I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10926

Network Effects on Labor Contracts of Internal Migrants in China: A Spatial Autoregressive Model

Badi H. Baltagi
Ying Deng
Xiangjun Ma

JULY 2017

DISCUSSION PAPER SERIES

IZA DP No. 10926

Network Effects on Labor Contracts of Internal Migrants in China: A Spatial Autoregressive Model

Badi H. Baltagi
Syracuse University and IZA

Ying Deng
University of International Business and Economics

Xiangjun Ma
University of International Business and Economics

JULY 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Network Effects on Labor Contracts of Internal Migrants in China: A Spatial Autoregressive Model*

This paper studies the fact that 37 percent of the internal migrants in China do not sign a labor contract with their employers, as revealed in a nationwide survey. These contract-free jobs pay lower hourly wages, require longer weekly work hours, and provide less insurance or on-the-job training than regular jobs with contracts. We find that the co-villager networks play an important role in a migrant's decision on whether to accept such insecure and irregular jobs. By employing a comprehensive nationwide survey in 2011 in the spatial autoregressive logit model, we show that the common behavior of not signing contracts in the co-villager network increases the probability that a migrant accepts a contract-free job. We provide three possible explanations on how networks influence migrants' contract decisions: job referral mechanism, limited information on contract benefits, and the "mini labor union" formed among co-villagers, which substitutes for a formal contract. In the sub-sample analysis, we also find that the effects are larger for migrants whose jobs were introduced by their co-villagers, male migrants, migrants with rural Hukou, short-term migrants, and less educated migrants. The heterogeneous effects for migrants of different employer types, industries, and home provinces provide policy implications.

JEL Classification: O15, R12, J41

Keywords: contract, co-villager network, spatial autoregressive logit model, internal migrants

Corresponding author:
Badi H. Baltagi
Department of Economics and Center for Policy Research
Syracuse University
426 Eggers Hall
Syracuse, NY, 13244-1020
USA

E-mail: bbaltagi@maxwell.syr.edu

* This project is supported by National Natural Science Foundation of China [71503045], [71401038], and Beijing Social Science Funding [15JGC158]. All errors are the responsibility of the authors.

1 Introduction

Why would people accept a job without the protection of a labor contract? This is a decision made by around 37 percent of internal migrants in China based on a nationwide survey in 2011.[1] Generally speaking, an "irregular" job without a contract is not attractive due to lower wages, longer work hours, less insurance, and so forth. According to our data, these "irregular" jobs pay more than 20 percent lower hourly wages than the "regular" jobs that offer contracts, whereas they demand almost 10 percent longer weekly work hours than "regular" jobs at the mean. Among the migrants with "regular" jobs, 35 percent have work-related insurance and 48 percent receive on-the-job training, while the figures are 4 percent and 18 percent for migrants with "irregular" jobs (see Table 1B). There are around 271 million internal migrants in China in 2011, which accounts for one-third of total urban population.[2] Therefore, the welfare and return of migrant workers are important for China to move from an economic pattern heavily dependent on exports to one in which the demand of the country's own internal market plays a larger role in economic growth (Becker and Elfstrom, 2010).

In addition to the average low education level of internal migrants, which leads to less competitive human capital and prevents them from obtaining better "regular" jobs, another important factor that affects their acceptance of "irregular" jobs is the co-villager network.[3] This paper shows that if most co-villagers (from the same home province residing in the same host city), who form the social network of an individual migrant, work for jobs without labor contracts, it is more likely that the individual migrant will also accept a job without a contract. Previous literature has documented the role that social networks play in migrants' work decisions, welfare participation, and migrating destinations (Calvo-Armengol and Jackson, 2004; Montgomery, 1991; Carrington, Detragiache, and Vishwanath , 1996). Studies on "Guanxi" (Chinese expression of social networks) in China have also received much attention (Zhang and Li, 2003; Lovett, Simmons and Kali, 1999). Chen, Jin and Yue (2010) find that one's migration decision is influenced by her co-villagers because co-villagers help each other in moving cost and job search at the destination.

From the perspective of migrant employees, the job referral mechanism, limited information on contract benefits, and the "mini labor union" formed among co-villagers are the main reasons we emphasize for why internal migrants' decisions of accepting "irregular" jobs are affected by their co-villager networks.

First of all, many migrants found their jobs through friends or relatives in their co-villager networks. Calvo-Armengol and Jackson (2004) build an explicit network model on the transmission of job opening information, which leads to positive correlation between the employment status of agents who are directly or indirectly connected in the network of relatives, friends or acquaintances. Therefore, a contagion effect or network externality of employment status is observed.

[1] Although the Contract Law enacted on January 1, 2008 in China requires employers to provide written labor contracts for any position they offer, the law is not strictly implemented, especially during the global financial crisis right after the law was first launched. Becker and Elfstrom (2010) cited from the Beijing Federation of Trade Unions that only 32.8 percent of migrant workers had signed contracts by September 2008. They point out that the law specified the contract coverage but did not effectively monitor the process of signing these contracts. In practice, costs of employers certainly rise if they obey the requests on wage and hour provisions as well as injury and social insurance. To avoid it, employers prefer to hire people who do not require contracts.

[2] Source: "Report on China's Migrant Population Development 2012" edited by the National Population and Family Planning Commission of China.

[3] In our sample, 69 percent of the migrants have an education level of middle school or below (see Table 1A).

Ioannides and Loury (2004) conduct a comprehensive review on the effects of networks on job search.[4] Recently, Cingano and Rosolia (2012) show that the increase in the employment rate of the network of former fellow workers reduces unemployment duration of individuals. However, the above research does not specify the co-villager network in China. In our data, more than 42.2 percent jobs of migrants are obtained through co-villager networks, while the corresponding figure in the literature is around 50%, summarized by Patacchini and Zenou (2012). They also show that members of a particular ethnic group concentrate in specific jobs and when new employment opportunities become available at their workplace, they pass this information along to contacts of the same race and ethnic background.

In spite of the easier job-searching process, the quality of jobs introduced by contacts is not necessarily high. Elliott (2001) finds that the use of informal contacts results in significantly lower wages for less-educated workers. Battisti, Peri and Romiti (2016) also find that immigrants initially located in places with larger co-ethnic networks are more likely to be employed at first, but earn lower wages in the long run. Loury (2006) claims that jobs obtained through contacts are better than those found through formal methods only when the contact is a prior-generation relative. For our analysis, if most co-villagers of an internal migrant work in enterprises that do not offer contracts, it is more likely that this migrant will be introduced to similar "irregular" jobs. On the contrary, if she searches for jobs independently, the probability of entering an enterprise offering "regular" jobs with a contract could be higher.

Secondly, migrants living and socializing in a large network acquire most information and knowledge from their co-villagers with similar background. As a result of information blocking, they do not fully or clearly understand the benefits (such as insurance and training) and legal protection bundling in a labor contract. Hence they imitate the behavior of their co-villagers. Co-villagers' higher propensity of accepting contract-free jobs increases the propensity for this migrant to accept the same type of jobs. In Bertrand, Luttmer and Mullainathan (2000)'s work on the usage of welfare programs, they conclude that social networks affect an individual's behavior and preference through two important channels: information and norms. With a randomized experiment, Duflo and Saez (2003) analyze the positive effects of information and social interactions on employees' decisions to enroll in a retirement plan.

Lastly, collective bargaining from co-villager networks could substitute for labor contracts with respect to employment protection. To understand this channel, we first ask "if a person could choose between jobs with and without contracts, what can she benefit from the latter?" The answer is flexibility. Aguirregabiria and Alonso-Borrego (2014) find that after restrictions on the more flexible fixed-term or temporary contracts were removed in the Spanish labor market, permanent workers were replaced by temporary workers.[5] Booth, Francesconi and Frank (2002) point out that temporary contracts in Europe are an important component of labor market flexibility, especially in countries characterized by high levels of employment protection. Measuring job flexibility by

[4]For example, there is a positive correlation between getting assistance from a fraternity or sorority contacts and obtaining prestigious high-paying jobs (Marmaros and Sacerdote, 2002). Bayer, Ross and Topa (2008) find that a network, defined by residing in the same block, increases the probability of working together by over 33 percent. Other literature on this issue includes Hellerstein, McInerney, and Neumark (2008), Falcón (2007), Munshi (2003), Frijters, Shields and Price (2005), Battu, Seaman and Zenou (2011), to name a few.

[5]In fact, fixed-term or temporary contracts are both considered as "regular" jobs with contracts in our sample. Nevertheless, the contract-free "irregular" jobs in our case are even more flexible but offer less protection for the employees.

the years a migrant has worked for her current employer, we find in our sample that the average flexibility for jobs with and without contracts are 3.35 and 3.03 years, which implies that the latter is more flexible than the former. In other words, migrant workers prefer to access flexibility at the cost of lower wage and less employment security. This is particularly true for short-term job seekers, such as those who still have to do farm work at certain seasons of the year or many women who return home occasionally to take care of their children or old parents. Note that most internal migrants in China are from the rural area with rural Hukou (85 percent in our sample, see Table 1A).[6]

On the other hand, collective bargaining makes it possible for a migrant not to worry much about job security if she works with many co-villagers in the same workplace. In fact, a large proportion of employees could threaten the employer with a strike or to quit all together, which may prevent the institution from functioning. In this sense, the co-villager network can be regarded as a "mini labor union" which raises the bargaining power of the migrants. The insecurity of a contract-free job is thus reduced and the willingness of accepting it increases.

Besides the above 3 channels, from the perspective of the labor demand side, employers might also make use of the co-villager network to strategically recruit migrant workers. Hensvik and Nordstrom Skans (2016) verify that firms use referrals of productive employees in order to attract workers with better qualities. In our case, employers may utilize co-village networks of the provinces from which more migrants have accepted contract-free jobs to recruit workers who are willing to accept the same type of irregular jobs. This possibility is also in line with the job-searching channel from the labor supply side.

