NBER WORKING PAPER SERIES

SOCIAL TIES AND THE JOB SEARCH OF RECENT IMMIGRANTS

Deepti Goel
Kevin Lang

Working Paper 15186
http://www.nber.org/papers/w15186

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2009

We thank the Canadian Labour Market and Skills Researcher Network for funding this research. We also thank its reviewers and participants at the CLSRN conference and the Growth and Development conference, ISI Delhi, for helpful comments. We are grateful to Miles Corak for enabling access to the data housed in Statistics Canada, Ottawa, Canada. Much of this research was conducted while Goel was a Fellow at Statistics Canada. The views expressed herein are those of the authors and do not necessarily reflect the views of the Canadian Labour Market and Skills Researcher Network or of Statistics Canada. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Social Ties and the Job Search of Recent Immigrants
Deepti Goel and Kevin Lang
NBER Working Paper No. 15186
July 2009
JEL No. J31,J61,J64

ABSTRACT

We show that increasing the probability of obtaining a job offer through a network should raise the observed wages of workers in jobs found through formal channels relative to those in jobs found through the network. This prediction holds at all percentiles except the highest and lowest. The largest changes are likely to occur below the median of the offer distribution. We test and confirm these implications using a survey of recent immigrants into Canada. We develop a simple structural model consistent with the theoretical model and show that it can replicate the broad patterns in the data. Our results are consistent with the primary effect of network strength being to increase the arrival rate of offers rather than to alter the distribution from which offers are drawn at least among recent immigrants.

Deepti Goel
Institute for Financial Management and Research
24 Kothari Road, Nungambakkam
Chennai, TN, 600 034, India
dgoel.india@gmail.com

Kevin Lang
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and NBER
lang@bu.edu

1 Introduction

While it is plausible that networks play an important role in the labor market assimilation of immigrants, there is little compelling evidence that they play a positive role and even less about the mechanisms by which networks affect the labor market outcomes of immigrants. In this paper we use a very different approach from those found in the literature. We develop a simple theoretical model of the role of networks, test its comparative static properties and estimate a simple structural version of our model. Our results are consistent with the principal effect of strong networks for recent immigrants being to increase the arrival rate of offers rather than to change the distribution from which offers are drawn.

We build on Montgomery (1992) who shows that among workers whose network is *stronger* than formal channels are (in other words whose network provides a greater probability of an offer than do formal channels), those who find a job through the network should have *lower* wages than those who find a job through formal channels. We show that this wage differential is increasing in absolute value in network strength and that the increase occurs at all percentiles of the wage distribution except the highest and lowest. We provide an intuitive argument that the effect is likely to be largest at percentiles below the median of the offer distribution.

To test this prediction, we hypothesize that the probability of finding a job through the network is higher if the immigrant has strong ties to someone near whom he lives. Therefore, we capture *network strength* by the presence of at least one relative or close friend in the recent immigrant's area of residence in Canada at the time of arrival.[1] We find that a stronger network helps a recent immigrant find his first job through his social network and some evidence that it raises wages, at least at lower quantiles. We test the implication of the theoretical model by examining the interaction between network strength and job finding method. The predicted negative coefficient on the interaction term is confirmed for jobs towards the lower end of the observed wage distribution. At this end of his wage distribution, finding a network job is associated with higher wages for those with weak networks, and the interaction between network strength and finding a job through the network is negative.

The larger coefficient on the interaction term at lower quantiles is an expected result based on the theoretical model. However, the theory gives little guidance as to plausible magnitudes. Are the effects we find too large or too small to be consistent with a plausible specification of the theoretical model? To address this question, we estimate a very simple structural model in which wages are drawn from two log normal distributions, one for formal jobs and one for network jobs,

[1] It is important that the recent immigrant is asked about the presence of a strong social tie in his neighborhood, *just upon arrival*. This makes network strength exogenous to his subsequent labor market experience.

and ask whether the model can produce parameters that are comparable to those we find in the data. In fact, the model does a reasonable job of fitting the parameters, suggesting that a model in which the primary role of strong network ties is to increase the flow of job offers, rather than to change the wage distribution from which offers are drawn, is consistent with the data.

In contrast with our emphasis on network strength, the literature on the relation between social networks and immigrants' labor market outcomes has focused on the effects of living in ethnic enclaves or in areas with large numbers of immigrants from their home country. Thus, this literature is concerned with network size. There are many reasons why network size may be important. Employers within an enclave may prefer to hire individuals from their own country (Borjas 2000). By employing their compatriots, they reduce frictions at work arising from differences in language and work habits. Such employees may also better understand the tastes of consumers in the enclave, helping the firm serve its market more effectively. On the other hand, by providing jobs targeted towards its members and steering them into certain occupations, enclaves may limit their search horizons. They could therefore preclude jobs in the broader labor market, that may have been better matches. Living in an enclave may also lower the speed with which new immigrants learn host country skills, e.g. language, which may reduce their chances of moving to better jobs ((Lazear 1999)). Whether new immigrants benefit from such segregation may depend on the quality of the enclave, (e.g. the stock of human capital) ((Edin, Fredriksson, and Aslund 2003); (Borjas 1992); (Borjas 1995)).

The causal effect of enclaves or networks, more generally, is difficult to determine because of the likelihood of omitted variables bias. Unmeasured factors may make immigrants from a particular country more suitable for certain jobs that are concentrated in particular areas. There may be location specific factors that result in good labor market outcomes. For example, new immigrants to areas where existing immigrants have a low unemployment rate may also have a low unemployment rate, not because the existing immigrants are better able to help the new arrivals, but because labor market conditions are generally favorable there. Moreover, there may be important unmeasured differences between individuals who choose to locate near other members of their ethnic group and those who do not.

Some papers address omitted variables bias by using instrumental variable techniques. Munshi (2003) studies Mexican migrants in the United States. He proxies the individual's network by the proportion of the sampled individuals from his community who are located in the U.S. in that year. To avoid endogeneity problems, he uses lagged rainfall in the origin-community as an instrument for the size of network at the destination. He finds that the same individual is more likely to be employed and to hold a higher paying non-agricultural job when his network is exogenously larger. Edin, Fredriksson, and Aslund (2003) study the effects of ethnic enclaves on earnings using data from an immigrant policy initiative in Sweden, when government authorities distributed refugee immigrants across locales based on the availability of housing. They argue that this provides a natural experiment which allows them to estimate the causal effect of living in enclaves. They instrument current location attributes with attributes of the initial assigned location and find that enclaves improve labor market outcomes for less skilled immigrants. Although these papers are well-executed, they examine quite atypical immigrants, Mexican immigrants to the United States and refugees in Sweden, and their findings may not extend to other immigrants.

Although it is not the principal focus of our paper, we also examine *network size,* measured by the share of the local area population from the immigrant's country of birth. We adopt a difference in differences approach to estimate the network effect. This approach addresses some of the bias arising from unobservable group/location characteristics, but may still suffer from bias arising from unobservables that vary across both group and location or from individual-specific unobservables. We undertake this exercise primarily because it permits comparison of our results with the existing literature. We find that a larger network is associated with a higher probability of finding a job and also a higher probability of finding the first job using the social network, but the magnitudes of the effects are small. We find no evidence of a positive effect of network size on wages in first jobs.

In section 2 we develop the theoretical model of networks and derive its implications. Section 3 describes the empirical framework. In section 4 we provide a brief description of the data. The main empirical results are presented in section 5. Section 6 presents the structural models, while section 7 concludes.

2 THEORETICAL MODEL

The model draws heavily on Montgomery (1992). The result regarding the expected wage conditional on job-finding method can be found there in general form by translating variables appropriately.

Consider a recent immigrant looking for jobs. He faces two sources of job offers, the network source, and the formal/non-network source. Suppose that with probability p_n he receives an offer through the network, and with probability p_f he receives a job offer through the formal source. Also assume that he can receive at most one offer from each source. In each case, the wage offer is drawn from a common distribution function $F(w)$. Thus, we assume that the distribution of wage offers is independent of the source (relaxed later), and immigrants are homogenous.

There is a single time period. The immigrant worker accepts an offer if he receives at least one offer greater than his reservation wage. If he receives two offers, he chooses the higher offer, provided that it is higher than his reservation wage. For the moment, there is no loss in generality in treating wage offers below the reservation wage as non-offers, and defining $F(w)$ over the range of wages greater than the reservation wage, and p_n and p_f as the probabilities of receiving an offer greater than this cutoff.

2.1 Network Strength

With probability $(1 - p_n) * (1 - p_f)$, the worker receives no offers, with probability $(1 - p_n) * p_f$, he receives only a formal offer, with probability $p_n * (1 - p_f)$, he receives only a network offer, and with probability $p_n * p_f$, he receives both types of offers.

The expected wage conditional on receiving at least one job offer is

$$E(w) = \frac{(p_f + p_n - 2p_f p_n)E(w|N = 1) + p_f p_n E(w|N = 2)}{(p_f + p_n - p_f p_n)} \tag{1}$$

where N is the number of job offers received. It is straightforward to show that $E(w)$ is increasing in p_n (and p_f), provided the wage offer distribution is nondegenerate.

