

# Social Networks and the Dynamics of Labor Market Outcomes:

## Evidence from Refugees Resettled in the U.S. \*

Lori A. Beaman <sup>†</sup>

June 2009

### 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, the vintage of network members and labor market outcomes is non-monotonic. 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, 2005; Munshi, 2003; Topa, 2001; Laschever, 2005). Networks are

---

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

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

thought to be important in the labor market because of market imperfections, such as firms' difficulties screening applicants (Montgomery, 1991). However, relatively little attention has been paid to labor market dynamics 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.<sup>1</sup> 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. They are important since the United States has a long history of refugee resettlement, having accepted over 2.4 million refugees and asylum seekers since 1975. A large number of refugees are also resettled in Europe every year. There is also 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,<sup>2</sup> and Edin et al. (2003) argue that there is heterogeneity in the effects 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 that understanding dynamics and social network structure can help explain these results using a systematic framework.

Studying refugees is also useful when investigating how social networks matter in a dynamic

---

<sup>1</sup>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.

<sup>2</sup>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.

context because the relocation process facilitates 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; refugees resettled by this agency without family already in the U.S. cannot choose where to reside within the U.S. Coupled with data on all of the individual characteristics of the refugees known at the time of placement, I argue that this institutional feature 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 theoretical framework is one where individuals share job information with social network members within an overlapping generations model. The overlapping generations structure highlights the importance of the negative correlation in employment status demonstrated in the model by Calvo-Armengol and Jackson (2004) for short-run outcomes. 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.<sup>3</sup> 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 and average hourly wage of cohorts who arrive close in time to the large cohort, but will improve outcomes for those cohorts that arrive sufficiently later. The relationship between the size of a social network and labor market outcomes is therefore non-monotonic and dynamic.

The main challenge in identifying network effects is separating the causal impact of the 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.<sup>4</sup> To address the identification problem and disentangle the competition and information effects empirically,

---

<sup>3</sup>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.

<sup>4</sup>This is one of the “Reflection” problems articulated by Manski (1993).

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 my data set. By defining a refugee's social network as refugees of the same nationality resettled in the same city, 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 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 4.4 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. This result suggests that by providing job information, networks affect wages both through the employment rate and job quality.

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 empirical results for both employment and wages are presented in section 5. Section 6 addresses alternative explanations for the empirical findings, 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.<sup>5</sup>

The basic structure and timing of the model is as follows. Each agent lives and works for  $S$  periods, and each cohort  $c$  has  $N_c$  agents. In the empirical context of this paper, a cohort is defined by time of arrival in the U.S. labor market. 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  $s_c^t$ . 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 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 assumption therefore is that the size of the network is small compared to the entire economy. The advantage of this approach is that it enables the model to isolate the network effect directly. This assumption also reflects the empirical setting in which the predictions are tested.<sup>6</sup>

In this simple model, an individual can become employed by either receiving job information

---

<sup>5</sup>This paper therefore also abstracts away from the distinction made by Granovetter (1973) on weak versus strong ties.

<sup>6</sup>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.

directly or through a network member passing information, captured by the term  $r^t$  defined below.

This structure is formalized below:

$$s_c^t = a + r^t \quad \text{if } c = t \quad (1)$$

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:

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

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:

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

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.<sup>7</sup> Two

---

<sup>7</sup>Proof for a finite  $S$  available from the author upon request.

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  $c$ .<sup>8</sup>*

*Proof:* See appendix.

That is, the immediate effect of an increase in the size of the arriving cohort is to decrease 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 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. The assumption that each individual faces a constant rate  $a$  of hearing about a job directly ensures that labor demand is held fixed, so that the labor supply effect is not driving the model's prediction. 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 between  $k = j$  and  $j + S - 1$ .<sup>9</sup>*

*Proof:* See appendix.

Propositions 1 and 2 show that despite an initial negative effect on all cohorts, cohort  $j$ 's negative impact on employment on subsequent cohorts is mitigated over time, as the cohort gains experience in the labor market. Numerical analysis of the model shows that  $\frac{\partial s_k^k}{\partial N_j} > 0$  for at least

---

<sup>8</sup>This claim holds for all values of  $a$  and  $b$  such that  $s_c^j \neq 1$  for all  $c$  and  $j$ .

<sup>9</sup>As refugees will be in the U.S. labor market for many years and to simplify the analysis of the model, I restrict the analysis to the first period cohort  $j$  is in the market until its second to last period in the market.

one cohort in particular (the cohort that enters in the period before the large cohort  $k = j + S - 1$ ) and usually earlier cohorts. 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. Numerical analysis shows this is true at all parameter values. 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 where there is 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 A Model of Employment Rates and Wages

Calvo-Armengol and Jackson (2007) analyzes 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 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,  $w_{ict}^o$ . 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  $w_{ict}^o > w_{ict}$ . Alternatively, if  $w_{ict}^o < w_{ict}$ , the offer is passed to a randomly selected unemployed network member.<sup>10</sup> Wages are *iid* draws from the uniform distribution  $w \sim U[\underline{w}, \bar{w}]$ .  $w_c^e$  denotes the average wage for employed network members in cohort  $c$  in period  $c$ .<sup>11</sup> The analysis of the model is done by simulation.

One pattern that emerges in this model is that 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 and discard low wage offers.

There are two ways in which a change in network size affects wages in this model. First, a change in  $N_j$  affects the number of job offers an individual receives, and therefore, the wage is affected as well. Second, a change in  $N_j$  affects the proportion of the employed in the network who receive job information directly versus indirectly. Job offers available to the unemployed from within the network are those with wages which are sufficiently low that the employed network member who initially receives the job information rejects the offer. The average wage from an offer received indirectly is lower than the average wage from an offer received directly. A change in the proportion of the network becoming employed through the indirect channel therefore affects average wages within the cohort.