This paper applies a spatial autoregressive (SAR) model to examine how migrant workers' acceptance of contract-free jobs is affected by their co-villager networks. The SAR model, also known as the Cliff and Ord (1973) model, has been widely used in the empirical analysis of social networks.[7] Lin (2010) applies the SAR model to identify the peer effects in students' academic achievements. Baltagi and Yen (2014) allow spatial correlation among neighboring hospitals and estimate the effects of externalities generated by competition and knowledge spillovers on hospital treatment rates. In our case, the dependent variable of the SAR model is a binary variable indicating whether a migrant accepts a job without a contract.

We employ the Dynamic Monitoring Survey of Internal Migrants in 2011 from the National Population and Family Planning Commission of China. This is a nationwide survey with 128,000 individual migrants residing in 326 host cities in all 31 provinces of China, which has detailed demographic and work information of each migrant. In particular, we know whether the migrant has a contract as well as her home province and host city. In the benchmark SAR linear probability and logit regressions, we find that the probability of a migrant's acceptance of a job without a contract increases with the co-villager migrants' acceptance of contract-free jobs. The effects become larger for migrants whose jobs were introduced by their co-villagers, for male migrants, migrants with rural Hukou, short-term migrants, and less educated migrants. We also find that the network effects are only significant for some industries (manufacturing, construction, and lodging and catering) and some employer types (e.g. private and self-employed business). Policy implication drawn from these facts is that more attention on the implementation of the labor contract law should be

[6]Hukou is a social identity status in China, categorized into two types: rural or urban.

[7]LeSage and Pace (2009) have a detailed introduction of the SAR model. Its applications are widespread from studying spatial interactions at the macro level (countries, cities, etc.) to investigating social interactions at the micro level (households, individuals, etc.).

paid in these areas. The network effects also vary across home provinces. In the provinces of large network effects and a relatively high percentage of contract-free jobs (such as Chongqing, Sichuan, Gansu, and Heilongjiang), there is dynamic concentration of job-contract types among migrants from the same home province. If migrants from some certain home provinces always find "bad" low-paid jobs that do not offer contracts, their consumption in the host city as well as the remittance they send back to their home province will be lower than migrants from other home provinces, which will lead to income inequality across regions over time.

To our knowledge, this is the first paper that uses a nationwide data set to study the interplay of co-villager networks and the decisions of accepting a contract-free irregular job offer for internal migrants in China. The remainder of the paper is organized as follows. Section 2 presents the SAR linear probability model and the SAR logit model. Section 3 describes data sources and reports summary statistics. Section 4 discusses the empirical results and analyzes the network effects in different subsamples. Section 5 concludes.

2 Methodology

In the introduction, we emphasized three channels from the labor supply side that explain why migrant workers are more likely to accept irregular jobs without contracts if more people in their co-villager networks accept such jobs: the job search process, information blocking on contract benefit, and collective bargaining of the mini labor union formed by co-villager networks at work. To test the hypothesis that when the probability of co-villager migrants' acceptance of jobs without contracts increases, the probability increases for the targeting migrant to accept a job without a contract, we assign an outcome dummy of job types equal to 1 if an individual has a contract-free job, and 0 otherwise in the spatial autoregressive (SAR) model.[8]

The spatial weight matrix is specified based on the co-villager network, which is defined as people flowing out of the same home province and residing in the same host city. Only people within the network will make an impact on each other, as they are more likely to share information and provide job protection for each other. It is assumed that each migrant is equally affected by all others in the co-villager network. Hence, in the spatial autoregressive (SAR) model, we assume that each migrants in the same co-villager network has an identical influential weight which is the inverse of the number of co-villagers in the network. For example, if there are 40 migrants from Zhejiang province working in Shanghai, then for each migrant, her Zhejiang-Shanghai network consists of 39 people and each person is assigned a weight of 1/39 so as to satisfy the row-normalized condition of the spatial weight matrix. As a result, the spatial lagged dependent variable is the weighted average of the contracting decision of these 39 migrants in the network with the same weight of 1/39. Other migrants in the sample do not belong to the Zhejiang-Shanghai network and assigned zero weights. Using this spatial weight matrix, we estimate a spatial autoregressive (SAR) linear probability model as well as a SAR logit model.

We caution that people from the same home province in the same host city in our survey data do not necessarily know each other.[9] Nevertheless, since the data is collected in a random way,

[8]We do not consider the migrant workers who are unemployed, because there are only 1.4 percent migrants who report unemployed in our sample. Moreover, most migrants in our sample are from the rural area as mentioned. If they do not have a job in the host city, they will return home to do farm work.

[9]There are other applications of the spatial methodology in the literature, where people in the same group do not

the co-villager network defined in the data is representative of the true network of the population, which means the size of the networks in the sample is also proportional to the size of the true co-villager networks. In other words, we assume that if there are more people from Sichuan province working in Beijing in the true population, for example, we would see more Sichuan people in Beijing in our data. Further, we also assume that if there are more people in the defined network, there will be more interconnection occurring among them in the network. We are aware that the second assumption is a little strong as people from different regions may have different co-villager culture and there may not be sufficient job-related interconnection among people in some co-villager networks. As a result, we run subsample regressions for people whose jobs are introduced by co-villagers in that these people are indeed using their true co-villager interconnection to find jobs.

2.1 The SAR Linear Probability Model

We start with a standard SAR linear probability model, which is given by

$$Contract_{i_s} = \lambda \sum_{j_s=1}^{N_s} w_{i_s,j_s} \cdot Contract_{j_s} + X_{i_s}\beta + e_{i_s}, \quad i_s = 1,2,...,N_s; s = 1,2,...,n_0. \tag{1}$$

The index i_s represents migrant i in network s that consists of N_s migrants. There are a total number of n migrants that can be divided into n_0 networks. $Contract_{i_s}$ is a binary variable of migrant i's choice on accepting an irregular job without a contract ($Contract_{i_s} = 1$) versus a regular job offering a contract ($Contract_{i_s} = 0$). The predicted value can be interpreted as the probability that a migrant accepting a contract-free job. The spatial lagged dependent variable $\sum_{j_s=1}^{N_s} w_{i_s,j_s} \cdot Contract_{j_s}$ is a weighted average of the decisions on whether or not to accept a contract-free job of all other migrants in the same network s as migrant i. As a typical element of an $n \times n$ predetermined spatial weight matrix W, w_{i_s,j_r} relates to migrant i in network s and migrant j in network r, respectively. We assign a positive weight to w_{i_s,j_r} if migrants i and j are in the same network; otherwise, zero. More specifically, we define $w_{i_s,j_s} = \frac{1}{N_s-1}$ if $i \neq j$ and $w_{i_s,j_r} = 0$ if $s \neq r$. In addition, the diagonal elements of the spatial weight matrix is assumed zero, that is, $w_{i_s,i_s} = 0$. In other words, W is a block-diagonal matrix and the sum of each row is normalized to 1, as $\sum_{r=1}^{n_0} \sum_{j_r=1}^{N_r} w_{i_s,j_r} = \sum_{j_s=1}^{N_s} w_{i_s,j_s} = 1$ for $i_s = 1,2,...,N_s; s = 1,2,...,n_0$. The coefficient λ captures the network effects on the probability of accepting a contract-free job.

X_{i_s} is a $1 \times k$ vector of exogenous characteristics of migrant i, including the log value of hourly wage, gender, (rural or urban) Hukou status, marital status, age, years the migrant has stayed in the residential city as well as the fixed effects of the current residential city, home province, education, industry, occupation, type of employer, and the product of home province and education. e_{i_s} is an independent error with mean 0 and variance $\sigma_{e,i}^2$. β is a $k \times 1$ vector of coefficients of these exogenous regressors. We regard the fixed effects of the employer types as the controls from the demand side of the migrant labor market.

necessarily know each other. For instance, Lin, Wu and Lee (2006) use occupation and township as a way to identify connectivity between individuals in a study of national identity formation in Taiwan using spatial regressions. In this paper, people may not know each other but they belong to the same occupational group or township group.

Since each migrant belongs to only one home province, there is no overlap in the co-villager network and the subscript s that indicates co-villager network can be suppressed for notation simplification. Thus, Equation (1) can be simplified to

$$Contract_i = \lambda \sum_{j=1}^{n} w_{ij} \cdot Contract_j + X_i\beta + e_i, \quad i = 1,2,...,n. \tag{2}$$

An endogeneity issue arises as the decisions of taking contract-free irregular jobs of individuals in the network are also influenced by the decision of the objective individual in a symmetric way. In other words, $\sum_{j=1}^{n} w_{ij} \cdot Contract_j$ is correlated with the error term e_i. To achieve consistent estimates, we follow a generalized spatial two-stage least squares (GS2SLS) approach suggested by Kelejian and Prucha (1998).[10] The estimation procedure involves two steps.

- Step 1: We estimate Equation (2) by the two-stage least squares (2SLS) procedure using the instrument set $Z = (X, WX)$, where $X = (X_1^{'}, X_2^{'}, ..., X_n^{'})'$ is an $n \times k$ matrix of all exogenous regressors, to obtain the residuals $\hat{e}_i$. Define $\hat{\Omega} = diag(\hat{e}_1^2, \hat{e}_2^2, ..., \hat{e}_n^2)$ and L is a Cholesky array that satisfies $L\hat{\Omega}L' = I$.
- Step 2: We premultiply regression (2) by L and estimate the transformed regression by 2SLS using $(L')^{-1}Z$ as the instruments. The variance-covariance matrix is estimated by the heteroskedasticity-consistent variance-covariance procedure suggested by White (1980), which is given by $[R'Z(Z'\hat{\Omega}Z)^{-1}Z'R]^{-1}$, where $R = (W \cdot Contract, X)$ and $Contract$ is the $n \times 1$ vector of $Contract_i$.