What about wages conditional on the method through which the job was found? The expected wage conditional on accepting a job through the network is,

$$E(w|n) = (1 - p_f)E(w|N = 1) + p_f E(w|N = 2) \tag{2}$$

which is independent of p_n.

The expected wage conditional on accepting a job through the formal source is,

$$E(w|f) = (1 - p_n)E(w|N = 1) + p_n E(w|N = 2) \tag{3}$$

which is increasing in p_n.

It follows immediately that the gap between the expected wage conditional on finding a job through the network and the expected wage conditional on finding a job through the formal mechanism is decreasing in network strength, as defined by p_n.

Finally, consider the level of the difference in earnings conditional on job finding method:

$$E(w|n) - E(w|f) = (p_f - p_n)(E(w|N = 2) - E(w|N = 1)). \tag{4}$$

The sign of this difference depends on the relative probability of finding a job through the formal method and the network. If the network is less likely to produce a job than the formal source ($p_n < p_f$), then *workers who find jobs through networks will have higher wages, on average, than those who find them through formal methods.* Note that, if the network is more likely to produce a job than the formal source, then those finding jobs through networks will have lower wages than do those finding them through formal methods. This is the key insight in Montgomery (1992).

The intuition for this counterintuitive result is straightforward. Consider an extreme example. Suppose the network almost never generates a job offer (p_n is close to 0) while formal search almost always yields an offer (p_f is close to 1). In this scenario, almost all recent immigrants receive an offer from the formal source while very few receive an offer from the network. Therefore, those who accepted network jobs almost definitely chose between *two* offers, while those who accepted formal jobs, almost all chose the *one* offer they had. Therefore, those in network jobs have higher wages compared to those in formal jobs, even though the network is weaker than the formal source.

The result on the sign of the difference in earnings conditional on job finding method is sensitive to the assumption that the distribution of wages is the same for the two job sources. When the network distribution stochastically dominates, or is a mean preserving spread of the formal distribution, it is more likely that expected wage conditional on finding the job through the network is higher than expected wage conditional on finding the job through the formal source irrespective of the relation between p_n and p_f.[2]

2.2 Differing Wage Distributions

Let the distribution of wages received through the network conditional on receiving an offer be $F_n(w)$, and similarly, the distribution of the formal wages conditional on receiving an offer be $F_f(w)$.

The expected wage conditional on receiving a job offer is,

$$E(w) = \frac{p_f(1-p_n)E(w_f) + p_n(1-p_f)E(w_n) + p_f p_n E(w|N=2)}{(p_f + p_n - p_f p_n)}. \tag{5}$$

As before, improvements in the strength of either the formal or network domains will raise expected wages.

[2] However, in a slightly different context, Montgomery (1992) provides examples to show that even when both sources are equally strong and the network distribution stochastically dominates or is a mean preserving spread of the formal distribution, expected wage conditional on network job could be lower than expected wage conditional on formal job. Thus, the sign of the difference in expected wage conditional on job finding method can go in either direction when the network and formal distributions are different.

The expected wage conditional on accepting a job through a network is,

$$E(w|n) = \frac{(1-p_f)E(w_n) + p_f E(w_n|w_n > w_f)P(w_n > w_f)}{1 - p_f + p_f P(w_n > w_f)} \tag{6}$$

which, as in the simpler model, is independent of p_n.

The expected wage conditional on accepting a job through formal means is,

$$E(w|f) = \frac{(1-p_n)E(w_f) + p_n E(w_f|w_f > w_n)P(w_f > w_n)}{1 - p_n + p_n P(w_f > w_n)} \tag{7}$$

which, as before, is increasing in p_n. Therefore, the gap between the expected wage conditional on finding a job through the network and the expected wage conditional on finding a job through formal mechanisms, is decreasing in network strength, as defined by p_n.

2.3 Effects at Percentiles of Observed Wage Distribution

It is important to note that while economists often focus on differences in means, our argument applies equally to percentiles of the observed wage distributions**.**

The cdf of the observed network wage distribution is independent of p_n because conditional on receiving a network offer, the probability that the offer will be better than the formal offer is independent of the probability of a network offer.

In contrast, as described in the following theorem, the cdf of the observed formal sector wage distribution $F_f(w|f)$ is decreasing in p_n.

Theorem 1 *Let $F_f(w)$ be continuous on $[a, b]$ with $F_f(a) = 0$ and $F_f(b) = 1$. Then $d(F_f(w|f))/dp_n < 0$ for $a < w < b$ and $d(F_f(w|f))/dp_n = 0$ for $w = a, b$.*

Proof.

$$\begin{aligned} F_f(w|f) &= \frac{\int_a^w (1-p_n+p_n F_n) f_f dx}{\int_a^b (1-p_n+p_n F_n) f_f dx} \\ &= \frac{F_f - \int_a^w p_n(1-F_n) f_f dx}{1 - \int_a^b p_n(1-F_n) f_f dx}. \end{aligned}$$

$$\frac{d}{dp_n}\left(\frac{F_f - \int_a^w p_n(1-F_n) f_f dx}{1 - \int_a^b p_n(1-F_n) f_f dx}\right) = \frac{\int_a^w F_n f_f dx - F_f \int_a^b F_n f_f dx}{\left(1 - \int_a^b p_n(1-F_n) f_f dx\right)^2} \tag{8}$$

Inspection of the numerator proves the second part of the theorem.

Now,

$$\frac{\int_a^w F_n f_f dx}{\int_a^b F_n f_f dx} = \frac{\int_a^w F_n f_f dx}{\int_a^w F_n f_f dx + \int_w^b F_n f_f dx}$$

Therefore from the first mean value theorem of integration, there exists weights ω_1.and ω_2, such that

$$\frac{\int_a^w F_n f_f dx}{\int_a^b F_n f_f dx} = \frac{\omega_1 F_f(w)}{\omega_1 F_f(w) + \omega_2(1 - F_f(w))}$$

where

$$0 < \omega_1 < F_n(w) < \omega_2 < 1$$

for $a < w < b$, from which it follows that

$$\frac{\int_{-\infty}^w F_n f_f dx}{\int_{-\infty}^\infty F_n f_f dx} < F_f.$$

and that

$$\frac{\int_{-\infty}^w F_n f_f dx - F_f \int_{-\infty}^\infty F_n f_f dx}{\left(1 - \int_{-\infty}^\infty p_n(1 - F_n) f_f dx\right)^2} < 0.$$

■

The theorem establishes that the percentile associated with any wage, except the highest and the lowest, in the observed formal sector wage distribution (wage conditional on formal sector employment) is reduced when the probability of a network offer increases. The intuition is straightforward. Any network offer beats a formal offer if it is greater than the formal offer and has no effect on the acceptance of formal offers above it. Most network offers will beat a very low formal offer but will not beat a very high formal offer. On average, therefore a network offer reduces the probability that the worker accepts a low formal offer by more than it reduces the probability that the workers accepts a high formal offer. The distribution of accepted offers shifts to the right. From this intuition, it should be clear that the effect on the percentiles does not depend on the continuity of F_n although the math will be slightly messier if the distribution has mass points.

Since the percentile of the conditional distribution associated with each wage goes down, the wage associated with each percentile goes up. Moreover, since the observed formal sector wage distribution is independent of the probability of receiving a formal offer, it is an immediate corollary that the difference between any percentile (except the very highest and very lowest) of the observed formal sector wage distribution and the same percentile of the observed network wage distribution will increase when the probability of a network wage offer increases.

The effect of network strength on the difference in the observed network and formal wages at each percentile suggests a potentially more powerful test of the model. Since there is no effect at the highest and lowest percentiles, there must be some percentile at which the effect is larger than the mean effect, and this difference might be sufficient to outweigh the reduced efficiency of estimating a percentile rather than the mean.

We anticipate that the effect of network strength on the formal/network wage differential will be largest somewhere below the median. We offer two arguments for this intuition.

First, suppose that the (log) formal wage offer distribution is $N(0,1)$. And assume that the network sector distribution is degenerate at some value, say 0, for concreteness. If the network is very weak so that the probability of a network offer is close to 0, then the observed wage distribution in the formal sector is very close to the formal offer distribution, and its 2.5 percentile is very close to -1.96, its median to 0, its 97.5 percentile to 1.96. Now assume that when the network is strong, the probability of a network offer is close to 1. Then the observed formal distribution is very close to a truncated standard normal, and its 2.5 percentile is very close to .031, its median to .67 and its 97.5 percentile to 2.24. It will be apparent that the effect on the observed formal sector wage distribution will generally be larger at lower centiles (although the effect goes to zero as the centile being considered goes to zero).[3] This need not hold for all distributions, but it will frequently be true, so that we expect larger effects at lower centiles.

Second, numerical simulations show that if f_n and f_f have the same normal distribution and if the probability of a network offer is .25 when the network is weak and .5 when the network is strong, values roughly consistent with our data, then the biggest wage change occurs at roughly the 25th percentile.

