Social Networks and the Labour Market Outcomes of Rural to Urban Migrants in China

Corrado Giulietti[†], Martin Guzi[§], Zhong Zhao[◊] and Klaus F Zimmermann[‡]

Draft October 2010 - PLEASE DO NOT QUOTE

ABSTRACT

In this paper, the role of social networks on the labour market outcomes of rural migrants in China is investigated. Information on the size and quality of the network are derived using data from a unique survey and are used to estimate regression models of wages on social networks. The main findings indicate that for migrants who found a job through informal channels, having a larger network and an employed closest tie is associated with higher wages. These results are corroborated testing the same models for individuals who found the job through more formal channels. Potential threats to the internal validity of the results are discussed and evidence is provided to support the fact that endogeneity issues do not affect the results.

[†] IZA - Institute for the Study of Labor, P.O. Box 7240, 53072 Bonn, Germany. E-Mail: giulietti@iza.org, phone: +49 228 3894 401

[§] IZA - Institute for the Study of Labor, P.O. Box 7240, 53072 Bonn, Germany. E-Mail: guzi@iza.org, phone: +49 228 3894 162

[◊] Renmin University of China, E-Mail: mr.zhong.zhao@, phone: +86 10 82502205

[‡] IZA - Institute for the Study of Labor, P.O. Box 7240, 53072 Bonn, Germany. E-Mail: director@iza.org, phone: +49 228 3894 200

1. INTRODUCTION

Social networks are an important informal channel through which information about job opportunities is transmitted to individuals. The network bridges the informational gap between the worker and the firm by providing details on vacancies to the former and information on workers' attributes to the latter, therefore reducing uncertainty and improving the match. The mitigation of informational asymmetries implies that network characteristics, such as its size and quality influence labour market outcomes of individuals who use social networks in looking for jobs, as theorized by Jackson (2001) and Calvò-Armengol and Jackson (2004, 2007). Networks, as represented by relatives, friends and acquaintances, are particularly important for migrants, since they typically lack information about the host local labour market and about the characteristics of the jobs offered.

The scope of this paper is to explore the relationship between social networks and wages of Chinese rural to urban migrants. Although it has been established that informal networks are helpful in the phase of looking for wage work (Granovetter, 1973; Bayer et al. 2008), the impact of social network on wages could be ambiguous. For example, Montgomery (1991) shows that unemployed who are connected with individuals in well-paid jobs have higher wages than unemployed with ties to less well-paid employees. On the contrary, Datcher Loury (2006) demonstrates that if workers use social networks as a last resort when they cannot obtain better job offers through formal channels, they command lower wages.

The focus on China is relevant not only due to the size of the workforce migrating from rural to urban areas, but also to the fact that in this, as in many other developing countries, the institutional environment is full of uncertainty and frictions, and hence relying on friends and relatives to look for a job is often a more effective way than using formal channels. The data

used in this study show in fact that more than 90 percent of migrants use family and friends to look for their first jobs after migration.

To explore the research questions, detailed measures of size and quality of the social network are introduced. This is achieved by using a unique survey of urban migrants in China, where individuals are asked about the number of contacts (and their location) and the characteristics of the closest contacts (such as employment status and education).

The analysis is carried out using linear regression models. The results indicate that size and quality of the network are positively correlated with wages. The correlation is stronger when controls for industry, occupation, firm size and city are introduced.

Several threats to causality might affect the correlation between social networks and outcomes. In order to address potential endogeneity, several approaches are explored in the paper. First, the results of the benchmark model are compared with those obtained by estimating the same regression on the group of individuals who did not obtain a job through the social networks and for which it is expected that size and quality of the network should not affect wages. Second, evidence is provided to corroborate the fact that measuring the network at a time different than that when the job was obtained does not substantially impact the results.

The paper is organised as follows. Section 2 reviews the major studies on social networks. Section 3 contains a description of the data and summary statistics. The empirical strategy is outlined in Section 4, while in Section 5 results from OLS analysis are presented. Section 6 discusses potential endogeneity of the social networks. Finally, concluding remarks are contained in Section 7.

