

TITLE: SOCIAL NETWORKS AND THE LABOUR MARKET OUTCOMES OF RURAL
TO URBAN MIGRANTS IN CHINA

AUTHORS: CORRADO GIULIETTI,
MARTIN GUZI (SUBMITTER),
ZHONG ZHAO,
KLAUS F. ZIMMERMANN

KEYWORDS: SOCIAL NETWORKS, WAGE, MIGRANTS, CHINA

JEL CLASSIFICATION: J31, J61, O15

E-MAIL: GUZI@IZA.ORG

SOCIAL NETWORKS AND THE LABOUR MARKET OUTCOMES OF RURAL TO URBAN MIGRANTS IN CHINA

CORRADO GIULIETTI[†], MARTIN GUZI[§], ZHONG ZHAO[◇] AND KLAUS F ZIMMERMANN[‡]

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 social networks on wages. The main findings indicate that for migrants who found a job through informal channels, having a large 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 causal interpretation of the results are discussed and evidence is provided to support the fact that endogeneity issues do not affect substantially the findings.

Keywords: social networks, wage, migrants, China

JEL classification: J31, J61, O15

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

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

[◇] Renmin University of China, E-Mail: mr.zhong.zhao@gmail.com

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

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, who 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. This finding is robust to the inclusion of controls for industry, occupation, firm size, city and to a series of alternative specifications that are estimated. Unobservable factors might however affect the correlation between social networks and outcomes, and hence several potential channel of endogeneity are discussed 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: 1) measuring the network at a time different than that when the job was obtained does not substantially impact the results; 2) estimates are not particularly sensitive to the migration choices of the individuals in the sample.

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. 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 social ties in transmitting information on vacancies to unemployed individuals and in producing job referrals to the employers¹. A corollary of these theories is that both size and quality of the network 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 usually 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). Regression models are used to correlate labour market outcomes with a proxy for the social network capable of capturing geographical or ethnic proximity of individuals (Topa 2001, Clark and Drinkwater 2002, Munshi 2003, Bayer et al 2008, Patacchini and Zenou 2008).

Very few studies use direct measures of the network. One exception is Cappellari and Tatsiramos (2010), who draw information on the employment status of the best friends using the British Household Panel Survey. They find that transitions from unemployment into employment and re-employment wages are positively correlated with the number of employed friends.

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 governorates. They find that density is positively correlated with the probability of finding a job through the network. The quality of networks, as captured by the unemployment rate, is negatively correlated with the probability of finding a job through the social networks.

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

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 provides a detailed analysis of the effect of social ties on wages.

3. Data

The empirical analysis is based on two waves of a large scale household survey conducted in China between 2008 and 2009 within the Rural to Urban Migration in China and Indonesia project (RUMiCI). Each wave includes approximately 5,000 migrant households who have moved from rural to urban areas. Migrants are randomly chosen in the fifteen top migrant destination cities in China². The dataset includes detailed information about socio-demographic characteristics, labour market outcomes, health conditions, major life events and, essential for the analysis, information on the size and the quality of social network.

Although the survey contains questions related to all household members, information about social network is only provided by the respondent household head, and hence only these

² Due to the global recession, in 2009 many migrants returned back to their rural villages of origin. For this reason, less than 30% of the observation can be matched over time, compromising the exploitation of the longitudinal aspect.

individuals have been selected. The final sample consists of 3,350 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. The focus on this subset allows capturing better the mechanisms of the social networks which operate in the current area of residence of migrants and helps mitigating issues related to the self-selection of migrants³. For completeness, however, results that include migrants who are still in their first job after migration will be presented too.

The size of the social network (henceforth *NS*) is measured by the number of greetings that each individual has sent during the Chinese New Year to persons residing in urban areas⁴. Observations with more than hundred contacts have been excluded, and *NS* has been normalised to 1. 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 relationship to the respondent (friend, relative, etc...). The quality of the network (henceforth *NQ*) is represented by an indicator for the employment status of the closest contact. Using information from all closest contacts, an alternative measure is also defined; in this case, quality corresponds to the share of employed individuals among all closest ties.

³ The decisions of migrating and working in the first job in the city are 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.

⁴ The question on the number of contacts is designed in a way that it is possible to distinguish between those living in rural areas and those residing in cities.

The analysis focuses 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) are used as comparison group to corroborate the fact that the social network only affects the outcomes of individuals who have obtained a job through it.

