Refugee Resettlement: The Role of Social Networks and Job Information Flows in the Labor Market [1]

Lori Beaman, Yale University [2]

August 18, 2006

PRELIMINARY DRAFT.

COMMENTS ARE WELCOME.

Abstract

As of 2005, there were over 9.2 million refugees worldwide and hundreds of thousands asylum seekers. Resettlement of refugees to North America and Europe is the primary strategy used when repatriation and local integration into the country of first asylum are not possible or undesirable. An important aspect of the resettlement process is how to distribute refugees within the new host country, and essential to that question is the role of social networks in facilitating job information to new arrivals. This paper provides empirical evidence of information flows regarding job opportunities within social networks among refugees resettled in the U.S. An adapted version of the model developed by Calvo-Armengol and Jackson (2004) provides the intuition that competition can exist between network members for job referrals and investigates the dynamics between the size of a social network, the tenure of network members and labor market outcomes. This model is tested empirically using a sample of refugees resettled in the U.S. from 2001-2005. The size of the social network is measured by the number of individuals from the same place of birth who reside in the same metropolitan area. The econometric specification controls for individual characteristics at the time of arrival as well as metropolitan area and nationality group fixed effects. The empirical analysis is consistent with the predictions of the model: a larger number of network members who have arrived in the U.S. one year ago lowers the probability of employment and the wage while more tenured network members improve labor market outcomes for recently arrived refugees. This paper therefore provides empirical evidence of the theoretical work by Calvo-Armengol and Jackson and isolates both costs and benefits of participation in social networks.

[1] I am indebted to Bob Carey, Christine Petrie, and especially Chris Bujara at the International Rescue Committee for providing access to the data for this project and teaching me about the refugee resettlement process. I also thank Joe Altonji, Pat Bayer, Fabian Lange, Rohini Pande, Mark Rosenzweig, Chris Udry and seminar participants at Yale's Labor and Development Lunches for feedback and encouragement.

[2] The author can be contacted at lori.beaman@yale.edu.

1 Introduction

Due to prolonged and protracted conflicts throughout the world, there are a substantial number of refugees who are either unable to return to their home country or are unable to stay safely in their first country of asylum. Often security concerns in refugee camps go hand in hand with dire living conditions. The primary solution to improving the lives of such refugees is resettlement in a third country. During 2004, for example, 676,400 people applied for asylum and in addition over 83,000 refugees were permanently resettled to third countries through UNHCR resettlement programs, mostly to European and North American countries.[18] However, there is no consensus on the optimal method of resettlement within the new destination country. Policies vary widely from the dispersal policies in some European countries to the clustering method used by at least some American resettlement agencies. While there are many factors which determine whether a location is a good match for a refugee, the ability of the refugee to integrate into the local labor market is essential. At the core of the debate between dispersing versus concentrating refugees geographically is the role of ethnic networks in facilitating access to the local labor market. This paper seeks to empirically estimate the effect of ethnic networks in providing job information to new refugees resettled in the U.S.

While there are many nations which accept refugees on a temporary basis immediately after the rupture of a crisis, less than 20 nations run UNHCR resettlement programs and accept refugees regularly on an annual basis. The countries who host the bulk of resettled refugees are: the United States, Canada, Australia, Sweden, Norway, Finland, New Zealand, Denmark, and The Netherlands.[18] Of these, the Netherlands, Norway, Denmark, and Sweden have all implemented and experimented with dispersal policies. In Denmark, for example, refugees are distributed throughout the country to municipalities in inverse proportion to the existing percentage of ethnic minorities. Given that local authorities are legally responsible for providing housing and other social services to refugees, the dispersal policy was implemented out of growing concern that the financial burden of refugee resettlement was falling disproportionately on the capital and the larger cities.[8] The other explicit objective of the program is to further integration. The ideas of "spreading the burden" and decreasing segregation are the underlying motivation behind the implementation of dispersal policies in other European countries as well, such as Sweden and the

Netherlands.[9] While the U.S. does not have a centralized resettlement program, at least some resettlement agencies in the U.S. follow a policy of clustering refugees in geographic locations which have pre-existing ethnic communities. Despite significant federal funding given to the resettlement agencies, individual states in the U.S. bare some of the financial burden of resettlement in the form of TANF, food stamps, and other social services. The concentration policy in the U.S. stands in stark contrast to the system used widely in Europe.

An important question in this debate is how social networks create economic incentives and impact the economic decision making of recently arrived refugees. There are numerous hypotheses on this topic in the literature. First, the existence of ethnic enclaves may diminish incentives for refugees to invest in host country-specific human capital. A number of empirical papers seek to provide evidence on the impact of enclaves or residential segregation on economic outcomes, pointing to this mechanism as theoretical motivation. Borjas (2000) provides some evidence that residential segregation has a negative impact on the assimilation of refugees.[3] By contrast, Edin et al. (2003) find that ethnic enclaves lead to earnings gains for less skilled refugees who were placed according to a settlement policy in Sweden.[10] However, as the differing empirical findings suggest, neither paper is able to identify the specific way that enclaves influence behavior. In particular both papers are unable to isolate the negative incentive effect but instead capture a net effect encompassing many different factors.

Social networks may also provide information on alternatives to employment which may in fact inhibit economic assimilation. Bertrand et al. (2000) find that larger networks and networks whose members use welfare more intensively encourage welfare use among individuals in the U.S. whose native language is not English.

Finally, networks can facilitate access to the labor market by providing job information or employee referrals. There is an extensive literature which documents the importance of informal job referrals in the U.S. labor market. Numerous studies report that at least 50% of jobs are obtained through family and friends, for example. Montgomery's seminal theoretical work emphasizes the role of social networks in helping to overcome the problem of imperfect information about a unemployed individuals' ability.[14] If members of a social network have better information about other members' ability, then firms will use informal employee referrals to make hiring decisions. Munshi (2002) develops a model similar to that of Montgomery and then provides empirical evidence that social networks enchance employment outcomes among Mexican migrants in the U.S.[15] Using exogenous variation from rainfall shocks in Mexico to predict network size, he finds a larger network increases the probability of employment and the probability of being employed in a higher paying occupation for network members. In addition to employee referrals, the services provided by the social network, as reported in the Mexican Migration Project data, also include financial assistance and housing. While the empirical results are therefore consistent with the model of job referrals presented by Munshi, the additional benefits the network provides to newly arrived migrants are likely to be contributing to the estimated network effect. Since refugees are provided with housing and financial assistance by a formal resettlement agency, identifying the role of networks in providing labor market access directly is important. Munshi also finds that it is senior network members who are influencing the labor market outcomes of others, and the effect of recently arrived network members could not be distinguished from zero. Given that these estimates are noisy, the author is unable to draw strong conclusions from this particular finding. It is suggestive, however, that size is not the only relevant measure to capture how a network influences the economic outcomes of its members, and that social networks may have a more complex way of affect the labor market depending on its structure.

Calvo-Armengol and Jackson (2004) show theoretically that the structure of a social network can influence the dynamics of employment.[4] In a model where job information is distributed randomly and can be shared with network members, the authors show that employment outcomes are positively correlated across all individuals in a network in the steady-state. There is, however, negative correlations between individual network members in certain states. This implies that there are in fact costs to having a larger social network if that network has a particular structure. The authors then allow labor market participation to be endogenous and finds that differing initial conditions across networks can lead to long-run inequalities across groups.

This paper seeks to isolate one specific mechanism through which social networks affect labor market outcomes for refugees by providing empirical evidence consistent with a model of job information transmission within a social network. The model, based on the work done by Calvo-Armengol and Jackson (2004), provides a framework in which individuals have a random probability of receiving job information. This information is either used to obtain a job or passed on to an unemployed member of the individual's social network. The model predicts that having a larger network can in some cases lead to a deterioration in labor market outcomes. More specifically, competition can exist between network members for job information, thus creating a non-linear relationship between the size of a social network and labor market outcomes depending on the tenure of network members. In fact a larger social network may differentially influence labor market outcomes over time: first lowering the employment rate and eventually improving it. Using data from the internal records of the International Rescue Committee (IRC) on refugees resettled in the U.S. from 2001-2005, I test the hypotheses from this model to look at the dynamic relationship between social network size and labor market outcomes.

The institutional environment of refugee resettlement provides arguably exogenous variation in the size and structure of social networks in order to identify the role of networks in job information transmission. The resettlement of refugees in the U.S. is implemented by voluntary resettlement agencies who have been contracted by the State Department to provide all initial services such as housing, financial assistance and job training/job referrals. The sample of refugees analyzed in this study are those who do not have family members in the U.S. at the time of their arrival to assist in their resettlement. In this case, it is the sole responsibility of the contracted resettlement agency to choose a geographic location for these individuals. This precludes individuals from sorting into localities based on unobservable individual characteristics, a common source of bias when attempting to identify social network effects. Furthermore, all individual characteristics used by the IRC when placing refugees into particular cities are available in the data. Since the IRC resettles refugees from a wide range of origin countries to 16 regional offices, the econometric specification also controls for unobserved metropolitan area and nationality/ethnic group characteristics through fixed effects.

The empirical analysis provides evidence supporting the model. A larger number of network members who arrived in the year prior to newly arrived refugees lowers the likelihood of employment and wages, while a larger network of individuals who arrived 2 years prior or more improves the labor market outcomes of fresh refugees. This provides clear evidence that social networks impact labor market outcomes of newly arrived refugees by providing job information. It also shows that network effects are more complex than the previous empirical literature has shown, involving both costs and benefits in a dynamic relationship. Finally, this paper provides empirical support of the model developed by Calvo-Armengol and Jackson and can therefore (weakly) lend support for the other non-testable implications of the model regarding the role of social networks in sustaining inequality.

The paper is organized as follows: Section 2 discusses the framework of the theoretical model of information flows within social networks, Section 3 provides details on the institutional background and data and section 4 covers the empirical strategy. The results of the empirical analysis will be presented in section 5 and finally section 6 concludes.

2 Theoretical Framework

2.1 A Model of Employment Rates

This paper seeks to provide empirical evidence of a model of job information transmission within social networks. The theoretical framework presented here is an adaptation of the model developed by Calvo-Armengol and Jackson (2004). In their model, information about jobs arrive randomly to agents, and this information leads directly to employment in that job. Thus, if an individual who is unemployed hears about a job, he will take the position. However, if the agent who receives the job information is already employed, then he passes along the information to a direct connection within his social network. At the end of each period, there is an exogenous break-up of jobs. The objective of their model is to show that in the steady-state, there is positive correlation of employment outcomes across time and across all agents within a network. They authors also incorporate the possibility of agents dropping out of the labor market, which due to a contagion effect, can lead to persistent levels of inequality across different groups.

I provide empirical evidence of an adapted version of this model. To do this, I first make the simplifying assumption that all individuals within a network are connected, thereby eliminating the distinction made by Calvo and Jackson between direct and indirect connections. Furthermore I incorporate the model into an overlapping generations framework which corresponds well with the empirical setting of refugee resettlement.

