Social Networks and Employment Performances: Evidence from Rural – Urban Migration in Vietnam

Dang, Duc Anh

Australian National University

26 October 2015

Social Networks and Employment Performances:
Evidence from Rural – Urban Migration in Vietnam

Duc Anh Dang

This version: September, 2015

Abstract

This paper considers the effects of social network on income and employment dynamics of rural-urban migrants in Vietnam. Estimation of a causal effect is challenging because unobserved factors affects both employment performances and social networks. I address this endogeneity problem by using instrumental variable method. The results suggest that social networks improve migrant's earnings and make wage earners willing to change their jobs.

Keywords: migration, social network, employment
JEL: D02, J61

1. Introduction

Social networks are considered an important informal mechanism through which information about job opportunities is transmitted. By solving information and commitment problems in environments where markets are inefficient, social networks bridge the informational gap between the worker and the firm by providing information on both sides, therefore reducing uncertainty and improving the match (Munshi, 2011). Social networks even are much more important to migrants from rural to urban regions who found difficult to adapt with new environment. Because developing economies are associated with poorly functioning markets, networks through 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. Therefore, it is not surprising that much of the empirical literature on social networks has been situated in these economies.

The significant role of social contacts in obtaining employment has long been recognized. However, what less well understood are the possible effects on subsequent wages or the decision of workers to change their employment position of using such networks. While positive wage effects from using social networks are sometimes found, this is far from universal and has led to consideration being given as to why using social networks may appear to lower wages (Delattre & Sabatier, 2007). One explanation is due to training costs (Pellizzari, 2010). Firms may want to spend extra effort in using formal means of filling posts (for example, advertising or using recruitment agencies), rather than informal means (relying on referrals from incumbent workers for hiring new workers). It happens more common where the costs of the posts remaining unfilled are high, perhaps due to high training costs and consequent high wages. A second explanation bases on the argument of job seeker impatience. Those keen to find employment quickly may use job contacts, sacrificing potentially higher wages from better matched posts for quicker entry into work (Bentolila et al., 2010). Given the ambiguous theoretical predictions about the impact of using job contacts on subsequent wages, it is worthwhile to test the hypothesis empirically.

In addition, empirical analyses on the effects of social networks have been plagued by various conceptual and data problems. Many commonly used datasets lack information on the structure and composition of individuals of social networks. Analyses are further complicated by various endogeneity issues such as the refection problem and selection bias. A refection problem arises when individual and network members' outcomes are determined

simultaneously, which inherently confounds the measure of networks' influence. Selection bias leads to a correlated unobservable problem if people tend to associate with others based in part on some unobserved group characteristics they favor. In such a case, an observed positive association between an individual's outcome and those of their associated network members may not be causal but rather due to some unknown factors that affect both social links and individual's own labor market outcomes (see Munshi, 2011 for a review).

The literature seeks to control for a number of individual characteristics and economic conditions that could be correlated. However, the obvious concern is that unobserved variables remain unaccounted for. Observed individual characteristics such as age, education, and occupational experience may not capture traits such as initiative and diligence that play a critical role in determining the individual's market outcomes. Empirical studies on migrant networks in the labor market, for example, use the number of friends or relatives in cities to measure the strength of the individual's network. If individuals with greater ability at the destination have a greater propensity to migrate, then social groups with higher ability will have larger migration rates. The number of friends and relatives in the city could then be contaminated by the migrant's unobserved ability, giving rise to a spurious network effect. Using received help or the extent of social interaction to measure the network suffers from potential selectivity bias as well, since we would expect more able individuals to receive more help and to be better connected in equilibrium. Using fixed effects will fully capture constant individual characteristics but will fail to account for unobserved factors that vary over time.

This study seeks to improve our knowledge about the relationship between social networks and income dynamics and decision to change the job position of rural to urban migrants within a developing country through the use of a novel source of internal migration data. The focus on Vietnam is relevant due to the fact that as in many other developing countries, its institutional environment is full of uncertainty and frictions, and hence relying informal information to get a better job could be a more effective way than using formal channels. From my knowledge, this paper could be one of the first studies investigating the importance of social network in shaping migrants' earning dynamics.

To explore the research questions, I take a question asking about the number of urban callings that migrant made during the Lunar New Year as the proxy for social network. The analysis is carried out using linear regression models. I find that people who have more urban callings

seem to get better paid job. In addition, wage-earners with more social network exhibit their willingness to change their employment to run their own business.

Several threats to causality might affect the correlation between social networks and outcomes. Omitted individual or destination factors would bias the estimates. In addition, the suspected causal relationship between wages and migrant networks can be mutual, leading to simultaneity problem. In order to address potential endogeneity, I use the historical weather variation as an instrument for social networks. The inherent characteristics of weather variation at the departure of migrants provide a basis for the instrument's exogeneity. Weather variation at rural regions would decrease the expected agricultural output of the households exposed to it, and thereby encourage members of these households to migrate to cities. Therefore, it is unlikely that weather variation could impact on migrant's income at the destination, other than through the social network.

The results from the instrumental variable (IV) approach suggest that the social network helps to improve labour incomes and make migrants be willing to change their job. To confirm the findings from IV approach, I carry out some sensitivity tests on the validity of instrumental variable. To address the concern of whether the exclusion restriction is satisfied, I perform a number of falsification tests that examine the reduced-form relationship between weather variation and incomes. The results show that I still find that social networks estimated by the IV approach have positive effects on income dynamics.

I begin, in Section 2, by reviewing literature on the impact of social networks on incomes. Section 3 then documents the data used. The identification strategy employed is discussed in detail in section 4. Section 4 also reports OLS estimates of the relationship between social networks and dynamics of income and employment. I then turn to the issues of causality OLS estimates that control for an extensive set of observable characteristics, as well as the IV estimates. Section 5 offers concluding thoughts.

2. Literature Review

There are a numerous studies on the relationship between the social networks and labor market outcomes (see Jackson (2010) for a comprehensive survey). 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. For example, Granovetter (1995)

argues that many people find their jobs through social relations, not only through formal channels. Social networks allow people looking for work to gather better information about the availability of jobs as well as the characteristics of the job.

However, there are no common agreements on the possible effects on subsequent wages. While Granovetter confirms the positive effects of social networks on incomes and the appropriateness of the work is found through the relationship, other studies have not found the differences in incomes between those who seek their job in the formal channels and others through social networks (such as Lin, 1999; Mouwn, 2003; Franzen & Hangartner, 2006). In addition, one of the principal challenges for the empirical analysis is that the network is not observed. Therefore, literature seeks to approximate the social network by 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 rather than networks (Topa, 2001; Clark & Drinkwater, 2002; Munshi, 2003, Bayer, Ross & Topa, 2008, Patacchini & Zenou, 2008).