Descriptive statistics are presented in Table 1. The first column summarises the main variables of the sample. Each individual has on average 12 urban contacts; in more than ninety percent of the cases the closest contact is employed. Migrants are relatively young, have above nine years of education and have left home roughly 8 years before the survey. The network characteristics for the groups *REF*=1 and *REF*=0 are fairly similar, while the remaining covariates exhibit some differences. One striking aspect is the wage differential between the two groups. Individuals who have obtained jobs through family and friends usually earn (unconditionally) 8% less than those who found the job through more formal channels. In the OLS results, the role of the social network in “compensating” this penalty will be discussed.

Tab 1 - Summary statistics

	Total sample	REF=1	REF=0
<i>NS</i>	0.12 (0.16)	0.13 (0.18)	0.12 (0.15)
<i>NQ</i>	0.93 (0.26)	0.92 (0.27)	0.93 (0.25)
Age	28.93 (9.39)	27.77 (8.26)	29.82 (10.0)
Share of females	0.31 (0.46)	0.33 (0.47)	0.29 (0.45)
Years of education	9.51 (2.38)	9.94 (2.48)	9.18 (2.24)
Years since migration	8.32 (6.17)	7.47 (5.64)	8.97 (6.47)
Hourly wage	6.82 (3.94)	7.12 (4.27)	6.60 (3.64)
N	3350	1447	1903

Source: RUMiCI 2008 and 2009

4. Empirical framework

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

$$w_i = \alpha + \beta_1 NS_i + \beta_2 NQ_i + \gamma X_i + \varepsilon_i, \quad (1)$$

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 includes year and gender dummies, years of education, a quadratic on age, tenure and years since migration. In the preferred specification, indicators for industry, occupation, firm size and cities will be included as well. The parameters of interest are β_1 and β_2 , 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 it 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 matter only for those who obtained a job through the social network.

Potential problems with the estimation of equation (1) are related to the endogeneity of the network measures⁵. Unobservable factors might be correlated with both wages and the network characteristics, leading to biased estimates. The direction and magnitude of this bias depends on the partial correlation of the potentially omitted variable(s) with the error term. For example, if more productive individuals are more likely to have a larger network, then the estimates of β_1 will be upward bias. The same problem might affect the estimates of β_2 . A

⁵ Another potential 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.

standard approach to address unobserved heterogeneity is to make use of panel data techniques; this solution, however, cannot be implemented due to data limitation.

Another source of endogeneity is related to the timing of the survey. Respondents are typically required to give information on characteristics of their network at the time of the survey, but not at 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 coefficients β_1 and β_2 will be biased. For example, finding a new job could increase the opportunities of enlarging the network. Similarly, the characteristics of the network could be correlated with the migration patterns of the individuals. For example, migrants might have a “stronger” network if they migrated relatively close to their home village; similarly, their contacts could be less stable if they changed several cities, because this requires creating new relationships.

All the potential channels of endogeneity discussed above will be explored in the appropriate section.

5. Results

Table 2 presents the results for five models. Columns I-III are related to the group $REF = 1$, IV to the group $REF = 0$ and V to both groups. In the first model, only human capital variables are included. The estimates for network size and quality are positive, significant at the 1% level. The point estimate for size implies that 10 more contacts are associated with a 1.8% increase in wages. Similarly, having an employed close tie correlates with a wage premium of above 12%.

The remaining estimates have all the expected signs and are significant⁶. 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
<i>NS</i>	0.184 *** (0.078)	0.248 *** (0.075)	0.309 *** (0.069)	0.077 (0.070)	0.105 (0.068)
<i>NQ</i>	0.124 *** (0.040)	0.118 *** (0.038)	0.125 *** (0.035)	0.027 (0.042)	0.018 (0.043)
Years of education	0.038 *** (0.005)	0.037 *** (0.005)	0.035 *** (0.004)	0.035 *** (0.005)	0.036 *** (0.003)
Age	0.044 *** (0.007)	0.041 *** (0.007)	0.035 *** (0.007)	0.023 * (0.012)	0.031 *** (0.006)
Age squared	-0.066 *** (0.010)	-0.066 *** (0.010)	-0.058 *** (0.009)	-0.041 *** (0.018)	-0.053 *** (0.008)
Years since migration	0.010 (0.006)	0.011 * (0.006)	0.015 *** (0.006)	0.020 *** (0.008)	0.016 *** (0.004)
Years since migration squared	-0.032 (0.021)	-0.034 * (0.020)	-0.036 * (0.019)	-0.066 *** (0.028)	-0.042 *** (0.015)
Tenure	0.056 *** (0.010)	0.053 *** (0.009)	0.046 *** (0.008)	0.049 *** (0.010)	0.047 *** (0.006)
Tenure squared	-0.204 *** (0.061)	-0.199 *** (0.058)	-0.175 *** (0.052)	-0.137 *** (0.056)	-0.167 *** (0.037)
Female	-0.191 *** (0.023)	-0.149 *** (0.024)	-0.123 *** (0.021)	-0.041 * (0.024)	-0.09 *** (0.016)
<i>REF</i>					-0.125 *** (0.054)
<i>NSxREF</i>					0.191 *** (0.093)
<i>NQxREF</i>					0.109 *** (0.055)
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	1903	1903	1903	1447	3350
\bar{R}^2	0.18	0.24	0.37	0.37	0.37