2.2 The SAR Logit Model

Since the predicted probabilities in the SAR linear probability model do not always fall between 0 and 1, we consider an SAR logit model which could better accommodate the binary nature of the dependent variable and always provide meaningful probability predictions. However, the nonlinear transformation of a logit model adds complexity in the estimation procedure as the typical maximum likelihood estimation (MLE) often involves n integrals in the likelihood function which can be burdensome when the sample size is large. Several approaches have been proposed to produce consistent estimates for the SAR model with a limited dependent variable. McMillen (1992) suggests an expectation–maximization (EM) algorithm to estimate the coefficients of the spatial probit model. Pinkse and Slade (1998) provide conditions of employing the generalized methods of moments (GMM) estimation to the spatial probit model. LeSage (2000) proposes Bayesian simulation approaches for SAR models.[11] Yet the computation intensities of these estimators depend highly on the sample size n, since they require the inversion of an $n \times n$ matrix. In fact, our data set includes 28,614 individuals (after data cleaning), which makes it difficult to apply the estimation approaches mentioned above, even with strong computational power.

We estimate the SAR logit model using a linearized version of the generalized method of moments (LGMM) suggested by Klier and McMillen (2008), which is particularly designed for

[10] See Lee (2003) and Kelejian, Prucha and Yuzefovich (2004) for instrumental variable estimation with different sets of instruments.

[11] See LeSage and Pace (2009) and Smirnov (2010) for detailed reviews of the estimation methods for the spatial discrete choice models.

large samples. The estimation is reduced into two steps - standard logit followed by 2SLS. We adopt their LGMM approach and consider the following SAR logit model

$$Contract_i^* = \lambda \sum_{j=1}^{n} w_{ij} \cdot Contract_j^* + X_i\beta + u_i, \quad i = 1,2,...,n, \tag{3}$$

$$Contract_i = 1(Contract_i^* > 0),$$

where $Contract_i^*$ is a latent continuous variable measuring the propensity of migrant i accepting a contract-free irregular job versus a regular job offering contracts. Migrant $i's$ propensity depends upon the spatially weighted average of propensities of other migrants in her network, which is expressed as $\sum_{j=1}^{n} w_{ij} \cdot Contract_j^*$, where w_{ij} is the $(i,j)th$ element of the spatial weight matrix W, defined in the same way as in the last section. In fact, $Contract_i^*$ is unobservable. Instead, we can only observe a binary variable $Contract_i$ which takes the value of 1 when migrant i accepts an irregular job without contract, and 0 otherwise. $1(\cdot)$ is an indicator function that takes the value of 1 if $Contract_i^* > 0$, and 0 otherwise. When the coefficient of the network effect $\lambda > 0$, it implies that higher propensities of accepting jobs without contracts of migrants in the co-villager network increase the propensity of accepting a contract-free irregular job of migrant i. u_i follows an independently distributed logistic distribution with mean 0 and variance σ_u^2. Consider the matrix form of Equation (3)

$$Contract^* = \lambda W Contract^* + X\beta + u, \tag{4}$$

$$Contract = 1(Contract^* > 0),$$

where $Contract^*$ and u are the $n \times 1$ vectors of $Contract_i^*$ and u_i, respectively. X is an $n \times k$ matrix of k exogenous regressors. The reduced form can be written as

$$Contract^* = (I - \lambda W)^{-1} X\beta + (I - \lambda W)^{-1} u. \tag{5}$$

The variance-covariance matrix of the reduced form error term $(I-\lambda W)^{-1}u$ is proportional to $\Sigma = [(I-\lambda W)'(I-\lambda W)]^{-1}$, which implies heteroskedasticity and autocorrelation when spatial dependence exists, i.e. $\lambda \neq 0$. Denote the ith diagonal element $of\, \Sigma\; as\; \sigma_i^2$, we define $X_i^* = \frac{X_i}{\sigma_i}$ and $X^{**} = (I-\lambda W)^{-1} X^*$, where X^* is an $n \times k$ matrix of X_i^*.

2.3 The Linearized GMM Estimation

As in Pinkse and Slade (1998), the generalized logit residual can be represented by $\varepsilon_i = Contract_i - P_i$, where $P_i = \frac{exp(X_i^{**}\beta)}{1+exp(X_i^{**}\beta)}$. Define the gradient terms as $G_i = (G_{\beta,i}, G_{\lambda,i})$, where $G_{\beta,i} = \frac{\partial P_i}{\partial \beta} = P_i(1-P_i)X_i^{**}$ and $G_{\lambda,i} = \frac{\partial P_i}{\partial \lambda} = P_i(1-P_i)[H_i\beta - \frac{X_i^{**}\beta}{\sigma_i^2}\Lambda_{ii}]$. H_i is the ith row of matrix $H = (I - \lambda W)^{-1} W X^{**}$, Λ_{ii} is the ith diagonal element of $\Lambda = \frac{1}{2}\Sigma[W'(I-\lambda W) + (I-\lambda W)'W]\Sigma$. When $\lambda = 0$, $(I-\lambda W)^{-1}$ degenerates to an identity matrix I. Therefore, β can be consistently estimated by a standard logit model ignoring the spatial structure. The notation can also be greatly simplified as $X_i^{**} = X_i$. Let $\Gamma = (\beta', \lambda)'$ and $\Gamma_0 = (\hat{\beta}_0', 0)'$, where $\hat{\beta}_0$ is the estimate of β in the standard logit model. The gradient terms reduce to $G_{\beta,i} = P_i^0(1-P_i^0)X_i$ and $G_{\lambda,i} = P_i^0(1-P_i^0)H_i^0\beta_0$ when $\lambda = 0$, where H_i^0 is the ith row of an $n \times k$ matrix $H^0 = WX$.

Linearizing ε_i around the initial estimates of parameter Γ_0, we have

$$\varepsilon_i \approx \varepsilon_i^0 - G_i(\Gamma - \Gamma_0), \tag{6}$$

where $\varepsilon_i^0 = Contract_i - P_i^0$ and $P_i^0 = \frac{exp(X_i\hat{\beta}_0)}{1+exp(X_i\hat{\beta}_0)}$. Re-organizing Equation (6), we can obtain

$$\varepsilon_i^0 + G_i\Gamma_0 \approx G_i\Gamma + \varepsilon_i. \tag{7}$$

Define Z_i as the ith row of a matrix of instruments Z. Consider a theoretical moment condition $E(Z_i'\varepsilon_i) = 0$, the corresponding sample moment is thus

$$m(\beta,\lambda) = \frac{1}{n}\sum_{i=1}^{n} Z_i'\varepsilon_i. \tag{8}$$

A GMM estimator can be achieved by minimizing $\varepsilon'ZMZ'\varepsilon$ with respect to β and λ, where M is a positive definite weight matrix and ε is defined as the $n \times 1$ vectors of ε_i. Note that the computation load depends on the sample size n as it involves the inversion of an $n \times n$ matrix. Hence, the GMM estimator becomes infeasible when n is large. However, if M is set to $(Z'Z)^{-1}$, minimizing $\varepsilon'ZMZ'\varepsilon$ is equivalent to conducting a 2SLS estimation of a regression with ε as the error term and Z as the set of instruments. From Equation (7), instead of minimizing $\varepsilon'Z(Z'Z)^{-1}Z'\varepsilon$ with respect to β and λ, the GMM estimator can be achieved by performing 2SLS estimation of $\varepsilon_i^0 + G_i\Gamma_0$ on G_i, using a matrix of instruments Z.

In sum, the LGMM estimation procedure can be conducted in the following two steps:

Step 1: Estimate a standard logit model of $Contract$ with respect to all the exogenous variables X to obtain a consistent estimate of β_0, $\hat{\beta}_0$. Then calculate the residuals $\hat{\varepsilon}_i^0$ as well as the gradient terms $\hat{G}_i = (\hat{G}_{\beta,i}, \hat{G}_{\lambda,i})$.

Step 2: Denote $\hat{\varepsilon}^0 + \hat{G}_\beta\hat{\beta}_0$, $\hat{G}_\beta$ and $\hat{G}_\lambda$ as the matrix counterparts of $\hat{\varepsilon}_i^0 + \hat{G}_{\beta,i}\hat{\beta}_0$, $\hat{G}_{\beta,i}$, and $\hat{G}_{\lambda,i}$. Conduct 2SLS estimation of $\hat{\varepsilon}^0 + \hat{G}_\beta\hat{\beta}_0$ on $\hat{G}_\beta$ and $\hat{G}_\lambda$ using Z as a set of exogenous instruments. More specifically, the 2SLS estimation involves the following two stage regressions:

- Stage 1: Regress $\hat{G}_\beta$ and $\hat{G}_\lambda$ on Z, respectively, to obtain the predicted values $\hat{G}_\beta^p$ and $\hat{G}_\lambda^p$, where $Z = (X, WX)$ is the instrument set.
- Stage 2: Regress $\hat{\varepsilon}^0 + \hat{G}_\beta\hat{\beta}_0$ on $\hat{G}_\beta^p$ and $\hat{G}_\lambda^p$. The corresponding coefficients of $\hat{G}_\beta^p$ and $\hat{G}_\lambda^p$ are the estimates of β and λ, respectively.

The advantage of the LGMM method is that no matrix needs to be inverted, because it requires only the standard logit and linear 2SLS estimation. The linearization significantly reduces the computation time and load as long as λ is small and the true structure is given by Equation (3). See Klier and McMillen (2008) for a detailed discussion of the finite sample properties of LGMM estimation.

3 Data and Summary Statistics

The main data source is the Dynamic Monitoring Survey of Internal Migrants in 2011 from the National Population and Family Planning Commission of China.[12] This is a nationwide survey with 128,000 individual migrants (who are 16-59 years old) residing in 326 host cities in 31 provinces all over China. This survey includes individual demographic, work, and family information such as age, gender, education, rural or urban Hukou status, marriage status, family members, current residential city, home province, wage, insurance, on-the-job training, industry, occupation, type of employer, how many years the migrant has stayed in the residential city, etc.[13] In particular, we know whether or not the migrant has signed a contract for the current job and whether their jobs were introduced by their relatives, classmates, or friends from hometown.

