language. This is suggestive that the IRC is able to balance staff resources across groups. A second possibility is that certain groups are best at certain types of jobs or industries, and a large arriving cohort from a given country taxes the number of contacts the IRC has in that industry. This is unlikely since the jobs refugees hold initially, as shown by a mean sample wage of $7.48, are low wage and low skilled jobs which any group could do. The robustness checks in section 6 will also address this concern. A final piece of evidence, not shown for brevity, is that a refugee's use of ESL or employment services is not reduced if he arrived in a larger cohort, though this data is only available for a subset of the sample.

Another alternative interpretation is that the competition effect is not due to social network dynamics but merely reflects an increase in labor supply, driving down wages as in Cortes (2008). However, the number of new arrivals in a given network is very small each year, around 30 people on average. It is therefore unlikely that such a small addition to the labor market in cities such as

[23]The results are similar when using discrete variables indicating whether the refugee was placed in an office with at least one staff member or volunteer who speaks at least one of the languages spoken by the refugee.

Atlanta and Salt Lake City influences the unemployment rate or the equilibrium wage level even if labor markets are segmented.[24]

The refugee context facilitates the identification of both the negative and positive network effects described in the model: I argue that sorting into localities based on unobservables is minimized, and refugees are a small group that will not generate general equilibrium effects. In a more general setting, however, both of these factors will affect the dynamics of labor market outcomes. First, the general equilibrium effects may exacerbate the negative effect and make it persist for longer. Second, other types of migrants and non-migrants can choose the city or neighborhood they live in. Therefore, sorting into cities may minimize the competition effect.[25] It is hard to predict which of the two would prevail. One piece of evidence is from Munshi (2003). He finds that among Mexican migrants in the U.S., the number of established migrants has a positive effect on the probability of employment among new arrivals to the U.S. This suggests that even for the largest immigrant group in the U.S., the positive effects are likely to persist.

Static Approach

To contrast the econometric strategy focusing on dynamics with the static approach, I estimate the effect of the stock of network members on the probability of employment. Columns 5 and 6 of table 3 show that the static analysis produces contradictory results. 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 as in Column 6, the

[24] A more formal approach to differentiating the two different "competition" effects uses a shock to the job arrival rate a induced by September 11, 2001. The network model would predict a reduction in the competition effect when a declines while said reduction in a would exacerbate the negative impact in a classic model of labor supply. The data is consistent with the network model: the competition effect is muted post-9/11. The results are available from the author upon request.

[25] There may still be a significant competition effect even with sorting: as highlighted by Calvo-Armengol (2004), endogenously formed networks can be too large, driving up unemployment within the network. Galeotti and Merlino (2010) also find that individuals over-invest in network contacts because the negative externality affecting other network members is not internalized.

effect becomes insignificant and the point estimates are negative. The static analysis would be inconclusive regarding the existence of social networks providing job information to newly arrived refugees, highlighting the importance of considering dynamics when analyzing the role of social networks in labor markets. By properly structuring the network variables to reflect the dynamic relationship between network size and labor market outcomes, it is easy to detect network-based job information transmission, and the results are not sensitive to the specification used. Furthermore, the static approach fails to identify the presence of the short-run competition effect. The dynamic approach therefore sheds new light on how networks function and how they affect the labor market outcomes of network members.

5.2 Job Quality: Wages

The effect of changes in network size on wages is an empirical question. Proposition 3 showed that there should be a positive correlation between $N_{jk(t)}$ and wages, conditional on being employed. However, this only holds if the network is perfectly efficient. If the network does not achieve efficiency, there two countervailing effects. On the one hand, an increase in $N_{jk(t)}$ will decrease wages since an individual will receive fewer 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 rise. We may observe one of three patterns demonstrated in Figure 3.

Columns 1 and 2 of table 4 show that the size of the network in periods $t-2$, $t-3$ and $t-4$ are positive and statistically significant. The coefficient on the number of network members who arrived in year t is positive but not statistically significant. The point estimate on and $t-1$ is close to zero but again imprecisely estimated. This is broadly consistent with the model for intermediate values of the job arrival rate, a. As Figures 3 and 4 show, the model can generate a U shape relationship when the network is not perfectly efficient. This highlights the importance of relaxing the assumption of perfect efficiency in social networks.

The model also shows that variation in network size has heterogeneous effects on hourly wages (unconditional on employment) due to the dynamic relationship between network size and wages, as already shown for employment status. To confirm this, I estimate equation 6 with the full male sample and impute wage offers as zero for the unemployed. The selection bias problem induced by the imputation of wages with OLS is discussed at the end of this section. Consistent with the model, column 3 of table 4 shows that a standard deviation increase in the number of network members who arrived in time $t-1$ decreases the wage by \$.75. An increase of one standard deviation in $N_{jk(t-2)}$ increases the hourly wage by \$.48. These results reflect both the effect of the network on employment and the direct effect on wages. Including a wider range of demographic and other control variables as in column 4 of table 4 leads to little change in the network coefficients. An alternative estimator, least absolute deviations (LAD), also produces similar results, as shown in column 5 table 4. LAD alleviates the concern of censured wage offers under the assumption that unemployed individuals receive wage offers below the median offer made to employed workers with comparable skills.[26]

6 Alternative Explanations and Robustness

This section considers alternative explanations for the empirical findings and shows robustness checks.