2. Literature review

The studies on the impact of social networks on labour market outcomes are numerous. Theoretical works have highlighted the role of the social ties in transmitting information on vacancies to unemployed individuals and in producing job referrals to the employers[1]. A corollary of these theories is that both size and quality of the network will influence labour market outcomes (Calvò-Armengol and Jackson (2004, 2007)).

The empirical literature is also flourishing. One of the principal challenges for the empirical analysis is that the network is not observed. Hence, the standard approach is to approximate the social network using information on groups which are known to be socially cohesive and clustered in areas (e.g., ethnic minority groups). As a consequence, regression models are used to correlate labour market outcomes with a proxy for the social network which is capable of capturing geographical or ethnic proximity of individuals (Topa 2001, Clark and Drinkwater 2002, Munshi 2003, Bayer, Ross and Topa 2005, Patacchini and Zenou 2008).

There are very few studies that use direct measures of the network. Cappellari and Tatsiramos (2010) draw information on the employment status of the best friends using the British Household Panel Survey. They find that transitions from unemployment into employment are positively correlated with the number of employed friends. Calvò-Armengol, Patacchini and Zenou (2009) create a network variable based on schoolmates using the National Longitudinal Survey of Adolescent Health; they find that the position in the network is strongly correlated with students' performance.

Similarly, very few works consider both size and quality of the networks. One exception is Wahba and Zenou (2005), who use local measures such as the density of population and unemployment rate to capture, respectively, size and quality of networks in Egyptian

---

[1] A comprehensive survey of theoretical contributions can be found in Jackson (2010).

governorates. They find that density is positively correlated with the probability of finding a job through the social networks. This is however happening up to a certain threshold, beyond which congestion effects exist and hence individuals in particularly dense areas are less likely to find a job through the social networks. The quality of networks, as captured by the unemployment rate, is negatively correlated with the probability of finding a job through the social networks.

Another study where a measure of quality (or strength) is considered is Goel and Lang (2010). The authors use data on recent arrivals to Canada to explore the impact of social network on wages. To approximate the size of the network, they use the share of previous immigrants from each country in metropolitan areas. They capture the network strength using information on close ties; in particular they use an indicator for immigrants who had a contact in Canada before migrating. They find that the impact of obtaining jobs through the social networks on wages is decreasing on their measure of network strength.

Building upon the studies above, this paper uses self-reported measures of social networks to capture both the size and the quality of the network and explore the effect of social ties on wages.

4. Data

The empirical analysis is based on a large scale household survey conducted in China in 2008 within the Rural to Urban Migration in China and Indonesia project (RUMiCI). The survey includes 5,000 migrant households who have moved from rural to urban areas. Migrants are randomly chosen from the fifteen top migration destinations in China. The dataset includes detailed information about socio-demographic characteristics, labour market outcomes, health

4

conditions, major life events and, essential for the analysis, information on the size and the quality of social network.

The survey is conducted at the household level; however, information about social network is only provided by the respondent head of household, and hence only these individuals have been selected. The final sample consists of 1797 migrants aged 16-70 who are formally employed in one occupation (i.e. excluding self-employed and second job holders). Unemployed migrants are not included, as they constitute a very small fraction of the sample. Finally, only migrants who changed their job after migration are included in the analysis. This is to better capture the mechanisms of social networks within each city and to mitigate issues related with the self-selection of migrants[2].

The size of the social network (NS) is measured by the number of greetings that each individual has sent during the Chinese New Year to persons residing in urban areas[3]. Survey respondents provide also information on their closest contacts. In particular, they are asked to give details of up to five persons and rank them according to their importance. Besides details about the closest persons' gender and educational level, the survey provides information on their employment status, urban residence, frequency of contacts and the relationship to the respondent (friend, relative, etc…). The quality of the network (NQ) is represented by an indicator for the employment status of the closest contact. As alternative measure of quality is defined by exploiting information from all closest contacts; in this case, the share of employed individuals among all closest ties is used.

---

[2] The decision of migration and to work in the first job in the city is in most of the cases simultaneous. When considering only individuals who changed their job after migration, the role of the urban social network is better captured.

