

# Entrepreneurship of the left-behind\*

Corrado Giulietti<sup>†</sup>, Jackline Wahba<sup>‡</sup> and Klaus F. Zimmermann<sup>§</sup>

(published in the Research in Labor Economics, 2013, 37, 65–92)

Accepted version: March 2013

## Abstract

While there is evidence that return migration promotes entrepreneurship and self-employment of those who migrated, previous studies have not focused on whether migration provides the same benefits to individuals who did not migrate. Using a unique dataset that provides information on both current and return migrants in rural China (RUMiC), we investigate the impact of migration on entrepreneurship among individuals with no migration experience. We explore the self-employment choices of individuals who live in households with return migrants and individuals who live in households that have migrants currently in the city, comparing them with individuals living in non-migrant households. Our methodology allows us to control for the potential endogeneity between the migration and self-employment decisions. The results show that return migration promotes self-employment among household members that have not migrated. However, left-behind individuals are less likely to be self-employed when compared to those living in non-migrant households.

**Keywords:** self-employment; rural to urban migration; RUMiC dataset

EconLit Subject Descriptors: J23, J61, O15

---

\*The authors would like to thank Konstantinos Tatsiramos and two anonymous referees for constructive comments. We are also grateful to participants in the 3<sup>rd</sup> joint Workshop of the China Institute for Employment Research, CIER and the Institute for the Study of Labor, IZA at Renmin University in Beijing, in research seminars at the Universities of Sydney and Melbourne, and at an International Symposium on Demographic Dividend and Socio-economic Development in Beijing organized by the Institute of Population and Labor Economics, CASS, Beijing. Jackline Wahba also acknowledges financial support from the ESRC grant no RES-238-25-0044. The Longitudinal Survey on Rural Urban Migration in China (RUMiC) consists of three parts: the Urban Household Survey, the Rural Household Survey and the Migrant Household Survey. It was initiated by a group of researchers at the Australian National University, the University of Queensland and the Beijing Normal University and was supported by IZA, which provides the Scientific Use Files. Financial support for RUMiC was obtained from the Australian Research Council, the Australian Agency for International Development (AusAID), the Ford Foundation, IZA and the Chinese Foundation of Social Sciences.

<sup>†</sup>IZA - Institute for the Study of Labor, P.O. Box 7240, 53072 Bonn, Germany - Giulietti@iza.org

<sup>‡</sup>Division of Economics, and ESRC Centre for Population Change, University of Southampton, Southampton, SO17 1BJ, UK - J.Wahba@soton.ac.uk

<sup>§</sup>IZA - Institute for the Study of Labor, P.O. Box 7240, 53072 Bonn, Germany - Zimmermann@iza.org

# 1 Introduction

In the economic literature there has been growing interest on the impact of migration on sending countries' economic development. One strand of the literature has examined the consequences of temporary/return migration on occupational choices and entrepreneurship of returnees (McCormick and Wahba, 2001; Dustmann and Kirchkamp, 2002; Mesnard, 2004). More recently, a few papers such as Piracha and Vadean (2010), and Démurger and Xu (2011) have compared the entrepreneurial activity of returnees with that of non-migrants, highlighting the benefits of temporary migration for the region of origin. Another strand of the literature has focused on the impact of migration and remittances on the left-behind (see for example, Woodruff and Zenteno, 2007 and Gibson et al., 2011). Antman (2013) provides a recent survey of studies on the consequences of migration on left-behind individuals.

By bringing together these two strands of the migration literature, we examine a topic that has not been studied before, namely the spill-over effects of migration on entrepreneurship. In particular, we explore whether individuals who did not migrate are more likely to become self-employed when other household members are currently migrating or have returned to the village.

Migration can contribute to the occupational choice of those left behind through several channels. On the one hand, individuals living in households with current migrants might receive remittances that would provide the required capital to set-up a business. On the other hand, it might be that current migrant households lose manpower and/or entrepreneurial skills that are important for self-employment, or that remittances provide the remaining family with the source of income to live without the need of extra earnings. There is a potential trade-off between the loss of human capital and gains in physical capital on the self-employment choice whilst the migrant is away. Once the migrant returns, while a household may still entail costs in the form of foregone urban wage, it might also benefit from having the returnee who has potentially accumulated physical and human capital in urban areas. Hence, the indirect impact of migration on the left-behind represents an empirical question.<sup>1</sup>

The effects of migration on the development of sending areas and the individuals left behind are

---

<sup>1</sup>Migration can also have general equilibrium effects. For example, if a large share of individuals leave the village, this can decrease competition and induce non-migrants to become self-employed; at the same time, however, migration to the city might reduce the aggregate demand in the village (say consumption) and potentially reduce self-employment. We thank one of the referees for having emphasised this point.

of particular concern for policymakers. Likewise, understanding the role played by credit constraints and entrepreneurial skills is relevant for policy purposes. If credit constraints represent the main obstacle towards entrepreneurship, policy should be focusing on micro-finance schemes, while if entrepreneurial skills are the main constraint, then providing training in setting-up and managing business would be more efficient.