Notes: robust standard errors in parentheses. * / ** / *** indicate significance at the 10% / 5% / 1% level. All models include a constant and a dummy for year 2009. Coefficients for age squared, years since migration squared and tenure squared are multiplied by 100 for purposes of representation.

⁶ 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.

In column II, indicators for occupation, industry and firm size are introduced; in column III city dummies are also added. While the estimates for the quality of the network are remarkably similar, the coefficient for the size increases substantially. An increase of the network size by 10 contacts is now associated with a 3.1% increase in wages. This sizeable growth in the magnitude of the coefficient suggests the existence of unobservable characteristics at city and at job level which are negatively correlated with the network. Henceforth, model III will be considered the best specification.

A first test to corroborate the causal interpretation of the results reported above is offered in column IV. Here, the model in column III is estimated for individuals who report finding the job through methods different than the social network. If the correlations estimated in models I-III 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. Results of column IV 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 with $REF=1$, while the magnitude is essentially zero. Moreover, most of the other covariates have similar estimates to the previous models, with the exception of the dummy for females, exhibiting a smaller estimate⁷.

Finally in column V, observations of models III and IV are pooled and an indicator for REF is interacted with network size and quality. The scope of this exercise is to investigate if social networks are helpful to reduce the wage gap existing between individuals finding jobs through formal and informal methods and that was observed in Table 1. In this model, the reference group correspond to individuals for which $\{NS, NQ, REF=0\}$. Estimates confirm that

⁷ In unreported results, the model in column V has been estimated using an indicator for the type of formal search channel used (direct application=1, assigned by the government=0). Results are essentially the same.

size and quality do not matter for individuals who found the job through formal methods. The dummy for REF approximates the wage gap: the large negative estimate is indicative of the fact even after controlling for individual, job and city characteristics the penalty is still substantial. However this value refers to individuals without network. The interaction terms reveal that the penalty is substantially reduced for a sufficiently large network and if the closest contact is employed. This is best seen through graphical representation. Figure 1 depicts the predicted wages for different cases. The solid line represents migrants for which $REF=0$ and either size or quality are set equal to zero. In both graphs, the dashed line shows a shift of the predicted densities to the left, corresponding to the penalty associated with $REF=1$. The dotted line, finally, shows that for a sufficiently large size of the network (an interval around the mean has been selected) the wage distribution shifts back to the right, somewhat compensating the initial penalty. In the case of network quality, the “premium” is even larger.