[3] The question of the RUMiC is designed in a way that it is possible to distinguish between contacts living in rural areas and in cities. Following Zhao et al (2009), observations with reported number of greetings above 200 are dropped.

The analysis will focus mainly on migrants who found a job through the social networks (REF=1). Individuals who obtained their current job through more formal channels (REF=0) will be used to compare the results.

Descriptive statistics are presented in Table 1. The first column summarises the main variables for the group who find the job through the social network; for comparison, figures for the whole sample are also shown. In general migrants are relatively young and have about nine years of education and they have left home roughly seven years before the survey. The percentage of female is quite small, due to the fact that only household head, usually males, are considered in the sample under scrutiny. The average wage is around 6.23 RMB/hour. As for the network measures, each migrants has on average 14 contacts in urban areas; in more than ninety percent of the cases the closest contact is employed. Interestingly, the statistics are very similar for the total sample.

Tab 1 - Summary statistics

	REF=1	Total sample
Age	28.36	27.73
	(9.17)	(8.60)
Female	0.29	0.31
	(0.45)	(0.46)
Years of education	9.23	9.58
	(2.15)	(2.28)
Years since migration	7.44	7.03
	(4.67)	(4.56)
Hourly wage	6.23	6.46
	(3.40)	(3.64)
NS	14	15
	(20)	(22)
NQ	0.93	0.92
	(0.26)	(0.26)
N	1018	1797

Source: RUMIC wave 1, 2008

5. Empirical framework

The relationship between network characteristics and wages is modelled by the means of the following econometric specification:

$$w_{ij} = \alpha + \beta NS_i + \gamma NQ_i + \delta X_i + \tau_j + \varepsilon_{ij}, \qquad (1)$$

where w is the logarithm of wages, NS and NQ represent the size and the quality of the network, X is a set of covariates which include a quadratic in age, years of education, a gender dummy, time since migration and industry, occupation and firm size dummies; τ is an indicator for each city. The parameters of interest are β and γ, which inform about the correlation between wages and, respectively, the size and the quality of the network.

Equation (1) is estimated both for individuals who obtained a job through the social networks and those who obtained through other channel (such as direct application, job agencies, etc...). The latter estimation serves as a robustness check to corroborate the fact that network variables should matter only for those who obtained a job through the social networks.

Potential problems with the estimation of equation (1) are related with the measurement and exogeneity of the network measures. Unobservable individual factors might be correlated with both wages and the network characteristics, leading to biased estimates. The direction and magnitude of this bias depend on the partial correlation of the omitted variable with the error term. For example, if more productive individuals are more likely to have a larger network, then the estimates of β will be upward bias. The same problem might affect the estimates of γ. The standard approach to address unobserved heterogeneity would make use of panel data techniques. This aspect will be explored when data from the second and third wave of RUMiC

will be made available. Controlling for individual effect will however still hinge on the assumption that unobservable characteristics are time unvarying. Another source of endogeneity is related with the timing of the survey. Respondents are typically required to give information on characteristics of their network that refer to the time of the survey, but not to the period when individuals searched or obtained their job. To the extent that size and quality of networks are affected by labour market events, the correlation coefficient will be biased. For example, finding a new job will increase the opportunities of enlarging the network. In this case, controlling for individual fixed characteristics will not address this type of bias. An alternative approach is then to use instrumental variables. For the regression of interest, one would need to find an instrument which is correlated with network characteristics but have no impact on wages[4]. The potential presence of endogeneity will be explored after presenting the OLS results.

6. Results

Table 2 presents the results for five models. In model I, only human capital variables are included. The estimate for network size and quality are positive, significant at 10%. The point estimate for size implies that 10 more contacts are associated with a 1.6% increase in wages. Similarly, having an employed closest tie correlates with a wage premium of nearly 10%. The remaining estimates, have all the expected signs and are significant. In particular, each year since

---

[4] Another problem with equation (1) is related to the measurement error of the network. This would affect particularly size, and has to do with the imperfect recall and with the reporting of rounded numbers of contacts; nevertheless, if anything, measurement error is expected to generate downward bias in the estimates.

migration is associated with something less than a 1% increase in wages[5]. The estimated wage

penalty for females is very similar to other studies based on RUMiC (Deng and Li, 2010).