2.4 Summary of Predictions

In sum, in the simple case of one offer from each source, the model has the following predictions regarding network strength, (the probability of finding the job through the network):

1. The expected wage is increasing in network strength.
2. If the distribution of wage offers in the formal and network sectors are identical, the expected wage conditional on finding a job through the network is higher than the expected wage conditional on finding a job through formal methods if and only if $p_n < p_f$.

[3] The effect at the 2.5, 50 and 97.5 percentiles is 1.991, 0.67 and 0.28 respectively.

3. The difference between the expected wage conditional on finding a job through the network and the expected wage conditional on finding a job through formal methods, is decreasing in network strength.

4. At any given centile, the difference between the wage conditional on finding a job through the network and the corresponding centile of the wage distribution conditional on finding a job through formal methods is decreasing in network strength. The magnitude of this effect varies across centiles.

In this simple case, implications (1), (3) and (4) hold even when the network distribution is not the same as the distribution of formal job offers.[4]

3 Empirical Framework

Our empirical strategy begins by verifying that the proposed measures of network strength and network size do predict network use. Having validated these measures in this way, we then ask whether or not these measures predict wages. Finally, we examine whether network strength interacts with network use as predicted by the theoretical model.

We propose that network size be measured by the log of the share of working age population in the locality who are from the new immigrant's country of birth. Network strength be proxied by, whether or not the new immigrant had at least one close friend or relative in the locality when he first arrived. Since we will focus on labor market outcomes six months after arrival, this is a plausible proxy.

Validating the Network Measures: To validate these proxies, we use multinomial logit to examine how they are related to the individual's employment outcome (unemployed, in job found through network [network job], in job found through formal means [formal job]). Thus we have

$$\ln \frac{P^l_{ijk}}{P^U_{ijk}} = \delta^l_1 F_{ijk} + \delta^l_2 S_{jk} + \beta^l X_{ijk} + \omega^{1l}_j + \lambda^{1l}_k \tag{9}$$

[4] Montgomery derives implication (2) above for the case where a worker has M formal contacts each with probability p_f of producing an offer and N network contacts each with probability p_n of producing an offer, and each offer is drawn from the same wage distribution. Unfortunately, we cannot prove implications (3) and (4) for the case of multiple contacts. The problem is that with multiple contacts, an increase in the probability of getting an offer through each member of the network not only increases the overall probability of getting an offer, but also changes the distribution of the highest wage offer received through the network. Therefore, the proofs in subsections 2.2 and 2.3 above do not apply. However, the fact that when networks are stronger than formal sources, conditional wages are higher for formal jobs than network jobs, and that, when networks are weaker than formal sources, the opposite is true, suggests that predictions (3) and (4) will apply generally, albeit not universally.

for $l = NJ, FJ$ i.e. network job and formal job respectively; U stands for unemployment, F is a dummy variable for having at least one friend/relative in the locality upon arrival, S is the network size variable, X is a set of additional controls that are likely to influence use of social networks in finding a job, and ω and λ are country of birth dummies and locality dummies. The subscripts i, j and k, refer to individual i, country of birth j, and locality k.

We view the multinomial logit as a simple way to assess the relation between our network measures and job finding.[5] Better networks should increase the probability of network employment relative to formal employment ($\delta^{NJ} > \delta^{FJ}$) and they should increase the probability of network employment relative to unemployment ($\delta^{NJ} > 0$). In a purely *ad hoc* sense, it is not obvious whether a better network should have a larger or smaller proportional effect on unemployment than on formal sector employment. Therefore, when validating our measures, we make no prediction about the sign of δ^{FJ}. However, it is straightforward to show that in the context of our model, network strength should have a larger effect on unemployment than on formal sector employment and that therefore $\delta_1^{NJ} > \delta_1^{FJ} > 0$.[6]

Network Structure: Once the network measures are validated, we regress various wage measures on the validated network measures, country of birth dummies, locality dummies, and other controls that are likely to affect labor market outcomes. This approach is given by,

$$\ln w_{ijk} = \alpha_1 F_{ijk} + \alpha_2 S_{jk} + \gamma X_{ijk} + \omega_j^2 + \lambda_k^2 + \upsilon_{ijk}. \tag{10}$$

where the variables are defined as in equation (9) above. The model predicts that workers with better networks will have higher wages.

In part, we carry out this estimation to permit comparison of results with the existing literature. The evidence on the effects of ethnic enclaves is mixed.[7] As discussed in the introduction, enclaves

[5] In principle, we could derive the multinomial logit specification as given by (9) by specifying the value of a job outcome . If we normalize the value of unemployment to be zero, assume that the unobservable is extreme value and that draws from the network and formal sectors are independent, this specification justifies the use of multinomial logit. However, this justification implies that network strength shifts the distribution of network (and possibly formal) offers. In contrast, our theory assumes that network strength only increases the arrival rate of network offers.

[6] The probability of unemployment is $(1-p_n)(1-p_f)$ while the probability of formal employment is $p_f(1-\lambda p_n)$ where λ is the probability that the the worker chooses the network job if he has both network and formal offers. So the log ratio of formal employment to unemployment is $\log(p_f) - \log(1-p_f) + \log(1-\lambda p_n) - \log(1-p_n)$ and its derivative with respect to p_n is $(1-\lambda) / [(1-p_n)(1-\lambda p_n) > 0$. Therefore our model implies that $\delta_1^{FJ} > 0$. However, in a more general model, network strength might also increase λ, in which case the derivative would be unsigned.

[7] Munshi (2003) and Edin, Fredriksson, and Aslund (2003) find that networks improve the labor market outcomes of Mexican immigrants to the United States and of refugees in Sweden. For three major cities in Canada, Hou and Picot (2003) find only a weak effect of exposure to own-group neighbours on immigrants' employment probability and annual earnings. In contrast, Lazear (1999) argues that immigrant enclaves reduce the rate at which immigrants

may benefit or hurt immigrants. Also, many studies have not fully addressed the omitted variables bias issues that typically plague the estimation of network effects. Since we control for both location and country of birth, our network measures are unlikely to be correlated with location characteristics or group characteristics. Thus, α_2 is a *difference in differences estimator*. It is identified through variations in network size between, for example, Indians in Toronto and Indians in Ottawa, and comparing this difference with variations in network size between the Chinese across the same two cities. Bias of this form would arise only if Indian immigrants were more likely than Chinese immigrants to locate in Toronto, because the industrial structure of the city benefits Indians more than it does the Chinese immigrants. This cannot be ruled out completely, but our greater concern is that where the immigrant locates may tell us something about the immigrant: a Russian immigrant who locates where there are few established Russian immigrants, may be quite different from one who seeks out a Russian immigrant enclave. Nevertheless, this approach is useful because it is straightforward and addresses directly the effect of network structure on immigrant labor market outcomes.

Testing the Role of Networks: Our primary focus is to test whether the wage difference between those who found their jobs using networks and those who found them using formal mechanisms is related to the validated measure of network strength. The equation above is augmented with an interaction between whether the individual found his first job through the social network, and whether he had at least one friend or relative in the locality when he first arrived.

$$\ln w_{ijk} = \theta_1 F_{ijk} + \theta_2 S_{jk} + \theta_3 NJ_{ijk} + \theta_4 (NJ_{ijk} * F_{ijk}) + \pi X_{ijk} + \omega_j^3 + \lambda_k^3 + \zeta_{ijk} \qquad (11)$$

where NJ is a dummy for whether the individual found his first job through the social network.

As explained in the theory section, when the immigrants' networks are stronger than formal channels (more likely to happen when they have a friend/relative close by), the effect of finding a job through the network should be negative, while when they are weaker, it should be positive. However, this result could also reflect other factors, such as differences between the wage distributions in the two sectors. The testable implication of the model is that θ_4 is negative, that is the difference between the expected wage conditional on finding a job through the network and the expected wage conditional on finding a job through the formal mechanism (in other words the network premium), is decreasing in network strength.

An important advantage of this approach over the standard method of regressing labor market

---

learn the host-country language, while there is a large literature indicating that knowledge of host-country language raises wages (see Berman, Lang, and Siniver, 2003, and the references therein).

outcomes on measures of network structure, is that it mitigates problems associated with omitted variables. New immigrants are likely to share the unmeasured characteristics of established immigrants. Thus, the outcomes of new and established immigrants are likely to be positively correlated even if there is no causal relation. If a locality is especially conducive to good labor market outcomes for a particular immigrant group, this is likely to generate a positive correlation between the outcomes of new and established immigrants. In contrast, there is little reason to expect that, in areas where an immigrant group has a particular advantage, or where the group is particularly favorably selected, the bias will depend on the method through which the new immigrant finds a job. What about the interaction term? One would have to think of an unobservable that would affect the network-formal wage differential differently for those with and without strong social ties. We do not see an obvious reason to expect such a relation. The most plausible case might be that we expect immigrants to be more likely to move to a locality where they have a strong tie if the network is particularly valuable in that area. If so, the coefficient on the interaction term will be biased in a positive direction, and we will be less likely to observe the negative coefficient our theoretical model predicts. The main motivation for including the interaction term comes from our theoretical model and the prediction that it have a significant negative sign in equilibrium. Thus, our main focus is to test the sign on θ_4.