The basic structure and timing of the model is as follows: each agent works for S periods, so that in the steady-state there are S cohorts in the network at any point in time. Each cohort c has N_c agents. It is this variation in cohort size which will provide estimable predictions from the model. If agent i in cohort c is employed at the end of period t, then $s_{ic}^t = 1$ and accordingly $s_{ic}^t = 0$ if agent i is unemployed. Since all agents within a cohort are identical, it is preferable to work with the employment rate within the cohort, s_c^t. Period t begins with some agents being employed and others not, so s_c^{t-1} describes employment rate of cohort c from the previous period. Information about job openings then arrive: any agent hears about a job opening with probability a between 0 and 1. It is important to note that the job arrival process is assumed to be independent across agents. If an agent is unemployed, he will fill the position. However, if the agent is already employed then the information will be passed along to a randomly selected network member who is unemployed. Job information can be shared with any unemployed member in the network, irrelevant of which cohort he belongs to. Accordingly, an 'older' network member can receive job information from a 'younger' member if the former is unemployed. Once job information arrives and is referred to unemployed members where suitable, jobs are immediately accepted. Finally, there is a positive probability for any employed agent to lose his job at the very beginning of the next period at the exogenous breakup rate b which varies between 0 and 1.

This structure can can be formalized in the following way:

For $t \geq S$:

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

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

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

where r^t represents the probability of receiving job information through an employed network member. For simplification of notation, the above expressions are equivalent to:

$$s_c^t = \frac{a\sum_{k=t-S}^{t} N_k}{\sum_{k=t-S}^{t} N_k - \sum_{k=t-S}^{t-1}(1-b)s_k^{t-1}} \qquad \text{if } c = t$$

$$s_c^t = (1-b)s_c^{t-1} + (1-(1-b)s_c^{t-1})(\frac{a\sum_{k=t-S}^{t} N_k}{\sum_{k=t-S}^{t} N_k - \sum_{k=t-S}^{t-1}(1-b)s_k^{t-1}}) \qquad \text{if } c \leq t \leq c+(S-1)$$

This simple model can be used to show a couple of predictions which can be tested empirically.

Claim 1 *The probability of employment at the end of period t for cohort j, s_j^t, is nondecreasing in t for reasonable values of a and b.*

Since each agent has an opportunity to receive information on jobs directly every period, being in the market for more periods will lead to a higher probability of employment. In other words, the longer a cohort has been in the market, the more periods in which cohort members would have accumulated information about job openings, increasing directly their probability of employment.

In the simplest case, in a model in which job information can not be passed, this is very clear. The employment probability for a given period t is expressed as:

$$s_c^1 = a$$

$$s_c^2 = (1-b)s_c^1 + a(1-(1-b)s_c^1) = a + a(1-b-a(1-b))$$

$$s_c^3 = (1-b)s_c^3 + a(1-(1-b)s_c^3) = a + a(1-b-a(1-b)) + a(1-b-a(1-b))^2$$

$$\vdots$$

$$s_c^t = \sum_{r=0}^{t-1} a(1-b-a(1-b))^r = \frac{a}{b+a(1-b)}[1-(1-b-a(1-b))^t]$$

which is clearly increasing in t since $0 \leq (1-b-a(1-b)) \leq 1$ for $0 \leq a \leq 1$, $0 \leq b \leq 1$. In fact, the derivative of s_c^t wrt t is:

$$\frac{\partial s_c^t}{\partial t} = -t\frac{a}{b+a(1-b)}(1-b-a(1-b))^t ln(1-(1-b-a(1-b))) > 0$$

since $ln(1-(1-b-a(1-b)))$ is negative.

However, in a model with information passing and changing cohort size, this claim may not hold. For example, if b is high and the size of an entering cohort p increases to a sufficiently high level, the decrease in the number of jobs available through the network to cohort $p-1$ due to the cohort size change can dominate the effect of being in the market one additional period. That is, $s_{p-1}^{p-1} > s_p^{p-1}$ in this circumstance.

Claim 2 *For all values $0 < a < 1$ and $0 < b < 1$, an increase in cohort size N_j decreases s_c^j for all c.*

Proof of Claim 2:

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

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

Diff wrt N_j:

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

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

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

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

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

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

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

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

The intuition is that since s_k^{c-1} does not change, increasing N_c only increases the number of unemployed individuals seeking job info from network members while leaving the number of employed members unchanged.

Claim 3 *For certain values* (a, b)*, an increase in cohort size* N_j *decreases* s_{j+1}^{j+1} *while increasing* s_k^k *for* $k > c+1$*.*

Diff wrt N_j gives:

$$\frac{\partial s_{j+1}^{j+1}}{\partial N_j} = \frac{a(1-b)}{D}[(\bar{N} - N_j)s_j^j - \sum_{k=j-2}^{j-1} N_k s_k^j + \bar{N}(N_j \frac{\partial s_j^j}{\partial N_j} + \sum_{k=j-2}^{j-1} N_k \frac{\partial s_k^j}{\partial N_j})]$$

where $D = 1/[N_j + \sum_{k=j-2}^{j-1} N_k(1-(1-b)s_k^j)]^2$

The idea is that an increase in N_j creates more competition for job information within the network, decreasing the employment probability for cohort $j+1$. However, as cohort j gains experience in the labor market and has higher rate of employment, the larger size becomes an asset.

As an example of how this model leads to the predictions outlined in Claims 1, 2, and 3, Figure (1) compares the employment rates of a control network in which cohort size is constant and that of a network in which the size of cohort j was doubled, keeping the other cohorts of constant size. The treated cohort, j, experiences a lower employment rate in their first period in the market, but by period 4, the larger cohort size leads to a slightly higher employment rate. s_{j+1}^{j+1} is represented as Time Period 1 in the figure entitled "Cohort 1 after Treated Cohort". Similar to the pattern displayed by cohort j, the initial employment rate is lower than it would have been in the absence of the cohort size shock, but this effect is largely gone by cohort $j+1$'s second period in the market. In fact by Time Period 3, corresponding to period $j+3$, the cohort reaches a higher employment rate than the counterfactual cohort. The following cohorts, $j+2$ and $j+3$, both receive gains in the employment rates for all 4 periods these cohorts are in the market.

However, as can be seen in the derivative above, the claim does not hold for all values of a, b. Since $\frac{\partial s_k^j}{\partial N_j} < 0$ for all k from Claim 2, all terms in the above derivative are negative except $(\bar{N} - N_j)s_j^j$. There are some values of a, b such that this positive term can dominate the other terms. For example, for high values of b (.8 and above), the derivative is positive. Such high levels of b, however, lead to very low average employment rates of less than .20 in all periods for all cohorts, which does not match well with the data available on average employment rates.[3]

[3]The range of values for which Claim 3 holds is also a function of the number of time periods an agent

2.2 A Model of Employment Rates and Wages

An implication from subsequent work by Calvo-Armengol and Jackson (forthcoming 2006) is that we should see the same pattern in wages as in employment. In this model, job information that arrives exogenously also includes a wage, so that an employed individual may act on the information and switch jobs if the offer wage is higher than his current wage. The role of referrals of job information is the same as above. The implications of this more general framework is that in the steady-state, information passing leads to positive correlation between the employment and wage status of agents who are connected by a social network. There is, however, the possibility of a negative correlation in wages across certain agents within a given period.

In order to empirically analyze the implications from such a model in the data, I first add in wages into the overlapping generations framework used above in the following way. With probability a, an individual receives job information which now also contains a wage. If the individual who receives the job info is unemployed, he takes the job. However, if the individual is employed, he accepts the job if $w^o_{ict} > w_{ict}$, where w^o_{ict} denotes the offer wage from the new job information randomly received by employed individual i. Alternatively when $w^o_{ict} < w_{ict}$, then the offer is passed to a randomly selected unemployed network member. Wages are *iid* draws from the uniform distribution $w \sim U[\underline{w}, \bar{w}]$. w^c_c will denote the average wage for employed network members in cohort c in period c.

In this model the effect on wages and employment is more subtle than in the simpler model without wages. For a given employment rate, the job information available in the network for unemployed members is diminished since employed network members with low wages are unlikely to provide information to the unemployed. This is because the only jobs which are passed are those which have sufficiently low wages such that the employed network member who initially received the job information would reject the offer. This also implies that individuals who become employed through job information that was passed to them by employed network members will have wages that are lower than the average.

While general results for this model are still in progress, I present here one numerical example which is consistent with the intuition that a negative relationship exists between own cohort size and $t-1$ cohort size and an individual's wage while a positive one exists

is in the market, S. Future drafts of this paper will contain a more complete discussion of the ranges of a, b, N_j, S such that Claim 3 holds.

for cohorts $k : k < t-1$. More precisely, Figures (2) and (3) reflect the results of simulating the model with $a = .40$ and $b = .05$ where agents work in the market for 5 periods, i.e. $S = 5$. Wages are distributed $w \sim U[5.15, 45.15]$ where $\underline{w} = 5.15$ reflects minimum wage law. The thought experiment here is to triple cohort size N_c and evaluate the effect on s_c^c and w_c^c and employment rates and wages of subsequent cohorts $c+1$, $c+2$, $c+3$ and $c+4$ (which are of normal size). The figures present both the results of the simulated model with the shocked cohort c and the counterfactual where cohort size remained constant. For cohorts c and $c+1$, both the employment rates and the average wage are lower in the first period than the levels that would have been achieved under the counterfactual of no change in cohort size. The effect on cohort $c+2$ in its first period in the market, however, is close to zero while cohorts $c+2$ and $c+3$ show initial gains from the increase in cohort c. This simulation exercise will provide guidance on interpreting the empirical results analyzing wages and network size and structure in section 5.

3 Institutional Environment and Data

The United States has a long history of refugee resettlement, having accepted around 2.4 million refugees and asylees since 1975. Since 1996, over 500,000 refugees and asylees have been admitted. Refugees come from a wide variety of countries and flee their homes for widely varying reasons, from war-related violence to religious persecution to retribution for political views. The process in which refugees gain access to the U.S. creates a unique opportunity to look at the role of ethnic networks. Refugees are a well-defined group. According to Immigration and Nationality Act (INA) Section 101: a refugee is

> any person who is outside any country of such person's nationality...who is unable or unwilling to return to...that country because of persecution or a well-founded fear of persecution on account of race, religion, nationality, membership in a particular social group, or political opinion.

Refugees are distinct from asylees in that refugees' status determination occurs overseas. Asylees, by contrast, travel by their own means to the United States and then apply for protected status upon arrival.