There are very few studies that use direct measures of the network. Cappellari & 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 & 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. Wahba & Zenou (2005), who use density of population to capture size of networks in Egyptian, show 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. Another study by Goel and Lang (2010) using data on recent arrivals to Canada find that the impact of obtaining jobs through the social networks on wages is decreasing on their measure of network strength.

In this paper, I follow Giulietti et al. (2010) to use direct measure of social network, which is self-reported of migrants on the number of greetings that they make during the Lunar New Year to urban people, to explore the effect of social ties on income dynamics and temptation to change jobs.

3. Data Sources and Description

Data is taken from the Vietnam Rural Urban Migration Survey (VRUMS) in 2013. The VRUMS is conducted by the Central Institute for Economic Management of the Ministry of Planning and Investment, Vietnam under the technical support from Research School of Economics at the Australian National University. The objective of VRUMS is to gather sample information on rural-urban migration in Vietnam anchoring to 2012 Vietnam Household Living Standards Survey (VHLSS2012). The survey intends to help estimating the effects of large scale rural-urban migration in the process of economic development and to be able to assist the government in formulating the right economic and social policies to facilitate the rural-urban migration and urbanization process.

This survey collects data and information from 869 migrant households and individuals from rural areas of the whole Vietnam to urban areas mainly in Ha Noi and HoChiMinh City and surrounding areas for work. Migrant households are defined as a household whether members of the migrant worker must have family relationship or relatives, live with the respondent, share incomes and expenditures at the time of interview. These households or individuals come from the rural household base of Vietnam Household Living Standard Survey (VHLSS) in 2012, whereas VHLSS is national representaive survey undertaken biannually by the General Statistical Office of Vietnam (GSO).

Because VHLSS2012 used 2009 Population and Housing Census as a sampling frame, we expect that VHLSS2012 will be representative of population. In addition, the information on rural-urban migration taken from VHLSS2012 rural households will be well representative of rural-urban migration in Vietnam. The procedure of survey is implemented as followings: Calling the rural households in the VHLSS2012 to see whether they have migrant workers to either HaNoi, HoChiMinh City or other surrounding urban areas. If the households/families have migrants, ask and get information and contact details of the migrants (including both long-term (over 6 months) and short term (6 months and shorter migrants) originating from the rural households; get the permissions of rural households for contacting their household member migrants in the cities; then calling these migrants or migrant households and making appointments for interviews based on a VRUMS2013 standard questionnaire.

Due to the complex procedure of survey process, we expect that there are some reasons that may make the sample biased. ***First,*** there are quite a lot of rural phone numbers that have been missing from Vietnam's Household Living Standard Survey. ***Second,*** there are some rural households with incorrect numbers. ***Third,*** some rural households refused to provide or provide wrong phone number of migrants in the cities. ***Fourth,*** selected migrants may refuse to participate. All of above reasons will affect the quality of survey data and may lead to sample biases due to selection problems. For example, we may expect the survey may miss out rural poor households who do not have landline or mobile phones. Therefore, our survey may be well representative of richer rural households rather than overall migrant population. In addition, the individuals who choose not participating may systematically different from those who do. These sampling errors may make the final sample be biased.

To check the potentials of sample selection biases, we compared some main characteristics of rural households who have household member migrating for work in Hanoi or HoChiMinh City but not be appeared in two surveys (VRUMS2013 and VHLSS2012). The detail strategy is as followings:

First, we took information and characteristics of both long-term and short-term migrants of rural household to HN and HCM small sample of VHLSS2012 (income and expenditure survey). In VHLSS2012, there are sections asking information about two types of migrations: long-term and short-term migrants. Long-term migrants are defined as former members of rural households "who (1) left the household within the last 10 years or (2) may have left at an earlier time but household still considers important for your household, in terms of obligations to old parents living with your household or financial support to your household". We picked people who take up work in the Hanoi or HCM within the first 6 months and did not include people attending school or being economically inactive. There are 307 migrants who moved to work in HN or HCM and took up a job within the first 6 months.

Short-term migrants are regarded as people who are member of rural households and leaving the household for "at least 6 month for work since 2002". The short-term migrant included both returned migrants and unreturned migrants who are still working far way. Therefore, VHLSS2012 does not cover households who have people who migrate to those cities less than 6 months. For short-term migrants, we only took people who have not returned to households. In total, we have 319 migrants, of which 307 long-term migrants (who used to be rural household members) and 12 unreturned short-term migrants (who are currently rural household members) who are currently working in HN or HCM.

Second, we then match rural households of these 319 short-term and long-term migrants in VHLSS2012 with those of 243 migrants in VRUM2013. We found that only 2 short-term and 70 long-term migrants in VHLSS2012 coming from the same rural households in

VRUMS2013. Therefore, 172 households with migrants in VRUMS2013 have not covered by VHLSS2012. In other words, VHLSS2012 may underreport the number of migrants, especially who are very short-migrants leaving households for less than 6 months. However, at the same time, there are 244 migrants reported in VHLSS2012 but have not been included in VRUMS2013. This number of migrants may come from sampling errors as explained above.

We then assess whether two groups of rural households who have missing migrants from each survey to see whether they are balanced on some observable characteristics. Such information would provide tentative evidence how selection biases are. The table below provides balanced tests of some main demographic indicators. The first two columns show the simple means of two samples in these indicators. The two samples are indifferent in means in most indicators, except age which is marginally different. However, the distribution of age in two groups of rural households is almost similar[1]. The balanced tests confirmed that although there are sampling errors in VRUMS2013, the sample still assures its randomness and representativeness.

Difference-in-Means by VHLSS2012 and VRUMS2013

Variables	Means		Difference in means	
	Unmatched VHLSS2012	Unmatched VRUMS2013	T-statistics	P- value
Minority	0.08	0.06	0.49	0.62
Age of HH Head	54.36	52.27	1.74	0.08
Gender (Male=1)	0.76	0.81	1.18	0.24
Marital Status of HH Head (Married=1)	0.99	0.98	0.99	0.32
Proportion of HH members under 16	0.05	0.06	0.99	0.32
Proportion of HH members over 65	0.14	0.12	0.81	0.42
Years of schooling of HH Head	7.08	7.27	0.53	0.60
Rural household locations	2.01	2.11	1.21	0.23
Hour wage rate in the past 30 days of HH Head	16.05	17.90	0.95	0.34
Number of Observations	244	156		

Note: Rural household locations are coded as 1 if in the North, 2 if in the Centre and 3 if the South, respectively. Individual information is from rural household heads. The Unmatched VRUMS2013 are rural households who have migrants found in VRUMS2013 but those migrants have not been reported in VHLSS2012. The Unmatched VHLSS2012 are rural households who report having migrants in VHLSS2012 but those migrants have not been found in VRUMS2013. The third column is the T-statistics of means differences of two samples.