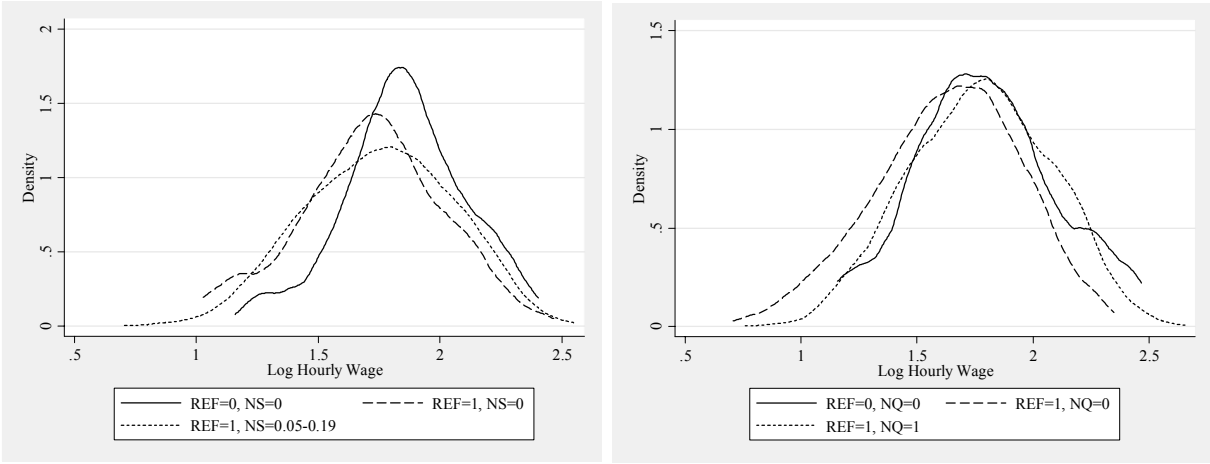


Fig 1 – Density of wages by network success and quality

5.1 A closer look at wages

In this subsection, the effect of the social network along the wage distribution is explored. Following an approach similar to Goel and Lang (2010), quantile regression for model III in Table 2 is estimated. The top panel of Table 3 reports the estimates for the 0.25, 0.5 and 0.75 quantiles, respectively.

Table 3 – Quantile regression

	Q1	Q2	Q3
<i>NS</i>	0.299 *** (0.090)	0.337 *** (0.075)	0.308 *** (0.097)
<i>NQ</i>	0.108 *** (0.039)	0.131 *** (0.046)	0.156 *** (0.052)
<i>NS</i>	0.295 *** (0.094)	0.320 *** (0.101)	0.304 *** (0.095)
<i>NQ</i>	0.119 *** (0.042)	0.146 *** (0.043)	0.158 *** (0.066)

Notes: robust standard errors in parentheses. * / ** / *** indicate significance at the 10% / 5% / 1% level. Models include all variables of Table 2, col. III.

While the estimate for size is not particularly sensitive to the location in the wage distribution, the magnitude of quality is somewhat larger in the higher quantiles. One possible explanation is that there are factors positively correlated with both wages and the quality of the network, such as omitted characteristics of contacts (Cappellari and Tatsiramos 2010). Hence in the second panel, the same model is estimating using additional information about the closest contact, such as marital and urban status, education and gender. However, controlling for these characteristics produces even higher estimates. One hypothesis is that unobserved ability is at work. High ability individuals are more likely to have high ability contacts, determining a relatively higher correlation between quality and wages. While this supposition is not testable with the data used in this paper, it remains an interesting question which requires further research.

5.2 Sensitivity analysis

In Table 4, the sensitivity of the results to the definition of social networks and of the selected sample is investigated. In column I, the preferred specification is estimated using alternative definitions of network size and quality. For size, the total number of contacts (i.e., including individuals outside the urban area) is considered. Quality is defined as the average employment rate of all closest contacts. The estimate for size is still strongly significant, but its magnitude substantially weaker than in Table 2. This confirms 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, the estimate for quality is somewhat larger.

Table 4 – Robustness checks

	I	II	III	IV
<i>NS</i>	0.153 *** (0.043)	0.414 *** (0.081)	0.260 *** (0.055)	0.365 *** (0.068)
<i>NQ</i>	0.142 *** (0.038)	0.129 *** (0.043)	0.102 *** (0.027)	0.112 *** (0.036)
Years of education	0.036 *** (0.004)	0.029 *** (0.006)	0.033 *** (0.003)	0.032 *** (0.005)
Age	0.034 *** (0.007)	0.039 *** (0.008)	0.043 *** (0.005)	0.037 *** (0.007)
Age squared	-0.057 *** (0.009)	-0.063 *** (0.010)	-0.066 *** (0.007)	-0.059 *** (0.010)
Years since migration	0.015 *** (0.006)	0.016 *** (0.006)	0.010 *** (0.004)	0.012 * (0.006)
Years since migration squared	-0.037 *** (0.019)	-0.042 *** (0.020)	-0.023 (0.015)	-0.029 (0.022)
Tenure	0.049 *** (0.008)	0.052 *** (0.009)	0.040 *** (0.005)	0.048 *** (0.009)
Tenure squared	-0.188 *** (0.052)	-0.191 *** (0.055)	-0.117 *** (0.025)	-0.179 *** (0.063)
Female	-0.120 *** (0.021)		-0.104 *** (0.016)	-0.123 *** (0.022)
N	1903	1351	3245	2111
\bar{R}^2	0.37	0.35	0.37	0.30

Notes: robust standard errors in parentheses. * / ** / *** indicate significance at the 10% / 5% / 1% level. All models include a constant and a dummy for year 2009. Coefficients for age squared, years since migration squared and tenure squared are multiplied by 100 for purposes of representation.