How does one become a refugee? The president, after consulting Congress, sets designated nationalities and processing priorities each year which fit into the predetermined ceiling for total refugee admissions levels. The Bureau of Population, Refugees, and Migration (PRM) of the State Department develops the application criteria and specific admission levels while INS officers adjudicate individual cases in refugee processing centers around the world. Often these centers are within refugee camps, although individuals can also apply for refugee status in the local U.S. embassy. Once the INS designates an individual as having refugee status, the PRM is responsible for overseas processing and transportation to the U.S.[4]

The PRM's final role in the resettlement process is to allocate all accepted cases to one of twelve contracted voluntary resettlement agencies. The resettlement agencies are responsible for acquiring housing, providing initial benefits including cash assistance and in-kind support, as well as providing access to resources such as ESL training and job assistance. This makes estimating the effects of social networks on labor market outcomes among refugees resettled in the U.S. a particularly interesting case since the mechanism through which these networks operate can be pinpointed. Since refugees are provided with housing and some initial financial assistance, the potential intervention by the social network is more limited than the case of Mexican migrants. I use data from one voluntary resettlement agency, the International Rescue Committee (IRC), who resettles approximately 12 percent of all refugees and asylees. In this paper I look specifically at individuals who are granted refugee status directly, excluding both asylum seekers and refugees who attained admittance via family reunification. For these individuals, the IRC has the sole discretion in determining where the refugee will be resettled among its 16 regional offices. The IRC receives information from the State Department about individual characteristics of each refugee including basic information such as country of citizenship plus demographic information including age, gender, marital status and education. With this information, the IRC decides to send each refugee or refugee family to one of its 16 regional offices. It is important to note is 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 home country or country of first asylum overseas to the chosen IRC regional office within the U.S.

[4]Transportation of refugees to the U.S. is usually contracted out to the International Organization for Migration.

One remaining question is how are refugees distributed between IRC's 16 regional offices. The IRC does not have an explicit placement rule, although they do follow a few general guidelines. First, the IRC seeks to place refugees in locations where there is the presence of a pre-existing ethnic or nationality-based community. They also attempt to choose a regional office based on language competencies. The goal is to send each refugee to an office which has either a staff member of a volunteer with competency in a language spoken by refugee. Individual refugees or refugee families who have special medical problems, such as HIV or severe mental health concerns, are only sent to particular offices which specialize in such cases.

In addition to policies oriented towards achieving a good match between an individual refugee and a city, the IRC also budgets for the total number of refugees expected to arrive in each regional office. To do this, each regional office is budgeted a total number of people per year plus a target for non-family reunification refugees. These numbers are estimated using projected numbers for how many refugees are expected to be admitted to the U.S. from each region of the world as provided by the State Department. The department of the IRC responsible for placing refugees therefore attempts to match these numbers. Often the actual numbers can vary substantially from those anticipated since the actual number of refugees who arrive from a region can be volatile. There is also a great deal of uncertainty about the number of family reunification cases arriving each year. Since family reunification cases are predestined for particular offices, this shifts the allocation of non-family reunification cases away from budgeted numbers. Finally, the overall number of refugees sent to a particular office is also a function of employment statistics at the regional office level.

There are also three special groups whose placement do not follow the usual system outlined above: the Somali Bantu, Meskhetian Turks, and the Lost Boys (Sudanese youth from the Kakuma Refugee Camp in Kenya). The decision on which localities would be selected as sites for each of these groups was not made exclusively by the IRC. Particularly for the Somali Bantu and the Lost Boys, there was collaboration between all of the voluntary resettlement agencies leading to a coordinated placement policy. While it is not clear that any additional information was used in selecting the sites (or the distribution of refugees from each group across these sites), there is a particular worry about unobservable characteristics of these groups and how each group matches with city characteristics. The econometric analysis will therefore be done both including and excluding these groups to rule out concerns about endogenous placement of individuals in these social groups.

As for the remaining information provided to the IRC by the PRM, the IRC reports using a limited amount of this information in the allocation process. Given that this is difficult to verify, the data set used in this analysis fortunately includes all information given to the IRC prior to each refugee's arrival. In fact, the data was compiled from the very forms provided to the IRC from the PRM. I can therefore control for individual characteristics which the IRC uses in the allocation process.[5] This is important since it removes the problem of sorting based on unobserved characteristics which exists in other studies estimating social network effects.[6]

The data from the IRC comprises of approximately 4,700 individuals from "free" cases, where a free case is one where there are no family members in the U.S. to assist in the resettlement of the case. There are three components to this data. A fairly rich set of demographic variables which were compiled by the INS and the PRM prior to the refugee's arrival in the U.S. is available, including ethnicity, date of birth, country of first asylum, the size of the family being resettled, initial English language level and education received in the home country. This data is comprehensive of all individual characteristics known by the IRC at the time of placement and were manually entered from the paper forms the IRC received from PRM. Labor market outcomes, in particular employment status and hourly wage, were collected by the IRC at 90 days after each refugee's arrival. For the period 2001-2003, industry and occupation codes are available for those employed.[7] Finally, data on the total number of individuals (inclusive of all ages) placed in each IRC regional office by nationality from 1997 through 2004 were retrieved from archived aggregate reports. Unfortunately, individual-level data prior to 2001 are currently unavailable.

[5] I make the distinction here between individual characteristics and those characteristics which will be shared by an entire ethnic group, for example. This issue will be discussed in the next section.[2]

[6] Bertrand et al. (2000), for example, evaluate the role of networks in welfare participation. This study uses a similar empirical strategy with neighborhood and language group fixed effects, but there remains the possibility of differential selection of individuals into metropolitan areas based on unobserved preferences for work and welfare participation.

[7] Unfortunately as of 2004, the IRC changed the 90 day report format and eliminated the field to report employer information.

There is a wide variety of ethnic groups and nationalities in the data. The largest groups are from Afghanistan, Bosnia, Liberia, Somalia, and the Sudan, although there are in total 38 different ethnic groups represented. The IRC has 16 offices where they resettle "free" cases. Fortunately this structure creates variation in the size of local networks available to the resettled refugees while enough clustering to produce statistically distinguishable results.

In order to get an estimate of the size of each ethnic group's network in a given geographic space, I will be using two different measures. The primary analysis will define the social network as refugees from the same nationality who were resettled in the same regional office. Since the aggregate data is available from 1997 onwards, this measure of network size for an individual will include fellow refugees resettled in the four years prior to that individual's arrival. One limitation to this data is that I am unable to distinguish between adults and children so the size of the network inappropriately includes the number of children. To the extent that this may impact the empirical results, given that the model pertains exclusively to working age adults, the network size variable is an inflated measure of the actual number of network members available to provide employment information.

The second measure of the size of the social network comes from the 2000 Census data available through IPUMS. I calculate the size of the network at level of the metropolitan statistical area (MSA). Unfortunately the Census does not allow for as detailed ethnicity information as the IRC data. In many cases matching ethnicity is more akin to matching nationality. For example, the ethnic group "Somalian" is composed of over 11 ethnicities: Banjuri, Benadir, Darod, Mushanguli, Tigryan, Asharaf, Majindo, Midgan, Manyasa, Rahweyn, and Tumale. While ideally each ethnicity would be treated as their own ethnic group, due to data limitations in the Census, I must use a more aggregate measure of ethnicity. On the other hand, the social network is not exclusively by nationality either since I can identify some ethnic groups which cover multiple countries, such as the Kurds. 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. It is important to note that since the Census does not obtain information on the foreign born's visa type or residency status/citizenship, this measure will include all immigrant types, ranging from illegal immigrants to permanent residents and naturalized citizens.

Supplemental information is also available from a survey of refugees and asylees collected by the Department of Health and Human Services' Office of Refugee Resettlement (ORR). The survey is designed to be a panel study, where each respondent is interviewed for 5 years, and is intended to be representative of all refugees and asylees who were admitted to the U.S. in a given year. The panel structure is such that all resident household members are interviewed in each round, but only the primary respondent is tracked over time. I currently have access to this data from 1993-2004. There are on average around 2,000 individuals interviewed each year, although the numbers vary substantially from round to round. The strength of this data is its depth of information on the usage of public resources such as TANF, food stamps, and refugee cash assistance (RCA), as well as labor market outcomes including employment status and wages. There is also some information available on the job search process such as the use of informal referrals. The agency which facilitated the respondent's resettlement and whether the respondent was a "free" case, however, is not known, and therefore this sample may not be precisely comparable to the IRC sample used in the majority of the analysis.

4 Econometric Specification

The objective of this paper is to empirically test the predictions of a simple model of job-related information flows in social networks in order to better understand the best way to maximize the labor market success of newly resettled refugees. The model corresponds nicely to the empirical setting. Claim 2 predicts that having a larger number of network members who arrived in the prior year, corresponding to the N_{j-1} cohort, will decrease the probability of a new refugee obtaining employment within the first 90 days. More senior cohorts, conversely, will have a positive effect on employment. Drawing upon Calvo-Armengol and Jackson's full model also suggests that information flows within social networks will generate this differential pattern across network member cohorts in wages .

The specific collection of people who constitute an individual's true social network is not known, therefore I take advantage of two different data sources as described above to create two network measures, one of which comprises exclusively of refugees and the other which includes all adult individuals from the same country of origin/ethnic group. Since the data structure of these two sources differ, the empirical specifications vary as well.

While using the aggregate data on IRC placements from 1997-2005, the empirical specification will be as follows:

$$Y_{ijkt} = \alpha + \gamma_1 N_{jk(t)} + \gamma_2 N_{jk(t-1)} + \gamma_3 N_{jk(t-2)} + \gamma_4 N_{jk(t-3)} + X_{ijkt}\beta + \delta_j + \phi_k + \lambda_t + \epsilon_{ijkt} \quad (1)$$

where Y_{ijkt} represents either employment status or wages for individual i, and N_{jkt} is the number of refugees from country of origin j resettled by the IRC in regional office k in fiscal year t. Therefore $N_{jk(t-1)}$ is the number of refugees who arrived during the entire fiscal year prior to i's arrival. Since information prior to 2001 is only available at the fiscal year level, I can not create a variable which is specific to the particular month or date of i's arrival. Therefore the network variables $N_{jk(t)}$, $N_{jk(t-1)}$, $N_{jk(t-2)}$, and $N_{jk(t-3)}$ are the same for all refugees who arrive in the same fiscal year, are resettled in the same regional office and share the same country of origin/ethnicity. The number of refugees who are resettled in the same year as individual i, N_{jkt} is particularly problematic since the entire cohort does not arrive in the U.S. at the same time. The variable N_{jkt} therefore includes individuals who had not yet arrived in the U.S. and would not be competitors for job information from the network. This problem is addressed in the following section. According to the model, we would anticipate γ_1 and γ_2 to be negative while γ_2 and γ_3 would be positive.