The survey covers both the migrants who currently and used to members of rural households who are working in these cities. From my knowledge, this is one of the best surveys available

---

[1]We also try to take only rural households who have migrants found in VRUMS2013 but have not been reported in VHLSS20112 *since 2002* to make two groups more consonant in the definition on rural-urban migration. The number of these rural households is 144 and the results are similar.

on rural-urban migration in Vietnam in the sense that it captures the most complete range of working migrants in the cities: both long-term (over 6 months) and short term (6 months and shorter such as temporary migrants and seasonal) migrants that have not been fully taken into account in other surveys.

Another novel of the survey is that it includes questions asking about both the current incomes and the first job income that allow me to investigate the income dynamics of migrants. The surveys also includes comprehensive information on household and personal characteristics, detailed health status, employment, training and education of adults and children, social networks, family and social relationships, life events, and mental health measures of the individuals.

To investigate the social network of migrants, I takes of the number of urban callings that migrant made during the Lunar New Year as a proxy for 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 exact wording of the question is as follows: "During the last Lunar New Year, how many people in total did you send your greetings. Among them, _____person(s) is (are) currently living in the city".

To investigate the impact of social network on employment transition, I use information from the question asking about migrants who are currently wage earners and reported that they are thinking to change their jobs by running their own business. The exact question is: "Have you ever thought of running a business of your own?". Respondents could either answer never think about it or they already had. They also had the option of answering that they never think about it seriously "because it would be very difficult". I construct a measure that takes on the binary value of 0 and 1, where 0 corresponds to the response 'Never' and 'Never seriously' and 1 to the response 'Yes, I have'. I then use OLS to estimate a linear probability model. Another strategy is to estimate a logit model. As I discuss below, the estimates are qualitatively identical if I pursue this alternative strategy.

The summary statistics for migrant's household heads are presented in Table 1. On average, migrants have about ten years of education and they have left home more than eight years before the survey. The percentage of female is quite small because only household head, usually males, are considered in the sample under scrutiny. Current jobs of migrants have

been better paid than their first job in the cities. There are more than one-third of current wage earners are willing to change their job and run their own business. Regarding the network measure, each migrant has on average ten contacts in urban areas.

4. Estimating Equations and Empirical Results

4.1 OLS Estimates

I begin by estimating the relationship between social network and employment dynamics. My baseline estimating equation is:

$$Outcome_{ic} = \alpha_i + \beta Social\_network_i + X_i^{'}\Gamma + Z_i^{'}\Pi + \varepsilon_{ic} \text{ (1)}$$

where i indexes individual, c original community in rural regions. The variable $Outcome_{ic}$ denotes measures of change income of migrants from first job in the cities and the current job of migrant i, originating from commune c in rural regions. It also measures the willingness of current wage earners in changing their job. $Social\_network_i$ represents the number of callings to urban people that migrants make in the Lunar New Year. β is the coefficient of my main measures that indicates the relationship between the social network and change in the individual's outcomes. I expect β to be positive and statistically significant. α_i measures the individual's ability, which directly determines his outcome in the destination labor market ε_{ic} is an exogenous labor demand shock, which reflects the idea that individual migrants from given location could be endowed with specific skills that channel them into particular segments of the labor market even when networks are absent.

The vector $X_i^{'}$ controls a set of individual-level covariates, which includes the characteristics of household head, such as age, age squared, years of education, a gender variable indicator, a dummy variable for people who are ethnic minorities and dummies for being employed in state or foreign sectors. The vector $Z_i^{'}$ consists of other variables, such as duration of staying in the cities and change in job types.

In addition, the main explanatory variable, $Social\_network_i$, in Equation (1) may have similar effects on people coming from the same commune in rural areas. Given the potential

for within-group correlation of the residuals, I clustered the standard errors for a potentially arbitrary correlation between individuals in the same original commune in rural areas.

Table 2 reports OLS estimates of the impacts of social networks on change in migrant's incomes between the current and first-job incomes. Of which, the income from the first-job is adjusted for inflation because migrants in the city arrived in various points of time. In all estimations, standard errors are adjusted for clustering of observations of the same commune. In Column 1, I estimate the relationship between the number of urban callings with migrant's earnings. The estimates show that the number of urban callings has a positive impact on change in migrant's incomes. This is consistent with the hypothesis that the social network positively affects individual's earnings. At the same time, the coefficient is statistically significant. Realizing that there are some outliers, that may drive the results (see Figure 1), in Column 2, I exclude migrants with more than 100 callings. The effect of social networks is three times higher. The result indicates that on average one more contacts on average is associated with nearly one percent increase income changes.

Estimates of other variables are also consonant with the results from other studies and expectations. People working in the foreign sector have better change in incomes. While the age and age square variables both significantly influence change in wages at the 0.05 significance level; however, the directions of the two effects are different. This implies a diminishing marginal effect of the age.

Columns 3-5 of Table 2 report estimates of Equation (1) with the additional controls included. In Column 3, I control for duration of migrants in the cities. Network effects will depend on both their size and their vintage, since migrants who have been in the cities longer are more established and better social networks. I also add the variables that measure the change in the job-types of current and first job of migrants in the cities. I classify the jobs with contract as formal ones and code them as 1 and the others as 0. In the two last columns, I control for the interaction between the number of callings and different cohorts of duration of staying the cities. The results indicate that the number of urban callings bring more benefits to more established migrants, especially to ones who stay in the cities less than 3 years and from 5 to 8 years. Based on the estimates from Column 4, the point estimate for this cohort implies that

one more contact on average is associated with a 0.9 percent increase in income changes, which is equal to 10.6 per cent of the sample average for log of migrant's current incomes.[2]

OLS estimates examining the relationship between the number of urban callings and wage-earners' willingness to run their own business are reported in Table 3[3]. The specification reported includes the similar control invariables as in the income equation except I include log of current incomes which may affect the job decision of migrants. The estimates indicate a positive and significantly statistical relationship between social network and willingness to run their own business by current wage earners except the last column. However, the coefficient of the number of urban callings in the last column has the same sign and magnitude with that in the first column. An increase in standard errors may reflect a loss of precision arising from significant attrition of observations when I add more control variables. I also check for robustness to alternative estimation methods. Because the responses for thoughts in running their own business are restricted in range, they may not be normally distributed. To overcome this problem, I use a logit model instead. The results from the logit model in Appendix A.1 are qualitatively identical to our OLS estimates. The marginal effects are consistent with those estimated by OLS and statistically significant.