We first delete the observations migrating within a province from the sample in that the co-villager network is defined as people from the same province and currently residing in the same host city, which means for within-province migrants, all people in the host city are the co-villagers of the targeting individual including the local residents. After further dropping the observations for whom contract information is missing as well as the outliers with the largest and smallest 1 percent of wage, we end up with 28,614 observations and 300 host cities in the sample.[14] These internal migrants come from 31 home provinces in China. In spite of the large population of provinces in China, the average size of the network is not very big in the sample. The median and the mean of the network size are 51 and 118 co-villagers, respectively.

The dependent variable is a dummy variable $Contract_i$ equal to 1 if the individual migrant has a job without a contract, and 0 if the individual has a job with a contract, either fixed-term or non-fixed-term. In the full sample, 49 percent of the internal migrants have signed fixed-term contracts, 14 percent have signed non-fixed-term contracts, and 37 percent do not have any labor contract. The first column of Table 1A reports the percentage of the population who hold contracts, while the second column reports the percentage of those who do not hold contracts in various subsamples. The last column presents the percentage of the type of migrants in the full sample. As shown, women and long-term migrants (who have stayed in the host cities for more than 1 year) are a little more likely to take regular jobs with contracts. Almost 40 percent of the migrants with rural Hukou do not hold contracts, while those with urban Hukou are much less likely to accept contract-free jobs. Interestingly, people who found their job through networks, which means their jobs were introduced by their relatives, classmates, or friends from their hometown, are more likely to accept contract-free jobs than others. Moreover, the less educated the person is, the less likely she will hold a contract. A larger percentage of very old or very young migrants do not hold contracts. Lastly, single migrants have a lower probability of holding contracts.

Table 1B presents summary statistics for job-related variables in the samples with or without contracts. For contract-free jobs, the mean hourly wage is around 10.09 RMB. This is 20.34

[12]The survey is conducted from 2010 until present year. We use 2011 because the information on contracts is only available for this year.

[13]There are 15 industries as categorized in the survey including manufacturing, mining, agriculture/forestry/pasture/fishing, construction, electric/coal/water supply, wholesale and retail, lodging and catering, social service, finance/insurance/real estate, transportation/storage/communication, health care/sport/public welfare, education/culture/radio/movie/television, R&D/technology service, government/political organizations/social groups, and other industries; 10 types of employers including land contractor, government or public institution, state owned enterprise, self-employed small business, private business, Hong Kong/Macao/Taiwan enterprise, Japanese/Korean enterprise, European/American enterprise, sino-foreign joint venture, and other enterprises; and 18 types of occupations including technical professionals, business related, etc.

[14]Most missing observations correspond to the self-employed migrants who do not have contract information.

percent lower than the mean hourly wage for the jobs with contracts (12.65). The corresponding mean weekly work hours are 9.34 percent higher (59.23 versus 54.17). The percentage of the population with insurance (including pension, medical insurance, injury insurance, unemployment insurance, maternity insurance, and housing funding) is significantly larger in the sample of regular jobs. The average percentage is 34.79, compared to 4.06 in the sample of contract-free jobs. In addition, only 18.44 percent of the workers with contract-free jobs have received skill training, in comparison with 47.85 percent of contract-holders. These figures confirm that job openings that provide contracts are much better in terms of payments, weekly work hours, insurance, and training to employees.

Table 1C reports summary statistics for job flexibility measured by the years a migrant has worked for her current employer in the samples with or without contracts. Job flexibility is reported as 0 when a migrant has worked for her current employer for less than 12 months. Apparently, migrants with contract-free irregular jobs hold their current jobs for a shorter period on average, and thus the mobility of contract-free jobs is higher. Table 1D further reports the percentage of migrants with different work experience in the current job in the contract and non-contract samples. For example, the percentage in the non-contract sample (27.4%) is larger than that in the contract sample (19.6%) when the period a migrant has worked for the current employer is less than 1 year, the 25 percentile of job flexibility in the full sample. However, when the work period is between 1 and 2 years (the median = 2), between 2 and 4 years (the 75 percentile = 4), or between 4 and 41 years (the maximum value = 41), the opposite case shows up. To conclude, more migrants with contract-free irregular jobs stick to their current jobs for a shorter time period, whereas more migrants with regular jobs that offer contracts hold their jobs for a longer time period. This demonstrates the revealed flexibility of the irregular jobs.

4 Empirical Results

Tables 2-10 provide estimation results for the SAR linear probability and the SAR logit models for the full sample as well as different subsamples. "SAR Linear Model" reports the GS2SLS estimation coefficients and heteroskedasticity-robust standard errors for the SAR linear probability models, while "SAR Logit Model" reports the marginal effects of LGMM results for the SAR logit models. The empirical analysis is primarily based on the SAR logit models, however, the SAR linear probability results are also provided for a robust check. All regressions are conducted in Stata14.

4.1 Basic Results of the Full Sample

Table 2 presents the estimation results for the full sample controlling for different sets of fixed effects. As specified in Section 2, we use $Z = (X, WX)$ as a set of instruments to deal with the endogeneity problem resulting from the spatial lagged dependent variable. X represents the set of all exogenous variables. Specifically, Columns (1) and (4) only control for individual demographic variables including the log value of hourly wage, gender (a dummy equal to 1 if gender is male, and 0 otherwise), Hukou status (a dummy equal to 1 if rural, and 0 if urban), marital status (a dummy equal to 1 if married or have married, and 0 if single), age, years the migrant has stayed in the residential city. In Columns (2) and (5), we add the fixed effects of the current residential city, industry, occupation, type of employer. Columns (3) and (6) also control for the product of home province and education so as to control for the unobserved heterogeneous factors at the home-province-education level.

Both the SAR linear probability model and the SAR logit model show significant positive effects of the spatial lagged dependent term. This implies that a migrant's probability of accepting a job that does not provide a contract increases with the decisions of other migrants in her co-villager network who accept jobs without contracts. Taking Column (6) as the baseline outputs with the most strict controls, the result suggests that when the percentage of migrants in the co-villager network accepting jobs without contracts increases by 10%, the probability of the targeting migrant accepting such jobs increases by about 0.13%. We understand that besides network, there might be other possibilities that may explain the results we find. For instance, it may, to some extent, represent the magnitude of information or skills collected by people migrating from the same province to the same host city.

Our estimates also imply that a higher wage is associated with a lower probability of taking irregular jobs that do not offer a contract. Migrants with rural Hukou, married migrants, and younger migrants are also more likely to accept contract-free irregular jobs. For married migrants, one possible explanation is that they may have families back in their home province, and thus they value flexibility more as they need to visit their families more often. The reason why young migrants find bad contract-free jobs may be due to their lack of working experience.

4.2 Migrants Who Found Jobs through Networks

As mentioned in Section 2, we identify the more effective co-villager networks by dividing the sample into two subsets: (1) migrants who found their jobs through the co-villager networks, that is, their jobs were introduced by their relatives, classmates, or friends from hometown; (2) migrants who found their jobs through other channels, such as local friends in the host city, the internet, job-searching agents, and so on. There should be more job-related interconnections in the network of the former case. We anticipate that migrants who found their jobs through their co-villager networks tend to work together with their co-villagers and there is a larger network effect.

We utilize the same network weight matrix constructed in the full sample and run regressions in both subsamples. Table 3 only reports the estimates of the specification with the same controls as in Columns (3) and (6) in Table 2 using the GS2SLS and LGMM methods. As predicted, the network effects in Subsample (1) are more than twice as those in Subsample (2) regardless of the estimation methods. Recall that the percentage of contract-free jobs in Subsample (1) is also higher than that in Subsample (2). The estimation outcome is consistent with the literature that the quality of jobs introduced by contacts is usually low especially for low-educated workers.[15]

4.3 Network Effects in Different Home Provinces

Next, we run regressions for people from different home provinces, controlling for the demographic variables as well as the host province, industry, occupation, type of employer and education fixed effects. In these regressions, we do not include the fixed effects of the host city and the product of home province and education, because the number of observations in some subsamples is not large enough to have sufficient degree of freedom. Table 4 demonstrates the heterogeneous network effects for migrants from Sichuan, Henan, Anhui, and Hunan, which are the 4 largest provinces

[15] Recall again that 69 percent of the migrants have an education level of middle school or below in our sample.

from which migrants move out. The proportion of migrants in our data from these provinces are 13.13%, 10.66%, 10.57%, and 6.47%. The SAR logit model shows significant network effects in Sichuan and Henan provinces, and the effect is stronger in Sichuan. We also draw Figure 1 to illustrate the heterogeneous network effects in all home provinces on a map, with the darker color represents the larger magnitude of network effects. The blank provinces mean that the network effects are insignificant there. As we can see, Chongqing and Jiangxi have the largest effects, followed by Shandong, Guangxi, Sichuan, Zhejiang, Heilongjiang, Gansu, and Henan. Among these 9 provinces with significant effects, two-thirds (Gansu, Sichuan, Heilongjiang, Chongqing, Zhejiang, and Jiangxi) have higher percentage of contract-free jobs than the mean percentage in the full sample. Liaoning is an exception that has a negative coefficient. Different culture and norms across home provinces may explain the various network effects estimated.

The heterogeneous effects across home provinces provide important policy implication. In the province of large network effects and a relatively high percentage of contract-free jobs, there is dynamic concentration of job-contract types among migrants from the same home province over time. If migrants from certain home provinces always find "bad" low-paid jobs that do not offer contracts, their consumption in the host city as well as the remittance they send back to their home province will be lower than migrants from other home provinces, which will further contribute to income inequality across regions.