As discussed in the theory section, the prediction that θ_4 is negative applies not only to the conditional mean (OLS estimation) but also to all conditional quantiles except the highest and lowest. However, we would expect θ_4 to be most negative at quantiles below the median. We therefore also use quantile regression to estimate equation (11) at the 25th, 50th and 75th percentiles.

4 Data and Descriptive Analysis

4.1 Data

We use a 20% 2001 Census of Canada sample to calculate characteristics of immigrant populations by country of origin and location within Canada. The sample is restricted to the working age population (those between 24 and 64 years old). According to the 2001 Census, immigrants constitute 18 percent of the Canadian population, and 21 percent of the labor force. They come from more than 200 source countries. In order to ensure that there are sufficient observations in each cell to calculate network measures with reasonable precision, source countries with fewer than 500 immigrants in the census sample are dropped. The geographic unit used to characterize local networks is the Census Metropolitan Area, CMA, or the Census Agglomeration, CA.[8] Using the 2001 Census, we calculate the share of working age population in each CMA/CA from each source country, which is our measure of network size. Measures of the wage distribution of the employed immigrant population from a particular country, residing in a particular CMA/CA, are also obtained from the Census.

Our remaining data come from the Longitudinal Survey of Immigrants to Canada (LSIC), collected by Statistics Canada, and Citizenship and Immigration Canada. The LSIC sample consists of immigrants who arrived in Canada between October 1, 2000 and September 30, 2001 and were 15 years or older. We refer to this population as *recent immigrants*. The LSIC is a longitudinal survey with three waves: six months, two years and four years after arrival in Canada. It provides data on the recent immigrant's characteristics, such as, sex, age, education, languages spoken, country of birth and geographic location in Canada, and also his job history, which includes labor force status, weekly wage if employed, etc. We restrict the sample to those respondents between 24 and 64 years old who are in the labor force. In the first wave, 74 percent of this age group were in the labor force. We drop individuals whose first job was arranged before they migrated to Canada, or who were either self employed, or in a family business,[9] and those who change CMA/CA.[10] We limit the analysis to observations from CMA/CAs and from source countries with at least ten immigrants in the LSIC sample, and those from source countries retained in the census sample.[11] The final wave one LSIC sample consists of 5103 recent immigrants, from 51 different source countries and residing in 16 different CMA/CAs across Canada.[12]

[8] A census metropolitan area (CMA), or a census agglomeration (CA), is formed by one or more adjacent municipalities centered on a large urban area, known as the urban core. The census population count of the urban core is at least 10,000 to form a census agglomeration, and at least 100,000 to form a census metropolitan area. To be included in the CMA or CA, other adjacent municipalities must have a high degree of integration with the central urban area, as measured by commuting flows derived from census place of work data. In the 2001 Census, there are 27 CMAs and 113 CAs across Canada.

[9] In the first wave of the LSIC, 8.5 percent of the recent immigrants in the labor force are in pre-arranged jobs. 2.6 percent report being self-employed and 0.6 percent report being involved in a family business, when asked about their first jobs.

[10] A mover is dropped because it is not clear whether one should consider his network to be the relevant group in the new location or in the old one. For example it could be the case that a person's network in his previous location helped him find a job in his current location. In this case, it would be incorrect to characterise the relevant group at the current/interview location as his social network. In order to simplify matters and present clean results movers are excluded from the sample. At the time of the first wave, 7.7 percent of recent immigrants in the labour force had changed CMA/CA.

[11] We lose only 8% of the remaining LSIC sample due to these location and source country restrictions.

[12] Of the 16 CMA/CAs only one (Guelph) is a CA. Therefore, henceforth they will be referred to as CMAs.

4.2 Descriptive Analysis

Although we drop a large number of countries and localities, the largest sending countries and the largest receiving localities account for the vast majority of immigrants. According to the 2001 Census, the top twenty source countries account for 68% of the working age immigrant population. In recent years there has been a change in the composition of source countries, with an increase in share of immigrants from Asia, and a decline in the share from Europe. According to the LSIC, China followed by India are the top two source countries for recent immigrants, constituting 22 percent and 15 percent of the working age recent immigrant population. Recent immigrants are settling in areas where there is an already large concentration of both native and immigrant population. Toronto, Montreal, Vancouver, Ottawa, Calgary have 52 percent of Canada's working-age population, 75 percent of its working-age immigrants and 83 percent of its recent working-age immigrants.

Table 1 shows the means for the variables in the first wave of the LSIC. By the time of the first wave, 69 percent of the sample had held a first job. Of these 42 percent (29 percent of the entire sample) reported that they found this job through a friend or relative, which we define as a network job. The remainder used other methods, such as, contacting the employer directly, responding to newspaper advertisements, employment agencies, the internet, referral from another employer or a union. We refer to these as formal jobs.

We proxy network strength by a binary variable. It takes the value 1, if the individual reports that he already had at least one relative or friend in the city where he resides, when he first arrived in Canada. While 89 percent had at least one relative or friend in Canada on arrival, 82 percent had one in the city where they reside. By this measure, most recent immigrants have strong networks.

We capture the size of the recent immigrant's network by the natural logarithm of the share of working age CMA population from his country of birth. Note that, since CMA dummies are included, this is isomorphic to including the natural logarithm of the number of immigrants from his country of birth. We cannot distinguish between the roles of absolute and relative size.

Two things must be noted at this point. First, finding the first job through the social network does not necessarily imply the presence of a relative or friend in the same city of residence on arrival. Immigrants may have found their first job through a friend made after migrating to Canada, a relative or friend elsewhere, or through a compatriot who is not a relative or close friend. Thus, having a network job does not imply having a strong network. Second, to the extent that job search is complex, the dichotomous measure of the "use of the social network" and the theoretical concept it wishes to capture are not perfectly correlated. For example, if a friend tells me that there are job openings where he works, and I apply and get a job there, do I report that I found the job through a friend, or that I applied directly to the employer? Thus, admittedly, the measure of use of network (i.e. having a network job) is imperfect.

5 Results

5.1 Validating the Network Measures

Table 2 shows the relation between network size, network strength and various background characteristics, on the one hand, and whether the recent immigrant found a network job or a formal job or did not find a job. The table gives multinomial logit coefficients. Although the coefficients are not derivatives, they are directly related to the probability of the employment outcome relative to unemployment. Finding the effect of a variable on network employment relative to formal employment requires comparing the two coefficients.

The first two columns give the determinants of first job type after six months. Our key variable, network strength is strongly related to the probability of getting a network offer within the first six months. The derivative at the mean probabilities is about 11 percentage points.[13]

As noted earlier, for the network strength measure to be valid, its coefficient in the network job equation should be greater than its coefficient in the formal job equation. This is confirmed. Our model implies that the coefficient in the formal job equation should be positive. While positive values lies within the confidence interval, the point estimates are negative.

In contrast, the network size measure increases the value of both formal and network offers, and the effect is statistically significant in both cases although the coefficient in the network equation is noticeably larger. As a result a larger network is associated with a high probability of having a network job but has little effect on the probability of having a formal job.

This pattern holds for first jobs reported within the first two and first four years. Network strength continues to have a large positive effect on the value of a network job while having no effect on the value of a formal job. The effect of network size on first jobs declines with time in Canada. By the third wave of the LSIC the effect of network size is no longer statistically significant at conventional levels although its magnitude has dropped only slightly.

It is not surprising that the results for first jobs are consistent across the three pairs of columns.

[13]For any individual, the derivative of the probability of network sector employment with respect to a variable x is $P_n\left(b_n P_u + (b_n - b_f)P_f\right)$ where b_i is the coefficient on the variable in the network/formal employment equations and P_i is the probability of being in sector i. In the text we describe derivatives calculated at the mean probabilities.

Most of the immigrants found a first job within six months. The fact that even among those who have been in Canada for four years, networks are important for finding a first job does not tell us whether networks continue to be important once immigrants have established themselves in the labor market. To address this question, we turn to Table 3 which provides estimates of the effects of networks on current jobs.

At six months the results are similar to those obtained for first jobs, presumably because many of the immigrants' current jobs are also their first jobs. The magnitude and significance of both the strength and size terms are similar to those in Table 2. Strong ties increase the value of network jobs but not formal jobs while network size has positive effects on both, albeit a stronger effect on network jobs than on formal jobs.

The results for the current job at the interview at the end of the second year after arrival in Canada are similar except that the effect of network size on the formal sector loses significance even though its magnitude is greater than in the estimates at six months.