These two different channels work in opposite directions, creating an ambiguity in the model.

---

<sup>10</sup>This could produce an incentive for network members to behave strategically since it may be optimal for an unemployed network member to refuse the job offer and wait another period to be eligible for job information from other network members. This assumption can be relaxed so that information is passed randomly to anyone in the network. The magnitude of the impact from a change in network size decreases rather sharply but still demonstrates the pattern of first increasing competition and then improving outcomes for later entering cohorts.

<sup>11</sup>The search literature has argued that either Pareto or exponential distributions are appropriate as wage distributions (Lancaster and Chesher, 1983; Lynch, 1983). I tested the sensitivity of the model's predictions by assuming wages are exponentially distributed with  $w \sim E(\lambda = 3) + 5.15$ . Analysis of all other parameter values shows results consistent with the uniform distribution: the prediction holds for the employment for all parameter values, and wages conditional on employment continue to be ambiguous for certain values, despite the support of the wage distribution being much smaller.

Wages conditional on employment do not necessarily follow the same pattern as in section 2.1. For example, an increase in  $N_j$  leads to a smaller number of offers available to cohort  $j$  in period  $j$ , lowering wages. However, it also reduces the proportion of jobs acquired through the network. Since these jobs have lower wages on average, the average wage of those who are employed may actually increase. It is therefore possible to have within-period increases in average wages among the employed due to an increase in  $N_j$ , despite the decline in the employment rate.

There is no general prediction with regards to wages conditional on employment. However, an analogous Claim to the Propositions in Section 2.1 holds, according to simulations of a grid of the state space, for employment rates and unconditional wages.

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

where  $w_j^j$  represents the average (unconditional) hourly wage of the entire network, including a wage of 0 for those who are unemployed. The offsetting effect from changes in the composition of network members obtaining jobs through direct versus indirect channels is not strong enough to change the prediction for unconditional wages.

The model therefore predicts that employment and wage rates are inversely correlated with the number of recently arrived refugees, but positively correlated with the number of senior network members.<sup>12</sup>

<sup>12</sup>A key difference between the theoretical model presented in section 2.1 and this model is how efficient the network is in passing along job information, with the former being perfectly efficient at passing an offer on until it reaches an unemployed network member. The main effect of increasing efficiency is an increase in the average employment rate. The predictions of the model are robust to different levels of efficiency.

## 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.<sup>13</sup> 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.<sup>14</sup> 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.

---

<sup>13</sup>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.

<sup>14</sup>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 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.<sup>15</sup> 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 members until the allocation process has been completed, which is generally within one week of the State Department contacting the agency. The refugee travels directly from his place of residence overseas to the chosen IRC regional office within the U.S.

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. The belief is that refugees are better off working immediately and pursuing ESL and other training courses in the evening. Some regional offices provide additional long-term support,

---

<sup>15</sup>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.

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). These numbers are estimated using projected numbers 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 the actual number of refugees who arrive from a region can be volatile. There is also a great deal of uncertainty about the number of family reunification cases arriving each year. Since family reunification cases are predestined for particular offices, this shifts the allocation of non-family reunification cases and often the total number of refugees sent to each city away from budgeted numbers. Finally, the overall number of refugees sent to a particular office is also a function of overall employment statistics at the regional office level. Overall, the IRC

employee who is in charge of such decisions 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 still some cases that were delayed due to heightened September 11, 2001 security requirements.

As for the remaining information provided to the IRC by the PRM, the IRC reports using a limited amount of this information in the allocation process. Given that this is difficult to verify, the 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.<sup>16</sup> 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 fairly rich 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 of birth, country of first asylum, the size of the family being resettled, initial English language level, and education received in the home country. Individual characteristics which the IRC observed

---

<sup>16</sup>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.

during the allocation process are therefore also available to the econometrician.

Labor market outcomes, in particular employment status and hourly wage, were collected by the IRC at 90 days after each refugee's arrival. Finally, data on the total number of individuals (inclusive of all ages) placed in each IRC regional office by nationality from 1997 through 2004 were retrieved from archived aggregate reports. Unfortunately, individual-level data prior to 2001 are 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.<sup>17</sup> 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 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 "reset-

---

<sup>17</sup>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.

bled” by the IRC. By 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 5.1.2.

Two additional data sources are used. The 2000 census available through IPUMs is used to construct an alternative network measure as discussed in the Appendix. Supplemental information is also available from 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. family reunification cases, however, so the ORR sample is therefore not precisely comparable to the IRC sample.

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, it is not the case that it is too short of a time period for refugees to have found jobs. In fact, it is a particularly interesting time period to examine the role of network effects as it is an intense job search period for refugees.

## 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:

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

for individual  $i$  from country of origin  $j$  in city  $k$  who arrived at time  $t$ .  $Y_{ijkt}$  represents either employment status or wages for individual  $i$ .  $N_{jk(t-1)}$  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.  $N_{jk(t-3/t-4)}$  is analogously the number of refugees who arrived three and four years prior.  $N_{ijkt}$  is the number of refugees from country of origin  $j$  resettled by the IRC in regional office  $k$  who arrived in fiscal year  $t$  up to  $i$ 's specific date of arrival. Note the subscript  $i$ : refugees who arrived after  $i$  are excluded from  $N_{ijkt}$  since they would not be acting as competitors nor providers of job information to individual  $i$ .<sup>18</sup>

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. The competition effect diminishes with the length of time elapsed between the arrival of the two cohorts, with the positive information effect eventually dominating. Therefore, negative point estimates of  $\gamma_1$  and  $\gamma_2$  and positive estimates of  $\gamma_3$  and  $\gamma_4$  are consistent with the model.