Since the IRC resettles multiple ethnic groups across multiple cities, both geographic and ethnicity-specific factors can be controlled for using fixed effects. Unobservable factors at the city level are controlled for using metropolitan-area fixed effects, ϕ_k. Thus ϕ_k would, for instance, control for variations in the local labor market which affect all ethnic groups equally. Additionally δ_j is an ethnic group fixed effect. Thus if one particular ethnicity has lower human capital on average or if the types of people who become refugees vary across sending countries, this effect common to all refugees in a group is captured. λ_t controls for differences across arrival years for all refugees. This is an important control given the large changes in the resettlement process which took place after September 11, 2001. Resources available to the IRC diminished dramatically and according to the IRC, many employers became more reluctant to hire refugees, particularly those from Muslim countries. The additional control variables X_{ijkt} include the individual's age, age squared, gender, and the number of individuals who were resettled together (approximately family size). Additional controls include initial English ability, initial education level, marital status, religion, and health status. The error term is corrected for clustering at the nationality group/regional office/year of arrival level since this is precisely the level at which the network data vary.

To further test for the pattern predicted by the model, I also use Census data to construct a measure of network size which includes all individuals from a country of origin group in a given metropolitan area.[8] This measure will include all immigrants groups, not only refugees. 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.

$$Y_{ijkt} = \alpha + \phi_1 N_{jk(t=1999)} + \phi_2 N_{jk(t=1999)} * \lambda_{2001} + X_{ijkt}\beta + \delta_j + \phi_k + \lambda_{2001} + \epsilon_{ijkt} \qquad (2)$$

Y_{ijkt}, X_{ijkt}, δ_j, ϕ_k, and ϵ_{ijkt} are defined as above. As described above, $N_{jk(t=1999)}$ is the size of the network for those immigrants who arrived in 1999 according to the Census, and λ_{2001} is an indicator for those refugees who arrived in 2001 (as opposed to 2002). We would expect ϕ_1 to be positive and ϕ_2 to be negative. The differential network effect across the two cohorts is therefore captured by ϕ_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. According to the model, by 2002 the network members would have acquired additional job information, becoming employed themselves, such that they would be able to provide referrals to newly resettled refugees. These network members would, however, be more likely to be competitors for job information with those who arrived more closely to them in time, namely refugees in the 2001 cohort. In this specification, the error term is corrected for clustering at the nationality group/regional office level.

Both of the above equations will be estimated by a linear probability model for the probability of employment.[9] However, there remains a problem in estimating the wage equation. The model predictions imply that the effect of network size should have an effect on *offer* wages. However, the data provided by the IRC only provides wages for

[8]Most networks are defined at the level of the MSA, however some include multiple MSAs. For example, refugees resettled in the New York office can be resettled in either New York-Northeastern NJ MSA or the Nassau Co., NY MSA. Thus the network size includes both MSAs since there is likely to be contact between individuals across this geographical space.

[9]Both probit and logit models provide similar results.

those individuals who are employed, and as such offer wages for those who are unemployed are unknown. The interpretation of $\hat{\gamma}_1$, $\hat{\gamma}_2$, $\hat{\gamma}_3$, $\hat{\phi}_1$ and $\hat{\phi}_2$ when equations (1) and (2) are estimated by OLS for wages when the sample restricted to those who are employed is unclear. Since it is unlikely that the wages of employed workers are a random subset of the wage offers to all workers, these parameter estimates are not necessarily consistent.[10] In particular, when the wage regression is conditional on employment, it does not control for how the network impacts labor market participation.

The classic solution to this problem is to estimate a structural model of wage offers and labor market participation. Without a suitable exclusion restriction, however, classic selection models are not necessarily identified.[11] One alternative solution is to impute unobserved wages as zero and estimate the wage equation using least absolute deviations (LAD). Following Johnson, Kitamura and Neal (2000), consider the following model:

$$w_i = X_i'\beta + \epsilon_i$$

where w_i is wage offer, X_i are observed characteristics and ϵ_i are unobserved traits for individual i. However, w_i is unobserved if i is unemployed. Let I_i denote individual i's employment status, where $I_i = 1$ implies that i is employed. We can therefore create another variable y_i such that $y_i = w_i$ if $I_i = 1$ and $y_i = 0$ if $I_i = 0$. The key assumption is that all unemployed individuals receive wage offers below the median offer made to employed workers with comparable skills:

$$w_i < X_i'\hat{\beta} \qquad if \quad I_i = 0$$

Under this assumption, LAD estimation is unaffected by imputing unobserved wage offers as zero. Johnson, Kitamura and Neal show that in the NLSY, the assumption is confirmed in the vast majority of cases. They use panel data to follow up on those individuals who were unemployed in 1990 and 1991 to show that this method is a fairly accurate way to get unbiased estimates in the face of selection problems. The primary concern with the assumption, as expressed by Altonji and Blank (1999), is that those with missing wages may be those with high-wage offers who are temporarily not working.

[10] The model with wages presented above does allocate wages randomly. Individuals who receive job information from other network members do not receive random wages though. Additional implications of this model regarding selection are currently being investigated.

While it is difficult to provide direct evidence on the validity of this assumption for the sample of refugees used in this study without panel data, I will note that the majority of refugees in the IRC sample come to the U.S. with very low levels of education and often little to no English skills. As can be seen in Table (1), 45% of men in the sample arrived in the U.S. with no English ability. 34% had only received some primary school education or less. The refugees in sample generally find employment in low skilled service positions, such as housekeepers. In terms of industry of employment, 24% were employed by the traveller accommodation industry and an additional 9% in restaurants.[11] Even for those individuals who arrive in the U.S. with prior skills, the IRC explicitly encourages them to gain immediate employment in a low-skilled position and later attempt to transition into a position in their field instead of remaining unemployed to search for a position in their prior occupation. It is therefore likely that those who are unable to gain employment in the initial 90 days after arrival are those with limited skills, beyond which is observed by the econometrician, who would otherwise have low wage offers.

5 Empirical Results

The first piece of evidence in support of the model can be found in Table (2). Here the ORR data is used to test Claim 1. There is a strong correlation between the length of time since resettlement in the U.S. and the probability of employment even after controlling for a number of demographic information including age, marital status, education prior to arrival in the U.S., resettlement state, country of citizenship, and year of the survey. Of course, this is not a causal parameter since length of tenure in the U.S. will be correlated with a number of factors which would increase the likelihood of employment, including English language acquisition and other U.S.-specific human capital accumulation.[12] Nonetheless, an extra year of residence in the U.S. since resettlement is associated with a 3.4% increase in the probability of being employed. This information therefore supports Claim 1 that network members who have a longer tenure in the U.S. are more likely to employed and consistent with the idea that these members will be in a better position to provide job information to

[11]See Table (14).

[12]Ideally I would also include individual fixed effects to exploit the panel nature of the survey, however I am currently unable to fully match individuals within the sample across rounds due to data quality issues. Therefore, these results use pooled OLS.

new arrivals. Those refugees with the lowest levels of tenure, conversely, are the least likely to be employed and therefore could serve as potential competitors for job information from other network members.

The current resettlement pattern of IRC refugees is central both to the question of the optimal resettlement strategy of refugees as well as the empirical strategy. Table (3) confirms that the IRC is following the clustering strategy they state as policy. The table presents the correlation matrix of the number of people the IRC allocated to each nationality/regional office pair, i.e. the size of each cohort across 4 year periods from 1997-2005. Indeed, there is a positive correlation of the numbers of refugees by nationality sent to a given regional office over all time periods. The largest correlation is between the numbers resettled in the current year and the one year prior.

5.1 Probability of Employment

The results of estimating equation (1) confirm the predictions of the information transmission model. Table (4) shows that a larger number of network members who arrived in the prior year strongly decreases the probability of employment for a new entrant. A one standard deviation increase in $t-1$ network size decreases the probability of employment by 3.2%. Given that the mean level of employment in the sample is 64%, this constitutes a decline of 5%. Recall from the analysis done with the ORR data that an additional year in the U.S. lead to an increase in the employment rate of 3%. Therefore this negative network effect is an economically significant factor in determining refugee labor market unemployment. This component would be lost if I only estimated the effect from the total size of the network instead of looking for the dynamic component of social network behavior. This is an important step in understanding how networks function beyond estimating a composite reduced form effect.

As is consistent with the model, however, a larger number of refugees with two to four years of experience living in the U.S. at the time of arrival of a new refugee 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 3.8%. In this specification, the number of refugees who arrived in the prior 3 and 4 years were combined, although the individual coefficients are both positive and jointly significant at the 5% level when estimated separately. It is fairly surprising that the coefficient on the number of refugees who arrived in years $t-3$ and $t-4$ is smaller than that of the $t-2$ network, although it is still positive and statistically significant. One reason for this is that out migration is likely to be higher for refugees who had been resettled 3 or more years prior to the new arrival.[13] Out migration within the first 90 days is 6.96%, and therefore it is quite plausible that the smaller coefficient reflects the fact that this variable has more measurement error in representing the true number of network members currently available to the new arrival. Attenuation bias would then push down the size of the coefficient compared to that of the $t-2$ cohort.

The coefficients on the control variables are as expected, although the interpretation is unclear given that the coefficients are a mixture of the causal relationship and the selection rule used by the IRC. Age displays a concave relationship with the employment rate, increasing at a decreasing rate. Household size is negative, reflecting that a larger household may also contain more potential workers, thereby diminishing the incentive to work or providing the opportunity to pursue education or full-time language training for any given individual. The coefficients on the year indicators, available upon request from the author, are consistent with the IRC's intuition that the economic opportunities of newly arrived refugees diminished dramatically after September 11th, only recently recovering in 2005 and beyond.

The coefficient on the size of own cohort, i.e. the number of refugees who arrived in year t, is positive and not statistically significant. Recall that Claim 2 predicted that all cohorts active in the market in time t would be negatively effected by an increase in the size of the entering cohort in that time period. This prediction is not held in the data. However, the variable for the size of the current cohort used in Columns 1 and 2 is not the appropriate measure. As discussed in the previous section, it contains the total number of refugees who arrived in the entire fiscal year t. It therefore includes refugees who were resettled after the arrival and reporting of the 90 day employment outcomes for many sample respondents. These individuals are included as network members even though they could not have directly influenced the sample respondents' outcomes. For example, a refugee i who arrived in January of year t has his labor market outcomes reported in late

[13]The data used to measure network size is the total number of refugees who were placed in a given city in a given year, and I do not know if those individuals continue to live in their initial location.

March or April. A refugee who then arrived in July would neither compete nor provide job information to individual i and should be excluded from the variable $N_{jk(t)}$. The variable used in Columns 1 and 2 of Table (4) therefore mismeasures the relevant network size. As long as this measurement error is classical, the attenuation bias would upwardly bias the results. The ideal variable would include only those individuals who arrived in year t up to individual i's specific date of arrival. Since individual records with date of arrival are not available prior to 2001, I am unable to construct the correct variable.