[26]The model predictions imply that the effect of network size should have an effect on *offer* wages, but wage offers to an unemployed individual below his reservation wage are unobserved, creating a censoring problem. Without a suitable exclusion restriction to use a classic selection model, I follow Neal and Johnson (1996) and Johnson, Kitamura, and Neal (2000) by imputing a wage of zero to unemployed individuals and use LAD estimation (Heckman, 1974). 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.

6.1 Falsification Test

The indicator variables for city-year, nationality-city and city-nationality group assuage many concerns about endogenous placement of refugees into cities. As described in section 4, however, the empirical specification can not directly address the possibility that there are shocks to unobserved match quality between a nationality group and a city. Such shocks could be the result of changes in the returns to skills across cities or compositional changes within a group over time, for example. There therefore may be a systematic pattern to IRC placement of refugees which is confounded with the social network mechanism of interest.

I offer several pieces of evidence against this alternative explanation. First, as already mentioned, if the IRC used information on changes to the returns to skills, there would be a positive relationship between network size in time t and outcomes which we do not observe in the data. Second, there were significant lags - up to multiple years - between when the IRC made the decision of where to place a refugee and when that refugee was given final security clearance to arrive in the U.S. The uncertainty of arrival times was heightened by the procedural changes implemented after September 11, 2001. This would make exploiting such patterns extremely difficult. Third, strategic placement would produce much less regular patterns in the flows of refugees from a particular group, which is not what we observe in table 1.[27]

The structure of the data also facilitates a "falsification test" using the 2001-2004 sample. I test whether the number of refugees who arrive in year $t+1$ impacts the probability of employment. Since there is no possible interaction between sample refugees and $t+1$ refugees, there should be no significant relationship. A systematic placement pattern used by the IRC may, however, generate a significant relationship between N_{t+1} and employment in period t.

[27] The positive correlation across cohorts highlights that network flows are not independent shocks across time, as was considered in the model. It is unlikely that the co-linearity of the network size variables affects the results since separate regressions including each network variable, $N_{jk(t)}$ to $N_{jk(t+4)}$, create similar estimates.

Columns 1 through 4 in table 5 are consistent with the maintained identification assumptions. 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 estimates are quite similar to columns 1 through 4 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 that time-variant match quality shocks correlated with network size explain the results. It also casts further doubt on the concern that there is a dynamic relationship between IRC resources and the number of refugee arrivals.

6.2 Definition of Social Network

The above analysis defines the relevant social network as the number of non-family reunification refugees resettled in a city from the same country of origin. This section checks relaxes this assumption using two alternative specifications. I first use the 2000 Census to construct a measure of the network which includes all immigrants types. I then return to defining the network as including only refugees but include family reunification refugees.

6.2.1 Alternative Network Measure: Census Data

The 2000 census data available through IPUMS provides an alternative measure of the network in which network members are those who come from the same country of origin or ethnic group irrespective of immigration status. I construct a measure of network size which includes all individuals from a country of origin group in a given metropolitan area. This measure includes all immigrants, not only refugees. I calculate the size of the network at the level of the metropolitan statistical area (MSA).[28] I define a social network by either nationality or ethnicity, depending on the availability

[28] Most are defined by a single MSA but some networks are defined using multiple MSAs, such as New York.

of the relevant code in IPUMS which most closely matches the information available on the IRC refugees. IPUMS data also provides the age of each network member as well as the year of arrival in the U.S. Therefore the network size variable is specific to the year of arrival of each network member and is restricted to only prime age adults. Since the census does not contain information on the foreign born's visa type or residency status/citizenship, this measure includes all immigrant types, ranging from illegal immigrants to permanent residents and naturalized citizens.

Since the data structure differs from the network measure used above, I use an alternative empirical specification. In order to test the hypothesis using the 2000 census data, I restricted the size of the network 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 IRC refugees who arrived in 2001 and 2002.

$$Y_{ijkt} = \alpha + \phi_1 N_{jk(t=1999)} + \phi_2 N_{jk(t=1999)} * \lambda_{2001} + X_{ijkt}\beta + \delta_j + \phi_k + \lambda_{2001} + \epsilon_{ijkt} \tag{7}$$

Y_{ijkt}, X_{ijkt}, δ_j, ϕ_k, and ϵ_{ijkt} are defined in section 4. $N_{jk(t=1999)}$ is the number of immigrants from group j who arrived in city k in 1999 according to the census, and λ_{2001} is an indicator for those refugees who arrived in 2001. Estimates showing ϕ_1 to be positive and ϕ_2 to be negative would be consistent with the model. ϕ_2 captures the differential network effect across the two cohorts: 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. Network members who arrived in 1999 may be competitors for job information with those who arrived in 2001. However, by 2002, they would be better able to provide referrals to the newly resettled refugees having acquired additional job information over time.

Using the 2000 census to create a second network measure allows for flexibility in the definition of the social network. However, some individual network members in the sample may have self-selected their preferred location based on a number of unobserved factors. While this measure of the network is more susceptible to selection bias, it does test the generality of the job information sharing effect across two independently constructed network measures.

Table 2 shows the average network size in 1999 as measured by the Census. Although the average is higher than in the IRC sample, this masks heterogeneity: network size in 1999 is 0 for almost half of the sample and the median value is 27. Of course, many immigrants from a given country are missing since the IPUMS data is based on a 5% sample of the Census and information which could lead to identification of individuals is aggregated. Nevertheless, the Census data suggests that the IRC network may be the principal source of individuals from their nationality group for many refugees in the sample.