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 labor market outcomes. In the case of refugee resettlement, the institutional environment provides a strategy to mitigate this problem of correlated unobservables.

---

<sup>18</sup>For example, an individual who arrived in December could not influence the 90 day labor market outcomes of a refugee who arrived in January.

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  $\epsilon_{ijkt}$  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.<sup>19</sup> In the preferred specification, I include only  $\delta_{jt}$  and  $\phi_k$  controls. Time variant heterogeneity at the nationality group level is captured by  $\delta_{jt}$ . The specification allows groups to differ, for example, if one nationality group has lower human capital on average or the types of people who become refugees vary across sending countries. Since this term varies across years 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,  $\phi_k$ , are therefore included. There are three additional potential sources of bias. Shocks at the city level ( $\xi_{kt}$ ), match quality between nationalities and cities ( $\tau_{jk}$ ), and finally shocks to match quality ( $\nu_{jkt}$ ). 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. Finally, I will argue in Section 5.1.1 that time varying shocks to a nationality-city pair cannot explain the pattern observed.

---

<sup>19</sup>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.

## 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.<sup>20</sup> 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%. To further put the magnitude of this effect in perspective, 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 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 4.4 percentage points. The numbers of refugees who arrived in the prior three and four years are combined, and the estimate is significant at the 5% level.<sup>21</sup> One-

---

<sup>20</sup>A linear probability model is used, though results are similar in a discrete choice framework, and the error term is clustered at the nationality group-city level.

<sup>21</sup>These two years are pooled for statistical power. When included individually, both point estimates are close to

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 coefficient on the number of refugees who arrived in years  $t - 3$  and  $t - 4$  is positive and significant, it is smaller than that of the  $t - 2$  network. One reason for this is that 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.<sup>22</sup> The out-migration rate within the first 90 days is 6.96%, but I do not observe in-migration. It is therefore plausible that the smaller coefficient may reflect the fact that  $N_{jk(t-3/t-4)}$  has more measurement error in representing the true number of network members currently available to the new arrival compared to  $N_{jk(t-2)}$ . If the measurement error is classical, attenuation bias pushes down the size of the coefficient on  $N_{jk(t-3/t-4)}$  compared to the coefficient on  $N_{jk(t-2)}$ . The possibility of selective migration, and therefore non-classical measurement error, is discussed in section 6.2.

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.<sup>23</sup> To ensure that 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

---

0.4 but not precisely estimated. Moreover, the point estimates are similar to those in column 1 even if other network variables are excluded; this indicates that multi-collinearity is not creating a spurious result.

<sup>22</sup>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.

<sup>23</sup>IRC Exemption from Employment is an indicator for those refugees who were not required to look for work. Exemptions were generally granted due to health problems or the presence of dependents in the household which prevented the individuals from working.

as predicted by the model? This “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 skill in placement decisions, network size would be endogenous. While there are a number of theoretical reasons this is unlikely to generate the pattern observed in table 3,<sup>24</sup> 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-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.<sup>25</sup> The estimates are larger than in the specification in columns 1 and 2, although only the coefficient on  $N_{jk(t)}$  is statistically different across specifications. Also, the effect of  $N_{jk(t-3/t-4)}$  is less precisely estimated. The effect of having a network member with three or four years’ experience in the U.S. may be highly collinear with having a permanently larger social network, resulting in an insignificant effect of  $N_{jk(t-3/t-4)}$ . The inclusion of the nationality group-city fixed effects rules out the alternative hypothesis of unobserved time invariant match quality directly. 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

---

<sup>24</sup>First, if the IRC used information on shocks to match quality, there would be a positive relationship between network size in time  $t$  and outcomes. This is in contrast to table 3. Second, this would create a uniformly upward bias, not a differential effect between recently arrived refugees and tenured refugees. Third, strategic placement would produce an oscillating pattern in the flows of refugees from a particular group, which is inconsistent with table 1.

<sup>25</sup>There are 198 nationality-city, 72 city-year and 115 nationality-year pairs.

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 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. To investigate this, I included two discrete variables indicating whether the refugee was placed in an office with either a staff member or a volunteer who speaks at least one of the languages spoken by the refugee. In results available upon request, the coefficients of interest are not significantly altered. The level effect of having a volunteer who speaks the same language positively impacts employment while there is no impact from the number of staff members speaking the refugee's language. The other major services provided by the IRC are ESL courses and employment services. While data is only available for a subset of the data, a refugee's use of these services is not reduced if he arrived in a larger cohort.

### **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  $t$  through  $t - 4$  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 effect becomes insignificant and the point estimates are negative. The static analysis would be inconclusive on 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, the presence of network-based job information transmission is easily detected and 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 affect the labor market outcomes of network members.

### 5.1.2 Robustness Analysis

#### Falsification Test

The structure of the data 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. This exercise serves to test whether there is a cyclical pattern to IRC placement of refugees which may be confounded with the social network mechanism of interest.

Columns 1 through 4 in table 4 are consistent with the maintained identification 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 drive the results. It also casts doubt on the concern that there is a dynamic relationship between IRC staffing resources and the number of refugee arrivals.

#### 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 the sensitivity of the analysis to this assumption. 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?

The inclusion of the number of family reunification refugees in addition to the number of social network members does not alter the main findings, as seen in column 5 of table 4. The number of family reunification refugees does not significantly affect sample refugees' labor market outcomes, and the estimates of the effect of the number of non-family reunification refugees in years  $t$  to  $t - 4$  on sample refugees' outcomes are not significantly impacted. Analysis using the census, which incorporates all immigrant types, also shows evidence consistent with the model. Refugees who arrived in 2002 benefited more from having a larger number of network members who arrived in 1999, according to the census, than refugees who arrived in 2001. The census results are presented and discussed in section A.3 of the Appendix.