To investigate the hypothesis that measurement error is driving the insignificant effect of own cohort on employment outcomes, Columns 3 and 4 of Table (4) restrict the sample to refugees who arrived between 2003 and 2005. In this case, the number of refugees who arrived in time t, $t-1$ and $t-2$ can be constructed correctly. This comes a cost though since the sample size is reduced significantly. This analysis is therefore useful in showing that the coefficients on $t-1$, $t-2$ and $t-3/t-4$ can be interpreted as causal and unaffected by the mismeasurement of the year t variable. It is also suggestive of the true relationship between network size in year t and the probability of employment. Column 3 shows that the correctly measured variable for network size in time t is estimated to be negatively associated with employment. Inference is limited though by the large standard errors. The coefficients on $N_{jk(t-1)}$, $N_{jk(t-2)}$ are both negative although only $N_{jk(t-2)}$ is precisely estimated. In order to provide evidence that the large standard errors are driven by the decline in sample size and not sample selection, Column 4 uses the same variables as in Columns 1 and 2 but uses only the restricted sample.[14] Here both $\hat{N}_{jk(t)}$ and $\hat{N}_{jk(t-1)}$ are also statistically insignificant as in Column 3 but $\hat{N}_{jk(t)}$ is again positive. This suggests that the positive coefficient is a consequence of measurement error in the variable for $N_{jk(t)}$ used in Columns 1,2 and 4.

The specification estimated in Column 1 contains a limited number of demographic covariates. There are accordingly individual characteristics which may have been used by the IRC when choosing an individual's location, thereby being correlated with network size in ϵ_{ijkt}. Column 2 addresses this potential concern by including a wide range of individual

[14] The number of observations differ in Columns 3 and 4 because Column 3 uses the calendar year to define $N_{jk(t)}$ for all t instead of the fiscal year definition used in Columns 1,2 and 4. This was done so as not to lose observations at the end of year 2005 which are thrown out in Columns 1,2 and 4 since fiscal year 2006 data are not currently available. The fiscal year runs from September to August.

characteristics as well as interaction terms between household size and the resettlement city. The estimates of $N_{jk(t)}$ for all t are robust to the inclusion of these variables as can be seen in Column 2 of Table (4). The coefficients remain largely the same and continue to be significant at the 5% level. This set of variables span the information which is available to the IRC at the time of placement. This is important since it removes any correlation between unobserved individual characteristics and the network size variables. Any other characteristic affecting employment outcomes would not have been known by the IRC when making decisions over placements and therefore uncorrelated with $N_{jk(t)}$ for all t. The interaction terms between household size and individual regional offices control for the IRC's placement rule sending large families to cities with less expensive housing. These terms also capture any bias induced by the IRC taking into account differential welfare policies, since larger families are likely to benefit more from TANF, food stamps and public housing.

To return to the correlation matrix presented in Table (3), if any correlation exists between network size in a given period and ϵ_{ijk}, then the covariance structure between the network variables themselves could cause the observed pattern. However, the positive correlation across all periods fortunately indicates that this is not the case. Further confirmation that the correlation between the network variables is not creating a spurious pattern will be shown in section 5.3 where an alternative measure of network size which is out of the control of the IRC is used.

5.2 Probability of Employment: Robustness Analylsis

While I am able to control for individual characteristics which may impact the way the IRC distributes refugees, I can not definitively rule out the possibility of unobservables in the ϵ_{ijk} term which vary at the jk, jt or jkt level and are correlated with network size. For example, adding on a bit of structure onto the error term in equation (1) gives:

$$Y_{ijkt} = \alpha+\gamma_1 N_{jk(t)}+\gamma_2 N_{jk(t-1)}+\gamma_3 N_{jk(t-2)}+\gamma_4 N_{jk(t-3)}+X_{ijk}\beta+\delta_j+\phi_k+\nu_{ijkt}+\mu_{jk}+\upsilon_{jkt}+\tau_{jt}$$

In this case, if μ_{jk}, τ_{jt} or υ_{jkt} are correlated with N_{jkt} in any t, then the estimates of γ will be biased. I will therefore include a wide range of additional variable to evaluate the stability of the estimates of γ and to argue that unobservable factors in μ_{jk}, τ_{jt} and υ_{jkt} are not responsible for the results.

First, the way the INS adjudicates individual cases and therefore selects people to become refugees is not well known. There is also a great deal of variation in the time delay between when an individual is granted refugee status and when he is allowed to travel to the U.S. If individuals are able to influence this process, there could be a selection problem. For example, individuals who enter the U.S. after a large cohort may differ in an unobservable way from those who enter at the same time as the large cohort. In that case, τ_{jt} would be correlated with $N_{jk(t)}$ for all t. To address this concern, I estimate two different specifications of equation (1): one with fixed effects for nationality-year interactions and another directly including the total number of refugees who arrived in each year t of group j. The former can be found in Column 3 of Table (5). The coefficients on N_{jkt-1}, $N_{jk(t-2)}$ and $N_{jk(t-3,t-4)}$ are all of the expected sign. Table (6) shows the estimation results including $N_{j(t)}$, $N_{j(t-1)}$,$N_{j(t-2)}$ and $N_{j(t-3,t-4)}$. The coefficients on the parameters of interest are again similar to those found in the baseline results and $\hat{N}_{j(t)}$ for all t are insignificant, independently and jointly.

An alternative hypothesis and a common problem in identifying network effects based on geographic variation is that there may be city-ethnic group specific matches which are both unobserved and correlated with network size. This would arise if, for example, there are characteristics or skills which are common to all individuals in ethnic group j which receive a higher return in particular cities k. Thus if particular immigrant groups receive differentially higher wages in a particular city and the IRC uses this information while making decisions on how to distribute refugees, network size would be endogenous. There are two reasons this is unlikely to be the case. First, the IRC themselves state that they take no position on whether certain cities are preferable for particular ethnic groups. In fact, the employment outcome data which is used in this analysis has never before been analyzed by the IRC to gauge where groups perform better.[15] Not unlike the canonical pool player subconsciously calculating angles while playing, though, this alone does not definitively rule out the possibility that there are not other unobservable characteristics of a city-ethnic group pair which are being used to determine placement by the IRC. The second argument is that a comparative advantage story such as this would not generate a negative $\hat{\gamma}_2$ and positive $\hat{\gamma}_3$ and $\hat{\gamma}_4$ estimates.

[15]Given that the majority of the data used in the project were created from paper records, there is little possibility that any systematic review of this information was done by the IRC.

To provide further evidence against the idea that there is comparative advantage for some groups in particular cities, I use data on the number of refugees who arrived via family reunification and were resettled by the IRC. Since these refugees are immediately reunited with their family already in the U.S., their placement in the U.S. is chosen by their family members. If there were specific match qualities at either the jk or jkt level, then the family members of these refugees would exploit this and self-select into preferable cities. If that is the case, a larger number number of family reunification refugees would be associated with higher employment rates for all refugees from that ethnic group j in city k at time t. Table (7) shows the results of estimating equation (1) with the additional variables for the number of family reunification refugees arriving from time t to time $t-4$. The coefficients of γ are left qualitatively unchanged, with a negative $\hat{\gamma}_2$ and positive $\hat{\gamma}_3$ and $\hat{\gamma}_4$. There is also no significant effect of the number of family reunification refugees on the employment outcomes of sample refugees. None of the individual coefficients are significant and the joint test does not reject the null hypothesis of no effect. Thus this test suggests that comparative advantage is not a strong mechanism affecting refugee employment outcomes.

The earlier section on the institutional environment of refugee resettlement explained the key placement strategies used by the IRC. I will now try to rule out the possibility that any of those factors are leading to a bias in the γ estimates. One of the placement criteria used by the IRC is to attempt to send each refugee to an office where someone can speak the same language as the refugee. I therefore include 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. As can be seen in Table (8), these variables do not significantly affect the probability of employment. In fact, having a staff member in the office who speaks the same language is negatively associated with employment. However, this coefficient is estimated with a great deal of noise which is likely to be the explanation for this unintuitive result. The estimates of γ continue to be as predicted by the model of information transmission. Column 2 of Table (5) rules out concern over endogeneity of placement of the Somali Bantu, Meskhetian Turks and the Lost Boys. Excluding those groups actually leads to an increase in the size of the γ coefficients. Finally, both the budgeted numbers and the employment rates at the office level could reflect city-level characteristics which are time variant. The first column in Table (5) shows the results from including for city-year interaction fixed effects. This specification also controls for other time varying city characteristics which may influence refugee employment outcomes. The coefficients of interest continue to reflect the pattern predicted by the model, although $\hat{\gamma}_{jk(t-2)}$ is no longer statistically significant. Given that the city-year controls dramatically increases the number of covariates in the estimating equation, this could easily lead to larger standard errors. The fact that the results remain stable qualitatively supports the baseline analysis.

Finally, social networks play an important role in many aspects of a refugee's life. In particular, network members may be helpful in providing translation services either during the job search process or on-the-job. This effect could be influencing the coefficients on $N_{jk(p)}$, for $p < t-1$ in particular. A larger number of network members to translate at the workplace site is likely to increase the probability of employment of recently arrived refugees. This benefit of the network would be stronger for those individuals who arrive in the U.S. with no or low English knowledge. I therefore test to see if there are heterogeneous network effects depending on refugees' initial English levels. Table (9) shows results of including $N_{jk(t)}$ for all t and $N_{jk(t)}$ for all t interacted with initial English ability. The level effects are consistent with the previous analysis and the interaction terms are insignificant. This indicates that the coefficients on $N_{jk(t)}$ for all t do not reflect the ability of social networks to provide English language services but is instead more consistent with the model of job information transmission.

5.3 Probability of Employment: Census Data Specification

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 statuses. In this case, individual network members have self-selected into their preferred location based on a number of unobserved factors. So again, while this measure of the network does not rule out the possibility of a selection bias due to comparative advantage on a *prima facie* basis, it instead supports the generality of the job information sharing effect across two independently constructed network measures. Table (12) 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 average network size from those individuals who arrived in 1999 according to the 2000 Census is larger than that of the refugee community alone, with a median of 21. Therefore by increasing network size from the median to the 75th percentile, 95, increases the probability of employment for the 2002 cohort by 2.3%. This is a very similar effect to that estimated when the network was defined as exclusively within the refugee community. The interaction term between the 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 not statistically significant at any reasonable level. This is consistent with the referral model: those refugees who arrive less than 2 years after the network members do not gain from an increase in network size while those who arrived sufficiently later do experience the positive influence of the network in terms of job information. While not shown in Table (12), these results are also robust to the inclusion of a richer set of demographic variables.

5.4 Wages

Both measures of network size provide complementary evidence on the importance of job information flows within ethnic networks among refugees resettled in the U.S. for employment. I now turn to the role of networks on wages. The OLS estimates of equation (1) in Table (10) indicate that there is no significant network effect in predicting wages. The analysis is restricted to employed male refugees since wage offers to the unemployed are unknown. A specification where network size in years 3 and 4 are estimated separately with the inclusion of additional control variables provides similar results as to Column 1.