4.4 Network Effects with Different Types of Employers

We further categorize the full sample into 4 different types of employers: (1) private and self-employed business, (2) state-owned enterprises, (3) Hong Kong, Macao, Taiwan, or foreign enterprises, and (4) sino-foreign joint venture enterprises. We run regressions in these subsamples, controlling for the same fixed effects as in Section 4.3. Both SAR linear probability models and SAR logit models show significant network effects only for type (1) and (2) employers in Table 5. The effect is larger for the latter, presumably because migrants generally cannot enter these institutions through formal recruitment process and they only work there as temporary employees, which is exactly the group that relies more on co-villager networks to seek job. Moreover, temporary jobs are also relatively less likely to offer contracts. Foreign firms or sino-foreign joint venture firms usually obey the law more strictly and offer less contract-free jobs.

4.5 Network Effects in Different Industries

We also investigate network effects in different industries and only observe significant effects in 3 industries: manufacturing, construction, and lodging and catering. Table 6 reports the GS2SLS and LGMM estimates (controlling for the same fixed effects as in Section 4.3) for the 4 largest industries, among which only wholesale and retail is insignificant. As we can see, the network effect is largest in the construction industry, followed by lodging and catering and manufacturing industries. Take construction as an example. A construction project in China is often contracted out to a team in which the members are usually recruited from the same village. Therefore, the migration decision, job search process, and the contract decision of these people are strongly influenced by the co-villager network. Moreover, it is easier for these co-villagers who work in the same workplace to form a defacto "mini labor union," which could take the place of a formal contract.

The heterogeneous effects across different employer types and industries suggest that to implement the labor contract law better, more attention should be paid to (1) the private and self-employed business as well as the state-owned enterprises, and (2) manufacturing, construction as well as lodging and catering.

4.6 Network Effects for Migrants of Different Education Levels

In Table 7, we investigate the network effects for migrants of different education levels: primary school or below, middle school, high school or secondary technical school, and junior college or above. Both the GS2SLS and LGMM estimates (controlling for the same fixed effects as in Section 4.3) reveal that the migrants with lower education levels rely more on co-villager networks. According to the SAR logit model, significant network effects only exist among migrants with education levels below high school. The magnitude is largest for middle school graduates while slightly lower for migrants with primary education or below.

4.7 Robustness Check for Other Subsamples

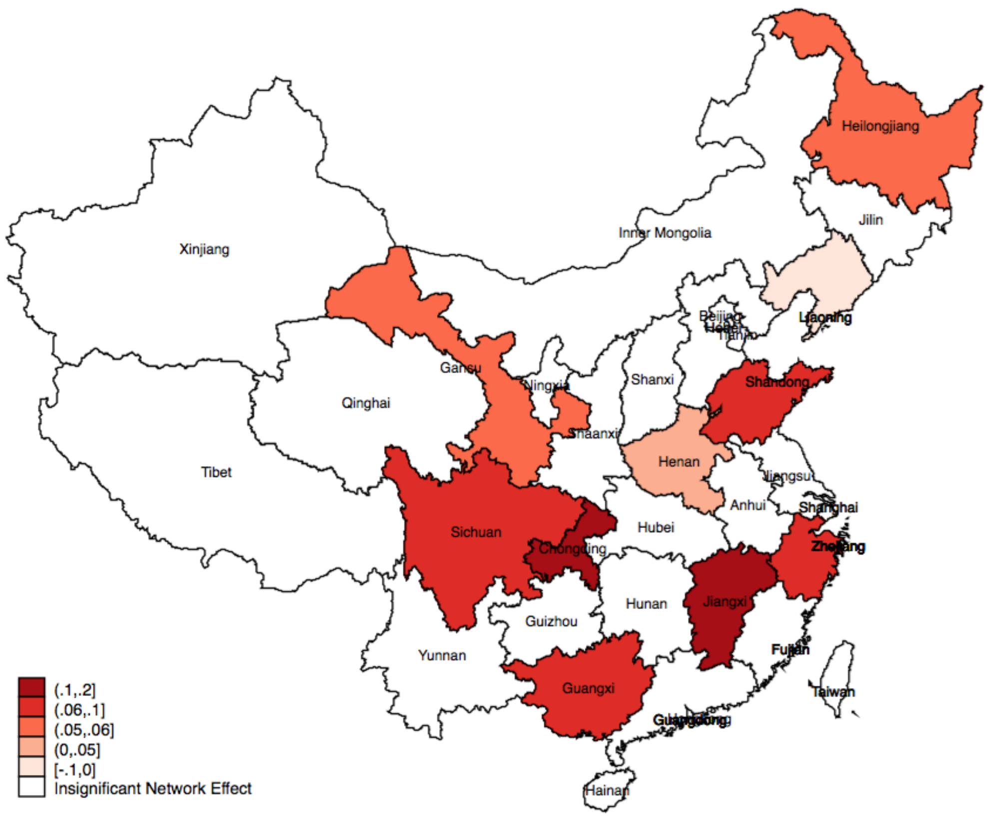
Lastly, we conduct sensitivity analysis in the following subsamples: (1) males versus females; (2) migrants with rural Hukou versus urban Hukou; (3) long-term migration (people who have stayed in the residential city for more than a year) versus short-term migration (people who have stayed in the residential city for a year or less than a year). Tables 8-10 only report estimates with the most strict controls as in Columns (3) and (6) in Table 2 using the GS2SLS and LGMM methods, respectively. According to the SAR logit model, network effects are slightly larger for short-term migrants and become insignificant for women and migrants with urban Hukou.

5 Conclusion

This paper employs a comprehensive data set from a nationwide survey to study how co-villager networks affect internal migrants' decision of accepting irregular jobs without contracts in the labor market of China. Surrounded by a network with a large amount of co-villagers work for business that do not provide contracts, the migrant not only has limited job choices, but also knows little about the benefits a regular labor contract would offer because of the lack of information. Additionally, the network may serve as a "mini labor union" that raises the bargaining power of the migrants and substitutes for the contract with respect to employment protection. As a result, the migrant tends to accept a contract-free job offer that features flexibility. We use a spatial lagged dependent term to capture the network effect, estimated by a GS2SLS method for the SAR linear probability model and a linearized GMM approach for the SAR logit model. Our empirical work confirms that the probability of a migrant's acceptance of jobs without contracts increases with the acceptance of jobs without contracts of other migrants in her co-villager network. We further find that the network effects are larger in the subsamples of migrants whose jobs were introduced by their co-villagers and less educated migrants and there are heterogeneous network effects across home provinces, employer types, and industries. Our findings offer policy implications from the aspects of income inequality and the implementation of the labor contract law.

Appendix

Figure 1: Heterogeneous Network Effects across Home Provinces



Note: Figure 1 illustrates the heterogeneous network effects across home provinces. The provinces with significant effects are colored in red and the darker color represents the larger magnitude of network effects.

Table 1: Summary Statistics

Table 1A: Percentage of Population with and without Contracts

		Has Contract (%)	No Contract (%)	Percentage in the Full Sample (%)
Full Sample		63.35	36.65	-
Gender	Male	62.24	37.76	59.00
	Female	64.95	35.05	41.00
Hukou	Rural	60.61	39.39	84.97
	Urban	78.84	21.16	15.03
Migrating Years	Equal to or Less than 1 Year	61.75	38.25	40.20
	More than 1 Year	64.43	35.57	59.80
Found Job	Yes	61.32	38.68	42.19
through Network	No	64.83	35.17	57.81
Education	School Dropouts	41.72	58.28	1.63
	Primary School	45.87	54.13	14.07
	Middle School	59.62	40.38	53.45
	High School	70.92	29.08	14.15
	Secondary Technical School	79.80	20.20	6.57
	Junior College	87.17	12.83	6.29
	University	93.30	5.70	3.55
	Graduate School	96.43	3.57	0.29
Age	[16, 18)	50.82	49.18	1.49
	[18, 30)	68.08	31.92	43.98
	[30, 40)	63.34	36.66	31.83
	[40, 50)	55.65	44.35	19.76
	[50, 59]	50.95	49.05	2.95
Marital Status	Married	67.64	32.36	69.56
	Single	61.48	38.52	30.44

Note: Table 1A reports the percentage of population who have contracts or not in different subsamples of (1) males or females, (2) rural or urban Hukou, (3) long-term migrants (staying in the host city for more than a year) or short-term migrants (staying in the host city for a year or less than a year), (4) migrants who found their jobs through networks or not, (5) different education levels, (6) different ages, and (7) different marriage status.

Table 1B: Statistics of Job-Related Variables in the Contract or Non-Contract Samples

		# of Obs.	Mean	Std. Dev.	Min.	Median	Max.
Hourly Wage	Full Sample	28,614	11.71	6.35	3.50	10.36	48.75
	Has Contract	18,128	12.65	6.83	3.50	10.63	48.75
	No Contract	10,486	10.09	5.03	3.50	8.93	43.13
Weekly Working Hours	Full Sample	28,614	56.02	13.51	2	56	112
	Has Contract	18,128	54.17	13.09	4	50	112
	No Contract	10,486	59.23	13.61	2	60	112

		Full Sample		**Has Contract**		**No Contract**	
		Yes (%)	**No (%)**	**Yes (%)**	**No (%)**	**Yes (%)**	**No (%)**
Insurance	Pension	29.04	68.66	44.62	53.43	3.83	94.98
	Medical Insurance	33.58	64.16	49.14	47.95	6.68	92.19
	Injury Insurance	38.21	58.21	54.00	41.56	10.90	86.99
	Unemployment Insurance	18.48	78.15	28.32	67.17	1.49	97.12
	Maternity Insurance	12.93	83.35	19.83	75.13	1.01	97.55
	Housing Fund	8.30	88.45	12.84	82.77	0.45	98.25
	Average	23.42	73.50	34.79	61.34	4.06	94.51
Work Skill Training		37.07	62.93	47.85	52.15	18.44	81.56

Note: Table 1B presents the summary statistics for the job-related variables in samples with or without contracts, separately. The job-related variables include hourly wages, insurance (i.e. pension, medical insurance, injury insurance, unemployment insurance, maternity insurance, and housing fund) as well as work skill training.