4.2 Identifying the Causal Relationship

To consistently estimate an OLS model, the explanatory variable of interest, the network size, should be uncorrelated with individual unobserved ability. This assumption, however, is very likely to be violated. These 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. In addition, the wages and the network size may be mutually determined, leading to the simultaneity bias. For example, high wage in the cities may assist migrants in expanding their network or people may encourage more friends and relatives to migrate that also enlarge their network. 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

---

[2]The mean of log of migrant's current income is 8.44. The effect is calculated as 0.9/8.44=0.106 or 10.6 per cent of mean.
[3]The main reason to use OLS rather other estimators such as logit is that the coefficients estimated by OLS are easier to be interpreted.

not to the period when individuals searched or obtained their job. To the extent that size of networks is affected by labor market events, the correlation coefficient will be biased. In this case, controlling for individual fixed characteristics will not address this type of bias. Another problem with equation (1) is related to the measurement error of the network. This would affect size of network, 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.

In this section, I try to assess whether the correlations documented to this point are causal by using an instrument for social networks. For the regression of interest, one would need to find an instrument which is correlated with network characteristics but have no direct impact on wages[4]. Origin characteristics that generate exogenous variation in the size of the migrant network, but are uncorrelated with labor demand shocks at the destination, could be valid instruments. This paper exploits the intensity of weather variations in the origins of migrants as an instrument. Under certain assumptions, the weather variation can be seen as an exogenous shock to the size of outflow migrants from the rural regions because the occurrence and destructive power of weather in a certain areas are random.

Although there are limited scholarly explorations on the association between migration behavior and the environment, the close relationship is intuitive. Weather variation would decrease the expected agricultural output of the households exposed to it, and thereby encourage members of these households to migrate. Consequently, being hit by natural hazards will trigger the outflow migration from rural regions. In other words, the higher the intensity of a natural disaster is, the more rural households are likely to move to urban areas.

Natural hazards are common in Vietnam. Rural households in Vietnam are exposed to many natural risks that could potentially threaten their livelihoods. In addition, since the majority of households in rural areas rely on agricultural activities, they will experience fluctuations in agriculturally derived income from exogenous natural shocks such as drought, floods, pest infestation and livestock disease (CIEM, DOE, ILSSA & IPSARD 2007). Here, I take rainfall variation as a proxy for the riskiness of natural environment. Literature indicates that the year-to-year rainfall variations capture the effects of hazardous natural environment such as floods, typhoons and storms in Vietnam reasonably well. For example, Benson (1997) shows that typhoons are typically

associated with heavy rainfall and strong winds. Each typhoon accounts for about 10 to 15 per cent, and sometimes even more, of annual rainfall and causes flash floods and landslides. In addition, heavy rainfall causes rivers to fill and potentially results in flooding. Therefore, I expect the more typhoons and storms or natural disaster in general one region suffers from, the more rainfall volatility it has.

The data on rainfall variability are obtained from weather stations in 87 districts produced by the Institute of Meteorology and Hydrology.[4] These stations are allocated to capture the best variation of weather within regions. For the remaining districts without stations, the weather conditions are assumed to be similar to districts sharing the same borders with them but have a weather station. The reason for this strategy is that stations are expected to gauge significant weather variations in different regions. Therefore, weather data from one station could be used to measure neighboring districts with similar conditions.

Monthly rainfall observations (from January to December) were available over 30 years for each station from 1975 to 2006. For each month, I calculated the standard deviation over the 30 years for each station, and obtained the average rainfall deviation of each station over 12 months to investigate year-to-year rainfall fluctuations. More specifically, consider rainfall variable x, station i, month m and year y, and define x_{imy} as the value of x in station i in month m in year y. For each month m, I compute the standard deviation of x_{imy} over all years (denoted s_{im}), which measures the month-specific variability of variable x in station i. To obtain a compound measure of year-to-year variability for station i I average s_{im} over the twelve months.

The reason to take weather variation for a long time is twofold: (i) the migration data covers both long-term and short-term migrants with the longest duration of migrants in the cities about 48 years. Therefore, the reasons to migrate may originate from historical natural shocks rather than the present ones and (ii) rainfall variation may closely relate with other bio-geographic conditions such as land quality and terrain ruggedness. All of these can have both direct and indirect effects on agricultural incomes and living conditions or rural people that create incentives for migration.

---

[4]On average, there are nearly 12 districts in one province. The area of each district ranges from 27.8 to 3677.4 square kilometres and the mean is 660 square kilometres. For the period 1975–2006, the data is taken from Thomas et al. 'Natural disasters and household welfare: evidence from Vietnam', Policy Research Working Paper, 2010, World Bank.

To be even more cautious about the exogeneity, the working assumptions are set up in such a way as to make the IV estimates as reasonable and cautious as possible: (i) Rainfall variation in the rural regions is assumed to not affect any labor market conditions at the destination; (ii) Unobserved individual heterogeneity (e.g., ability, preferences, health conditions, reservation wage), is assumed to be uncorrelated with the intensity of natural disasters. This is important especially because natural disasters may lower the reservation wage of a potential migrant. Given the substantial income gaps between rural and urban Vietnam, however, even the lowest wage levels in cities are still attractive to villagers. This makes the reservation wage argument inapplicable in this setting (see also Munshi, 2003).

Table 4 reports the results of the first-stage IV estimates. Because the distribution of the weather variation is highly left skewed, with a small number of observations taking large values, I report estimates using the natural log of the weather variation measure. All the coefficients have the expected sign. The further weather variation, the higher the number of urban callings. All weather variation coefficients are statistically significant.

In the second stage, the estimated coefficients for social networks are significant and positive. The magnitude of the coefficient in the IV estimation does not change substantially if other controls are included, ranging from 0.072 to 0.089. The results in Column 4 of Table 5 show that the result is still significant when all other variables are controlled. The magnitude of the IV estimates is higher than those from the OLS estimates. One explanation for this is that the attenuation bias, resulting from measurement errors, leads OLS estimates to be biased towards zero, and IV results in an increase in the magnitude of the coefficient.

The F-test for an excluded instrument is also reported. The F-statistics in Table 4 range from 5.72 to 9.47, suggesting that for some specifications there may be a potential concern about weak instruments. If a proposed instrument is not strongly correlated with the endogenous variables then the instrumental variable two-stage least squares (IV-2SLS) estimates may be somewhat biased toward ordinary least squares (OLS) estimates (Bound, Jaeger & Baker, 1995; Staiger & Stock, 1997). For this reason I also use the LIML Fuller Instrumental variable estimation method that is a bias-corrected limited information maximum likelihood estimator and provides the most unbiased estimates for inference purposes when the instrument is potentially weak (Stock, Wright & Yogo, 2002). The regression provides similar estimates.