As argued in section 4, however, conditioning on employment may have biased the OLS estimates. Columns 2 and 3 provide estimates of equation (1) with the full male sample using LAD with wage offers for the unemployed imputed as zero. The estimates are exactly what the model would predict. A one standard deviation increase in the number of network members who arrived in time $t-1$ decreases the wage by \$.16. Finding a small effect is not surprising given that most refugees are employed in low-skilled jobs and accordingly the distribution of wages is very tight. More importantly, it provides further evidence of the empirical validity of Claim 2. The size of the cohort who arrived two years earlier has no statistically significant effect on wages although the point estimate is positive. Both the size of the coefficient and the level of significance increases for network members who arrived in $t-3$ and $t-4$. This shows the pattern predicted by the model of network members becoming increasingly valuable to new arrivals as their exposure in the labor market in the U.S. increases.

Including a wider range of demographic and other control variables as in Column 3 of Table (10) leads to little change in the network coefficients. Additionally, the estimates on the control variables are consistent with their effects on wages. Age is again concave: wages increase with age at a decreasing rate as is observed in numerous settings in the U.S. labor market. Household size is again negative and the aggregate changes in refugees' wages over time display the pattern that late 2001-2004 was an increasingly difficult period for refugees with an improvement in 2005.

Given that the model's predictions are driven by the distinction between employed and unemployed network members, ideally the size of the network would be broken down along those lines. Unfortunately, since there are no individual records prior to 2001 and the IRC only collects employment and wage data as of 90 days after arrival, this is not possible. However, restricting the sample to those refugees who arrived between 2003 and 2005 allows for an analysis which is at least suggestive of this preferable specification. By making the assumption that there is persistence in employment outcomes over time, i.e. that the probability of employment at 90 days is a good predictor of whether that individual will spend most time on average employed, I can construct the number of network members who arrived during the previous two years who are likely to be employed at time t. Indeed, Table (11) confirms that an increase in the number of individuals who were unemployed as of 90 days after their arrival in the U.S. is negatively associated with employment rates of refugees who arrive in time t, up to two years after the arrival of the network members. Conversely the number of refugees who were employed as of 90 days after arrival is positively correlated with employment outcomes. This same pattern is found for wages for the sample who are employed using OLS.

5.5 Wages: Census Data Specification

Changing the estimation approach to use the Census data as in equation (2) provides qualitatively similar results. Again, as indicated by Column 1 of Table (13), the OLS estimates show no significant effect of network size on wages of those employed. The coefficients in Columns 2 and 3 from using the full sample with LAD estimation tell a very different story.

The network effect for refugees who arrived in 2002 is positive and statistically strong. An increase in the network size from the median to the 75th percentile improves wages by $0.10. The interaction of network size with the dummy indicating arrival in 2001 is negative and statistically significant as in the employment regression. The sum of the two network coefficients is negative but statistically insignificant. This closely parallels the results found in the employment results as well as those predictions of the theoretical model.

The results in this section depend critically on the assumption that unemployed refugees receive offer wages which are below the median wage of those employed with similar observable characteristics. Future work will try both to provide direct evidence on the relevance of the assumption for the data used in this study as well as look to other methods to estimate the wage equation while controlling for the selection bias created by the missing wage observation problem. In particular, a matching estimator and David Lee's (2002) trimming for bounds procedure seem particularly well suited for this application in order to further support the LAD estimates shown here.

6 Conclusion

This paper presents strong evidence on the importance of ethnic networks in facilitating access to local labor markets for refugees recently resettled in the U.S. The empirical results support a model of job-related information flows within a social network. Both the size and the structure of the network, as measured by length of tenure of network members in the U.S., influence the labor market outcomes of newly arrived refugees. This provides an important insight into the functioning of social networks and provides empirical evidence of the relevance of Calvo-Armengol and Jackson's theoretical work. As argued by the latter and further supported by the empirical findings in this paper, the positive benefits of a larger network dominate the short-run competition effects from unemployed network members. Nonetheless, my analysis shows that there are also costs to participation in a social network, a point which is lost when the dynamic relationship between employment, wages and social network structure is ignored.

The results in this paper also come from an arguably unlikely environment. Given that the refugees in this sample did not know anyone in the receiving community prior to resettlement, it is striking to find evidence that these networks function well and as predicted by a theoretical model within the first 90 days after the refugee's arrival. I would argue that this sample of refugees, therefore, teaches us about how networks function more generally, as the construction of the network is fairly analogous to "anonymous" networks which form due to geographic proximity within, say, a neighborhood.

What else can we learn from this analysis to guide policies related to refugee resettlement? The analysis indicates that there are gains and losses from clustering refugees. The specific pattern observed in the data suggests that the ideal placement method for resettlement agencies to increase short run employment outcomes would be one which involves oscillating groups from a particular country between cities from one year to the next. This strategy provides the benefit to new arrivals of having a large number of network members who have been in the U.S. for two years and are able to provide job information to them, while minimizing the number of competitors for this information. However, if the goal of refugee placement is to maximize labor market outcomes in the long-run, then it is better to increase the sizes of ethnic communities without oscillation since the competition costs dissipate over time.

It is relevant to note that while this paper provides evidence of the importance of social networks in providing labor market opportunities, and therefore points to one drawback of dispersal policies, these effects are estimated as of 90 days after arrival and the long run impact of social networks on integration and language acquisition remains an open question and an area of future research. Additionally, while social networks and job referrals have been shown to play an important role in influencing economic outcomes, there are numerous other location characteristics which need to be taken into account, including local labor market conditions and the political/social environment of a particular locality. Therefore the decision to place a refugee in a given location must take into account the local labor market characteristics in addition to the size of the available ethnic community, insofar as these factors are at time at odds with each other.

References

[1] Altonji, J. and R. Blank. "Race and Gender in the Labor Market," in Orley Ashenfelter and David Card, eds. *Handbook of labor economcs*, Vol. 3. Amsterdam: North-Holland, 1999, pp. 3144-3213.

[2] Bertrand, M., E. F. P. Luttmer and S. Mullainathan. "Network Effects and Welfare Cultures." *Quarterly Journal of Economics*, 2000, 115 (3), pp. 1019-1055.

[3] Borjas, G. J. "Ethnic Enclaves and Assimilation." *Swedish Economic Policy Review*,2000, 7, pp. 89-122.

[4] Calvo-Armengol, A. and M. O. Jackson. "The Effects of Social Networks on Employment and Inequality." *American Economic Review*, 2004, 94 (3), pp. 426-454.

[5] Calvo-Armengol, A. and M. O. Jackson. "Networks in Labor Markets: Wage and Employment Dynamics and Inequality." Forthcoming: *Journal of Economic Theory*, 2006.

[6] Conley, T. G. and G. Topa. "Socio-Economic Distance and Spatial Patterns in Unemployment." *Journal of Applied Econometrics*, 2002, 17 (4), pp. 303-327.

[7] Cortes, K. E. "Are Refugees Different from Economic Immigrants? some Empirical Evidence on the Heterogeneity of Immigrant Groups in the United States." *Review of Economics and Statistics*, 2004, 86 (2), pp. 465-480.

[8] Damm, A. P. and M. Rosholm. "Employment Effects of Dispersal Policies on Refugee Immigrants, Part II: Empirical Evidence."*Mimeo*, 2003.

[9] Edin, P.-A., P. Fredriksson and O. Aslund. 2001. "Settlement Policies and the Economic Success of Immigrants." CEPR Discussion Paper 2730.

[10] 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*, 2003, 118 (1), pp.329-357.

[11] Heckman, J. "Shadow Prices, Market Wages, and Labor Supply." *Econometrica*, 1974, 42 (4), pp. 679-694.

[12] Johnson, W., Y. Kitamura and D. Neal. "Evaluating a Simple Method for Estimating Black-White Gaps in Median Wages." *American Economic Review*, 2000, 90 (2), pp. 339-343.

[13] Manski, C. "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, Vol. 60, No. 3, pp. 531-542.

[14] Montgomery, J. D. "Social Networks and Labor-Market Outcomes: Toward an Economic-Analysis." *American Economic Review*, 1991, 81 (5), pp. 1408-1418.

[15] Munshi, K. "Networks in the Modern Economy: Mexican Migrants in the US Labor Market." *Quarterly Journal of Economics*, 2003, 118 (2), pp. 549-599.

[16] Neal, D. A. and W. R. Johnson. "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy*, 1996, 104 (5), pp. 869-895.

[17] Topa, G. "Social Interactions, Local Spillovers and Unemployment." *Review of Economic Studies*, 2001, 68 (2), pp. 261-295.

[18] United Nations High Commissioner on Refugees, *Refugees by Numbers (2005 edition)*.

7 Appendix

Table 1: Summary Statistics

	Mean	Std. Dev.	No. Obs
IRC Data:			
Age	31.81	11.131	2784
HH Size	2.850	2.261	2784
Employment rate	0.639	0.480	2599
Wage (conditional on employment)	7.334	1.313	2302
Married	0.459	0.498	2730
Spoke No English at Time of Arrival	0.445	0.497	2409
Unknown Technical Training	0.019	0.135	2784
Adult Education	0.013	0.115	2784
Other Technical Training	0.004	0.060	2784
No Education	0.087	0.282	2784
Primary School	0.244	0.430	2784
Secondary School	0.415	0.493	2784
Vocational School	0.019	0.135	2784
University	0.126	0.332	2784
Graduate School	0.014	0.116	2784
Unknown Education	0.061	0.239	2784
HIV Positive	0.009	0.096	2784
Num Refugees Resettled in Prior Year	34.14	38.72	2599
Num Refugees Resettled in 2 Years before	24.66	41.29	2599
Num Refugees Resettled in 3 Years before	26.86	50.89	2599
Num Refugees Resettled in 4 Years before	25.89	48.72	2599
2000 Census Data:			
Network Size which Arrived in 1999	112.70	223.86	1164

Table 2: Probability of Employment Using ORR Data

Number of years in U.S. since Resettlement	0.034 ***
	(0.002)
Female	-0.098 ***
	(0.005)
Age	-0.008 ***
	(0.000)
Married	0.156 ***
	(0.006)
Years of Schooling Prior to Arrival in U.S.	0.029 ***
	(0.001)
Constant	0.695
	(0.440)
No obs.	30,906

a Standard errors are in parentheses.

b Also included are fixed effects for nationality, survey year, and initial resettlement state.