Table 2 – OLS estimates of wages on network characteristics

	I	II	III	IV	V
NS	0.158 *	0.208 ***	0.246 ***	0.145 ***	0.048
	(0.089)	(0.080)	(0.086)	(0.060)	(0.062)
NQ	0.099 *	0.086 *	0.110 ***	0.113 ***	0.020
	(0.054)	(0.051)	(0.046)	(0.055)	(0.063)
Years of education	0.050 ***	0.049 ***	0.043 ***	0.044 ***	0.034 ***
	(0.007)	(0.007)	(0.006)	(0.006)	(0.007)
Age	0.056 ***	0.052 ***	0.041 ***	0.041 ***	0.039 ***
	(0.010)	(0.010)	(0.009)	(0.009)	(0.017)
Age squared	-0.083 ***	-0.080 ***	-0.068 ***	-0.068 ***	-0.065 ***
	(0.014)	(0.014)	(0.012)	(0.012)	(0.025)
Years since migration	0.008 *	0.010 ***	0.013 ***	0.013 ***	0.010 *
	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)
Female	-0.189 ***	-0.148 ***	-0.143 ***	-0.139 ***	-0.054
	(0.033)	(0.033)	(0.029)	(0.029)	(0.034)
Occupation controls	No	Yes	Yes	Yes	Yes
Industry controls	No	Yes	Yes	Yes	Yes
Firm size controls	No	Yes	Yes	Yes	Yes
City dummies	No	No	Yes	Yes	Yes
N	1018	1018	1018	1018	779
$\overline{R}^2$	0.15	0.22	0.39	0.39	0.37

Notes: robust standard errors in parentheses. * / ** / *** indicate significance at the 10% / 5% / 1% level. Coefficients of network size and age squared are multiplied by 100 for purposes of representation.

When controlling for occupation, industry and firm size, as in model II, the coefficient

for the size of the network increases in magnitude and is strongly significant. On the other hand,

a small decrease in the coefficient for quality is observed.

The third model introduces city dummies in order to control for unobservable city

characteristics. The estimated magnitude of size and quality is now slightly larger than in model

---

[5] It is important to recall, as noted by Borjas (1987) that with a single cross-section it is not possible to estimate simultaneously the parameters for years since migration, age and cohort effects.

II but estimated with better precision. An increase of the network size by 10 contacts is now associated with a 2.5% increase in wages. Similarly, the premium for quality is about 11%. The increase in magnitude of the estimated coefficients suggests the existence of unobservable city characteristics which are negatively correlated with the network. This is confirmed by Figure A1 in the Appendix, where residuals and network size from model II have been aggregated at city level and plotted in a graph. The graph shows a strong negative correlation between the residuals and network size. This is indicative of the fact that, failing to control for unobservable city characteristics, will yield biased estimates. Henceforth, model III will be considered the best specification.

The sensitivity of the results to the definition of social network is tested in model IV. For size, the total number of contacts (urban and non urban) is considered; for quality, the average employment rate among all closest contacts is considered. The estimate for size is still strongly significant, but its magnitude substantially weaker than in model III. This corroborates the fact that for migrants who change job in the city, the urban network is playing a stronger role than the whole range of contacts. On the contrary, there is no appreciable change in the estimate for quality – the t-test for the difference in the estimates being insignificant at 5% level.

Finally, in the last column, a robustness check to validate the findings is carried out. The same regression model is estimated for employed individuals who report finding a job through methods different than the social networks. If the correlations estimated in models I-IV are informative of a mechanism of the type "social networks help finding better paid jobs", then they should not be significant for individuals who did not found a job through the network. In model V this hypothesis is tested for the group $REF = 0$. Results clearly indicate that neither the size, nor the quality of the network are associated with higher wages. Interestingly, the standard errors

are very close to the models of $REF = 1$, while the magnitude of the estimate is essentially zero. Moreover, most of the other covariates have similar estimates to the previous models (the exception being the estimates for female dummy, which is now insignificant).