In addition, because the IV estimate mainly applies to the subgroup of individuals more affected by natural disasters, the IV estimate can be interpreted as a Local Average Treatment Effects (LATE) (Imbens & Angrist, 1994). If the IV estimate is to be interpreted as a class of LATE, we must raise the question about the mechanism that explains how natural disasters influence migration and why networks effects differ across individuals. One possible mechanism is that that less able people (in terms of earning ability at the destination) are more responsive to natural disasters since they have relatively lower ability to compensate for losses due to natural disasters. That is, people of lower earning ability are more likely to be pulled out from the rural regions by natural hazards. If this is the case, the IV estimate can be interpreted as a weighted average network effect and the weight for less able migrants is relatively higher.

The IV results for the impacts of social network on willingness to run their own business by wage-earners are also consistent with prediction. In Table 7, all estimated coefficients for social network are significant and positive, indicating that increase in social network makes wage-earners to be more willing to run their own business.

Sensitivity Tests

My IV strategy rest on the assumptions that the weather variation does not affect demand labor at destination and migrant's earning ability and preferences. The first assumption is likely to be satisfied. I have not found any literature that documents that weather variation may create a mass migration from rural regions that have big impacts on labor market condition in the cities over the last 30 years. However, the second assumption may be violated. Recent studies, such as (Durante, 2009; Dang, 2012; Cameron, 2011) shows that natural disasters may change individual behaviors. They find that people who live in more frequency of natural disasters and weather variation trust other people more. In addition, they seem more risk averse. If trustworthiness and risk attitudes correlate with migrant's incomes then the IV estimates will be biased and inconsistent. Natural conditions also may affect health's migrants. To test all of these possibilities, I control for several variables including individual trust, risk preferences and health conditions. The results show that the results are almost identical.

Another way to test this likelihood is to estimate the reduced form relationship between weather variation and migrant's incomes. The estimation results are reported in Table 8.

When I examine the reduced form, I find a strong positive and highly significant relationship between weather variation and change in migrant's incomes. This correlation is consistent with the first-and-second stage IV estimates in Tables 4 and 5, individuals who migrate from more weather variation regions tend to create more social networks and this in turn help them find a better jobs and incomes.

5. Conclusion

Despite the proliferation of research seeking to identify the mechanisms and measure the magnitude of internal migration, little emphasis has been placed on probing the direct causal effects of migrant networks on labour market outcomes at the destination. This paper explores the causal effects of the size of migrant networks on income and employment dynamics among migrants in Vietnam. It complements recent research on the labour market effects of migrant networks in other developing economies.

Controlling for the unobserved factors influencing migration decision, identification is achieved through instrumenting the network size by the intensity of weather variation occurring in the sending commune of the migrants. The empirical results show that the size of the migrant network significantly improves the wages of migrants and makes wage earners willing to run their own business.

The results of this paper suggest that social networks help overcoming some of the frictions present in the labour market. On possible channel is that the social network helps to reduce the asymmetric information between the employer and the employee, therefore improving the job match. The results also show that although there is a stronger formalization of job search channels in developing countries, personal contacts will remain an important channel to obtain better paid jobs.

References

Bayer, P., & Ross, S.L., & Topa, G., 2008. 'Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes,' *Journal of Political Economy*, University of Chicago Press, 116(6): 1150-1196.

Benson, C 1997. *The economic impact of natural disasters in Vietnam*, Working Paper, Overseas Development Institute.

Bentolila, S., Michelacci, C. and Suarez, J., 2010. `Social contacts and occupational choice', *Economica*, 77 (305): 20-45.

Bound, J., Jaeger, D.A & Baker, R.M., 1995. Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak, *Journal of the American Statistical Association*, 90 (430): 443-450.

Calvo-Armengol, A., Eleonora Patacchini & Yves Zenou, 2009. 'Peer Effects and Social Networks in Education,' *Review of Economic Studies*, Oxford University Press, 76(4): 1239-1267.

Cameron, L & Shah, M., 2011. Risk-taking behavior in the wake of natural disasters, Forthcoming, *Journal of Human Resources.*

Cappellari, L., & Tatsiramos, K., 2010. '*Friends' Networks and Job Finding Rates*,' CESifo Working Paper Series 3243, CESifo Group Munich.

Clark, K., & Drinkwater, S., 2002. 'Enclaves, neighbourhood effects and employment outcomes: Ethnic minorities in England and Wales,' *Journal of Population Economics*, 15(1): 5-29.

CIEM, DOE, ILSSA & IPSARD 2007. *Characteristics of the Vietnamese rural economy: evidence from a 2006 rural household survey in 12 provinces of Vietnam*, Statistical Publishing House, Hanoi.

Dang, A., 2012. *Cooperation makes Beliefs: Weather Variation and Sources of Social Trust in Vietnam*, ANU College of Business and Economics Working Paper.

Delattre, E. & Sabatier, M., 2007. 'Social capital and wages: An econometric evaluation of social networking's effects', LABOUR, 21(2): 209-236.

Durante, R., 2009. *Risk, Cooperation and the Economic Origins of Social Trust: an Empirical Investigation*, Job Market Paper, Brown University.

Goel, D., & Lang, K., 2009. '*Social Ties and the Job Search of Recent Immigrants*,' NBER Working Papers 15186, National Bureau of Economic Research, Inc.

Granovetter, M., 1995. '*Getting a Job: A Study of Contacts and Careers*', University of Chicago Press, Chicago.

Giulietti, C., Guzi, M., Zhao, Z., & Zimmermann, K., 2010. *Social Networks and the Labour Market Outcomes of Rural to Urban Migrants in China*, Working paper.

Franzen, A., & Hangartner, D., 2006. 'Social Networks and Labour Market Outcomes: The Non-Monetary Benefits of Social Capital', *European Sociological Review*, 22(4): 355-368.

Imbens, G., & Angrist, J., 1994. 'Identification and Estimation of Local Average Treatment Effects', *Econometrica*, 62 (2): 467-475.

Jackson, M. O., 2010. 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.

Lin, N., 1999. 'Social Networks and Status Attainment', *Annual Review of Sociology*, 25: 467-87.

Mouw, T., 2003. 'Social Capital and Finding a Job: Do Contacts Matter?', *American Sociological Review*, 68: 868-898.

Munshi, K., 2011. 'Labor and Credit Networks in Developing Economies', in Jess Benhabib, Alberto Bissin, and MatthewO. Jackson, eds. *Handbook of Social Economics*. New York: Elsevier.

Munshi, K., 2003. 'Networks in Modern Economy: Mexican Migrants in the U.S Labor Market', *Quarterly Journal of Economics*, 118(2): 549-599.

Patacchini, E. & Zenou, Y., 2004. '*Ethnic Networks and Employment Outcomes'*, IZA Discussion Papers, 3331.