The second column restricts the sample to males only. While the estimate for size is substantially larger, the magnitude associated with quality is remarkably similar to the benchmark model.

Column III includes individuals who are still in their first job after having migrated. These individuals were excluded to mitigate the potential confounding effect of the network at their place of origin, which might have played a role both for helping the individual migrating and finding a job. In this enlarged sample, both estimates for size and quality are slightly smaller than those in Table 2.

Finally, in column IV, the sample is expanded to include self-employed individuals. Although the mechanism with which the social network might operate for these individuals could be somewhat different than that of wage earners, it is interesting noting that by including them the estimate for size is somewhat larger, while that for quality slightly smaller.

To summarise, the robustness checks presented in this subsection show that although the estimates are to some extent sensitive to the definition of the network and to the sample of individuals selected, the pattern of the results is remarkably similar to that of the preferred specification.

6. Addressing 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. Estimates of the preferred model suggest that network characteristics only matter for individuals who found a job through the network; however 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_1 and β_2 will be biased upward. In the regression analysis this potential bias is partially mitigated by the inclusion of proxies for productivity, such as education, years of migration and tenure.

Similarly, 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 how this is related to wages. The graphs below show 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.

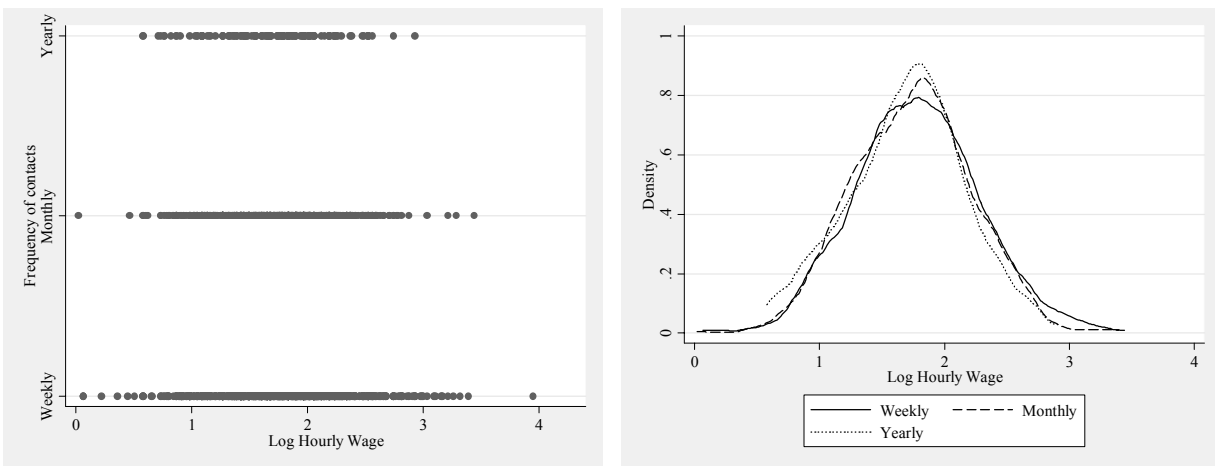


Fig 2 – Frequency of contacts of closest friends and wages

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 changed from the time of employment and if the reason of this change is correlated with wages. For example, if after becoming employed an individual is more likely to have contacts with employed individuals, OLS regression could overestimate the true impact of the network. Since all individuals in the sample have changed their job after migration, 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 the new employment status. To investigate this issue more in depth, potential channels that might affect the size or the quality of the network are investigated. Three such channels are explored: the first is related to the timing at which the individual obtained the job and at which the network is measured; the second and third explore the impact that migration patterns might have on network characteristics. To this aim, the benchmark model is re-estimated by interacting the network characteristics with an indicator NZ which captures, respectively, tenure, migration within the same province and the number of cities where the migrant lived before the current one.