Discussing causality issues

The estimated correlation is indicative of a strong relationship between the characteristics of the network and the wages of rural to urban migrants. Although the estimates of model V suggest that network characteristics only matter for individuals who found a job through the network, there are still potential issues that might affect a causal interpretation of the findings.

For example, if more productive individuals are more likely to have larger and better networks, the estimates of β and γ will be biased upward. This potential bias is partially mitigated by the inclusion, in the regression analysis, of proxies for productivity such as education and years of migration.

Similarly, it is plausible to assume that individuals with higher wages might want to invest more in leisure with friends and family. This could generate reverse causality between size and quality of the network and wages. An indirect assessment of this can be obtained by looking at the frequency of contacts and if this is related to wages. The graph below shows how the frequency of contacts varies with the wages of individuals in the sample. As it can be seen, there is no appreciable pattern, in the sense that higher wages are not associated with a higher frequency of contacts.
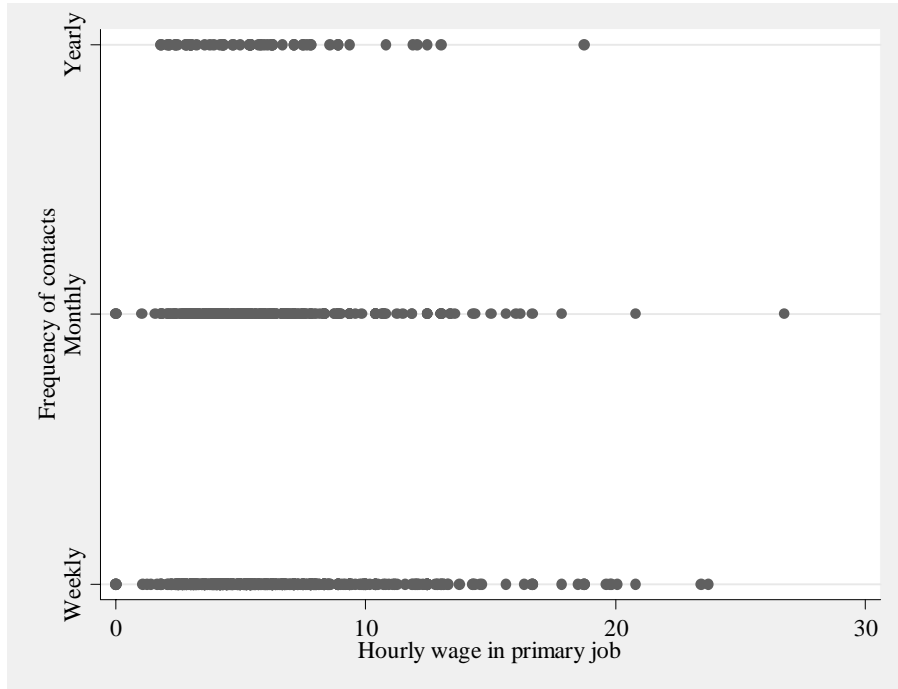
Fig 1 – Frequency of contacts of closest friends and wages

Another potential channel of endogeneity has to do with the timing of the survey. Equation (1) correlates wages and network measured at the time of the interviews. However, the network that should be used in the regression is the one when the job was obtained. A major problem would arise if the network changes from the time of employment and if the reason of this change is correlated with wages. Since all individuals in the sample have changed their job, it could be that their current network in the new working environment is different from the one that helped them finding the job. It must be noted that this problem, if present, would however affect size more than quality: it is unlikely that an individual changes her closest tie because of her own employment status. To investigate more in depth this issue, model III has been re-estimated only for individuals who obtained a job in the previous two years. By considering only migrants whose network is measured at a time closer to when the job was obtained, one could

12

assess how the issues mentioned above affect the results. The estimates for size and quality are 0.212 (s.e. 0.080) and 0.123 (s.e. 0.055). Although the point estimates are slightly different than those in model III, there is a clear indication that results for the subset with a more "stable network are qualitatively similar.

Conclusions

The relationship between social networks and wages of rural to urban migrants in China is explored in this paper. Using unique data collected during the year 2008, direct measures that capture the size and the quality of the networks have been constructed. The size of the network is defined as the number of urban contacts of each individual, while quality is defined by the employment status of the closest tie.