Probability of Employment

Column 1 in table 6 shows 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. 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 two 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 7 - provides qualitatively similar results. Column 2 of table 6 shows OLS estimates of the effect of network size on wages among those employed. There is no significant effect of network size on wages of those employed and the standard errors are quite large. The coefficients in column 3 for the full sample with LAD estimation are more informative. The network effect for refugees who arrived in 2002 is positive and the interaction term is negative; both are statistically significant at the 5% level. The sum of the two network coefficients is negative but statistically insignificant. This closely parallels the results found in the employment results and the theoretical model's predictions.

6.2.2 Family Reunification Refugees

I argued in section 3.3 that family reunification refugees are unlikely to participate in the same social networks as those who arrive without family. However, do family reunification refugees also influence labor market outcomes? Are there other immigrants from the same country of origin in the city who are not refugees who participate in the same network?

Column 5 of table 5 includes the total refugees the IRC resettled by nationality and resettlement city from years t to $t-4$. These measures are instrumented using the number of non-family reunification refugees using 2SLS. This of course requires the strong assumption that the non-family reunification refugees only affect outcomes through their effect on total network size. With that caveat, this specification provides a qualitatively similar picture to the main results in table 3. In results not shown, an alternative specification including the number of family reunification refugees in addition to the variables containing the number of non-family reunification refugees produces estimates of the effect of the number of non-family reunification refugees in years t to $t-4$ on sample refugees' outcomes which are not significantly altered. The number of family reunification refugees also shows no significant relationship with sample refugees' employment status. This provides additional evidence against the hypothesis that IRC employment services are driving the results: since the IRC also provides services to family reunification refugees, we would anticipate a negative relationship between the number of family reunification refugees and the sample refugees' employment status.

7 Conclusion

This paper presents evidence on the importance of social 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. While network architecture generates information transmission and can improve labor market outcomes, there is also a within-network competition effect that, empirically, has an economically sizable negative impact on labor market outcomes. The paper therefore tempers previous findings in the empirical literature which show the capacity for social networks to overcome market frictions.

A static analysis of the effect of total network size on labor market outcomes conflates the two opposite effects. This paper highlights that ignoring the dynamic relationship between employment, wages, and social network structure can erroneously create inconclusive results about the role of social networks in the labor market. This paper suggests that the tenure composition of social networks and other dimensions of structure are important in assessing the full network effect.

Resettlement policies vary widely; some European countries actively disperse refugees across the country (Edin et al., 2004), and there is an active policy debate on the impact of living in an enclave on a refugee or asylum seeker's labor market integration. Edin et al. (2003) and Borjas (2000) find contradictory evidence on the relationship between the proportion of individuals within a city from the same country and labor market performance for refugees in Sweden and in the U.S. respectively. This paper argues that a given network size may prove beneficial in some settings while negative in others.[29] Therefore, while the fact that refugee social networks provide labor market information to its members suggests a potential drawback to immigrant dispersal policies, dynamics are important to consider. Future research is needed to disentangle the long-run role of social networks in creating incentives or disincentives for integration and investments in host country-specific human capital.

References

Bandiera, O., I. Barankay, and I. Rasul (2008). Social capital in the workplace: Evidence on its formation and consequences. *Labour Economics, 15*, 725–749.

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

Bertrand, M., E. F. P. Luttmer, and S. Mullainathan (2000). Network effects and welfare cultures. *Quarterly Journal of Economics 115*, 1019–1055.

Bewley, T. F. (1999). *Why Wages Don't Fall During a Recession.* Cambridge: Harvard University Press.

Boorman, S. A. (1975). A combinatorial optimization model for transmission of job information through contact networks. *Bell Journal of Economics 6*, 216–249.

Borjas, G. J. (2000). Ethnic enclaves and assimilation. *Swedish Economic Policy Review 7*, 89–122.

Burdett, K. (1978). A theory of employee job search and quit rates. *American Economic Review 68*, 212–220.

Calvo-Armengol, A. (2004). Job contact networks. *Journal of Economic Theory 115*, 191–206.

Calvo-Armengol, A. and M. O. Jackson (2004). The effects of social networks on employment and inequality. *American Economic Review 94*, 426–454.

Calvo-Armengol, A. and M. O. Jackson (2007). Networks in labor markets: Wage and employment dynamics and inequality. *Journal of Economic Theory 132*, 27–46.

Cortes, K. E. (2004). Are refugees different from economic immigrants? Some empirical evidence on the heterogeneity of immigrant groups in the United States. *Review of Economics and Statistics 86*, 465–480.

Cortes, P. (2008). Effect of low-skilled immigration on u.s. prices: Evidence from CPI data. *Journal of Political Economy 116*, 381–422.

Edin, P.-A., P. Fredriksson, and O. Aslund (2003). Ethnic enclaves and the economic success of immigrants: Evidence from a natural experiment. *Quarterly Journal of Economics 118*, 329–357.

Edin, P.-A., P. Fredriksson, and O. Aslund (2004). Settlement policies and the economic success of immigrants. *Journal of Population Economics 17*, 133–155.

[29] Moreover, the heterogeneity found by Edin et al. (2003) across ethnic groups will also exist within an ethnic group and will vary over time.

Freidlin, M. and A. D. Wentzell (1984). *Random Perturbations of Dynamical Systems.* New York: Springer-Verlag.

Galeotti, A. and L. Merlino (2010). Endogenous job contact networks. *Mimeo, University of Essex.*

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology 78*, 1360–1380.

Heckman, J. (1974). Shadow prices, market wages, and labor supply. *Econometrica 42*, 679–694.