Table 1C: Statistics of Job Flexibility in the Contract or Non-Contract Samples

Years a Migrant Has Worked for the Current Employer		# of Obs.	Mean	Std. Dev.
Job Flexibility	Full Sample	28,580	3.23	4.16
	Has Contract	18,108	3.35	4.16
	No Contract	10,472	3.03	4.17

Note: Table 1C presents the summary statistics for job flexibility measured by the years a migrant has worked for her current employer in samples with or without contracts, separately. Job flexibility is reported as 0 when a migrant has worked for her current employer for less than 12 months.

Table 1D: Percentage of Migrants with Different Work Experience with the Current Employer in the Contract or Non-Contract Samples

Years a Migrant Has Worked for the Current Employer		Has Contract (%)	No Contract (%)
Job Flexibility	(0, 1) year: 25 percentile of job flexibility in the full sample = 1	19.58	27.40
	[1, 2) year: 50 percentile of job flexibility in the full sample = 2	23.87	23.46
	[2, 4) year: 75 percentile of job flexibility in the full sample = 4	24.85	21.64
	[4, 41) years: maximum of job flexibility in the full sample = 19	31.70	27.50

Note: Table 1D presents the percentage of migrants with different work experience in the current job in the contract and non-contract samples. For example, there are 19.58 percent and 27.40 percent of migrants who have worked for her current employer for less than 12 months in the samples with and without contracts, separately. The cutoffs are the 25 percentile, median, 75 percentile, and maximum of job flexibility in the full sample.

Table 2: Basic Regression with Full Sample

	Dependent variable: ***Contract***					
	SAR Linear Probability Model			SAR Logit Model		
	(1)	(2)	(3)	(4)	(5)	(6)
W·Contract	**0.5046*****	**0.1144*****	**0.2202*****	**0.1509*****	**0.0226*****	**0.0130****
	(0.0219)	(0.0188)	(0.0185)	(0.0098)	(0.0065)	(0.0058)
ln(Wage)	-0.1859***	-0.0952***	-0.0946***	-0.2568***	-0.0874***	-0.0702***
	(0.0063)	(0.0065)	(0.0065)	(0.0064)	(0.0123)	(0.0215)
Gender	0.0446***	0.0072	0.0077	0.0714***	0.0061	0.0039
	(0.0056)	(0.0054)	(0.0054)	(0.0052)	(0.0053)	(0.0043)
Hukou Status	0.0886***	0.0284***	0.0253***	0.1844***	0.0307***	0.0259**
	(0.0071)	(0.0077)	(0.0078)	(0.0076)	(0.0100)	(0.0102)
Marital Status	0.0285***	0.0242***	0.0237***	0.0469***	0.0169**	0.0134**
	(0.0073)	(0.0069)	(0.0069)	(0.0071)	(0.0071)	(0.0068)
Age	0.0028***	-0.0011***	-0.0011***	0.0031***	-0.0008**	-0.0006*
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0003)
Stay Years	-0.0003	0.0012*	0.0010*	-0.0001	0.0011*	0.0008
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0005)
Fixed effects						
Host City	No	Yes	Yes	No	Yes	Yes
Home Province	No	Yes	Yes	No	Yes	Yes
Industry	No	Yes	Yes	No	Yes	Yes
Occupation	No	Yes	Yes	No	Yes	Yes
Type of Employer	No	Yes	Yes	No	Yes	Yes
Education Level	No	Yes	Yes	No	Yes	Yes
Home Province*Education	No	No	Yes	No	No	Yes
# of Observations	28,614	28,614	28,614	28,614	28,614	28,614
R-Squared	0.3091	0.2482	0.2451	0.336	0.530	0.418

Note: Columns (1), (2), and (3) are the GS2SLS estimates from the SAR linear probability model with different sets of fixed effects, while Columns (4), (5), and (6) are the marginal effects of the LGMM estimates from the SAR logit model with the same controls. Standard errors are in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Table 3: Regressions in Subsamples of Migrants Who Found Jobs through Co-Villagers or Not

	Dependent variable: ***Contract***			
	SAR Linear Probability Model		SAR Logit Model	
	Job was Introduced by Co-Villagers	Job Was Not Introduced by Co-Villagers	Job was Introduced by Co-Villagers	Job Was Not Introduced by Co-Villagers
	(1)	(2)	(3)	(4)
W·Contract	**0.3546*** **	**0.1578*** **	**0.0230** **	**0.0083** **
	(0.0265)	(0.0223)	(0.0101)	(0.0042)
ln(Wage)	-0.0914***	-0.0974***	-0.0576***	-0.0692***
	(0.0105)	(0.0081)	(0.0189)	(0.0131)
Gender	0.0192**	-0.0039	0.0145*	-0.0047
	(0.0085)	(0.0068)	(0.0076)	(0.0051)
Hukou Status	0.0258*	0.0239**	0.0189	0.0221**
	(0.0136)	(0.0093)	(0.0124)	(0.0089)
Marital Status	0.0184*	0.0220**	0.0116	0.0100
	(0.0109)	(0.0085)	(0.0087)	(0.0068)
Age	-0.0016***	-0.0007	-0.0010*	-0.0003
	(0.0006)	(0.0005)	(0.0005)	(0.0003)
Stay Years	0.0011	0.0011	0.0008	0.0005
	(0.0010)	(0.0008)	(0.0007)	(0.0005)
Fixed effects	Host City, Home Province, Industry, Occupation, Type of Employer, Education, Home Province*Education			
# of Observations	12,072	16,542	12,072	16,542
R-Squared	0.2529	0.2517	0.445	0.422

Note: Columns (1) and (3) report the estimates in the subsample of migrants who found jobs through co-villager networks, while Columns (2) and (4) report the estimates for those who did not find job through co-villagers. Columns (1) and (2) are the GS2SLS estimates from the SAR linear probability model, while Columns (3) and (4) are the marginal effects of the LGMM estimates from the SAR logit model with the same controls. In all regressions, we control for the fixed effects of the host city, home province, industry, occupation, type of employer, education, and the interactions of home province and education. Standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level.

Table 4: Network Effects for Migrants from Different Home Provinces

	Dependent variable: ***Contract***							
	SAR Linear Probability Model				SAR Logit Model			
	Migrants from Sichuan	Migrants from Henan	Migrants from Anhui	Migrants from Hunan	Migrants from Sichuan	Migrants from Henan	Migrants from Anhui	Migrants from Hunan
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
W·Contract	**0.5315*****	**0.3201*****	**0.3206*****	**0.4423*****	**0.0659*****	**0.0151***	**0.0279**	**-0.0027**
	(0.0384)	(0.0660)	(0.0767)	(0.0590)	(0.0206)	(0.0090)	(0.0275)	(0.0096)
ln(Wage)	-0.1180***	-0.0887***	-0.0961***	-0.1395***	-0.0930***	-0.0381***	-0.0769	-0.0703
	(0.0181)	(0.0201)	(0.0215)	(0.0239)	(0.0179)	(0.0112)	(0.0491)	(0.0627)
Gender	0.0173	0.0192	0.0379**	-0.0278	0.0114	0.0115	0.0244	-0.0031
	(0.0154)	(0.0158)	(0.0169)	(0.0194)	(0.0127)	(0.0077)	(0.0236)	(0.0143)
Hukou Status	0.0560**	0.0201	0.0163	0.0070	0.0726***	0.0079	0.0212	0.0214
	(0.0256)	(0.0279)	(0.0291)	(0.0279)	(0.0248)	(0.0141)	(0.0455)	(0.0316)
Marital Status	0.0577***	0.0229	-0.0175	-0.0247	0.0521***	0.0163	-0.0105	0.0029
	(0.0216)	(0.0198)	(0.0224)	(0.0235)	(0.0186)	(0.0101)	(0.0284)	(0.0189)
Age	-0.0027***	-0.0012	-0.0001	-0.0006	-0.0020**	-0.0006	-0.0002	-0.0002
	(0.0010)	(0.0012)	(0.0013)	(0.0015)	(0.0008)	(0.0005)	(0.0014)	(0.0011)
Stay Years	-0.0005	0.0002	-0.0021	0.0018	-0.0008	0.0002	-0.0023	0.0008
	(0.0015)	(0.0020)	(0.0020)	(0.0023)	(0.0012)	(0.0009)	(0.0024)	(0.0018)
Fixed effects	Host Province, Industry, Occupation, Type of Employer, Education							
# of Observations	3,758	3,049	3,024	1,850	3,758	3,049	3,024	1,850
R-Squared	0.2647	0.2632	0.2655	0.2725	0.333	0.363	0.581	0.394

Note: Table 4 reports the co-villager networks for migrants from Sichuan, Henan, Anhui, and Hunan. Columns (1) - (4) are the GS2SLS estimates from the SAR linear probability model, while Columns (5) - (8) are the marginal effects of the LGMM estimates from the SAR logit model with the same controls. In all regressions, we control for the fixed effects of the host province, home province, industry, occupation, type of employer, and education. Standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level.