In contrast the importance of whether the immigrant had a close friend or relative in the locality on arrival seems to diminish by the time he or she has been in Canada for four years. We cannot tell whether this is because the importance of network strength declines once an immigrant has been in Canada for a few years or because the strength of networks on arrival is a poor proxy for strength of network after four years. To check if smaller sample size is driving the results for current/recent jobs at the four year interview, we carry out a robustness check by restricting the sample at the earlier interviews (six months and two years) to only those present at the four year interview (results not shown here). Network strength continues to predict finding current/recent jobs using networks at the six month and two year interviews, even when the sample is restricted as stated. Thus, it seems that the measures of network structure used here diminish in importance over time either because networks are less important or because the nature of networks changes with time spent in the host country. Interestingly, in the second and third waves, the coefficient on network strength in the formal job equation shifts to the positive value predicted by our model although it remains statistically insignificant at conventional levels.

We view the results in Tables 2 and 3 as strong confirmation of the validity of our network strength measure.

These tables also reveal that, recent immigrants with a high school or lower level of education are more likely to find their first jobs using networks, while those with a university degree are less likely to use networks to do so (reference category being immigrants with a college certificate). This conforms to the notion that the low skilled workers use networks much more than high skilled workers do.[14] Surprisingly the tables also show that university graduates are less likely to be employed. This could be because of more competition among highly educated immigrants in the labor market.[15] Also, immigrants fluent in English are less likely to use networks in finding their first jobs.

5.2 Network Structure and Wages

Table 4 presents estimates of the relation between network measures and wages on first jobs found by the first interview. In addition to our other network measures, we control for the median wage received by established network members. This captures possible differences in the types of jobs available to the immigrant. It generally has the wrong sign and is never positive and statistically significant.

The first column presents the results of an otherwise standard wage equation augmented with the network measures. Since it controls for both CMA and for country of origin, it can be interpreted as a difference-in-differences estimator. It reveals a small negative and statistically insignificant effect of network size on earnings. This is consistent with larger networks being associated with a faster arrival rate of offers (given the effect on employment found in tables 2 and 3), but with either little or no effect on the reservation wage, or increasing the arrival rate at the lower end of the wage distribution.

The presence of at least one relative or friend in the locality on arrival (network strength) enters with the right sign, and has a nontrivial point estimate (over 5 percent), but does not reach statistical significance at conventional levels.

This finding of an insignificant effect of network size on earnings is consistent with previous findings for Canada. Hou and Picot (2003) examine the association between living in a visible minority enclave and immigrants' labor market outcomes in Canada's three largest cities. They also find little association between exposure to own-group neighbors and immigrants' annual earnings. Since controls for language are included here, these results are not directly comparable to Lazear

---

[14] Departure from this conventional notion is examined in Saxenian (1999). The paper examines the extent to which the skilled Chinese and Indian immigrants are organizing ethnic networks in California's Silicon Valley to support the often risky process of starting new technology businesses. The author notes that Silicon Valley's new immigrant entrepreneurs are more highly skilled than their counterparts in traditional industries, but like those counterparts they have created a rich fabric of professional and associational activities that facilitate immigrant job search, information exchange, access to capital and managerial know-how and the creation of shared ethnic identities.

[15] As table 1 shows, 70 percent of recent immigrants in our sample hold a university degree. Thus, most Canadian immigrants today are highly educated. They may be competing for jobs amongst themselves, especially as finding good job matches takes longer for highly skilled workers.

(1999) in the United States, but they do not confirm an adverse effect of ethnic enclaves. As discussed before, there are some endogeneity concerns that are not addressed by the approach in table 4. Therefore, it is not clear whether these results differ from Edin, Fredriksson, and Aslund (2003) because of differences in the nature of immigrants to Sweden and Canada, or because of differences in approach.

The remaining columns present the results of quantile estimates. Because there is no simple cluster correction for quantile estimates, a clustered bootstrap method is used to calculate the standard errors. This approach is problematic because, since clusters rather than observations are resampled, the number of observations can vary across replications, and will typically be smaller than the number in the actual sample. This should, therefore, produce upwards biased standard errors for the coefficients of variables for which cluster has little or no explanatory power. Therefore, the result of the cluster bootstrap is only reported for network size, since it is measured at the level of the cluster, and ordinary standard errors are reported for the remaining variables which are measured at the level of the individual.

Columns (2) - (4) present quantile regression results for wages, conditional on having a wage. The results are similar to those obtained using OLS, although there is some evidence of an even larger effect of network strength on initial wages at the 25th percentile, and this effect is statistically significant at 0.1 level. A strong network results in a 9.4 percent increase in wages in first jobs at the 25th percentile of an individual's accepted wage offer distribution.[16] There is also evidence of an effect on wages at the median and 75th percentile of the worker's wage distribution. These effects are similar in magnitude to those obtained using OLS although only the effect on the median is statistically significant and then only at the .1 level.

Since networks affect the probability of being employed, looking only at the wages of the employed could give a misleading picture of their effect. For example, if networks are particularly effective in providing low-productivity workers with jobs, they might appear to lower wages even if they raise the offer distribution for all workers. To address this concern, columns (5) and (6) also include the unemployed immigrants (who are in the labor force) by assigning them very low wages. It is not possible to estimate the model for the 25th percentile because a high proportion of new immigrants are unemployed. There is continued weak evidence of a positive effect of network strength on wages, in that the coefficients are numerically, but not statistically, significant. In table 5, we extend the analysis to include individuals who found their first job after six months up to four

[16]Because there are controls for education and other factors that affect wages, the quantile regression results need to be interpreted carefully. The coefficient gives the effect of, for example, network strength on the 25th percentile of an individual's conditional acceptable wage offer distribution. This is quite different from saying the effect on the individual at the 25th percentile of the unconditional wage offer distribution.

years. The results are broadly similar. The effect of network strength is positive and statistically significant at conventional levels for the 25th percentile of wages in first jobs (conditional on having a job) obtained within two and within four years

5.3 Augmented Wage Model

As discussed above, the network structure approach is limited by the concern that individuals who locate near friends or family, or in areas where there are an unusually large number of established immigrants from their country of birth, differ in unmeasured ways from those who do not.[17] Therefore, we turn to testing the prediction of the theoretical model.

Table 6 shows the results for the wage equation augmented with method of finding the first job (whether or not it was found using the network) and the interaction between this variable and network strength.[18] It is easy to tell stories in which the use of a network is positively or negatively correlated with unobserved worker characteristics, but less easy to explain why this correlation should be noticeably different for those with and without strong networks. Therefore, our focus is on the interaction term.

Panel A shows the results for first jobs found within six months of arrival. Column (1) shows that among those who did not have a friend or relative nearby when they first arrived at their locality in Canada, finding a job through a network is associated with a trivial and statistically insignificant 0.2 percent lower wage. In contrast, among those with a friend or relative, the wage penalty associated with finding a job through a network is about 3.2 percentage points bigger although again not statistically significant.

However, as discussed earlier the theory section predicts that the coefficient on the interaction term will be negative at all quantiles, and there is strong reason to expect it to be more negative when we examine lower quantiles. Therefore in columns (2)-(4) we present quantile regressions for the 25th, 50th and 75th quantiles. As in earlier tables, only the standard error for network size is adjusted for clustering. As we should anticipate column (2), pertaining to wages at the 25th percentile (of the conditional wage distribution), conforms closely with the predictions of the

[17] Endogeneity of locating close to a friend or relative conditional on having one, does not seem to be a serious problem. Fully 92 percent of the recent immigrants who had at least one relative or friend in Canada chose to live close to their ties. One cannot completely rule out the possibility that individuals who choose to move to Canada without having a friend or relative present are different from other immigrants. However, we control for the presence of a relative or friend in Canada but not in the same locality as the recent immigrant. Such immigrants do not differ from those who do not have a tie in Canada.

[18] When interpreting the coefficients on network strength, network job and their interaction, it should be noted that the omitted group is that of immigrants in formal jobs and without strong social ties.

theoretical model. For wages at the 50th and 75th percentile, the interaction term is negative, as predicted, but is small and statistically insignificant.

Also at the 25th percentile, in the absence of strong networks (strong social ties), those finding their first jobs through networks have weekly wages that are 17.2 percent higher than those doing so using formal means. In other words, at the lower end of the wage distribution, among recent immigrants who do not have a strong social tie in their locality upon arrival, those who are in network jobs earn a wage that is 17.2 percent higher compared to those who obtained jobs through formal channels. The model predicts this difference if networks are less likely to relay job offers than are formal methods (more likely to happen in the absence of a relative or friend close by). However, as mentioned earlier, this finding is also consistent with other explanations such as unobserved differences between network users and non-users, or differences between the network and formal wage offer distributions. More importantly, as predicted by the model, the coefficient on the interaction term is negative and statistically significant. For those who have a strong social tie, those who are in network jobs earn a wage that in only 0.8 percent higher (17.2-16.4) compared to those who obtained jobs through formal channels. In other words, for those who have at least one relative or friend in their locality, the network premium (network-formal wage differential) is 16.4 percent lower (network premium for those with strong networks is 0.8 percent) compared to the network premium for immigrants without strong networks (network premium for those without strong networks is 17.2 percent). Therefore the network premium is decreasing in network strength as predicted by the theoretical model.