Ioannides, Y. M. and L. D. Loury (2004). Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature 42*, 1056–1093.

Johnson, W., Y. Kitamura, and D. Neal (2000). Evaluating a simple method for estimating black-white gaps in median wages. *American Economic Review 90*, 339–343.

Lancaster, T. and A. Chesher (1983). An econometric analysis of reservation wages. *Econometrica 15*, 1661–1676.

Laschever, R. (2009). The doughboys network: Social interactions and labor market outcomes of World War I veterans. *Mimeo, University of Illinois at Urbana-Champaign.*

Lynch, L. (1983). Job search and youth unemployment. *Oxford Economic Papers 35*, 271–282.

Manski, C. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economics Studies 60*, 531–542.

Mortensen, D. and T. Vishwanath (1994). Personal contacts and earnings. It is who you know! *Labour Economics 1*, 187–201.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. *Quarterly Journal of Economics 118*, 549–599.

Neal, D. A. and W. R. Johnson (1996). The role of premarket factors in black-white wage differences. *Journal of Political Economy 104*, 869–895.

Topa, G. (2001). Social interactions, local spillovers and unemployment. *Review of Economics Studies 68*, 261–295.

Wabha, J. and Y. Zenou (2005). Density, social networks and job search methods: Theory and application to Egypt. *Journal of Development Economics 88*, 443–473.

Young, H. (1993). The evolution of conventions. *Econometrica 61*, 57–84.

8 Tables

Table 1: Correlation Coefficients of Refugee Cohort Sizes: 1997-2005

	Current Year	Prior Year	2 Years Prior	3 Years Prior
# Refugees Resettled in Current Year	1			
# Refugees Resettled in Prior Year	0.5394	1		
# Refugees Resettled in 2 Years before	0.2859	0.4744	1	
# Refugees Resettled in 3 Years before	0.2711	0.3794	0.5892	1
# Refugees Resettled in 4 Years before	0.2399	0.3473	0.3971	0.5815

Table 2: Summary Statistics

	Mean	Std. Dev.	No. Obs
IRC Data:			
Age	33.99	11.05	1720
HH Size	2.76	2.04	1720
Employment rate	0.66		1720
Hourly Wage (conditional on employment)	7.48	1.36	1125
Spoke No English Upon Arrival	0.466		1453
Primary School	0.180		1720
Secondary School	0.464		1720
University or Above	0.202		1720
None, vocational or adult education	0.153		1720
Muslim	0.251		1720
IRC Exemption from Employment	0.059		1720
# Refugees Resettled in Year t	10.32	13.47	1720
# Refugees Resettled in Year $t-1$	29.47	34.13	1720
# Refugees Resettled Year $t-2$	25.16	43.82	1720
# Refugees Resettled Years $t-3$	31.70	58.21	1720
# Refugees Resettled Years $t-4$	34.10	55.66	1720
# Family Reunification Refugees Resettled in Year t	15.45	27.24	1720
# Family Reunification Refugees Resettled in Year $t-1$	18.77	48.25	1720
# Family Reunification Refugees Resettled in Year $t-2$	19.50	58.55	1720
# Family Reunification Refugees Resettled in Years $t-3$	31.70	58.21	1720
# Family Reunification Refugees Resettled in Years $t-4$	34.10	55.66	1720
2000 Census Data:			
Network Members who Arrived in 1999	150.83	267.84	753

a The mean of *# Refugees Resettled in Year p* is the number of refugees who arrived in year p averaged over all nationality groups in all cities.

Table 3: Employment Probability on Network Size

	1	2	3	4	5	6
# Refugees Resettled in Year t [e]	-0.222 **	-0.228 *	-0.346 **	-0.339 **		
	(0.110)	(0.118)	(0.145)	(0.162)		
# Refugees Resettled in Year $t-1$ [e]	-0.144 **	-0.119 *	-0.310 ***	-0.280 ***		
	(0.067)	(0.069)	(0.094)	(0.096)		
# Refugees Resettled in Year $t-2$ [e]	0.098 **	0.089 **	0.002	-0.015		
	(0.043)	(0.043)	(0.067)	(0.071)		
# Refugees Resettled in Year $t-3$ [e]	0.042	0.045	0.174 ***	0.162 **		
	(0.035)	(0.035)	(0.064)	(0.064)		
# Refugees Resettled in Year $t-4$ [e]	0.037	0.030	-0.068	-0.087		
	(0.039)	(0.039)	(0.066)	(0.068)		
# Refugees Resettled Years t to $t-4$ [e]					0.029 ***	-0.025
					(0.011)	(0.049)
Age	0.022 ***	0.021 ***	0.024 ***	0.024 ***	0.021 ***	0.023 ***
	(0.006)	(0.007)	(0.006)	(0.008)	(0.006)	(0.006)
Age^2	-0.0003 ***	-0.0003 ***	-0.0004 ***	-0.0003 ***	-0.0003 ***	-0.0003 ***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
HH Size	-0.017 ***	-0.015 **	-0.024 ***	-0.021 ***	-0.017 ***	-0.022 ***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)
P-value of $t-3$ and $t-4$	0.040	0.047	0.021	0.031		
P-value of education variables		0.410		0.225		
P-value of initial English level variables		0.004		0.001		
P-value of religion variable		0.191		0.393		
N	1720	1720	1720	1720	1720	1720

a # Refugees Resettled in Year t is the number of refugees with the same nationality as the surveyed individual who were resettled in the same year in the same city.

b SE are in parentheses & clustered by nationality-city. All specifications also include the following covariates: indicators for Matching Grant Enrollment, summer, delayed social security number, a special schooling program (Jubilee) and IRC exemption status, which the IRC usually gives either for health problems or due to a large number of dependents in the household which prevents the adult from working.

c Columns 1, 2 and 5 include fixed effects for nationality-year and city. Columns 3, 4 and 6 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 The network size coefficients and standard errors are multiplied by 100.