Table 5: Network Effects for Migrants with Different Employer Types

	Dependent variable: ***Contract***							
	SAR Linear Probability Model				SAR Logit Model			
	Private business or self-employed	state-owned firms	HK/Macao/Taiwan or foreign firms	Sino-foreign or joint venture firms	Private business or self-employed	state-owned firms	HK/Macao/Taiwan or foreign firms	Sino-foreign or joint venture firms
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
W·Contract	**0.1508*****	**0.2896*****	**-0.0134**	**-0.0101**	**0.0462*****	**0.1289*****	**-0.0451**	**0.0369**
	(0.0292)	(0.0369)	(0.0230)	(0.0281)	(0.0124)	(0.0491)	(0.0386)	(0.0856)
ln(Wage)	-0.1346***	-0.0391**	-0.0233***	-0.0187	-0.1031***	-0.0734*	-0.1782***	-0.1027
	(0.0085)	(0.0159)	(0.0083)	(0.0134)	(0.0342)	(0.0408)	(0.0496)	(0.1378)
Gender	0.0194***	-0.0168	0.0010	0.0026	0.0147*	-0.0090	-0.0164	0.0124
	(0.0074)	(0.0143)	(0.0057)	(0.0095)	(0.0089)	(0.0324)	(0.0237)	(0.0595)
Hukou Status	0.0390***	0.0396***	0.0123	0.0037	0.0248	0.0702	-0.0212	0.0896
	(0.0109)	(0.0153)	(0.0087)	(0.0127)	(0.0184)	(0.0510)	(0.0452)	(0.1355)
Marital Status	0.0273***	0.0298**	0.0081	-0.0225**	0.0208*	0.0507	0.0551*	-0.0784
	(0.0095)	(0.0151)	(0.0072)	(0.0101)	(0.0111)	(0.0383)	(0.0326)	(0.1352)
Age	-0.0014***	-0.0009	0.0009	-0.0004	-0.0012**	-0.0006	0.0009	-0.0029
	(0.0005)	(0.0009)	(0.0006)	(0.0007)	(0.0006)	(0.0017)	(0.0019)	(0.0059)
Stay Years	0.0014*	-0.0006	-0.0010	-0.0010	0.0012	0.0003	-0.0045	-0.0069
	(0.0008)	(0.0014)	(0.0010)	(0.0012)	(0.0008)	(0.0029)	(0.0034)	(0.0106)
Fixed effects	Host Province, Home Province, Industry, Occupation, Education							
# of Observations	19,477	2,911	2,715	1,374	19,477	2,911	2,715	1,374
R-Squared	0.2846	0.1398	0.3510	0.3466	0.570	0.292	0.484	0.551

Note: Table 5 reports the co-villager networks for migrants working for 4 types of employers: private and self-employed business, state-owned enterprises, Hong Kong, Macao, Taiwan, or foreign enterprises, and sino-foreign joint venture enterprises. Columns (1) - (4) are the GS2SLS estimates from the SAR linear probability model, while Columns (5) - (8) are the marginal effects of the LGMM estimates from the SAR logit model with the same controls. In all regressions, we control for the fixed effects of the host province, home province, industry, occupation, type of employer, and education. Standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level.

Table 6: Network Effects for Migrants in Different Industries

	Dependent variable: ***Contract***							
	SAR Linear Probability Model				SAR Logit Model			
	Manufacturing	Construction	Lodging & Catering	Wholesale & Retail	Manufacturing	Construction	Lodging & Catering	Wholesale & Retail
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
W·Contract	**0.2469*****	**0.2729*****	**0.2505*****	**0.1256****	**0.0162****	**0.0572*****	**0.0314***	**0.0155**
	(0.0300)	(0.0454)	(0.0618)	(0.0582)	(0.0075)	(0.0149)	(0.0162)	(0.0123)
ln(Wage)	-0.1027***	-0.0371**	-0.1653***	-0.1896***	-0.0442***	-0.0312	-0.1638***	-0.1622***
	(0.0103)	(0.0181)	(0.0250)	(0.0262)	(0.0086)	(0.0198)	(0.0264)	(0.0321)
Gender	0.0232***	-0.0452**	0.0122	0.0220	0.0104**	-0.0508**	0.0126	0.0126
	(0.0073)	(0.0214)	(0.0201)	(0.0239)	(0.0042)	(0.0246)	(0.0207)	(0.0217)
Hukou Status	0.0181	0.0431	0.0293	0.0100	0.0140	0.0566*	0.0186	0.0168
	(0.0115)	(0.0279)	(0.0289)	(0.0300)	(0.0100)	(0.0342)	(0.0305)	(0.0287)
Marital Status	0.0338***	0.0586***	0.0149	0.0106	0.0136**	0.0433*	0.0198	0.0143
	(0.0097)	(0.0222)	(0.0263)	(0.0302)	(0.0062)	(0.0225)	(0.0273)	(0.0268)
Age	-0.0021***	-0.0019**	0.0004	0.0021	-0.0008**	-0.0019**	0.0007	0.0017
	(0.0006)	(0.0009)	(0.0015)	(0.0018)	(0.0003)	(0.0009)	(0.0016)	(0.0017)
Stay Years	-0.0011	0.0093***	-0.0010	-0.0037	-0.0005	0.0122***	-0.0009	-0.0027
	(0.0010)	(0.0015)	(0.0026)	(0.0028)	(0.0005)	(0.0021)	(0.0027)	(0.0025)
Fixed effects	Host Province,Home Province, Occupation, Type of Employer, Education							
# of Observations	12,955	3,638	2,429	1,481	12,955	3,638	2,429	1,481
R-Squared	0.2799	0.3229	0.3032	0.3659	0.359	0.391	0.375	0.362

Note: Table 6 reports the co-villager networks for migrants working for 4 industries: manufacturing, construction, lodging and catering, and wholesale and retail. Columns (1) - (4) are the GS2SLS estimates from the SAR linear probability model, while Columns (5) - (8) are the marginal effects of the LGMM estimates from the SAR logit model with the same controls. In all regressions, we control for the fixed effects of the host province, home province, industry, occupation, type of employer, and education. Standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level.

Table 7: Network Effects for Migrants with Different Education Levels

	Dependent variable: ***Contract***							
	SAR Linear Probability Model				SAR Logit Model			
	Primary School or under	Middle School	High School or Equivalent	Junior College or above	Primary School or under	Middle School	High School or Equivalent	Junior College or above
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
W·Contract	**0.1645*****	**0.1845*****	**0.1813*****	**0.1789*****	**0.0134****	**0.0168****	**0.0063**	**0.0004**
	(0.0449)	(0.0282)	(0.0425)	(0.0338)	(0.0065)	(0.0072)	(0.0054)	(0.0034)
ln(Wage)	-0.0781***	-0.1073***	-0.1210***	-0.0534***	-0.0181**	-0.0678***	-0.0311***	-0.0432***
	(0.0184)	(0.0098)	(0.0129)	(0.0119)	(0.0085)	(0.0158)	(0.0119)	(0.0085)
Gender	-0.0055	0.0189**	0.0205*	-0.0136	-0.0020	0.0117*	0.0048	-0.0084
	(0.0155)	(0.0080)	(0.0111)	(0.0102)	(0.0047)	(0.0064)	(0.0082)	(0.0054)
Hukou Status	0.1453***	0.0198	0.0287**	0.0279**	0.0503**	0.0135	0.0075	0.0246***
	(0.0393)	(0.0145)	(0.0129)	(0.0121)	(0.0198)	(0.0111)	(0.0096)	(0.0067)
Marital Status	0.0363	0.0329***	0.0111	-0.0012	0.0082	0.0205**	0.0030	-0.0005
	(0.0263)	(0.0103)	(0.0134)	(0.0120)	(0.0084)	(0.0087)	(0.0096)	(0.0058)
Age	-0.0018**	-0.0016***	0.0010	0.0011	-0.0005	-0.0010**	0.0001	0.0004
	(0.0009)	(0.0005)	(0.0009)	(0.0010)	(0.0003)	(0.0005)	(0.0006)	(0.0005)
Stay Years	0.0018	0.0021**	-0.0024*	-0.0013	0.0006	0.0016**	-0.0001	-0.0019**
	(0.0013)	(0.0009)	(0.0013)	(0.0013)	(0.0004)	(0.0007)	(0.0010)	(0.0009)
Fixed effects	Host Province, Home Province, Industry, Occupation, Type of Employer							
# of Observations	4,492	15,293	5,929	2,900	4,492	15,293	5,929	2,900
R-Squared	0.2895	0.2656	0.2791	0.3349	0.382	0.382	0.382	0.382

Note: Table 7 reports the co-villager networks for migrants with different education levels: primary school or under, middle school, high school or secondary technical school, and junior college or above. Columns (1) - (4) are the GS2SLS estimates from the SAR linear probability model, while Columns (5) - (8) are the marginal effects of the LGMM estimates from the SAR logit model with the same controls. In all regressions, we control for the fixed effects of the host province, home province, industry, occupation, type of employer, and education. Standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level.

Table 8: Network Effects for Male and Female Subsamples

	Dependent variable: ***Contract***			
	SAR Linear Probability Model		SAR Logit Model	
	Male	Female	Male	Female
	(1)	(2)	(3)	(4)
W·Contract	**0.2290*** **	**0.2429*** **	**0.0080*** **	**0.0020**
	(0.0211)	(0.0269)	(0.0028)	(0.0069)
ln(Wage)	-0.0784***	-0.1232***	-0.0380***	-0.0861***
	(0.0083)	(0.0102)	(0.0081)	(0.0179)
Hukou Status	0.0229**	0.0213*	0.0172***	0.0261**
	(0.0101)	(0.0116)	(0.0064)	(0.0114)
Marital Status	0.0241***	0.0229**	0.0078*	0.0083
	(0.0089)	(0.0104)	(0.0045)	(0.0083)
Age	-0.0016***	-0.0005	-0.0005**	-0.0003
	(0.0005)	(0.0007)	(0.0002)	(0.0005)
Stay Years	0.0020***	-0.0008	0.0009**	-0.0005
	(0.0007)	(0.0010)	(0.0004)	(0.0007)
Fixed effects	Host City, Home Province, Industry, Occupation, Type of Employer, Education, Home Province*Education			
# of Observations	16,882	11,732	16,882	11,732
R-Squared	0.2509	0.2612	0.429	0.386

Note: Table 8 reports the co-villager networks for male or female migrants. Columns (1) and (2) are the GS2SLS estimates from the SAR linear probability model, while Columns (3) and (4) are the marginal effects of the LGMM estimates from the SAR logit model with the same controls. In all regressions, we control for the fixed effects of the host city, home province, industry, occupation, type of employer, education, and the interactions of home province and education. Standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level.