The correlation of wages and social networks has been estimated using OLS techniques, where the preferred specification contains also controls for industries, occupation, firm size and cities. The results of the estimation indicate a significant correlation between wages and both size and quality. The estimated impact of size on wages reveals a premium of 0.25% associated to each additional network member. Moreover, wages are about 11% higher for individuals whose closest contact is employed.

The results have been tested also for a "control" group for which it is expected that network measures should not matter in terms of wages. These are individuals who have obtained jobs through other channels. The results confirm the absence of any correlation between wages and network characteristics for this group. Potential channels that might affect a causal interpretation are discussed and some evidence is showed that the sample under scrutiny is not

substantially affected by endogeneity issues. Nevertheless, a more rigorous approach, such as controlling for time unvarying individual characteristics or instrumental variable techniques is necessary to further bring evidence of the robustness of these results. Further studies that want to extend the analysis of this paper should consider these two important aspects.

Internal migration in China is a growing phenomenon. Shading lights on the role of social networks on labour market outcomes has become essential, given the wide use of informal channel in this country[6]. The results of this paper are suggest that social networks help overcoming some of the frictions present in the labour market. The more likely channel is that of information: the social network helps to reduce the asymmetric information between the employer and the employee, therefore improving the job match. Although China is experiencing continuous reforms, and these will inevitably lead to a stronger formalization of job search channels, personal contacts will remain for long an important channel to obtain better paid jobs. Considering the large wage gap between migrant workers and urban residents largely determined by the restricted access to better jobs for migrants (Meng, 2009), the social network can also be considered as an important channel for reducing the inequality between similarly productive workers.

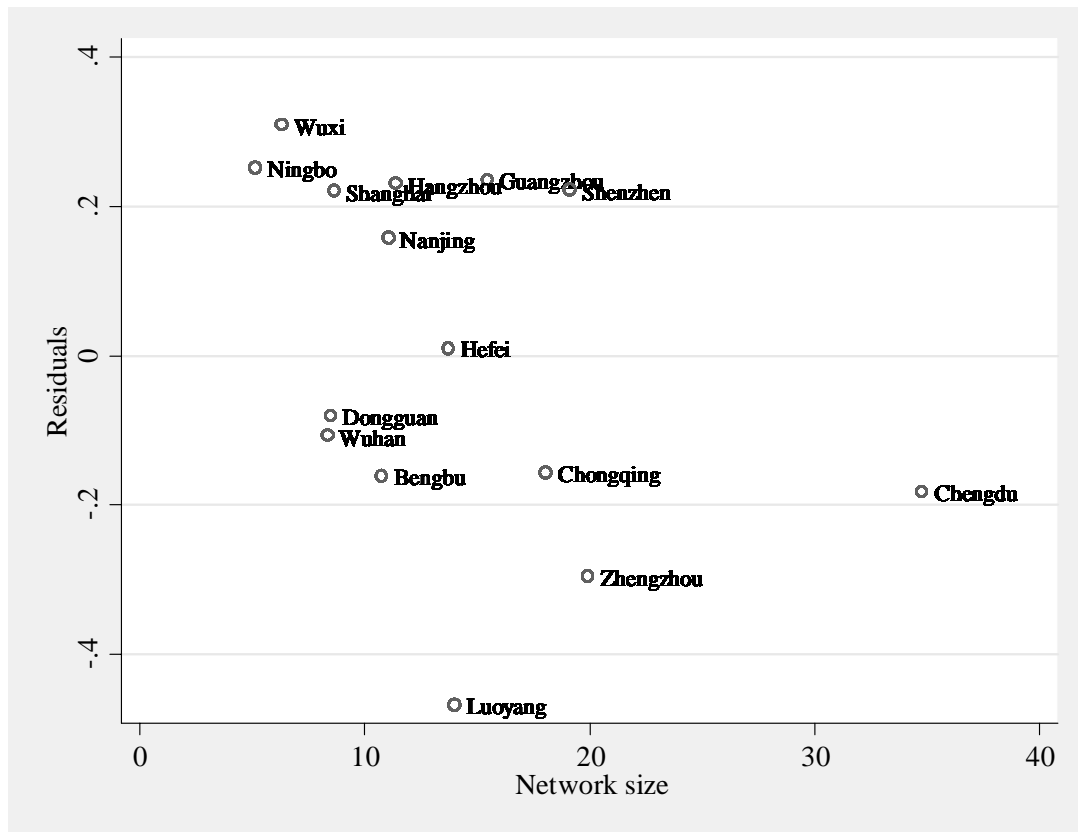[6] China can be defined as a "guanxi" society, where the term guanxi literally means connections.