In panels B and C of table 6, the estimates are replicated by adding in workers who found their first job after six months but within two and four years respectively. As seen here, the interaction term continues to be negative and significant albeit only at the .1 level in panel C.[19]

6 Structural Model

The results in the previous section are broadly consistent with the formal theoretical model presented earlier. However, it is not clear that the magnitudes of the effects can be reconciled with reasonable restrictions on the formal and network sector wage distributions. In this section, we ask whether a model with a single network wage offer distribution and a single formal sector wage offer distribution can fit the data.

[19] We also examined the network effect on wages by restricting the analysis to various sub-samples, according to gender and education and to immigrants who belonged to countries where English was *not* the lingua franca. The results were not demonstrably different and no interesting patterns across sub-groups were observed, possibly because standard errors get too large when sub-samples are used.

6.1 The Model

We assume that the immigrant receives offers with probability

p_f from the formal sector

p_w from the network sector if his network is weak

p_s from the network sector if his network is strong.

Let w_f, and w_n, where $n \in \{w, s\}$, represent (log) wage offer from the formal sector, a weak network (w_n) and a strong network (w_s). Then

$$w_f \tilde{} N(0,1)$$
$$w_n \tilde{} N(\mu_n, \sigma_n^2).$$

The imigrant receives zero, one or two offers. If he receives no offers, he is unemployed. If he receives one offer, he accepts that offer and is employed at the offered wage. If he receives two offers, he chooses the higher offer and is employed at that higher wage.

6.2 Estimation

The model has five free parameters: p_f, p_w, p_s, μ_n, and σ_n^2. We choose these parameters to match eight empirical parameters. The first four are employment probabilities: the probability of being employed in the formal sector when the network is strong, the probability of being employed in the formal sector when the network is weak, the probability of being employed in the network sector when the network is strong and the probability of being employed in the network sector when the network is weak. The next four are wage differentials between the network and formal sectors depending on whether the network is strong or weak. In each case we match both the mean differential and the 25th percentile differential.

The covariances among the empirical parameters are not derived easily. Therefore we treat them as zero and minimize

$$L = \Sigma \frac{(\widehat{m}_i - m_i)^2}{\sigma_i^2}$$

where $\widehat{m}$ is the value of the moment implied by our model and m is the empirical estimate of the moment and σ^2 is the variance of that empirical estimate.

6.3 Results

The parameter estimates suggest that the distribution of offers from the network and formal sectors are not all that different. The estimated mean (μ_n) of the log wage offer distribution in the network sector is -0.034 Log wages are also somewhat less dispersed in the network sector. The estimated standard deviation (σ_n) is 0.88, compared with the 1.0 assumed for the formal sector. The result is that we estimate that the average wage offer in the network sector is about 86% of the average wage in the formal sector.

We also estimate that immigrants are somewhat more likely to receive an offer from the formal sector (48.3%) than they are from the network sector when they have a strong network (44.1%), but they are much less likely to receive a network offer when their network is weak (23.9%).

Are these estimates consistent with observed distribution of wages among the employed recent immigrants to Canada? Table 7 shows the relation between key parameters estimated from the data and those derived from the model. The first eight parameters are the ones we tried to match. We can see that the model does a very good job of matching the distribution of employment, matching them to two significant digits. It also does a reasonable job of matching the average wage gap between the network and formal sectors, both when individuals have a weak network and when they have a strong network. The simulated **(**predicted**)** wage gap between the formal sector and a network job for an individual with a strong network is 5.2% compared with an actual value of 3.4%. The difference is about a half standard error. Similarly the model predicts that individuals with a weak network will earn 3.4% more if employed in a network job than in a formal job. The data reveal a gap of -0.2**%**, a difference equal to about three-quarters of the standard error on the estimate from the data.

The really large gap in the data occurs for individuals with a weak network who nevertheless end up in a network job. In this case, from the data we estimate a 17.2% wage gap while the model predicts a 12.1% gap, again a difference of about three-quarters of the standard error on the estimate from the data. The difference between the predicted (1.7%) and actual (0.7%) gaps when the network is strong is smaller. The value of our objective function at convergence is 2.18. Since this objective function does not take account of the covariances among the moments we are attempting to match, it does not provide a formal test of the model. It does, however, summarize the result that we are able to match the empirical results quite closely. We are less interested in a formal test of the model. Since it is quite implausible that both the network and formal sector wage distributions are normally distributed and that the wage distributions in the network sector are identical regardless of network strength, any failure to reject the model would merely reveal the lack of power of the test. The objective of the structural model is to show that the magnitudes we observe, particularly at different quantiles of the wage distribution are very broadly consistent with this model.

For this reason the last four rows of Table 7 show the wage gaps at the median and 75th percentiles, empirical values that we did not try to match directly. The results show that we match the wage gaps between network and formal jobs quite accurately when the network is weak. The predicted median gap is 3.7% compared with an actual gap of 2.3%; the corresponding figures at the 75th percentile are -5.0% and -3.1%. We also examine the difference between the network/formal gap between immigrants with strong and weak networks. At the median the predicted difference is -9.5% compared with an actual difference of -2.8%. This discrepancy is large but well within the confidence interval for the empirical estimate. At the 75th percentile, the predicted and actual differences are -7.7% and -2.6%. Although also large, this discrepancy is less than one standard error.

In sum, a model in which the main effect of a stronger network is to increase the rate at which offers arrive from the network rather than to draw on a different wage distribution is broadly consistent with the data.

7 Summary and Conclusions

We developed a theoretical model of the importance of networks for recent immigrants seeking jobs and derived the equilibrium results for immigrants with strong and weak networks. The model shows that among immigrants with networks that are stronger than formal channels, those who are in network jobs have lower wages than those in formal jobs. It also predicts that the network-formal wage differential is decreasing in network strength and that this effect should be most pronounced at lower quantiles. We tested these implications on a nationally representative sample of recent immigrants into Canada. The empirical strategy to carry out comparative statics, augments the difference in differences framework with an interaction term between network strength and finding a network job, and focuses on the coefficient of this interaction term. This strategy has an important advantage over the standard method of regressing labor market outcomes on measures of network influence, in that it mitigates problems associated with omitted variables. The model's prediction is not rejected in any of the specifications, and is strongly supported for wages at the lower end of an individual's acceptable wage distribution. This suggests that the presence of at least one strong social tie in the recent immigrant's immediate neighborhood upon his arrival increases the number of offers he receives from the network.

To test whether the magnitudes observed in the data were consistent with the theory, we estimated a simple structural model in which network and formal offers are drawn from two log-normal distributions. The model was able to produce parameters well within the confidence intervals of the empirical estimates. This suggests that a model in which the primary role of strong social ties is to increase the arrival rate of offers from the network distribution is consistent with the data. Our results also suggest that the offer distributions in the formal and network sectors differ only modestly so that Montgomery's (1992) model can be applied.

It is often argued that immigrants tend to cluster together because the presence of established immigrants facilitates assimilation of new arrivals, both in the labor market and in the social environment of the host country. We find that social networks help in the economic assimilation of recent immigrants. Our findings suggest that immigrants with strong social ties in their localities enjoy a faster arrival rate of jobs. We do not address other issues related to immigrant dispersion, including the longer term labor market effects of immigrant enclaves.

References

BERTRAND, M., E. F. P. LUTTMER, AND S. MULLAINATHAN (2000): "Network Effects and Welfare Cultures," *The Quarterly Journal of Economics*, 115(3), 1019–1055.

BORJAS, G. J. (1992): "Ethnic Capital and Intergenerational Mobility," *The Quarterly Journal of Economics*, 107(1), 123–150.

——— (1995): "Ethnicity, Neighborhoods, and Human-Capital Externalities," *The American Economic Review*, 85(3), 365–390.

——— (2000): "Ethnic Enclaves and Assimilation," *Swedish Economic Policy Review*, 7, 89–122.

——— (2002): "Homeownership in the Immigrant Population," *Journal of Urban Economics*, 52, 448–476.

CASE, A., AND L. KATZ (1991): "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youth," *National Bureau of Economic Research, Working Paper No. 3705*.

EDIN, P.-A., P. FREDRIKSSON, AND O. ASLUND (2003): "Ethnic Enclaves and the Economic Success of Immigrants - Evidence from a Natural Experiment," *The Quarterly Journal of Economics*, 118(1), 329–357.

GALARNEAU, D., AND R. MORISSETTE (2004): "Immigrants: Settling for less?," *Perspectives*, pp. 5–16.

GLAESER, E. L., B. SACERDOTE, AND J. A. SCHEINKMAN (1996): "Crime and Social Interactions," *The Quarterly Journal of Economics*, 111(2), 507–548.

GRANOVETTER, M. (1983): "The Strength of Weak Ties: A Network Theory Revisited," *Sociological Theory*, 1, 201–233.