Table 9: Regressions in Subsamples of Migrants with Rural/Urban Hukou

	Dependent variable: ***Contract***			
	SAR Linear Probability Model		SAR Logit Model	
	Rural Hukou	Urban Hukou	Rural Hukou	Urban Hukou
	(1)	(2)	(3)	(4)
W·Contract	**0.2207******	**0.1216******	**0.0085***	**0.0084**
	(0.0210)	(0.0255)	(0.0052)	(0.0076)
ln(Wage)	-0.0980***	-0.0821***	-0.0463**	-0.0653***
	(0.0074)	(0.0117)	(0.0224)	(0.0237)
Gender	0.0081	0.0005	0.0032	0.0039
	(0.0060)	(0.0101)	(0.0035)	(0.0092)
Marital Status	0.0218***	-0.0015	0.0099	0.0043
	(0.0077)	(0.0119)	(0.0062)	(0.0114)
Age	-0.0013***	0.0010	-0.0006*	0.0002
	(0.0004)	(0.0008)	(0.0003)	(0.0006)
Stay Years	0.0009	0.0027**	0.0005	0.0021*
	(0.0007)	(0.0012)	(0.0004)	(0.0013)
Fixed effects	Host City, Home Province, Industry, Occupation, Type of Employer, Education, Home Province*Education			
# of Observations	24,313	4,301	24,313	4,301
R-Squared	0.2493	0.2988	0.415	0.442

Note: Table 9 reports the co-villager networks for migrants with rural or urban Hukou. Columns (1) and (2) are the GS2SLS estimates from the SAR linear probability model, while Columns (3) and (4) are the marginal effects of the LGMM estimates from the SAR logit model with the same controls. In all regressions, we control for the fixed effects of the host city, home province, industry, occupation, type of employer, education, and the interactions of home province and education. Standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level.

Table 10: Regressions in Subsamples of Short-Term/Long-Term Migrants

	Dependent variable: ***Contract***			
	SAR Linear Probability Model		SAR Logit Model	
	Long-Term Migrants	Short-Term Migrants	Long-Term Migrants	Short-Term Migrants
	(1)	(2)	(3)	(4)
W·Contract	**0.2484*****	**0.2325*****	**0.0121****	**0.0160*****
	(0.0221)	(0.0251)	(0.0059)	(0.0060)
ln(Wage)	-0.1068***	-0.0717***	-0.0843***	-0.0424***
	(0.0080)	(0.0105)	(0.0220)	(0.0100)
Gender	0.0063	0.0083	0.0052	0.0036
	(0.0068)	(0.0084)	(0.0059)	(0.0060)
Hukou Status	0.0221**	0.0369***	0.0260**	0.0262**
	(0.0093)	(0.0119)	(0.0114)	(0.0118)
Marital Status	0.0268***	0.0206**	0.0202**	0.0155**
	(0.0089)	(0.0102)	(0.0095)	(0.0078)
Age	-0.0005	-0.0015**	-0.0005	-0.0009**
	(0.0005)	(0.0006)	(0.0004)	(0.0004)
Fixed effects	Host City, Home Province, Industry, Occupation, Type of Employer, Education, Home Province*Education			
# of Observations	17,112	11,502	17,112	11,502
R-Squared	0.2572	0.2607	0.465	0.422

Note: Table 10 reports the co-villager networks for long-term or short-term migrants. Columns (1) and (2) are the GS2SLS estimates from the SAR linear probability model, while Columns (3) and (4) are the marginal effects of the LGMM estimates from the SAR logit model with the same controls. In all regressions, we control for the fixed effects of the host city, home province, industry, occupation, type of employer, education, and the interactions of home province and education. Standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level.

References

[1] Aguirregabiria, Victor, and Cesar Alonso-Borrego. "Labor contracts and flexibility: evidence from a labor market reform in Spain." Economic Inquiry 52.2 (2014): 930-957.

[2] Baltagi, Badi H., and Yin-Fang Yen. "Hospital treatment rates and spillover effects: Does ownership matter?." Regional Science and Urban Economics 49 (2014): 193-202.

[3] Battisti, Michele, Giovanni Peri, and Agnese Romiti. "Dynamic effects of co-ethnic networks on immigrants' economic success." NBER Working Paper #22389 (2016).

[4] Battu, Harminder, Paul Seaman, and Yves Zenou. "Job contact networks and the ethnic minorities." Labour Economics 18.1 (2011): 48-56.

[5] Bayer, Patrick, Stephen L. Ross, and Giorgio Topa. "Place of work and place of residence: irregular hiring networks and labor market outcomes." Journal of Political Economy 116.6 (2008): 1150-1196.

[6] Becker, Jeffrey, and Manfred Elfstrom. "The impact of China's Labor Contract Law on workers." International Labor Rights Forum, May 12 (2010).

[7] Bertrand, Marianne, Erzo FP Luttmer, and Sendhil Mullainathan. "Network effects and welfare cultures." The Quarterly Journal of Economics 115.3 (2000): 1019-1055.

[8] Booth, Alison L., Marco Francesconi, and Jeff Frank. "Temporary jobs: stepping stones or dead ends?." The Economic Journal 112.480 (2002): F189-F213.

[9] Calvo-Armengol, Antoni, and Matthew O. Jackson. "The effects of social networks on employment and inequality." The American Economic Review 94.3 (2004): 426-454.

[10] Carrington, William J., Enrica Detragiache, and Tara Vishwanath. "Migration with endogenous moving costs." The American Economic Review 86.4 (1996): 909-930.

[11] Chen, Yuyu, Jin G. Zhe, and Yue Yang. "Peer migration in China." NBER Working Paper #15671 (2010).

[12] Cingano, Federico, and Alfonso Rosolia. "People I know: job search and social networks." Journal of Labor Economics 30.2 (2012): 291-332.

[13] Cliff, Andrew, and Keith Ord. Spatial Autocorrelation. Pion, London, 1973.

[14] Duflo, Esther, and Emmanuel Saez. "The role of information and social interactions in retirement plan decisions: evidence from a randomized experiment." The Quarterly Journal of Economics 118.3 (2003): 815-842.

[15] Elliott, James R. "Referral hiring and ethnically homogeneous jobs: How prevalent is the connection and for whom?." Social Science Research 30.3 (2001): 401-425.

[16] Falcón, Luis M. "Social networks and latino immigrants in the labor market: A review of the literature and evidence." Latinos in a Changing Society 11 (2007): 254-272.

[17] Frijters, Paul, Michael A. Shields, and Stephen Wheatley Price. "Job search methods and their success: a comparison of immigrants and natives in the UK." The Economic Journal 115.507 (2005): F359-F376.

[18] Hellerstein, Judith K., Melissa McInerney, and David Neumark. "Measuring the importance of labor market networks." IZA Discussion Paper 3570 (2008).

[19] Hensvik, Lena and Oskar Nordstrom Skans. "Social networks, employees selection and labor market outcomes." Journal of Labor Economics 34.4 (2016): 825-867

[20] Ioannides, Yannis M., and Linda Datcher Loury. "Job information networks, neighborhood effects, and inequality." Journal of Economic Literature 42.4 (2004): 1056-1093.

[21] Kelejian, Harry H., and Ingmar R. Prucha. "A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances." The Journal of Real Estate Finance and Economics 17.1 (1998): 99-121.

[22] Kelejian, Harry H., Ingmar R. Prucha, and Yevgeny Yuzefovich. "Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: Large and small sample results." Advances in Econometrics: Spatial and Spatio-Temporal Econometrics (2004): 163-198.

[23] Klier, Thomas, and Daniel P. McMillen. "Clustering of auto supplier plants in the United States." Journal of Business & Economic Statistics 26.4 (2008): 460–471

[24] Lee, Lung-fei. "Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances." Econometric Reviews 22.4 (2003): 307-335.

[25] LeSage, James P. "Bayesian estimation of limited dependent variable spatial autoregressive models." Geographical Analysis 32.1 (2000): 19-35.

[26] LeSage, James, and Robert Kelley Pace. Introduction to Spatial Econometrics. CRC Press, 2009.

[27] Lin, Xu. "Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables." Journal of Labor Economics 28.4 (2010): 825-860.

[28] Lin, Tse-Min, Chin-En Wu, and Feng-Yu Lee. "'Neighborhood' Influence on the Formation of National Identity in Taiwan: Spatial Regression with Disjoint Neighborhoods." Political Research Quarterly 59. 1 (2006): 35-46

[29] Loury, Linda Datcher. "Some contacts are more equal than others: Informal networks, job tenure, and wages." Journal of Labor Economics 24.2 (2006): 299-318.

[30] Lovett, Steve, Lee C. Simmons, and Raja Kali. "Guanxi versus the market: Ethics and efficiency." Journal of International Business Studies 30.2 (1999): 231-247.

[31] Marmaros, David, and Bruce Sacerdote. "Peer and social networks in job search." European Economic Review 46.4 (2002): 870-879.

[32] McMillen, Daniel P. "Probit with spatial autocorrelation." Journal of Regional Science 32.3 (1992): 335-348.

[33] Montgomery, James D. "Social networks and labor-market outcomes: Toward an economic analysis." The American economic review (1991): 1408-1418.

[34] Munshi, Kaivan. "Networks in the modern economy: Mexican migrants in the US labor market." The Quarterly Journal of Economics (2003): 549-599.

[35] Patacchini, Eleonora, and Yves Zenou. "Ethnic networks and employment outcomes." Regional Science and Urban Economics 42.6 (2012): 938-949.

[36] Pinkse, Joris, and Margaret E. Slade. "Contracting in space: An application of spatial statistics to discrete-choice models." Journal of Econometrics 85.1 (1998): 125-154.

[37] Smirnov, Oleg A. "Modeling spatial discrete choice." Regional Science and Urban Economics 40.5 (2010): 292-298.

[38] White, Halbert. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." Econometrica 48.4 (1980): 817-838.

[39] Zhang, Xiaobo, and Guo Li. "Does guanxi matter to nonfarm employment?." Journal of Comparative Economics 31.2 (2003): 315-331.