## 5.2 Job Quality: Wages

Turning to the role of networks in determining hourly wages, columns 1 and 2 of table 5 show the effect of network size on wages for the employed sample. It is an empirical question how changes in network size affect wages since the model highlights 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. Columns 1 and 2 of table 5 show that the size of the network in periods  $t - 2$  and  $t - 3/t - 4$  are positive and

statistically significant. There is no evidence that junior network members, those who arrived in years  $t$  and  $t - 1$ , impact average hourly wages conditional on employment. This result suggests, in the context of the information transmission model, that the effect of the network in changing the number of wage offers an individual receives is stronger than the compositional effect.

Claim 1 argues 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. In order to test Claim 1, I estimate equation 4 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. As shown in column 3 of table 5, 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. The estimates are consistent with the model's prediction, but may merely reflect the effect of networks on employment since the 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 5 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 5. LAD alleviates the concern of censored wage offers under the assumption that unemployed individuals receive wage offers below the median offer made to employed workers with comparable skills.<sup>26</sup>

---

<sup>26</sup>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 Alternative Explanations

This section considers two alternative explanations for the empirical findings: general labor market competition (6.1) and measurement error due to subsequent migration within the U.S. (6.2).

### 6.1 Labor Market Competition

The theoretical framework and empirical evidence sections both argue that an increase in the number of network members who arrived recently in the U.S. exacerbates competition *within* the social network. An alternative interpretation is that the competition effect has nothing to do with social network dynamics but merely reflects an increase in labor supply, driving down wages as in the work by Cortes (2008). The number of new arrivals in a given network is small each year, around 30 people on average. It is therefore unlikely that such a small addition to the total labor market in cities such as Atlanta and Salt Lake City influences the unemployment rate or the equilibrium wage level. However, if labor markets are very segmented, the increase in the number of refugees in a given group may be able to affect labor market conditions in a given industry or occupation.

A more formal way to separate these two different “competition” effects is a test using September 11, 2001 (9/11). Anecdotal evidence from the IRC suggests that the terrorist attacks on 9/11 had a large, negative impact on refugees’ labor market outcomes. The potential channels for the negative 9/11 shock include: an increase in xenophobia that decreased employment opportunities; or a negative shock to the tourism industry, where many refugees were employed.<sup>27</sup> The effect of this shock differs in the two models. The network model predicts, as will be explained below, a diminished competition effect after 9/11. Conversely, the classic supply and demand model predicts exacerbated competition after a negative shock in labor demand. This natural experiment also tests more generally how refugee social networks respond to an exogenous negative shock in

---

<sup>27</sup>A table with the distribution of industries from 2001-2003 is available from the author upon request.

the context of the job information transmission model.

The 9/11 shock can most simply be analyzed in the model in section 2 as a shock to the arrival rate,  $a$ . A decline in  $a$  dampens the effect of a change in the size of a cohort: both the initial competition effect and the positive information effects are muted. Figure 2 provides a graphical example of model predictions as  $a$  varies. The first panel shows how an increase in cohort  $c$  affects employment in that period, i.e.  $\frac{\partial s_c^c}{\partial N_c}$ . In particular, it maps out how  $a$  affects the difference between the employment rate of cohort “c” when all cohorts are constant size versus when cohort “c” is doubled (the treatment effect). As  $a$  increases, up to  $a = 0.4$ , the treatment effect becomes stronger, i.e. more negative. This means that for a given change in network size, an increase in  $a$  will exacerbate the competition effect. The second panel depicts the model’s prediction for the corresponding scenario where an increase in  $N_c$  raises the employment rate of a younger cohort. In the example presented here, the impact on the employment rate of cohort  $c + 2$  is positive for all values of  $a$ ,  $b$  and  $N_c$ , and an increase in  $a$  strengthens this effect.<sup>28</sup>

$$Y_{ijkt} = \alpha + \sum_{p=1}^4 (\gamma_p N_{ijk(t-p+1)} + \pi_p N_{ijk(t-p+1)} * Post9/11) + X_{ijkt}\beta + \delta_{jt} + \phi_k + \epsilon_{ijkt} \quad (5)$$

is the econometric specification used for the test. The hypothesis is that  $\pi_p$  should be of opposite in sign to  $\gamma_p$  for all  $p$ , reflecting the dampening of network effects after a decline in  $a$ .

Table 6 shows the test for a differential network effect after 9/11 from an increase in the number of network members who arrived in the same period,  $N_{jkt}$ , and two or more years prior,  $N_{jk(t-2/t-3/t-4)}$ .<sup>29</sup> The effect of senior network members is at most suggestive due to large standard

---

<sup>28</sup>The relationship between  $a$  and  $\frac{\partial s_c^c}{\partial N_c}$  is nonlinear; the treatment effect quickly converges to zero when the network is close to full employment. The intuition is that for low levels of  $a$ , there are few referrals available in the network; as  $a$  rises, more information is available to compete over. For high levels of  $a$  such that employment is near 100%, there is little competition since all are likely to become employed directly. Since the average employment rate in the sample is 66%, it is unlikely that the portion of the model in which an increase in  $a$  lowers competition is relevant empirically.

<sup>29</sup>The effect of an increase in  $N_{jk(t-1)}$  is theoretically ambiguous since, in the model, a change in  $a$  can lead to a change in where the treatment effect flips from negative to positive. Therefore, an interaction term for 9/11 and  $N_{jk(t-1)}$  is not included.

errors but the coefficient is the predicted sign. However, the coefficients of interest - reflecting the competition effect - are consistent with the social network model. While the negative competition effect is still present, there is a clear reduction in the magnitude of the competition effect after 9/11. This is in opposition to the prediction of the classic supply and demand model.