HOU, F., AND G. PICOT (2003): "Visible-Minority Neighbourhood Enclaves and Labour Market Outcomes of Immigrants," *Canadian Immigration Policy for the 21st Century*.

LAZEAR, E. P. (1999): "Culture and Language," *The Journal of Political Economy*, 107(6), S95–S126.

LIN, N., J. C. VAUGHN, AND W. M. ENSEL (1981): "Social Resources and Occupational Status Attainment," *Social Forces*, 59(4), 1163–1181.

MONTGOMERY, J. D. (1991): “Social Networks and Labor-Market Outcomes: Toward an Economic Analysis,” *The American Economic Review*, 81(5), 1408–1418.

——— (1992): “Job Search and Network Composition: Implications of the Strength-Of-Weak-Ties Hypothesis,” *American Sociological Review*, 57(5), 586–596.

MUNSHI, K. (2003): “Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market,” *The Quarterly Journal of Economics*, 118(2), 549–597.

OZDEN, C. (2006): “Educated Migrants: Is there Brain Waste?,” *International Migration Remittances and The Brain Drain*.

PICOT, G., AND A. SWEETMAN (2005): “The Deteriorating Economic Welfare of Immigrants and Possible Causes: Update 2005,” *Analytical Studies Branch Research Paper Series, Statistics Canada*.

PORTES, A. (1995): *The Economic Sociology of Immigration. Essays on Networks, Ethnicity and Entrepreneurship*. Russell Sage Foundation.

SAXENIAN, A. (1999): “Silicon valley’s New Immigrant Entrepreneurs,” *Public Policy Institute of California*.

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Observations
Dependent Variables			
Network Job	0.29	0.45	5103
Have a Job	0.69	0.46	5103
Weekly Wage	405	267.3	3390
Key Explanatory Variables			
Strong Network	0.82	0.38	5103
Network Size (not in logs)	0.022	0.018	5103
Additional Explanatory Variables			
Relative/Friend in Canada	0.89	0.31	5103
Female	0.41	0.49	5103
Age	35.5	8.1	5103
Married	0.83	0.37	5103
Number of children	0.94	0.94	5103
Speaks English Well	0.64	0.64	5103
Speaks French Well	0.12	0.12	5103
Lived in Canada Before	0.05	0.23	5103
Sponsored by Family Member	0.23	0.42	5103
Principal Applicant	0.72	0.45	5103
Economic Visa	0.77		5055
Family Visa	0.20		5055
Refugee Visa	0.03		5055
High School or Less	0.14		5075
Some University (no degree)	0.15		5075
University Degree or More	0.71		5075
Prior Occupation			
Manager	0.02		5058
Professional	0.39		5058
Paraprofessional	0.14		5058
Clerical	0.03		5058
Laborer	0.002		5058
Student/New Worker/None	0.42		5058

Table 2: Network Influence on Finding First Job by Time of Survey

	First Wave (6 months)		Second Wave (2 years)		Third Wave (4 years)	
	Formal	Network	Formal	Network	Formal	Network
Network Size	0.180**	0.320***	0.198**	0.313***	0.147	0.246**
	[0.074]	[0.090]	[0.079]	[0.095]	[0.090]	[0.101]
Network Strength	-0.073	0.473**	-0.019	0.512***	-0.000	0.442***
	[0.107]	[0.187]	[0.125]	[0.182]	[0.119]	[0.170]
Relative/Friend not close by	-0.069	-0.060	-0.043	0.022	-0.082	-0.071
	[0.168]	[0.270]	[0.211]	[0.272]	[0.192]	[0.228]
Female	-0.247**	-0.497***	-0.120	-0.321**	-0.070	-0.279*
	[0.107]	[0.122]	[0.115]	[0.156]	[0.124]	[0.151]
Age	-0.039***	-0.029***	-0.040***	-0.035***	-0.043***	-0.038***
	[0.005]	[0.009]	[0.007]	[0.008]	[0.007]	[0.007]
Married	0.028	-0.165	0.053	-0.124	0.175	-0.006
	[0.115]	[0.128]	[0.127]	[0.147]	[0.153]	[0.163]
Kids	-0.063	0.064	-0.036	0.073	-0.025	0.081
	[0.044]	[0.039]	[0.059]	[0.056]	[0.066]	[0.060]
High school or less[1]	0.073	0.463***	-0.106	0.422**	-0.136	0.393*
	[0.184]	[0.160]	[0.235]	[0.214]	[0.238]	[0.233]
University degree[1]	-0.122	-0.281***	-0.297**	-0.454***	-0.159	-0.316**
	[0.089]	[0.095]	[0.138]	[0.142]	[0.156]	[0.153]
Speak English well	0.077	-0.299***	0.144	-0.242**	0.042	-0.350***
	[0.095]	[0.107]	[0.127]	[0.122]	[0.105]	[0.108]
Speak French well	0.298	-0.169	0.284	-0.012	0.110	0.003
	[0.209]	[0.244]	[0.215]	[0.257]	[0.231]	[0.241]
Lived in Canada Before	0.298	0.068	0.255	0.035	0.166	0.042
	[0.187]	[0.233]	[0.235]	[0.270]	[0.219]	[0.252]
Principal Applicant	0.225	0.022	0.124	-0.034	-0.000	-0.171
	[0.158]	[0.151]	[0.193]	[0.205]	[0.196]	[0.204]
Sponsored by family	-0.110	0.043	0.231	0.466*	0.365	0.576*
	[0.192]	[0.233]	[0.238]	[0.277]	[0.301]	[0.329]
Family Visa[2]	-0.563**	-0.021	-0.831***	-0.514	-0.966***	-0.634
	[0.275]	[0.322]	[0.307]	[0.344]	[0.363]	[0.401]
Refugee Visa[2]	-1.007**	-0.130	-0.908**	-0.076	-0.903**	-0.066
	[0.392]	[0.325]	[0.434]	[0.387]	[0.435]	[0.397]
Observations	4982		4982		4982	
Log Pseudolikelihood[3]	-4845.83		-4533.34		-4452.69	
Clusters	383.00		383.00		383.00	

Robust standard errors corrected for group effects within CMA/CA-country cells in brackets;
* significant at 10%; ** significant at 5%; *** significant at 1%;
1 The omitted education category is Some University education but no degree
2 The omitted visa category is the Economic visa class
3 "Pseudolikelihood" because with clustered data we do not have independent observations
Also controls for occupation before coming to Canada, country of birth and CMA.

Table 3: Network Influence in Finding Current/Most Recent Job

	First Wave (6 months)		Second Wave (2 years)		Third Wave (4 years)	
	Formal	Network	Formal	Network	Formal	Network
Network Size	0.196***	0.297***	0.244	0.315*	0.309	0.392
	[0.074]	[0.090]	[0.153]	[0.173]	[0.305]	[0.320]
Network Strength	-0.072	0.468**	0.006	0.574**	0.046	0.208
	[0.101]	[0.192]	[0.246]	[0.253]	[0.312]	[0.286]
Relative/Friend not close by	-0.086	-0.044	-0.046	0.225	0.243	0.157
	[0.166]	[0.265]	[0.380]	[0.403]	[0.558]	[0.555]
Female	-0.282**	-0.439***	-0.373**	-0.476***	-0.773**	-0.961***
	[0.110]	[0.110]	[0.169]	[0.172]	[0.321]	[0.356]
Age	-0.039***	-0.027***	-0.031**	-0.031**	-0.036*	-0.040**
	[0.005]	[0.009]	[0.013]	[0.014]	[0.020]	[0.020]
Married	0.027	-0.11	0.010	-0.101	0.713*	0.520
	[0.119]	[0.111]	[0.220]	[0.246]	[0.398]	[0.436]
Kids	-0.056	0.052	-0.144	0.002	-0.263	-0.070
	[0.044]	[0.039]	[0.097]	[0.089]	[0.163]	[0.147]
High school or less[1]	0.104	0.445***	0.056	0.337	-0.274	0.150
	[0.183]	[0.159]	[0.358]	[0.373]	[0.481]	[0.568]
University degree[1]	-0.126	-0.284***	-0.366*	-0.772***	-0.249	-0.511
	[0.089]	[0.092]	[0.194]	[0.193]	[0.365]	[0.392]
Speak English well	0.097	-0.312***	0.516***	-0.115	0.252	-0.266
	[0.097]	[0.109]	[0.166]	[0.171]	[0.281]	[0.263]
Speak French well	0.274	-0.145	0.223	0.011	-0.249	-0.331
	[0.216]	[0.243]	[0.295]	[0.295]	[0.444]	[0.473]
Lived in Canada Before	0.300*	0.204	0.190	0.237	0.002	0.320
	[0.182]	[0.229]	[0.393]	[0.407]	[0.612]	[0.666]
Principal Applicant	0.2	0.104	0.309	0.217	-0.194	-0.516
	[0.166]	[0.148]	[0.241]	[0.258]	[0.312]	[0.333]
Sponsored by family	-0.127	0.076	-0.340	-0.004	-0.322	0.389
	[0.185]	[0.237]	[0.253]	[0.261]	[0.772]	[0.716]
Family Visa[2]	-0.548*	-0.122	-0.564	-0.416	-0.346	-0.498
	[0.288]	[0.323]	[0.354]	[0.390]	[0.945]	[0.881]
Refugee Visa[2]	-0.999**	-0.213	-1.010*	-0.076	0.632	1.015
	[0.397]	[0.318]	[0.544]	[0.533]	[0.809]	[0.748]
Observations	5003		3524		2585	
Log Pseudolikelihood[3]	-4881.94		-2845.99		-1850.08	
Clusters	385		312.00		246.00	

Robust standard errors corrected for group effects within CMA/CA-country cells in brackets;
* significant at 10%; ** significant at 5%; *** significant at 1%;
1 The omitted education category is Some University education but no degree
2 The omitted visa category is the Economic visa class
3 "Pseudolikelihood" because with clustered data we do not have independent observations
Also controls for occupation before coming to Canada, country of birth and CMA.