Table 4: Wages on Network Size: Conditional, Full Sample and LAD

	1	2	3	4	5
# Refugees Resettled in Year t [b]	0.163	0.202	-0.024 **	-0.023 *	-0.903 ***
	(0.349)	(0.336)	(0.011)	(0.012)	(0.291)
# Refugees Resettled in Year $t-1$ [b]	0.033	0.052	-0.024 ***	-0.021 ***	-0.553 ***
	(0.249)	(0.261)	(0.008)	(0.008)	(0.139)
# Refugees Resettled in Year $t-2$ [b]	0.641 **	0.550 **	0.004	0.003	0.912 ***
	(0.273)	(0.260)	(0.006)	(0.006)	(0.132)
# Refugees Resettled in Year $t-3$ [b]	0.425 **	0.427 **	0.014 ***	0.013 **	-0.108
	(0.175)	(0.167)	(0.005)	(0.005)	(0.109)
# Refugees Resettled in Year $t-4$ [b]	0.485 **	0.391 *	-0.002	-0.003	0.306 ***
	(0.231)	(0.232)	(0.005)	(0.005)	(0.095)
Age	0.082 ***	0.068 ***	0.229 ***	0.220 ***	0.168 ***
	(0.021)	(0.022)	(0.040)	(0.052)	(0.018)
Age^2	-0.0011 ***	-0.0010 ***	-0.0034 ***	-0.0032 ***	-0.0026 ***
	(0.0003)	(0.0003)	(0.0005)	(0.0007)	(0.0002)
HH Size	0.023	0.022	-0.156 ***	-0.132 **	-0.071 ***
	(0.023)	(0.023)	(0.050)	(0.054)	(0.018)
P-value of $t-3$ and $t-4$	0.003	0.004	0.030	0.057	0.005
P-value of education variables		0.046		0.150	
P-value of initial English level variables		0.077		0.000	
P-value of religion variable		0.640		0.287	
N	1125	1125	1706	1706	1706

a SE are in parentheses and clustered by nationality-city. Additional controls are included in all specifications as listed in Table 3.

b The network size coefficients and standard errors are multiplied by 100 in columns 1-2, and 5.

c Columns 1 and 2 are conditional on employment while Columns 3-5 use the full sample.

d Columns 1-4 include fixed effects for nationality-year, city-year and nationality-city.

e Columns 2 and 4 include additional individual covariates including: education, initial English level, religion.

f Column 5 shows LAD estimates with fixed effects for nationality group, year of arrival, and city.

Table 5: Employment: Falsification and Family Refugees

	1	2	3	4	5
# Refugees Resettled in Year $t+1$ [a]		0.085		0.051	
		(0.098)		(0.156)	
# Refugees Resettled in Year t [f]	-0.275 **	-0.299 **	-0.544 ***	-0.544 ***	-0.219 **
	(0.132)	(0.137)	(0.165)	(0.165)	(0.105)
# Refugees Resettled in Year $t-1$ [f]	-0.160 **	-0.156 **	-0.314 ***	-0.306 ***	-0.074
	(0.073)	(0.075)	(0.117)	(0.115)	(0.059)
# Refugees Resettled in Year $t-2$ [f]	0.074 *	0.074 *	0.032	0.032	0.127 ***
	(0.043)	(0.043)	(0.083)	(0.083)	(0.039)
# Refugees Resettled in Years $t-3$ [f]	0.029	0.032	0.167 **	0.171 **	0.029
	(0.038)	(0.038)	(0.073)	(0.076)	(0.037)
# Refugees Resettled in Years $t-4$ [f]	0.048	0.048	-0.049	-0.048	-0.001
	(0.041)	(0.041)	(0.091)	(0.092)	(0.030)
Age	0.026 ***	0.026 ***	0.027 ***	0.027 ***	0.021 ***
	(0.007)	(0.007)	(0.008)	(0.008)	(0.006)
Age^2	-0.0004 ***	-0.0004 ***	-0.0004 ***	-0.0004 ***	-0.0003 ***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
HH Size	-0.019 **	-0.019 **	-0.027 ***	-0.027 ***	-0.017 ***
	0.008	0.007	0.008	0.008	0.006
Types of Refugees Included in Network Measures	Non-Family	Non-Family	Non-Family	Non-Family	ALL
P-value of $t-3$ and $t-4$	0.037	0.027	0.062	0.064	0.504
N	1340	1340	1340	1340	1720

a SE are in parentheses & clustered by nationality-city. The network size coefficients and standard errors are multiplied by 100.

b Additional controls are included in all specifications as listed in Table 3. Columns 1-4 only contain years 2001-2004.

c Columns 1, 2 and 5 include fixed effects for nationality-year and city. Columns 3 and 4 include fixed effects for nationality-year, city-year and nationality-city.

d In column 5, the network size variables include both family and non-family reunification refugees. The specification is 2SLS with the number of non-family reunification refugees in time t back to $t-4$ included in the first stage as instruments.