## 6.2 Measurement Error and Secondary Migration

A limitation of the IRC data is that each refugee is observed at only one point in time. The empirical analysis relies on the assumption that the number of non-family reunification refugees from group  $j$  in city  $k$  in year  $t - 4$  are those who were placed in that city four years prior. As discussed in section 5, the imprecision and smaller than expected point estimate of  $N_{jk(t-3/t-4)}$  may be due to classical measurement error induced by refugees re-locating within the U.S. (secondary migration). In this section, I assess whether non-random out-migration can generate the pattern observed in section 5.

Consider first a theoretical framework where there are no social networks; refugees can choose to leave their initial resettlement city at any time, and the number of jobs available to refugees from a particular country is fixed. The probability of becoming employed is then simply the number of jobs available divided by the total number of refugees in the city in that year. In this simple model, a nonlinear pattern does emerge if a fraction of randomly selected refugees out-migrate each year. Using the notation from the empirical analysis, the effect of an increase in  $N_{jkt}$  is negative due to labor market competition, irrespective of the network, and the negative effect is mitigated over time as individuals leave the city. That is to say, the effect of an increase in  $N_{jk(t-2)}$  will be less negative than an increase in  $N_{jk(t-1)}$  or  $N_{jkt}$ . However, the out-migration model does not generate a positive relationship between employment and  $N_{jk(t-2)}$ . I also consider models where either employed or unemployed individuals are more likely to out-migrate. Simulation of

such models shows that the same pattern emerges as when randomly selected network members move out of the labor market. Recall that  $S$  is the number of periods an individual is in the market. During at least the  $S$  periods following the entry of the exogenously larger cohort, the average employment rate for the entering cohort is less than or equal to the rate when cohort size is constant. Therefore, it is unlikely that systematic out-migration alone can explain the empirical results in the paper.

Due to data limitations, the only direct evidence relating the decision to out-migrate to network size is restricted to decisions made during the first 90 days after arrival. While this is imperfect, such analysis, not shown, reveals no significant relationship between  $N_{jk(t)}$  through  $N_{jk(t-3/t-4)}$  and the binary variable indicating out-migration.

## 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. The paper highlights two opposing ways in which changes in network architecture affect outcomes: negative within-network competition and positive information transmission. These results provide insight into the functioning of social networks and provide empirical evidence that within-network competition over job information has an economically sizable negative impact on labor market outcomes. This result tempers previous findings in the empirical literature on social networks which show that networks play a beneficial role in overcoming 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 sheds some light on the contradictory conclusions found by Edin et al. (2003), who find a positive relationship between the stock of immigrants and labor market outcomes, and Borjas (2000), who finds a negative relationship between the proportion of individuals within a city from the same country and assimilation for refugees in the U.S. This paper suggests that the tenure composition of social networks and other dimensions of structure are important in assessing the full network effect.<sup>30</sup>

Funding for job training and ESL courses for refugees have declined significantly since 2001. Evidence that social networks provide labor market information suggests that there are spill-overs from policy interventions which provide job training to refugees. The returns to programs providing employment and training services to refugees would be underestimated if measured by program participants alone due to the positive externality created by job information sharing within social networks. The model's prediction applies to any social network providing job information to its members, not just political refugees. Though there are external validity limitations of the empirical evidence presented in this paper, we would expect there to be spill-overs in policy interventions which improve labor market performance more generally, particularly for other immigrant groups.

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

---

<sup>30</sup>Moreover, the heterogeneity found by Edin et al. (2003) across ethnic groups will also exist within an ethnic group and will vary over time.

widely from the dispersal policies in some European countries to the clustering method used by at least some American resettlement agencies (Edin et al., 2004). The empirical evidence in this paper showing that refugee social networks provide labor market information to its members suggests a potential drawback to immigrant dispersal policies. Sending refugees to areas with a community of tenured network members who have achieved relatively high employment rates could improve short-run labor market outcomes and ease the fiscal burden refugees put on local municipalities. Of course, sending a refugee to a place with many recently arrived refugees would likely further exacerbate the burden on towns and cities. However, this analysis looks only at short-run outcomes and therefore cannot provide an estimate of the total costs of dispersal policies. In the long run, there are additional considerations such as the way the network affects individuals' incentives to invest in learning the host country language or other types of human capital. 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

- Bayer, P., S. Ross, and G. Topa (2005). Place of work and place of residence: Informal hiring networks and labor market outcomes. *NBER Working Paper 11019*.
- 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.
- 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.
- Freidlin, M. and A. D. Wentzell (1984). *Random Perturbations of Dynamical Systems*. New York: Springer-Verlag.
- 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. (2005). The doughboys network: Social interactions and labor market outcomes of World War I veterans. *Unpublished*.
- 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.
- 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.
- UNHCR (2005 edition). Refugees by numbers.
- 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
<b>IRC Data:</b>			
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 $t - 3$ and $t - 4$	65.80	102.63	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 $t - 3$ & $t - 4$	48.89	149.99	1720
<b>2000 Census Data:</b>			
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 $t$ <sup>f</sup>	-0.222 ** (0.110)	-0.228 * (0.118)	-0.328 ** (0.147)	-0.320 * (0.163)		
# Refugees Resettled in Year $t - 1$ <sup>f</sup>	-0.145 ** (0.068)	-0.119 * (0.069)	-0.283 *** (0.100)	-0.253 ** (0.102)		
# Refugees Resettled in Year $t - 2$ <sup>f</sup>	0.100 ** (0.042)	0.094 ** (0.041)	0.112 ** (0.056)	0.100 * (0.056)		
# Refugees Resettled in Years $t - 3$ and $t - 4$ <sup>f</sup>	0.040 ** (0.016)	0.037 ** (0.016)	0.073 * (0.043)	0.058 (0.042)		
# Refugees Resettled Years $t$ to $t - 4$ <sup>f</sup>					0.029 *** (0.011)	-0.025 (0.049)
Age	0.022 *** (0.006)	0.021 *** (0.007)	0.024 *** (0.006)	0.024 *** (0.008)	0.021 *** (0.006)	0.023 *** (0.006)
Age Sq	-0.0003 *** (0.0001)	-0.0003 *** (0.0001)	-0.0004 *** (0.0001)	-0.0004 *** (0.0001)	-0.0003 *** (0.0001)	-0.0003 *** (0.0001)
HH Size	-0.017 *** (0.006)	-0.015 ** (0.007)	-0.023 *** (0.007)	-0.020 *** (0.007)	-0.017 *** (0.006)	-0.022 *** (0.007)
IRC Exemption from Employment	-0.497 *** (0.057)	-0.506 *** (0.058)	-0.523 *** (0.063)	-0.524 *** (0.065)	-0.494 *** (0.056)	-0.515 *** (0.063)
MG Enrollment	0.102 *** (0.027)	0.104 *** (0.028)	0.118 *** (0.031)	0.117 *** (0.032)	0.099 *** (0.026)	0.111 *** (0.030)
P-value of education variables		0.412		0.264		
P-value of initial English level variables		0.005		0.001		
P-value of religion variable		0.188		0.384		
No obs	1720	1720	1720	1720	1720	1720
Adjusted R squared	0.223	0.226	0.272	0.278	0.215	0.262