In Column I of Table 5 size and quality are interacted with a dummy for tenure, which equals one if the individual has been in the current job for more than 2 years. Hence there reference group corresponds to individuals for which the network is measured at a time closer to when the job was obtained. The estimate for size is somewhat larger than the benchmark model, while that of quality is slightly smaller; while these two estimates might be less affected by endogeneity bias than those in the benchmark model, the pattern is still very similar. Moreover, there are no remarkable differences with the group with longer tenure, as showed by the insignificance of the interaction terms.

Table 5 – Interaction models

	I	II	III
<i>NS</i>	0.382 *** (0.097)	0.334 *** (0.088)	0.202 *** (0.085)
<i>NQ</i>	0.111 *** (0.043)	0.149 *** (0.046)	0.134 *** (0.046)
<i>NSxNZ</i>	-0.166 (0.125)	-0.047 (0.135)	0.319 *** (0.129)
<i>NQxNZ</i>	0.038 (0.072)	-0.049 (0.070)	-0.026 (0.069)
<i>NZ</i>	0.016 (0.076)	0.083 (0.073)	-0.016 (0.068)
N	1903	1903	1903
\bar{R}^2	0.37	0.37	0.37

Notes: robust standard errors in parentheses. * / ** / *** indicate significance at the 10% / 5% / 1% level. Models include all variables of Table 2, col. III.

The second column explores whether migrating within or outside the province of residence affects the role of networks on wages. It could be that some unobserved factors such as cognitive skills or personality traits are correlated with the choice of the migration destination, in particular with migrating far from the origin village. To test this hypothesis, an indicator for migrating in a city within the same province of origin is introduced and interacted with the characteristics of the network. In this case, the reference group is constituted by individuals who migrated relatively close to their rural village. Interestingly, both estimates for size and quality are larger than those of the benchmark model in Table 3, suggesting that network effects are stronger for individuals that moved closer to their place of origin. The relatively large effect is confirmed by the negative sign of the interaction term which are, however, not significant.

The third column analyses the hypothesis that the effect of network on wages depends on the migration history of the individuals. The number of the city that a migrant moved before settling in the current location might have an ambiguous impact on the network. On the one hand, migrants who lived in many cities might have expanded their own network by knowing

more people; on the other one, they could also have a less stable network or could weaken the relationships with their close ties. To explore these hypothesis, an indicator for having migrated only in the current city is interacted with the network characteristics. The reference group, hence, corresponds to individuals who lived in more than one city. The coefficient for size is substantially smaller than the one in the benchmark model. Therefore, it seems that individuals that moved directly to the current location have a relatively higher “premium” for the dimension of their urban network. This is confirmed by the fact that the interaction term for size is significant; the sum of the two coefficients – main effect and interaction – is well above the one in the preferred specification. For quality, however, having changed more than one city is associated with a relatively larger premium than the average. The interaction term is negative, but insignificant.

Additional insight on the potential role of migration patterns is offered in Figure 3. Here the estimates and the standard error for size and quality related to the benchmark model are estimated as a function of the years since migration. If the characteristics of the network have changed substantially after migration and if this change is also correlated with the outcome variable, one would expect to see substantial changes over time. The figure shows very limited changes in the point estimates and with no particular monotonic pattern.

To summarise, in this section several channel of potential endogeneity have been explored, yielding supporting evidence of the causal interpretation of the results. Although estimates are slightly different across the models explored, the pattern outlined in the benchmark model is substantially unchanged. While panel data analysis or instrumental variable could have yield more unbiased results, they could not be implemented with the available data⁸.

⁸ Several instruments for network size and quality have been explored, including distance from the origin, village characteristics and personality traits. While these variables were not successful in explaining networks’