Table 6: Employment and Wage Effects Using Census Data for Network Measure

	Employment		Wages		
	1	2	3	4	5
Network size which arrived in 1999 [g]	0.0284 **	0.0220 *	0.013	0.249 ***	0.172 *
	(0.0121)	(0.0126)	(0.055)	(0.088)	(0.091)
Network size which arrived in 1999 * Refugee arrived in 2001 [g]	-0.0275 **	-0.0253 **	-0.043	-0.238 ***	-0.201 **
	(0.0114)	(0.0121)	(0.058)	(0.095)	(0.097)
Age	3.388 ***	3.020 ***	0.073 **	0.270 ***	0.229 ***
	(0.912)	(1.002)	(0.034)	(0.068)	(0.072)
Age Sq	-0.0530 ***	-0.0482 ***	-0.0009 **	-0.0042 ***	-0.0037 ***
	(0.0122)	(0.0129)	(0.0005)	(0.0009)	(0.0009)
HH Size	-1.789 *	-1.474	-0.017	-0.162 **	-0.122
	(1.093)	(1.046)	(0.028)	(0.078)	(0.076)
p-value of education variables		0.014			0.038
p-value of initial English level variables		0.066			0.003
p-value of religion variables		0.027			0.002
p-value of occupation variables		0.757			0.736
No obs	753	753	523	742	742
Adjusted R squared	0.187	0.199	0.300	0.183	0.207

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.