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.

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

d Columns 3, 4 and 6 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 Rows 1-4 are multiplied by 100.

Table 4: Employment: Falsification and Family Refugees

	1	2	3	4	5
# Refugees Resettled in Year $t + 1$ <sup>a</sup>		0.086 (0.099)		0.008 (0.179)	
# Refugees Resettled in Year $t$ <sup>a</sup>	-0.276 ** (0.131)	-0.300 ** (0.137)	-0.529 *** (0.167)	-0.529 *** (0.167)	-0.229 ** (0.111)
# Refugees Resettled in Year $t - 1$ <sup>a</sup>	-0.160 ** (0.074)	-0.156 ** (0.075)	-0.280 ** (0.128)	-0.279 ** (0.125)	-0.163 ** (0.073)
# Refugees Resettled Year $t - 2$ <sup>a</sup>	0.068 * (0.040)	0.069 * (0.039)	0.150 ** (0.062)	0.150 ** (0.062)	0.090 * (0.054)
# Refugees Resettled Years $t - 3$ and $t - 4$ <sup>a</sup>	0.039 ** (0.015)	0.041 *** (0.015)	0.097 (0.059)	0.097 (0.062)	0.051 ** (0.023)
Age	0.026 *** (0.007)	0.026 *** (0.007)	0.027 *** (0.008)	0.027 *** (0.008)	0.022 *** (0.006)
Age Sq	-0.0004 *** (0.0001)	-0.0004 *** (0.0001)	-0.0004 *** (0.0001)	-0.0004 *** (0.0001)	-0.0003 *** (0.0001)
HH Size	-0.019 ** (0.008)	-0.019 ** (0.007)	-0.026 *** (0.008)	-0.026 *** (0.008)	-0.017 *** (0.006)
IRC Exemption from Employment	-0.459 *** (0.076)	-0.458 *** (0.076)	-0.482 *** (0.086)	-0.482 *** (0.086)	-0.496 *** (0.057)
MG Enrollment	0.095 *** (0.031)	0.095 *** (0.032)	0.104 *** (0.038)	0.104 *** (0.038)	0.101 *** (0.027)
# Family Reunification Refugees Resettled in Year $t$ <sup>a</sup>					0.084 (0.100)
# Family Reunification Refugees Resettled in Year $t - 1$ <sup>a</sup>					0.001 (0.064)
# Family Reunification Refugees Resettled Year $t - 2$ <sup>a</sup>					-0.020 (0.074)
# Family Reunification Refugees Resettled Years $t - 3$ and $t - 4$ <sup>a</sup>					-0.012 (0.024)
No obs	1340	1340	1340	1340	1720

a SE are in parentheses & clustered by nationality-city. Rows 1-5 and 11-14 are multiplied by 100.

b 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.

Table 5: Wages on Network Size: Conditional, Full Sample and LAD

	1	2	3	4	5
# Refugees Resettled in Year $t$ <sup>b</sup>	0.160 (0.345)	0.205 (0.335)	-0.023 ** (0.011)	-0.022 * (0.012)	-0.951 ** (0.392)
# Refugees Resettled in Year $t - 1$ <sup>b</sup>	0.020 (0.240)	0.059 (0.255)	-0.022 *** (0.008)	-0.019 ** (0.008)	-0.610 *** (0.185)
# Refugees Resettled Year $t - 2$ <sup>b</sup>	0.616 (0.225) ***	0.565 (0.222) ***	0.011 ** (0.005)	0.010 ** (0.005)	0.715 *** (0.158)
# Refugees Resettled Years $t - 3$ and $t - 4$ <sup>b</sup>	0.450 (0.128) ***	0.413 (0.124) ***	0.007 ** (0.003)	0.006 * (0.003)	0.155 ** (0.072)
Age	0.082 (0.021)	0.068 (0.022)	0.229 *** (0.040)	0.220 *** (0.052)	0.178 *** (0.025)
Age Sq	-0.0011 *** (0.0003)	-0.0010 *** (0.0003)	-0.0034 *** (0.0005)	-0.0032 *** (0.0007)	-0.0027 *** (0.0003)
Case Size	0.023 (0.023)	0.022 (0.023)	-0.152 *** (0.051)	-0.127 ** (0.055)	-0.070 *** (0.024)
IRC Exemption from Employment	-0.423 (0.437)	-0.382 (0.279)	-3.985 *** (0.409)	-4.001 *** (0.430)	-5.327 *** (0.251)
MG Enrollment	0.025 (0.096)	0.043 (0.100)	0.820 *** (0.234)	0.804 *** (0.246)	0.469 *** (0.111)
P-value of education variables		0.045		0.167	
P-value of initial English level variables		0.077		0.000	
P-value of religion variables		0.636		0.281	
No obs	1125	1125	1706	1706	1706
Adjusted R squared	0.317	0.334	0.29	0.3	

a SE are in parentheses and clustered by nationality-city.