Pellizzari, M., 2010. `Do friends and relatives really help in getting a good job?', *Industrial and Labor Relations Review*, 63 (3): 494-510.

Stock, J., Wright, J., and Yogo, M., 2002. 'A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments', *Journal of Business and Economic Statistics,* 20: 518-529.

Staiger, D. and Stock, J., 1997. 'Instrumental Variables Regression with Weak Instruments', *Econometrica*, 65: 557-586.

Zhang, J. & Zhao, Z., 2013. 'Measuring the income-distance tradeoff for rural-urban migrants in China', *IZA Discussion Papers* 7160, Institute for the Study of Labor (IZA)

Topa, Giorgio, 2001. 'Social Interactions, Local Spillovers and Unemployment,' *Review of Economic Studies*, 68(2): 261-95.

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

Table 1. Summary Statistics

VARIABLES	Obs	Mean	Std. Dev.	Min	Max
Log of change in current and first job incomes	560	0.586	0.958	-4.94	6.42
Thought about running their own business	511	0.346	0.476	0	1
Switching in job types	555	0.110	0.477	-1	1
Years of schooling	546	9.656	2.953	0	12
Age	560	29.673	8.282	16	66
Age squared	560	948.973	588.115	256	4356
Gender	560	0.689	0.463	0	1
Minority	560	0.030	0.172	0	1
Working in state sector	559	0.118	0.323	0	1
Working in foreign sector	559	0.191	0.394	0	1
Number of urban calling	485	11.526	20.818	0	270
Duration of staying in the cities	558	7.050	5.232	0.167	35
Trust	540	0.129	0.336	0	1
Risk preferences	540	5.25	1.788	0	10
Health condition	560	0.954	0.211	0	1

Table 2. OLS Estimates. Relationship between number of urban callings and income dynamics

VARIABLES	(1)	(2)	(3)	(4)	(5)
	Log of change in current and first job incomes				
Number of callings to urban people	0.003*	0.009**	0.009**	0.009**	-0.010
	(0.002)	(0.004)	(0.004)	(0.004)	(0.008)
Years of schooling	-0.009	-0.014	-0.017	-0.016	-0.016
	(0.014)	(0.015)	(0.016)	(0.016)	(0.017)
Age	0.111***	0.106***	0.083**	0.082**	0.075**
	(0.031)	(0.031)	(0.036)	(0.036)	(0.035)
Age squared	-0.002***	-0.002***	-0.001***	-0.001***	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male:=1; w/o:=0	-0.089	-0.077	-0.071	-0.054	-0.054
	(0.094)	(0.094)	(0.095)	(0.098)	(0.100)
Minority:=1; w/o:=0	-0.377***	-0.393***	-0.401***	-0.377***	-0.395***
	(0.133)	(0.134)	(0.140)	(0.140)	(0.152)
State ownership:=1; o/w:=0	-0.118	-0.114	-0.105	-0.113	-0.136
	(0.113)	(0.112)	(0.113)	(0.114)	(0.116)
Foreign ownership:=1; o/w:=0	0.124	0.135	0.135	0.115	0.135
	(0.136)	(0.137)	(0.139)	(0.142)	(0.139)
Duration of staying in the cities			0.017	0.016	0.006
			(0.021)	(0.021)	(0.026)
Switching from informal to formal job types				0.107	0.088
				(0.094)	(0.096)
Number of urban calling X Less than 3 years staying in the cities					0.019**
					(0.007)
Number of urban calling X From 3 to less than 5 years staying in the cities					0.015
					(0.011)
Number of urban calling X From 5 to less than 8 years staying in the cities					0.028**
					(0.012)
Number of urban calling X From 8 to less than 12 years staying in the cities					0.023
					(0.014)
Constant	-0.994**	-0.928*	-0.584	-0.608	-0.417
	(0.478)	(0.473)	(0.544)	(0.546)	(0.515)
Observations	470	466	465	461	461
R-squared	0.047	0.053	0.060	0.061	0.074

Notes: ***, ** and * indicates significance level of 1%, 5% and 10% respectively against a two sided alternative. Clustered standard errors are in brackets.

Table 3. OLS Estimates. Relationship between number of urban callings and employment transition

VARIABLES	(1)	(2)	(3)	(4)	(5)
			Wage earners want to run a business		
Number of callings to urban people	0.005***	0.005***	0.005**	0.004*	0.006
	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)
Years of schooling	0.098**	0.099**	0.130***	0.103**	0.100**
	(0.040)	(0.040)	(0.045)	(0.047)	(0.047)
Age	0.009	0.006	0.007	0.001	0.001
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
Age squared	0.039***	0.025**	0.025*	0.020	0.021
	(0.012)	(0.013)	(0.014)	(0.013)	(0.014)
Male:=1; w/o:=0	-0.001***	-0.000**	-0.000**	-0.000*	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minority:=1; w/o:=0	0.148	0.155	-0.013	-0.003	-0.012
	(0.123)	(0.123)	(0.127)	(0.129)	(0.127)
State ownership:=1; o/w:=0	-0.025	-0.039	-0.011	-0.041	-0.035
	(0.055)	(0.054)	(0.063)	(0.063)	(0.063)
Foreign ownership:=1; o/w:=0	-0.009	-0.015	-0.043	-0.056	-0.061
	(0.051)	(0.051)	(0.054)	(0.055)	(0.055)
Duration of staying in the cities		0.011***	0.011***	0.008**	0.007
		(0.004)	(0.004)	(0.004)	(0.005)
Switching from informal to formal job types			0.058	0.045	0.046
			(0.050)	(0.049)	(0.049)
Log of total income per month				0.159***	0.161***
				(0.055)	(0.055)
Number of urban calling X Less than 3 years staying in the cities					-0.004
					(0.006)
Number of urban calling X From 3 to less than 5 years staying in the cities					-0.002
					(0.005)
Number of urban calling X From 5 to less than 8 years staying in the cities					-0.004
					(0.006)
Number of urban calling X From 8 to less than 12 years staying in the cities					-0.000
					(0.007)
Constant	-0.500**	-0.274	-0.276	-1.416***	-1.450***
	(0.196)	(0.209)	(0.230)	(0.462)	(0.469)
Observations	607	604	504	502	502
R-squared	0.052	0.070	0.081	0.095	0.096

Notes: ***, ** and * indicates significance level of 1%, 5% and 10% respectively against a two sided alternative. Clustered standard errors are in brackets.