9 Figures

Figure 1: Graphical Example of Model with Constant Wages

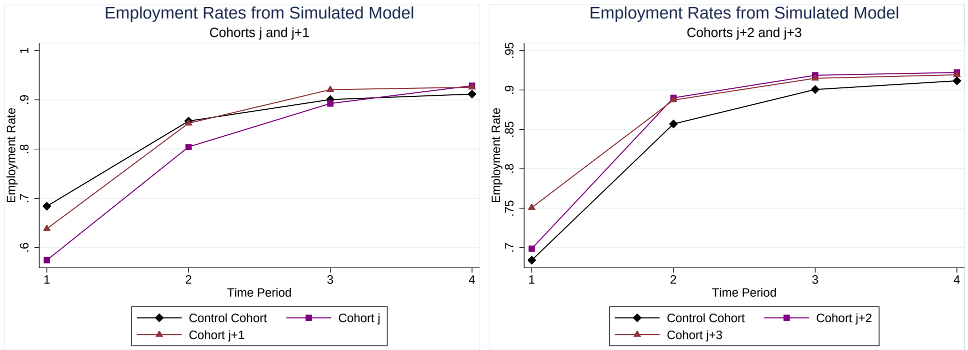


Figure 2: Graphical Example of Models with Uniformly Distributed Wages

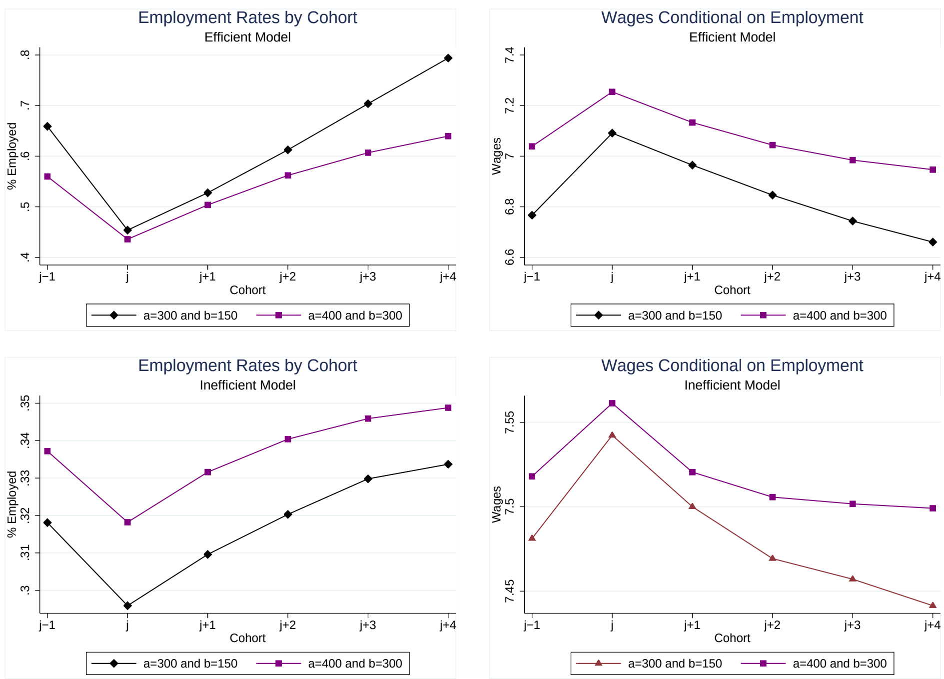


Figure 3: Effect on Employment and Wages in Inefficient Model: $b = 50$

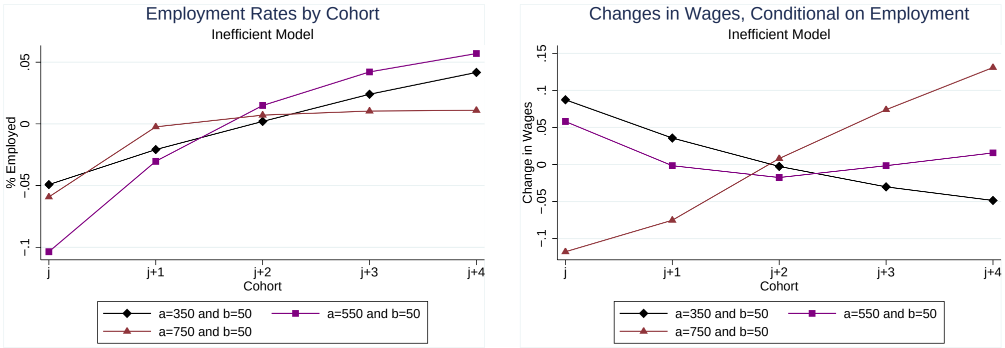
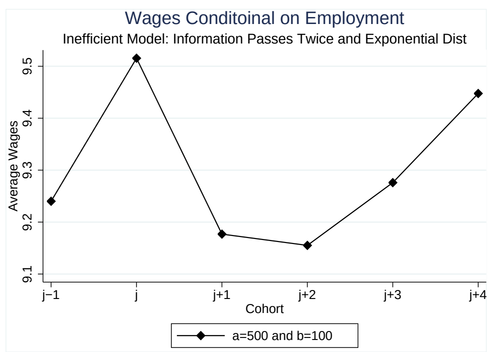


Figure 4: Effect on Wages using Exponential Distribution and Less Inefficiency



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_c^j *for all* c*.*[30]

Proof:

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

$$s_j^j = \frac{a(N_j + \sum_{c' \neq j} N_{c'})}{N_j + \sum_{c' \neq j} N_{c'}(1 - (1-b)s_{c'}^{j-1})} \tag{8}$$

[30]This claim holds for all values of a and b such that $s_c^j \neq 1$ for all c and j.

Differentiating with respect to N_j gives:

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

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

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

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

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

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

A.2 Proposition 2

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

Proof:

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

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

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

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

Therefore, need to show: $N_j s_j^j + s_{j-1}^j > s_{j-1}^{j-1} + s_{j-2}^{j-1}$

$$\frac{aN_j(2+N_j)}{2+N_j-(1-b)(s_{j-1}^{j-1}+s_{j-2}^{j-1})} + (1-b)s_{j-1}^{j-1} + \frac{a(2+N_j)(1-(1-b)s_{j-1}^{j-1})}{2+N_j-(1-b)(s_{j-1}^{j-1}+s_{j-2}^{j-1})} > s_{j-1}^{j-1} + s_{j-2}^{j-1} \tag{9}$$

Using the steady-state properties of the economy,[31] equation (9) will hold with equality if $N_j = 1$.

Since (9) holds with equality when $N_j = 1$, inequality holds if the expression is increasing in N_j.

$$\frac{a(2+N_j)(1+N_j-(1-b)s_{j-1}^{j-1})}{2+N_j-(1-b)(s_{j-1}^{j-1}+s_{j-2}^{j-1})} + (1-b)s_{j-1}^{j-1} \tag{10}$$

[31] 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.

Differentiating equation (10) with respect to N_j gives:

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

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

since $1+N_j-(1-b)s_{j-1}^{j-1}>0$. Since the denominator is greater than zero, if the numerator is greater than zero, then equation (9) holds.

Step 2: $s_{j+2}^{j+2}(N_j) > s_{j+1}^{j+1}(N_j)$

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

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

Need to show:

$$s_{j+1}^{j+1} + N_j s_j^{j+1} > N_j^j s_j^j + s_{j-1}^j$$

Let $N_j = 1+x$. Rearranging equation (10) gives:

$$(s_{j+1}^{j+1}-s_j^j)+(s_j^{j+1}-s_{j-1}^j)+x(s_j^{j+1}-s_j^j)>0 \tag{11}$$

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

$$s_j^{j+1} = (1-b)s_j^j + [1-(1-b)s_j^j]s_{j+1}^{j+1}$$

$$s_{j-1}^j = (1-b)s_{j-1}^{j-1} + [1-(1-b)s_{j-1}^{j-1}]s_j^j$$

$$s_{j-2}^{j-1} = (1-b)s_{j-1}^{j-1} + [1-(1-b)s_{j-1}^{j-1}]s_{j-1}^{j-1}$$

The left hand side (LHS) of equation (11) 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 $N_js_j^j + s_{j-1}^j > s_{j-1}^{j-1}+s_{j-2}^{j-1}$ as shown above, this implies

$$s_j^j - s_{j-1}^{j-1} > +s_{j-2}^{j-1} - s_{j-1}^j - xs_j^j > (s_{j-1}^{j-1}-s_j^j)(1-(1-b)s_{j-1}^{j-1}) - xs_j^j$$

Substituting the above gives:

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

since $s_{j+1}^{j+1} > s_j^j$ as shown above and $s_{j-1}^{j-1} > s_j^j$ as in Proposition 1.

A.3 The Worker's Dynamic Optimization Problem

The value of being unemployed in time t is:

$$\begin{aligned} U^t &= z + a\int_0^\infty \beta\{max[U^{t+1}, W(x)]\}dF(x) \\ &\quad +r^t\int_0^\infty \beta\{max[U^{t+1}, W(x)]\}dF(x)[1-G^t(x)] \\ &\quad +\left(1-a-r^t\right)\beta U^{t+1} \end{aligned} \tag{12}$$

The value of being employed at rate w at time t:

$$\begin{aligned} V^t(w) &= w + (1-b)\,a\int_0^\infty \beta\{max[U, V(x), V(w)]\}dF(x) \\ &\quad +(1-b)\,r^t\int_0^\infty \beta\{max[U^{t+1}, V^{t+1}(x), V^{t+1}(w)]\}dF(x)[1-G^t(x)] \\ &\quad +\left(1-a-r^t\right)(1-b)\,\beta V^{t+1}(w) \\ &\quad +ba\int_0^\infty \beta\{max[U^{t+1}, V^{t+1}(x)]\}dF(x) \\ &\quad +br^t\int_0^\infty \beta\{max[U^{t+1}, V^{t+1}(x)]\}dF(x)[1-G^t(x)] \\ &\quad +b\left(1-a-r^t\right)\beta U^{t+1} \end{aligned} \tag{13}$$

Note: r^t and G^t is a function of N which is the vector of cohort sizes. z are unemployment benefits (which are near zero for males in the U.S.). I assume the cost of search is zero if employed or unemployed.