Appendix

Figure A1 – Residuals from model I and network size

REFERENCES

Battu, H., P., Seaman and Y. Zenou (2004) "Job Contact Networks and the Ethnic Minorities", Research Institute of Industrial Economics Working Paper Series, 628.

Bentolila, S., C., Michelacci, and J. Suarez (2010) "Social Contacts and Occupational Choice, Economica, 77, 20-45.

Calvo-Armengol, A. and M. Jackson (2004) "The Effects of Social Networkson Employment and Inequality", American Economic Review, 94(3), 426-454.

Bian, Yanjie (1994). "Guanxi and the Allocation of Urban Jobs in China," China Quarterly 140, 971-999.

Calvo-Armengol, A. and M. Jackson (2007) "Networks in Labour Markets: Wage and Employment Dynamics and Inequality", The Journal of Economic Theory, 132(1), 27-46.

Datcher Loury, L. (2006) "Some Contacts Are More Equal than Others: Informal Networks, Job Tenure, and Wages", Journal of Labour, 24, 299-318.

Delattre, E. and M. Sabatier (2007), "Social capital and wages: an econometric evaluation of social networking's effects", Labour, 21, 209-236.

Dustmann, C. , A. Glitz and U. Schonberg (2010) "Referral-based Job Search Networks", UCL, mimeo.

Frijters, P. Shields, M.A. and S. Wheatley-Price (2005), "Immigrant job search in the UK: Evidence from panel data", Economic Journal 115, F359-F376.

Granovetter, M.S. (1974, 1995), Getting a Job: A Study of Contacts and Careers, Cambridge, MA: Harvard University Press.

Ioannides, Y. M. and L. Datcher Loury (2004) "Job Information Networks, Neighbourhood Effects, and Inequality", Journal of Economic Literature, 42, 1056-1093.

Jackson, M. O. forthcoming. "An Overview of Social Networks and Economic Applications" in the Handbook of Social Economics edited by J. Benhabib, A. Bisin, and M.O. Jackson, Elsevier Press.

Montgomery, J. D. (1991), "Social networks and labour-market outcomes: toward and economic analysis", American Economic Review 81, 1408-1418.

Mortensen, D.T. and T. Vishwanath (1994), "Personal contacts and earnings: it is who you know", Labour Economics 1, 187-201.

Munshi, K. (2003), "Networks in the Modern Economy: Mexican Migrants in the U.S. labour Market", Quarterly Journal of Economics, 118, 549-599.

Patacchini, E. and Y. Zenou (2004) "Ethnic Networks and Employment Outcomes", IZA Discussion Papers 3331.

Pellizzari, M. (2009) "Do Friends and Relatives Really Help in Getting a Good Job?" Industrial & Labour Relations Review, 63.

Wahba, J., and Y. Zenou (2005), "Density, social networks and job search methods: Theory and application to Egypt", Journal of Development Economics, 78, 443 - 473.

Xin, Katherine R. and Jone L. Pearce (1996). "Guanxi: Connections as Substitutes for Formal Institutional Support," The Academy of Management Journal 39, 1641-1658.

Yueh, Linda (2009). "Self-Employment in Urban China: Networking in a Transition Economy," China Economic Review 20, 471-484.

Zhang, Xiaobo and Guo Li (2003). "Does Guanxi Matter to Nonfarm Employment?" Journal of Comparative Economics 31, 315-331.

Zhao, Yaohui (2003). "The Role of Migrant Networks in Labor Migration: The Case of China," Contemporary Economic Policy 21, 500-511.