Table 4: (Log) Wages for First Jobs within Six Months: Difference in Differences

	OLS	Quantile regressions				
		Without Unemployed			With Unemployed	
		0.25	0.5	0.75	0.5	0.75
	(1)	(2)	(3)	(4)	(5)	(6)
Network Size	-0.036	-0.044	-0.020	-0.059	0.181	-0.046
	[0.029]	[0.058]	[0.038]	[0.041]	[0.205]	[0.050]
Network Strength	0.056	0.094*	0.047*	0.044	0.070	0.063
	[0.038]	[0.048]	[0.024]	[0.038]	[0.049]	[0.039]
Median wage in network	0.003	-0.009	-0.007*	-0.009	-0.004	-0.008
(000s per annum)	[0.007]	[0.007]	[0.004]	[0.006]	[0.007]	[0.006]
Relative/Friend not close by	0.075	0.024	0.028	0.114**	0.028	0.055
	[0.063]	[0.073]	[0.035]	[0.054]	[0.071]	[0.056]
Female	-0.214***	-0.246***	-0.117***	-0.164***	-0.231***	-0.182***
	[0.030]	[0.034]	[0.036]	[0.027]	[0.036]	[0.029]
Age	-0.003	-0.002	-0.002**	-0.004**	-0.010***	-0.007***
	[0.002]	[0.002]	[0.001]	[0.002]	[0.002]	[0.002]
Married	-0.017	-0.008	-0.028	-0.036	-0.020	-0.040
	[0.045]	[0.045]	[0.027]	[0.035]	[0.047]	[0.037]
Kids	-0.026**	-0.016	-0.001	-0.014	-0.008	-0.012
	[0.013]	[0.018]	[0.010]	[0.014]	[0.019]	[0.015]
High school or less[2]	0.143***	0.072	0.049*	0.080*	0.185***	0.100**
	[0.034]	[0.054]	[0.027]	[0.041]	[0.058]	[0.045]
University degree[2]	0.029	-0.014	-0.005	-0.007	0.009	-0.033
	[0.035]	[0.043]	[0.021]	[0.033]	[0.044]	[0.035]
Speak English well	0.068**	-0.035	0.040**	0.083***	-0.034	0.070**
	[0.034]	[0.037]	[0.018]	[0.028]	[0.037]	[0.029]
Speak French well	-0.162	-0.044	-0.179***	-0.106	0.010	-0.152**
	[0.098]	[0.090]	[0.046]	[0.070]	[0.090]	[0.070]
Lived in Canada Before	0.085	0.039	0.053	0.144**	0.096	0.131**
	[0.081]	[0.072]	[0.037]	[0.059]	[0.074]	[0.061]
Principal Applicant	0.055	0.060	-0.001	-0.025	0.131**	0.007
	[0.049]	[0.058]	[0.028]	[0.043]	[0.057]	[0.044]
Sponsored by family	0.070	0.045	-0.016	-0.075	-0.019	-0.045
	[0.055]	[0.079]	[0.039]	[0.058]	[0.078]	[0.056]
Family Visa[3]	-0.200***	-0.154	-0.018	0.009	-0.182*	-0.028
	[0.069]	[0.097]	[0.047]	[0.071]	[0.097]	[0.071]
Refugee Visa[3]	-0.251**	-0.181	-0.057	-0.122	-0.255**	-0.154
	[0.098]	[0.122]	[0.063]	[0.097]	[0.121]	[0.099]
Occupation before dummies	Yes	Yes	Yes	Yes	Yes	Yes
Country of Birth dummies	Yes	Yes	Yes	Yes	Yes	Yes
CMA dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3207	3207	3207	3207	4659	4659

For OLS, standard errors clustered by CMA/CA-country cells in brackets; number of clusters is 268
* significant at 10%; ** significant at 5%; *** significant at 1%
1. Very low wages were assigned to those without wages
2 The omitted education category is Some University education but no degree
3 The omitted visa category is the Economic visa class

Table 5: (Log) Wages, Network Size and Strength: Difference in Differences

	OLS	Quantile regressions		
		0.25	0.5	0.75
	(1)	(2)	(3)	(4)
	A. For first jobs within two years			
Network Size	-0.043	-0.024	-0.045	-0.059
	[0.028]	[0.057]	[0.035]	[0.042]
Network Strength	0.022	0.065*	0.033	0.046
	[0.036]	[0.036]	[0.029]	[0.036]
Observations	3816	3816	3816	3816
	B. For first jobs within four years			
Network Size	-0.036	-0.028	-0.048	-0.061
	[0.028]	[0.059]	[0.035]	[0.041]
Network Strength	0.022	0.073**	0.039*	0.039
	[0.038]	[0.035]	[0.023]	[0.039]
Observations	3946	3946	3946	3946

For other explanatory variables see table 4.
Standard errors clustered by CMA/CA-country
294 clusters in panel A and 298 in panel B.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: (Log) Wages, Method of Job Finding and Network Strength

	OLS	Quantile Regressions		
		0.25	0.5	0.75
	(1)	(2)	(3)	(4)
A. For first jobs within six months				
Network size	-0.036	-0.041	-0.021	-0.054
	[0.029]	[0.057]	[0.038]	[0.041]
Network Strength	0.068	0.128***	0.064**	0.052
	[0.045]	[0.049]	[0.031]	[0.044]
Network Job	-0.002	0.172**	0.023	-0.031
	[0.047]	[0.067]	[0.044]	[0.061]
Network Job*Network Strength	-0.032	-0.164**	-0.028	-0.026
	[0.053]	[0.072]	[0.047]	[0.065]
Observations	3207	3207	3207	3207
B. For first jobs within two years				
Network Size	-0.042	-0.024	-0.044	-0.061
	[0.028]	[0.057]	[0.035]	[0.042]
Network Strength	0.032	0.101**	0.044	0.054
	[0.045]	[0.045]	[0.037]	[0.043]
Network Job	-0.043	0.133**	-0.013	-0.069
	[0.049]	[0.063]	[0.051]	[0.057]
Network Job*Network Strength	-0.014	-0.142**	-0.022	-0.029
	[0.059]	[0.067]	[0.055]	[0.062]
Observations	3816	3816	3816	3816
C. For first jobs within four years				
Network Size	-0.036	-0.028	-0.049	-0.050
	[0.028]	[0.059]	[0.035]	[0.040]
Network Strength	0.031	0.117**	0.056*	0.055
	[0.047]	[0.048]	[0.030]	[0.045]
Network Job	-0.044	0.115*	-0.009	-0.074
	[0.050]	[0.065]	[0.042]	[0.061]
Network Job*Network Strength	-0.013	-0.129*	-0.030	-0.025
	[0.061]	[0.070]	[0.045]	[0.066]
Observations	3946	3946	3946	3946

For other explanatory variables see table 4.
For OLS, standard errors clustered by CMA/CA-country
268 in panel A, 294 in panel B, 298 in panel C
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7: Empirical and Simulated Parameters

	Empirical Parameters	Simulated
Parameters Match in Estimation		
Formal Employment	0.376	0.378
Strong Network	[0.007]	
Formal Employment	0.437	0.426
Weak Network	[0.016]	
Network Employment	0.332	0.332
Strong Network	[0.007]	
Network Employment	0.182	0.180
Weak Network	[0.012]	
Network/Formal Wage Gap	-0.034	-0.052
Strong Network	[0.033]	
Network/Formal Wage Gap	-0.002	0.034
Weak Network	[0.047]	
25th Percentile Wage Gap	0.007	0.017
Strong Network	[0.028]	
25th Percentile Wage Gap	0.172	0.121
Weak Network	[0.067]	
Parameters Not Used in Estimation		
Median gap in		
Network Formal Wage Differential	-0.028	-0.095
Strong Network - Weak Network	(0.047)	
Median Wage Gap	0.023	0.037
Weak Network	[0.044]	
75th Percentile gap in		
Network Formal Wage Differential	-0.026	-0.077
Strong Network - Weak Network	(0.065)	
75th Percentile Wage Gap	-0.031	-0.050
Weak Network	[0.061]	