Table 3: Correlation Coefficients of Refugee Cohort Sizes: 1997-2005

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

Table 4: Linear Probability Model of Employment Probability on Network Size

	Column 1	Column 2	Column 3	Column 4
# Refugees Resettled in Year t	0.00032	0.00029	-0.00123	0.00008
	(0.0005)	(0.0005)	(0.0012)	(0.0007)
# Refugees Resettled in Year $t-1$	-0.00086 **	-0.00086 **	-0.00019	-0.00063
	(0.0004)	(0.0004)	(0.0010)	(0.0006)
# Refugees Resettled Year $t-2$	0.00094 **	0.00089 **	0.00453 ***	0.00279 ***
	(0.0004)	(0.0004)	(0.0009)	(0.0008)
# Refugees Resettled Years $t-3$ & $t-4$	0.00039 **	0.00042 **		
	(0.0002)	(0.0002)		
Age	0.025 ***	0.024 ***	0.030 ***	0.030 ***
	(0.005)	(0.006)	(0.006)	(0.007)
Age Sq	0.0004 ***	0.0004 ***	0.0005 ***	0.0004 ***
	(0.0001)	(0.00008)	(0.0001)	(0.0001)
HH Size	-0.0189 ***	-0.0371	-0.017 ***	-0.015 **
	(0.006)	(0.038)	(0.006)	(0.006)
No obs	2409	2013	1487	1349
Sample Years	2001-2005	2001-2005	2003-2005	2003-2005

a SE are in parentheses and clustered by city-ethnicity-year.

b Columns 1 and 2 also include fixed effects for nationality and regional office effects.

c Column 2 also has additional controls: initial English level, education, religion, HIV status, and HH size/city interactions.

d Columns 3 and 4 also include age, age^2, HH size, fixed effects for nationality, month and regional office effects.

e For Current Year Network Size Column 3 uses number of refugees who arrived prior to date of refugee's arrival.

f Columns 1,2 and 4 uses all refugees who arrived during same fiscal year.

g Columns 1,2 and 4 uses network size variables and year dummies at the fiscal year.

Table 5: Robustness Analysis for Employment

	Column 1	Column 2	Column 3
# Refugees Resettled in Year t	0.00016	0.00078	0.00097
	(0.0005)	(0.0005)	(0.0007)
# Refugees Resettled in Year $t-1$	-0.00077 **	-0.00158 ***	-0.00171 **
	(0.0004)	(0.0005)	(0.0007)
# Refugees Resettled Year $t-2$	0.00055	0.00117 ***	0.00101 **
	(0.0004)	(0.0005)	(0.0005)
# Refugees Resettled Years $t-3$ & $t-4$	0.00037 **	0.00037 **	0.00046 **
	(0.0002)	(0.0002)	(0.0002)
Age	0.027 ***	0.033 ***	0.035 ***
	(0.005)	(0.006)	(0.006)
Age Sq	-0.00043 ***	-0.00050 ***	-0.00053 ***
	(0.0001)	(0.0001)	(0.0001)
HH Size	-0.019 ***	-0.020 ***	-0.020 ***
	(0.006)	(0.007)	(0.007)
City-year FE	Yes	No	No
Excluding Special Caseload Groups	No	Yes	Yes
Nationality-Year Controls	No	No	Yes
No obs	2409	1632	1632

a SE are in parentheses & clustered by city-ethnicity-year.

b Also included are age, age^2, HH size, fixed effects for nationality and regional office.

c Columns 2 and 3 also contain fiscal year dummies.

Table 6: Employment Robustness Analysis with Total Refugee Arrivals by Country

# Refugees Resettled in Year t	0.00070
	(0.0006)
# Refugees Resettled in Year $t-1$	-0.00131 **
	(0.0005)
# Refugees Resettled Year $t-2$	0.00108 **
	(0.0005)
# Refugees Resettled Years $t-3$ and $t-4$	0.00045 **
	(0.0002)
Total Refugees from Country j in Year t	0.00004
	(0.0001)
Total Refugees from Country j in Year $t-1$	-0.00007
	(0.0001)
Total Refugees from Country j in Year $t-2$	0.00005
	(0.0001)
Total Refugees from Country j in Years $t-3$ and $t-4$	-0.00005
	(0.0000)
Age	0.033 ***
	(0.006)
Age Sq	-0.00050 ***
	(0.0001)
HH Size	-0.020 ***
	(0.007)
P-value of joint sig of Total Refugee variables	0.73
No obs	1632

a SE are in parentheses & clustered by city-ethnicity-year.

b Also included are age, age^2, HH size, fixed effects for nationality, fiscal year and regional office.

c Excludes special caseloads.

Table 7: Employment Robustness Analysis with Family Reunification Refugees

# Refugees Resettled in Year t	0.00077
	(0.0005)
# Refugees Resettled in Year $t-1$	-0.00152 ***
	(0.0005)
# Refugees Resettled Year $t-2$	0.00143 ***
	(0.0005)
# Refugees Resettled Years $t-3$ and $t-4$	0.00046 **
	(0.0002)
# Family Reunification Refugees in Year t	-0.00023
	(0.0010)
# Family Reunification Refugees in Year $t-1$	-0.00005
	(0.0010)
# Family Reunification Refugees in Year $t-2$	-0.00106
	(0.0009)
# Family Reunification Refugees in Years $t-3$ and $t-4$	0.00037
	(0.0003)
Age	0.033 ***
	(0.006)
Age Sq	-0.00050 ***
	(0.0001)
HH Size	-0.020 ***
	(0.007)
P-value of joint sig of Family Reunification variables	0.76
No obs	1632

a SE are in parentheses & clustered by city-ethnicity-year.

b Also included are age, age^2, HH size, fixed effects for nationality, fiscal year and regional office.

c Excludes special caseloads.

Table 8: Employment Robustness Analysis with Staff Controls

# Refugees Resettled in Year t	0.00060
	(0.0005)
# Refugees Resettled in Year $t-1$	-0.00169 ***
	(0.0005)
# Refugees Resettled Year $t-2$	0.00086 *
	(0.0005)
# Refugees Resettled Years $t-3$ and $t-4$	0.00041 **
	(0.0002)
Staff Member Speaks Same Language	-0.01423
	(0.0294)
Volunteer who Speaks Same Language	0.06899
	(0.0474)
Age	0.033 ***
	(0.006)
Age Sq	-0.00050 ***
	(0.0001)
HH Size	-0.020 ***
	(0.007)
No obs	1632

a SE are in parentheses & clustered by city-ethnicity-year.

b Also included are age age^2, HH size, fixed effects for nationality, fiscal year and regional office.

c Excludes special caseloads.

Table 9: Employment Robustness Analysis with Heterogenous English Effects

# Refugees Resettled in Year t	0.00036
	(0.0005)
# Refugees Resettled in Year $t-1$	-0.00132 ***
	(0.0005)
# Refugees Resettled Year $t-2$	0.00137 **
	(0.0005)
# Refugees Resettled Years $t-3$ and $t-4$	0.00058 *
	(0.0003)
# Refugees Resettled in Year t * No English	-0.00023
	(0.0006)
# Refugees Resettled in Year $t-1$ * No English	0.00088
	(0.0006)
# Refugees Resettled Year $t-2$ * No English	-0.00089
	(0.0007)
# Refugees Resettled Years $t-3$ and $t-4$ * No English	0.00003
	(0.0003)
No English	-0.093 **
	(0.046)
Age	0.023 ***
	(0.006)
Age Sq	-0.00037 ***
	(0.0001)
HH Size	-0.016 ***
	(0.006)
P-value of F test for No English Interactions	0.41
No obs	2218

a SE are in parentheses & clustered by city-ethnicity-year.

b Also included are age, age^2, HH size, fixed effects for nationality, fiscal year and regional office.

Table 10: Wages on Network Size

	OLS Conditional on Employment	Median Regression: All	
	Column 1	Column 2	Column 3
# Refugees Resettled in Year t	0.00091	0.00047	0.00066
	(0.0029)	(0.0020)	(0.0006)
# Refugees Resettled in Year $t-1$	-0.00178	-0.00415 **	-0.00312 ***
	(0.0040)	(0.0020)	(0.0006)
# Refugees Resettled Year $t-2$	0.00194	0.00190	0.00064
	(0.0019)	(0.0019)	(0.0006)
# Refugees Resettled Years $t-3$ and $t-4$	0.00016	0.00354 ***	0.00268 ***
	(0.0011)	(0.0009)	(0.0003)
Age	0.144 ***	0.335 ***	0.239 ***
	(0.035)	(0.030)	(0.010)
Age Sq	-0.002 ***	-0.005 ***	-0.004 ***
	(0.001)	(0.0004)	(0.0001)
HH Size	-0.056 *	-0.183 ***	-0.121 ***
	(0.032)	(0.028)	(0.009)
No obs	1192	2469	2059

a SE are in parentheses and clustered by city-ethnicity-arrival year pairs.

b Also included are fixed effects for nationality and regional office.

c Column 2 also has additional controls: initial English level, education, religion, HIV status, and HH size/city interactions.

Table 11: Network Size Using Within Sample Employment Info

	Employment	Wage
# Refugees Unemployed during 2 Prior Years	-0.038 ***	-0.028 ***
	(0.005)	(0.009)
# Refugees Employed during 2 Prior Years	0.018 ***	0.014 ***
	(0.003)	(0.004)
Age	0.025 ***	0.068 ***
	(0.006)	(0.022)
Age Sq	0.000 ***	-0.001 ***
	(0.000)	(0.000)
HH Size	-0.013 **	0.032
	(0.006)	(0.021)
Constant	0.463 ***	6.702 ***
	(0.166)	(0.504)
No obs	1487	848

a Network variables indicate employment status as of 90 days after the network member's arrival. Therefore assuming persistence in employment outcomes over time.

b SE are in parentheses and are clustered by city-ethnicity-arrival year pairs.

c Also included are fixed effects for nationality, regional office and year of arrival.

d Sample includes only refugees resettled from 2003-2005.

Table 12: Employment Effects Using Census Data for Network Measure

Network size which arrived in 1999 * Refugee arrived in 2001	-0.0003 **
	(0.0001)
Network size which arrived in 1999	0.0003 **
	(0.0001)
Age	0.022 **
	(0.010)
Age Squared	-0.0004 ***
	(0.0001)
HH Size	-0.024 **
	(0.011)
2001 Year Dummy	0.080
	(0.057)
Constant	0.626 ***
	(0.188)
No obs	1156

a Standard errors are in parentheses and clustered by city-ethnicity pairs.

b Also included are fixed effects for nationality and regional office.

c The sample is restricted to refugees who arrived in 2001 and 2002.

d Network Size is number of individuals in the 2000 Census who arrived in 1999 by place of birth/MSA.

Table 13: Wage Effects Using Census Data for Network Measure

	OLS Conditional on Employment	Median Regression: All Observations
Network size which arrived in 1999 * Refugee arrived in 2001	-0.0003	-0.0015 ***
	(0.0005)	(0.0004)
Network size which arrived in 1999	-0.0001	0.0013 ***
	(0.0005)	(0.0004)
Age	0.089 ***	0.240 ***
	(0.030)	(0.021)
Age Sq	-0.0011 ***	-0.0037 ***
	(0.0004)	(0.0003)
HH Size	-0.029	-0.091 ***
	(0.027)	(0.021)
2001 Year Dummy	-0.070	0.569 ***
	(0.174)	(0.112)
Constant	6.914 ***	-4.033 ***
	(0.571)	(1.059)
No obs	833	1141

a Standard errors are in parentheses and clustered by city-ethnicity pairs.

b Also included are fixed effects for nationality and regional office.

c The sample is restricted to refugees who arrived in 2001 and 2002.

d Network Size is number of individuals in the 2000 Census who arrived in 1999 by place of birth/MSA.