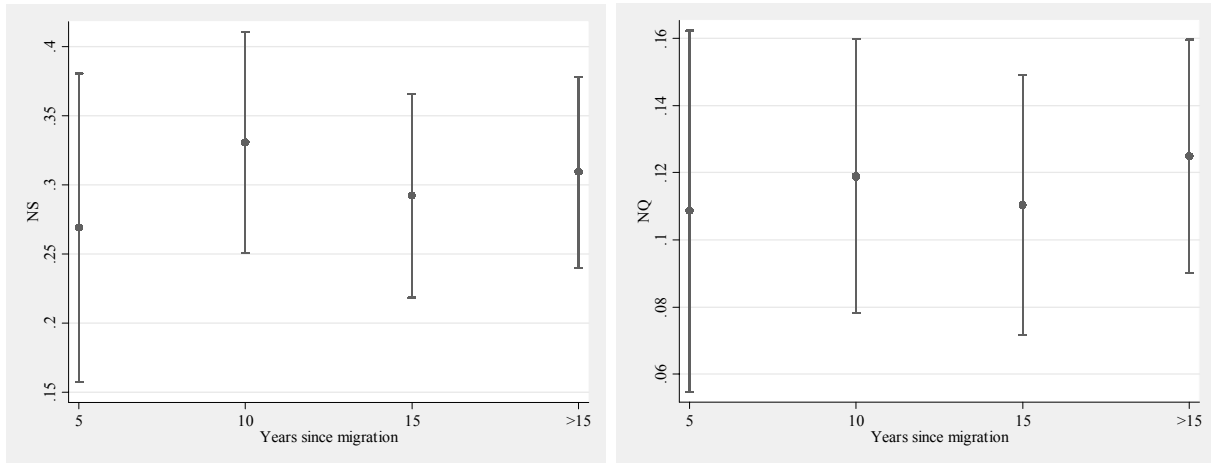


Fig 3 – The effect of network size by years since migration

7. 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 2008 and 2009, direct measures of size and the quality of the networks have been constructed.

The correlation of wages and social networks has been estimated using OLS techniques, where the preferred specification contains 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 about 3% associated to 10 additional urban network members. Wages are about 12% higher for individuals whose closest contact is employed.

The results have been tested also for a comparison group for which it is expected that network measures do not matter on wages. These are individuals who have obtained jobs through

characteristics, it is not possible to exclude that similar, unobservable factors, could be used in an instrumental variable approach.

more formal search channels. The results confirm the absence of any correlation between wages and network characteristics for this group. Robustness checks have been carried out and potential channels that might affect the causal interpretation of the results have been discussed, showing that the findings are not particularly affected by the selection of the sample or by endogeneity issues. In order to obtain results less prone to endogeneity bias, a more rigorous approach – not possible here because of available data – such as controlling for time unvarying individual characteristics or using instrumental variable techniques, is necessary. 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. Shining lights on the role of social networks on labour market outcomes has become essential, given the wide use of informal channel in this country⁹. The results of this paper 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 method 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 might also be an important channel for reducing the inequality between similarly productive workers.



ACKNOWLEDGMENTS

Collection of the Rural Urban Migration in China (RUMiCI) data used in this paper is financed by IZA, ARC/AusAid, the Ford Foundation and the Ministry of Labor and Social Security of China

⁹ China can be defined as a “guanxi” society, where the term guanxi literally means connections.

REFERENCES

- Bayer P, Ross S.L. and G. Topa (2008) “Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes”, *Journal of Political Economy*, 116(6), 1150-1196.
- Calvo-Armengol, A. and M. Jackson (2004) “The Effects of Social Networks on Employment and Inequality”, *American Economic Review*, 94(3), 426-454.
- 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.
- Cappellari L. and K. Tatsiramos (2010) “Friends' Networks and Job Finding Rates”, *IZA Discussion Papers* 5240
- Clark K. and S. Drinkwater (2002) “Enclaves, neighbourhood effects and employment outcomes: Ethnic minorities in England and Wales”, *Journal of Population Economics*, 15(1), 5-29.
- Datcher Loury, L. (2006) “Some Contacts Are More Equal than Others: Informal Networks, Job Tenure, and Wages”, *Journal of Labour*, 24, 299-318.
- Deng Q. and S. Li (2010), “Wage Structures and Inequality among Local and Migrant Workers in Urban China”, in *The Great Migration: Rural-Urban Migration in China and Indonesia*, Meng, Manning, and Li eds.
- Granovetter, M.S. (1974, 1995), *Getting a Job: A Study of Contacts and Careers*, Cambridge, MA: Harvard University Press.
- 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.

- Meng, X., Lee L., and P. Frijters (2010) “Jobs, Working Hours and Remuneration Packages for Migrant and Urban Workers”, in *The Great Migration: Rural-Urban Migration in China and Indonesia*, Meng, Manning, and Li eds.
- Montgomery, J. D. (1991), “Social networks and labour-market outcomes: toward an economic analysis”, *American Economic Review* 81, 1408-1418.
- 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 (2008) “Ethnic Networks and Employment Outcomes”, IZA Discussion Papers 3331.
- Topa, G. (2001) “Social Interactions, Local Spillovers and Unemployment”, *Review of Economic Studies*, 68(2), 261-95.
- 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.