Table 4. IV Estimates. Impacts of number of urban callings on income dynamics (First Stage)

VARIABLES	(1)	(2)	(3)	(4)
	Number of callings to urban people			
Log of rainfall variation	7.630***	6.825***	6.078**	6.078**
	(2.479)	(2.452)	(2.537)	(2.540)
Male:=1; w/o:=0	-0.842	-0.678	-1.275	-1.264
	(1.144)	(1.212)	(1.284)	(1.288)
Age	0.652**	0.178	0.145	0.150
	(0.299)	(0.327)	(0.352)	(0.348)
Age squared	-0.009**	-0.003	-0.003	-0.003
	(0.004)	(0.004)	(0.005)	(0.005)
Minority:=1; w/o:=0	1.840	1.547	1.184	1.119
	(3.810)	(3.477)	(3.554)	(3.484)
Years of schooling	0.949***	0.863***	0.861***	0.860***
	(0.144)	(0.146)	(0.148)	(0.148)
State ownership:=1; o/w:=0		0.548	0.915	0.987
		(1.971)	(2.004)	(2.005)
Private ownership:=1; o/w:=0		-0.567	-0.835	-0.767
		(1.230)	(1.294)	(1.295)
Duration of staying in the cities		0.288**	0.263**	0.265**
		(0.111)	(0.125)	(0.125)
Level of risk preferences			0.588*	0.594*
			(0.350)	(0.352)
Almost people can be trusted:=1			1.905	1.928
			(1.529)	(1.528)
Health condition				-0.928
				(2.236)
Constant	-55.309***	-42.789***	-40.463**	-39.690**
	(15.633)	(15.858)	(16.284)	(16.309)
Observations	467	465	447	447
F-test for excluded instrument	9.47	9.06	5.74	5.72

Notes: ***, ** and * indicates significance level of 1%, 5% and 10% respectively against a two sided alternative. Clustered standard errors are in brackets.

Table 5. IV Estimates. Impacts of number of urban callings on income dynamics (Second Stage)

VARIABLES	(1)	(2)	(3)	(4)
	Log of change in current and first job incomes			
Number of callings to urban people	0.072**	0.077**	0.089*	0.089*
	(0.035)	(0.039)	(0.047)	(0.047)
Male:=1; w/o:=0	-0.034	-0.017	0.065	0.064
	(0.120)	(0.128)	(0.151)	(0.151)
Age	0.067	0.071*	0.068	0.067
	(0.043)	(0.043)	(0.041)	(0.041)
Age squared	-0.001*	-0.001*	-0.001**	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)
Minority:=1; w/o:=0	-0.460	-0.488	-0.553	-0.543
	(0.287)	(0.305)	(0.361)	(0.355)
Years of schooling	-0.077**	-0.076**	-0.095**	-0.095**
	(0.036)	(0.037)	(0.043)	(0.043)
State ownership:=1; o/w:=0		-0.289	-0.379	-0.390*
		(0.215)	(0.232)	(0.235)
Private ownership:=1; o/w:=0		-0.098	-0.167	-0.177
		(0.164)	(0.163)	(0.164)
Duration of staying in the cities		-0.005	0.013	0.012
		(0.024)	(0.021)	(0.021)
Level of risk preferences			-0.039	-0.040
			(0.050)	(0.050)
Almost people can be trusted:=1			0.007	0.004
			(0.202)	(0.203)
Health condition				0.133
				(0.266)
Constant	-0.273	-0.299	0.011	-0.100
	(0.678)	(0.730)	(0.748)	(0.772)
Observations	467	465	447	447

Notes: ***, ** and * indicates significance level of 1%, 5% and 10% respectively against a two sided alternative. Clustered standard errors are in brackets.

Table 6. IV Estimates. Impacts of number of urban callings on employment transition (First Stage)

VARIABLES	(1)	(2)	(3)
	Number of callings to urban people		
Years of schooling	0.775***	0.651***	0.678***
	(0.123)	(0.131)	(0.133)
Male:=1; w/o:=0	-0.569	-1.087	-1.078
	(0.979)	(0.972)	(0.970)
Age	0.592**	0.263	0.254
	(0.259)	(0.296)	(0.297)
Age squared	-0.008**	-0.005	-0.005
	(0.004)	(0.004)	(0.004)
Minority:=1; w/o:=0	-2.424	-2.255	-2.397
	(2.172)	(2.107)	(2.133)
State ownership:=1; o/w:=0	0.882	0.829	0.965
	(1.814)	(1.975)	(1.977)
Private ownership:=1; o/w:=0	-0.762	-0.904	-0.916
	(0.961)	(1.038)	(1.039)
Current job through social network:=1; o/w:=0		-0.695	-0.711
		(0.930)	(0.928)
Duration of staying in the cities		0.191**	0.197**
		(0.087)	(0.087)
Level of risk preferences		0.238	0.227
		(0.252)	(0.252)
Almost people can be trusted:=1			1.694
			(1.189)
Log of rainfall variation	8.911***	9.102***	8.710***
	(1.994)	(2.100)	(2.106)
Constant	-59.772***	-55.434***	-53.407***
	(13.061)	(13.716)	(13.716)
Observations	607	568	568
F-test for excluded instrument	19.98	17.11	17.11

Notes: ***, ** and * indicates significance level of 1%, 5% and 10% respectively against a two sided alternative. Clustered standard errors are in brackets.

Table 7. IV Estimates. Impacts of number of urban callings on employment transition (Second Stage)