Figure 1: Graphical Example of Model

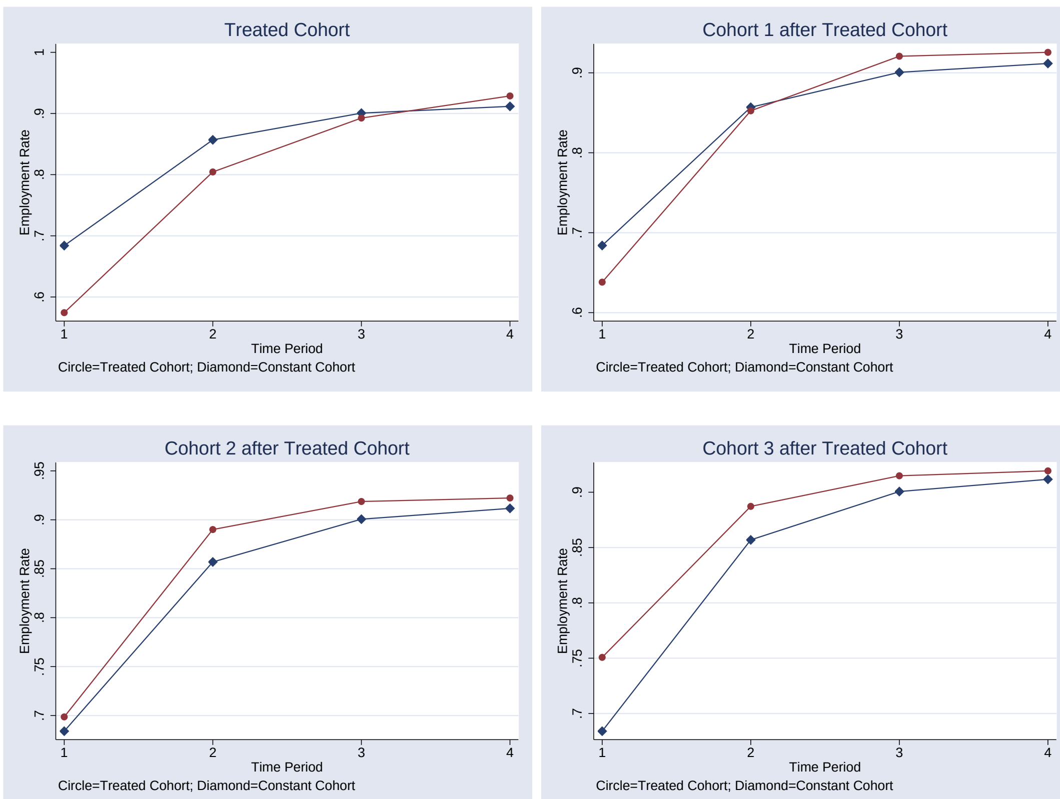

Figure 2: Graphical Example of Model with Wages: Effect on Average Wages

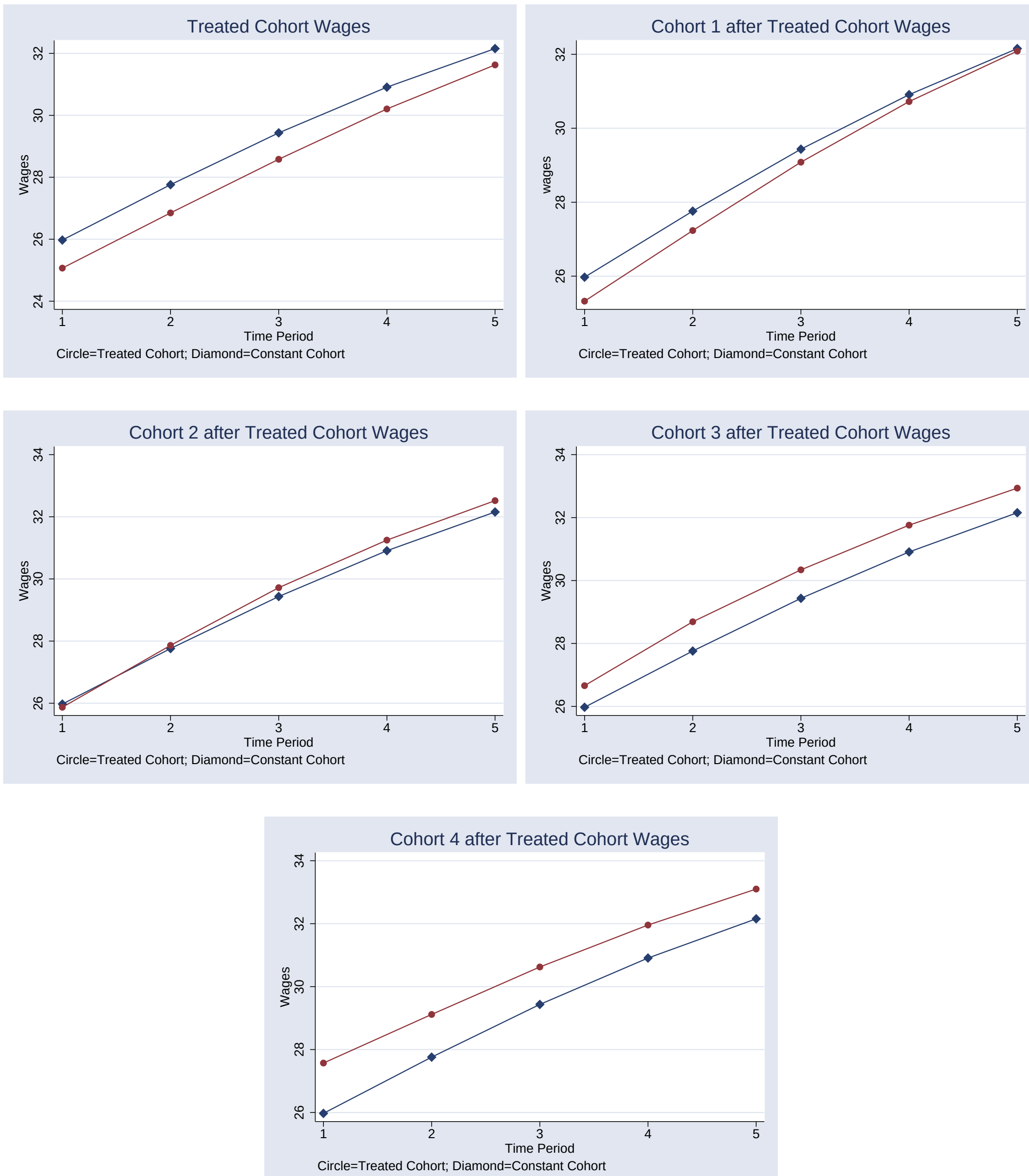

Figure 3: Graphical Example of Model with Wages: Effect on Employment Rates

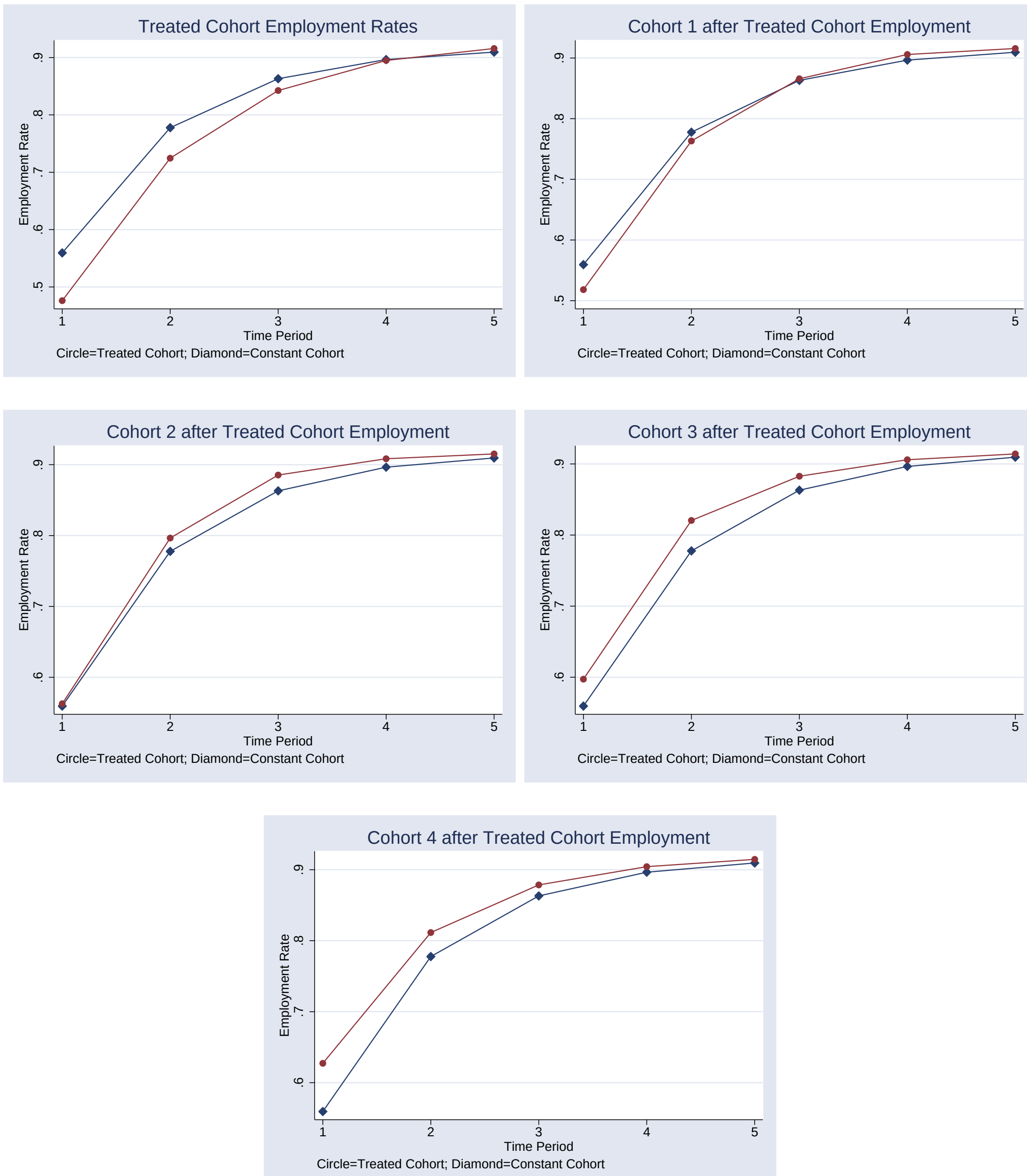

Table 14: Largest Industries from 2001-2003

Construction	.03
Animal slaughtering and processing	.05
Grocery Stores	.07
Misc general merchandise stores	.40
Misc Retail Stores	.04
Employment services	.04
Services to buildings and dwellings	.02
Colleges, including junior colleges	.04
Hospitals	.06
Traveller Accommodation	.24
Restaurants and other food services	.09
Other	.28

Above reflect the major industries in which the IRC sample gained employment from 2001-2003.