A.4 Lemma 1

Let the total number of people employed in the network in a given time period be denoted by $E^t = \sum_{k=t-(S-1)}^{t} N_k(1-b)s_k^t$. The timing convention is that, between periods, the oldest cohort retires. I therefore introduce $\hat{E}^t$ which captures the number of employed in the network who will continue to be employed at the beginning of the next period. That is, $\hat{E}^t = \sum_{k=t-S+2}^{t} N_k(1-b)s_k^t$. Accordingly we can rewrite r^t, the probability of becoming employed through an indirect (network) offer, from section 2.1 as $r^t = \frac{a\hat{E}^{t-1}}{\bar{N}^t-\hat{E}^{t-1}}$. $\hat{G}^{t-1}$ is the analogous distribution of accepted wage offers, adjusted for retirement. Now define $H(x)$ as the following:

$$1-H(x) = \frac{a\hat{E}^{t-1}(1-\hat{G}^{t-1}(x))(1-F(x))}{\bar{N}^t-\hat{E}^{t-1}(1-\hat{G}^{t-1}(x))} * \frac{\bar{N}^t-\hat{E}^{t-1}}{a\hat{E}^{t-1}} \tag{14}$$

Notice first that if everyone in the network were employed at the highest wage available, the $1-H(x)$ collapses into $1-F(x)$ for all values of x. Proposition 3 can be seen from the following:

$$1-F(x) \geq (1-F(x))\frac{(\bar{N}^t-\hat{E}^{t-1})(1-\hat{G}^{t-1}(x))}{\bar{N}^t-\hat{E}^{t-1}(1-\hat{G}^{t-1}(x))} \quad \forall x$$

If $(\bar{N}^t-\hat{E}^{t-1})(1-\hat{G}^{t-1}(x)) \leq \bar{N}^t-\hat{E}^{t-1}(1-\hat{G}^{t-1}(x))\ \forall x$, then the proposition holds. This is true since:

$$\begin{aligned}\bar{N}^t(1-\hat{G}^{t-1}(x)) - \hat{E}^{t-1}(1-\hat{G}^{t-1}(x)) &\leq \bar{N}^t - \hat{E}^{t-1}(1-\hat{G}^{t-1}(x)) \\ \bar{N}^t(1-\hat{G}^{t-1}(x)) &\leq \bar{N}^t \qquad (15)\end{aligned}$$

since $1-\hat{G}^{t-1}(x) \leq 1$ for all x.

A.5 Proposition 3

Consider first the fraction of the employed who earn a wage greater than x in period j: $\frac{1}{a+r^j}[a(1+F(x)) + \frac{a\hat{E}^{j-1}(1-\hat{G}^{j-1}(x))(1-F(x)}{\bar{N}^j-\hat{E}^{j-1}(1-\hat{G}^{j-1}(x))}]$. This can be re-written as $\frac{a}{a+r^j}(1-F(x)) + \frac{a}{a+r^j}\frac{\hat{E}^{j-1}(1-\hat{G}^{j-1}(x))(1-F(x))}{\bar{N}^j-\hat{E}^{j-1}(1-\hat{G}^{j-1}(x))}$.

Since $r^j = \frac{a\hat{E}^{j-1}}{\bar{N}^j-\hat{E}^{j-1}}$, and accordingly $\frac{a}{a+r^j} = \frac{\bar{N}^j-\hat{E}^{j-1}}{\bar{N}^j}$, the above expression can be written as:

$$\frac{\bar{N}^j-\hat{E}^{j-1}}{\bar{N}^j}(1-F(x)) + \frac{\bar{N}^j-\hat{E}^{j-1}}{\bar{N}^j}\frac{\hat{E}^{j-1}(1-\hat{G}^{j-1}(x)(1-F(x))}{\bar{N}^j-\hat{E}^{j-1}(1-\hat{G}^{j-1}(x))} = \frac{\bar{N}^j-\hat{E}^{j-1}}{\bar{N}^j-\hat{E}^{j-1}(1-\hat{G}^{j-1}(x))}(1-F(x))$$

Differentiating with respect to N_j:

$$\frac{\partial}{\partial N_j} = \frac{1-F(x)}{\bar{N}^j-\hat{E}^{j-1}(1-\hat{G}^{j-1}(x))} - \frac{(1-F(x))(\bar{N}^j-\hat{E}^{j-1})}{(\bar{N}^j-\hat{E}^{j-1}(1-\hat{G}^{j-1}(x)))^2}(1-F(x))[\frac{\hat{E}^{j-1}\hat{G}^{j-1}(x)}{(\bar{N}^j-\hat{E}^{j-1}(1-\hat{G}^{j-1}(x)))^2}] \geq 0$$

Therefore the fraction of those employed with a wage greater than x weakly increases for all x. This implies that the average wage of those employed increases weakly, as stated in the proposition.