b Rows 1-4 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 6: Employment: Shock of 9/11

	1	2
# Network Members Resettled in Year $t$ <sup>b</sup>	-0.706 *** (0.216)	-0.698 *** (0.219)
# Network Members Resettled in Year $t$ * Post 911 <sup>b</sup>	0.531 ** (0.267)	0.527 * (0.268)
# Network Members Resettled in Year $t - 1$ <sup>b</sup>	-0.312 *** (0.105)	-0.280 *** (0.108)
# Network Members Resettled in Years $t - 2, t - 3$ and $t - 4$ <sup>b</sup>	0.085 ** (0.042)	0.070 * (0.041)
# Network Members Resettled in Years $t - 2, t - 3$ and $t - 4$ * Post 911 <sup>b</sup>	-0.030 (0.026)	-0.030 (0.026)
Post 911	0.014 (0.169)	0.052 (0.170)
No obs	1720	1720

a SE are in parentheses and clustered by nationality-city.

b Coefficients are multiplied by 100.

c All columns include fixed effects for nationality-year, city-year and nationality-city.

d Column 2 includes additional individual covariates including: education, initial English level, religion.

## 9 Figures

Figure 1: Graphical Example of Model with Constant Wages

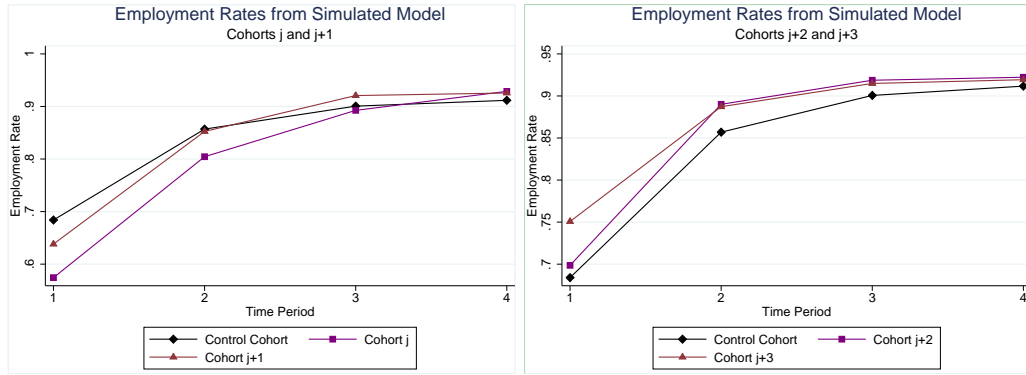
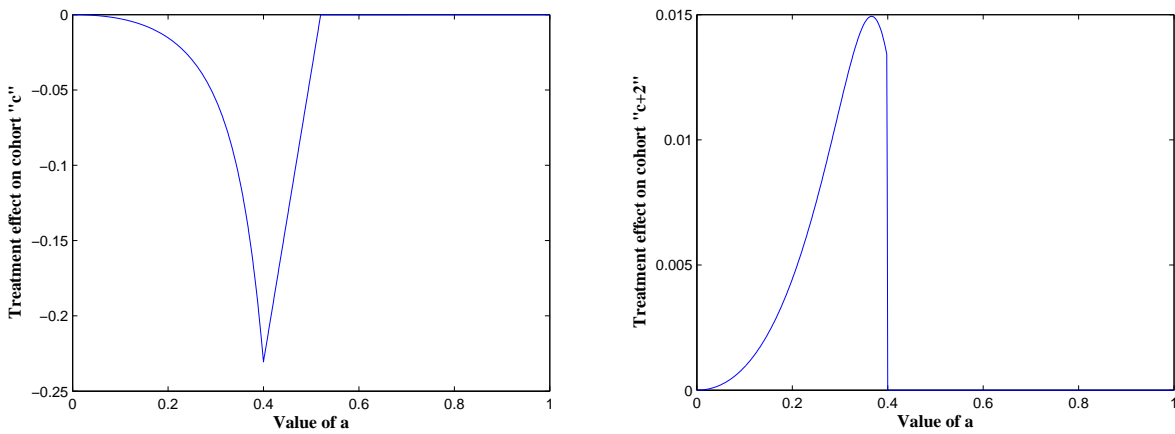


Figure 2: Graphical Example of Model Varying  $a$  when  $S = 4$  and Cohort  $c$  doubled



# 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  $c$ .<sup>31</sup>

*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:

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

Differentiating with respect to  $N_j$  gives:

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

**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  $c$ ,

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

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

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

<sup>31</sup>This claim holds for all values of  $a$  and  $b$  such that  $s_c^j \neq 1$  for all  $c$  and  $j$ .

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

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

$$\frac{aN_j(2 + N_j)}{2 + N_j - (1 - b)(s_{j-1}^{j-1} + s_{j-2}^{j-1})} + (1 - b)s_{j-1}^{j-1} + \frac{a(2 + N_j)(1 - (1 - b)s_{j-1}^{j-1})}{2 + N_j - (1 - b)(s_{j-1}^{j-1} + s_{j-2}^{j-1})} > s_{j-1}^{j-1} + s_{j-2}^{j-1} \quad (6)$$

Using the steady-state properties of the economy,<sup>32</sup> equation (6) will hold with equality if  $N_j = 1$ .