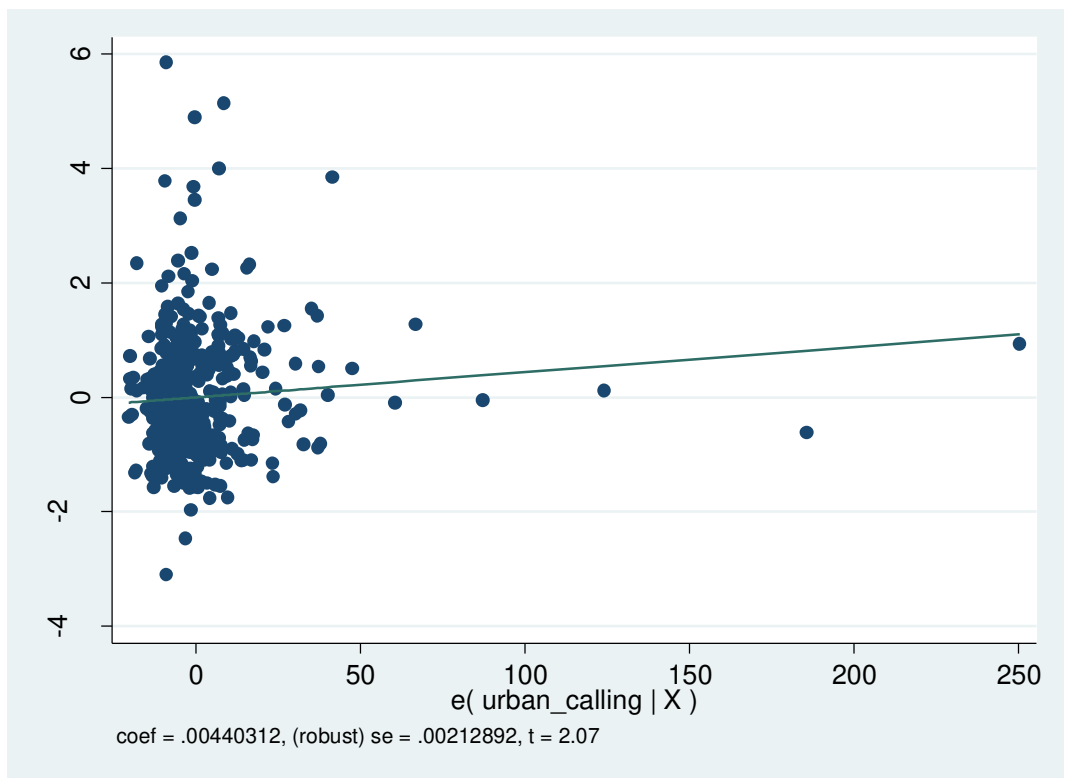| VARIABLES	(1)	(2)	(3)
	Wage-earners want to run a business		
Number of callings to urban people	0.046***	0.043***	0.041***
	(0.013)	(0.013)	(0.014)
Years of schooling	-0.022*	-0.020	-0.017
	(0.013)	(0.012)	(0.013)
Male:=1; w/o:=0	0.125**	0.146**	0.144**
	(0.057)	(0.058)	(0.057)
Age	0.015	0.011	0.012
	(0.016)	(0.017)	(0.017)
Age squared	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Minority:=1; w/o:=0	0.249*	0.225*	0.214
	(0.131)	(0.132)	(0.132)
State ownership:=1; o/w:=0	-0.046	-0.083	-0.075
	(0.093)	(0.096)	(0.094)
Private ownership:=1; o/w:=0	0.043	0.019	0.017
	(0.060)	(0.064)	(0.063)
Current job through social network:=1; o/w:=0		-0.006	-0.008
		(0.055)	(0.054)
Duration of staying in the cities		0.006	0.007
		(0.006)	(0.006)
Level of risk preferences		0.009	0.009
		(0.015)	(0.014)
Almost people can be trusted:=1			0.081
			(0.085)
Constant	-0.243	-0.221	-0.238
	(0.278)	(0.287)	(0.283)
Observations	607	568	568

Notes: ***, ** and * indicates significance level of 1%, 5% and 10% respectively against a two sided alternative. Clustered standard errors are in brackets.

Table 8. Reduced form. Relationship between weather variation and migrant's incomes

VARIABLES	(1)	(2)	(3)	(4)
	Log of change in current and first job incomes			
Log of rainfall variation	0.549**	0.526**	0.540**	0.540**
	(0.261)	(0.250)	(0.266)	(0.266)
Years of schooling	-0.008	-0.009	-0.018	-0.018
	(0.014)	(0.016)	(0.015)	(0.015)
Age	0.114***	0.085**	0.081**	0.081**
	(0.031)	(0.035)	(0.031)	(0.031)
Age squared	-0.002***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Male:=1; w/o:=0	-0.094	-0.070	-0.048	-0.049
	(0.091)	(0.096)	(0.094)	(0.094)
Minority:=1; w/o:=0	-0.328***	-0.369***	-0.447***	-0.444***
	(0.124)	(0.128)	(0.129)	(0.130)
State ownership:=1; o/w:=0		-0.246	-0.298*	-0.302*
		(0.161)	(0.153)	(0.155)
Private ownership:=1; o/w:=0		-0.142	-0.242**	-0.245**
		(0.139)	(0.117)	(0.119)
Duration of staying in the cities		0.018	0.036***	0.036***
		(0.021)	(0.013)	(0.013)
Level of risk preferences			0.013	0.013
			(0.025)	(0.025)
Almost people can be trusted:=1			0.176	0.175
			(0.113)	(0.114)
Health condition				0.051
				(0.168)
Constant	-4.252***	-3.596**	-3.584**	-3.627**
	(1.621)	(1.476)	(1.581)	(1.600)
Observations	467	465	447	447
R-squared	0.048	0.060	0.108	0.108

Notes: ***, ** and * indicates significance level of 1%, 5% and 10% respectively against a two sided alternative. Clustered standard errors are in brackets.

Figure 1. Relationship between income dynamics and number of urban callings

coef = .00440312, (robust) se = .00212892, t = 2.07

Appendix

A.1 Logistic regression

VARIABLES	(1)	(2)	(3)	(4)
	Wage-earners want to run a business			
Number of callings to urban people	0.017**	0.014*	0.033*	0.024
	(0.008)	(0.008)	(0.017)	(0.021)
Male:=1; w/o:=0	0.500**	0.378*	0.371	0.531**
	(0.219)	(0.227)	(0.228)	(0.249)
Years of schooling	0.034	0.003	0.004	0.009
	(0.041)	(0.040)	(0.041)	(0.041)
Age	0.206**	0.169*	0.182*	0.167*
	(0.103)	(0.099)	(0.101)	(0.099)
Age squared	-0.003**	-0.003**	-0.003**	-0.003**
	(0.002)	(0.001)	(0.002)	(0.001)
Minority:=1; w/o:=0	0.716	0.795	0.742	-0.079
	(0.533)	(0.543)	(0.549)	(0.680)
State ownership:=1; o/w:=0	-0.174	-0.313	-0.260	-0.143
	(0.265)	(0.274)	(0.278)	(0.299)
Foreign ownership:=1; o/w:=0	-0.085	-0.143	-0.169	-0.297
	(0.254)	(0.257)	(0.260)	(0.273)
Duration of staying in the cities	0.055***	0.041**	0.033	0.039*
	(0.019)	(0.018)	(0.021)	(0.024)
Log of total income per month		0.779***	0.810***	0.790***
		(0.250)	(0.256)	(0.286)
Switching from informal to formal job types				0.249
				(0.229)
Number of urban calling X Less than 3 years staying in the cities			-0.009	0.005
			(0.027)	(0.030)
Number of urban calling X From 3 to less than 5 years staying in the cities			-0.022	-0.013
			(0.025)	(0.028)
Number of urban calling X From 5 to less than 8 years staying in the cities			-0.021	-0.017
			(0.020)	(0.023)
Number of urban calling X From 8 to less than 12 years staying in the cities			-0.028	-0.017
			(0.020)	(0.024)
Constant	-4.89***	-10.32***	-10.738***	-10.397***
	(1.603)	(2.368)	(2.446)	(2.635)
Observations	605	601	601	503

Notes: ***, ** and * indicates significance level of 1%, 5% and 10% respectively against a two sided alternative. Clustered standard errors are in brackets.