Since (6) holds with equality when  $N_j = 1$ , inequality holds if expression is increasing in  $N_j$ .

$$\frac{a(2 + N_j)(1 + N_j - (1 - b)s_{j-1}^{j-1})}{2 + N_j - (1 - b)(s_{j-1}^{j-1} + s_{j-2}^{j-1})} + (1 - b)s_{j-1}^{j-1} \quad (7)$$

Differentiating equation (7) with respect to  $N_j$  gives:

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

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

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

Since denominator is greater than zero, if numerator is greater than zero, then equation (6) holds.

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

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

Need to show:

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

Let  $N_j = 1 + x$ . Rearranging equation (7) gives:

$$(s_{j+1}^{j+1} - s_j^j) + (s_j^{j+1} - s_{j-1}^j) + x(s_j^{j+1} - s_j^j) > 0 \quad (8)$$

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

---

<sup>32</sup>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.

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

The left hand side (LHS) of equation (8) is then:

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

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

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

Substituting the above gives:

$$\begin{aligned}
LHS &> (s_{j+1}^{j+1} - s_j^j)(2 - (1-b)s_j^j) + x(s_j^{j+1} - s_j^j) + (1-b)(1 - s_j^j)[(1 - (1-b)s_{j-1}^{j-1})(s_{j-1}^{j-1} - s_j^j) - x s_j^j] \\
&= (s_{j+1}^{j+1} - s_j^j)(2 - (1-b)s_j^j) + (1-b)(1 - s_j^j)(1 - (1-b)s_{j-1}^{j-1})(s_{j-1}^{j-1} - s_j^j) + x[(1 - (1-b)s_j^j)(s_{j+1}^{j+1} - s_j^j)] > 0
\end{aligned}$$

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

### A.3 Alternative Network Measure: Census Data

The main analysis assumes the social network is comprised only of non-family reunification refugees from the same country of origin in the same city. However, to relax that assumption, I also use 2000 census data available through IPUMS to 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.<sup>33</sup>

Since the data structure differs from the network measure used above, the empirical specification varies as well. In order to test the hypothesis using the 2000 census data, the size of the network is restricted to those who arrived most recently in the U.S., specifically those who arrived in 1999. I then look for a differential effect of this network for refugees who arrived in 2001 and 2002.

$$Y_{ijkt} = \alpha + \phi_1 N_{jk(t=1999)} + \phi_2 N_{jk(t=1999)} * \lambda_{2001} + X_{ijkt}\beta + \delta_j + \phi_k + \lambda_{2001} + \epsilon_{ijkt} \quad (9)$$

$Y_{ijkt}$ ,  $X_{ijkt}$ ,  $\delta_j$ ,  $\phi_k$ , and  $\epsilon_{ijkt}$  are defined in section 4.  $N_{jk(t=1999)}$  is the size of the network for those immigrants who arrived in 1999 according to the census, and  $\lambda_{2001}$  is an indicator for those refugees who arrived in 2001. Estimates showing  $\phi_1$  to be positive and  $\phi_2$  to be negative would be consistent with the model. The differential network effect

<sup>33</sup>I calculate the size of the network at level of the metropolitan statistical area (MSA); most are defined by a single MSA but some networks are defined using multiple MSAs, such as New York. I define a social network by either nationality or ethnicity, depending on the availability of the relevant code in IPUMS. IPUMS data also provides the age of each network member as well as the year of arrival in the U.S. Therefore I can create a network size variable which is specific to the year of arrival of the network members. The information on age also allows me to restrict the network to only prime age adults. Since the census does not obtain information on the foreign born's visa type or residency status/citizenship, this measure will include all immigrant types, ranging from illegal immigrants to permanent residents and naturalized citizens.

across the two cohorts is therefore captured by  $\phi_2$ : an increase in the number of network members who arrived in 1999 would have a smaller or negative impact on labor market outcomes for those who arrived in 2001 than for those who arrived in 2002. Network members who arrived in 1999 would be more likely to 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.

### **Probability of Employment**

Using the 2000 census to create a second network measure allows for flexibility in the definition of the social network. This measure expands the potential network members to those who come from the same country of origin or ethnic group but who may have different immigration status. In this case, individual network members have self-selected their preferred location based on a number of unobserved factors. While this measure of the network is more susceptible to selection bias due to comparative advantage, it does test the generality of the job information sharing effect across two independently constructed network measures. Column 1 in table A.1 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 A.1, 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 9 - provides qualitatively similar results. Column 2 of table A.1 indicates that the OLS estimates show no significant effect of network size on wages of those employed although the signs of the point estimates are as expected from the model. The coefficients in 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.

Table A.1: Employment and Wage Effects Using Census Data for Network Measure

	Employment		Wages		
	1	2	3	4	5
Network size which arrived in 1999 <sup>g</sup>	0.0284 ** (0.0121)	0.0220 * (0.0126)	0.013 (0.055)	0.249 *** (0.088)	0.172 * (0.091)
Network size which arrived in 1999 * Refugee arrived in 2001 <sup>g</sup>	-0.0275 ** (0.0114)	-0.0253 ** (0.0121)	-0.043 (0.058)	-0.238 *** (0.095)	-0.201 ** (0.097)
Age	3.388 *** (0.912)	3.020 *** (1.002)	0.073 ** (0.034)	0.270 *** (0.068)	0.229 *** (0.072)
Age Sq	-0.0530 *** (0.0122)	-0.0482 *** (0.0129)	-0.0009 ** (0.0005)	-0.0042 *** (0.0009)	-0.0037 *** (0.0009)
HH Size	-1.789 * (1.093)	-1.474 (1.046)	-0.017 (0.028)	-0.162 ** (0.078)	-0.122 (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.