The plan of the paper is as follows. Section 2 provides a background about China's migration and employment patterns after the economic reform of the 1970s, along with a review of the previous literature. Section 3 presents the empirical strategy, whilst Section 4 describes the data. Section 5 summarises the main findings, and Section 6 concludes.

## 2 Background and Related Literature

This paper focuses on China in order to explore the impact of migration on the entrepreneurship choice of the left-behind. China provides us with a very interesting case study for several reasons. First, the country is experiencing mass rural-urban migration, triggered by the economic reform which started at the end of the 1970s. Prior to that period, the combination of the household registration system (*hukou*) and the imposed quotas for per capita consumption considerably limited human mobility between rural and urban areas. Agricultural productivity increased with the beginning of the economic restructuring, yielding both an excess rural labour force and a more stable supply of food. Furthermore, these changes were accompanied by a rise in the inflow of foreign investment in urban areas, which itself created a high demand for low-priced labour force. Consequently, the combination of these vicissitudes progressively generated the largest movement of labour in human history. Recent estimates reveal that over 220 million people have left their permanent residential town for over 6 months (NBS China, 2011). Indeed, compared to the 2000 census, the migrant population has increased by around 100 million.

The second reason for our focus on China lies in the temporary nature of rural-urban migration, driven by the restriction of free labour mobility. At least until now, internal migration in China has been a temporary phenomenon, with the majority of migrants eventually returning back home to their village. This highlights the importance of return migration through the work

experience acquired by the migrants in the city and used once they return home.<sup>2</sup>

Another important feature is that the majority of migrants typically remit a substantial share of their income back to their family in the home village. There are several reasons why such large amounts are sent back home, including family arrangements (due to the persistence of hukou regulations, the spouse and children of migrants are often left behind in the village), cultural factors (such as the widespread moral obligation to take care of parents and elderly, embodied in the Chinese traditional virtue of the xiao) and the lack of social security in rural areas. Remittances are considered to have a positive effect on economic development, given that they contribute to poverty reduction through increasing consumption and insuring rural households in times of low agricultural production.

Finally, the employment structure in rural areas has profoundly changed in parallel with mass migration. Over recent decades, the workforce has progressively moved from the primary to the secondary and tertiary sectors, in a trend accompanied by the decline in state-owned enterprises. The share of individuals engaged in self-employment related activities (composed by both own-account and individuals employed in private enterprises) rose from around 22% of the total non-farming rural workforce in 1996 to 27% in 2010 (NBS China, 2011).

Indeed, entrepreneurship plays an important role in economic growth, innovation and competitiveness, and is therefore crucial for a country such as China transiting into a market economy. Our focus here lies on the entrepreneurship choice (mainly own account businesses) of the non-migrant individuals. To explore the indirect effect of migration on entrepreneurship of individuals in rural areas, we use a unique dataset that provides information on rural China (RUMiC). In contrast to the common approach adopted in the literature, we do not focus on the choices of the migrants themselves, but rather compare individuals left-behind who live in households with migrants currently in cities with those who live in households with return migrants and those who live in non-migrant households. Our identification strategy, based on a trivariate recursive Probit model, allows us to control for the potential endogeneity between three states: self-employment, living in a household with return migrants and living in a household with migrants currently in the city. Furthermore, this approach allows us to assess the “spill-over effects” of migration on non-migrants, and whether they foster or hamper entrepreneurship.

In developing countries, access to credit is seen as a major obstacle for entrepreneurship; mi-

---

<sup>2</sup>For a recent study on the self-employment choice of migrants temporarily residing in Chinese cities, see Giulietti et al. (2012).

gration allows such liquidity constraints to be overcome. Several studies have focused on the occupational choice of return migrants and particularly on self-employment and entrepreneurship amongst returnees. For example, Dustmann and Kirchkamp (2002) develop a model where migrants decide simultaneously about the optimal migration duration and their activities after return. They find that more than half of the Turkish returnees in their sample are economically active, and most engage in entrepreneurial activities. Mesnard (2004) models migration as a way of overcoming credit constraints in the presence of capital markets imperfections, finding that the majority of entrepreneurial projects started by Tunisian returnees are entirely financed through overseas savings. Using cross-sectional data from Pakistan, Ilahi (2002) finds that repatriated savings become a significant factor in the choice of self-employment versus wage employment. McCormick and Wahba (2001) contribute a different insight by showing that savings are more significant than human capital acquisition for the probability of entrepreneurship of illiterate Egyptian return migrants. However, for educated returnees, both access to credit through overseas savings and human capital accumulation are significant determinants of entrepreneurship. None of these studies control for the migration selectivity when studying the occupational choice of migrants, and indeed most of them only consider return migrants.

More recently, Piracha and Vadean (2010) examine the occupational choice of returnees in Albania, comparing them with non-migrants. In their work, they attempt to control for the selection bias associated with return migration, yet not for the emigration decision in the first place. Using data from a household survey conducted in one county in China, Démurger and Xu (2011) examine the impact of migration experience on individuals' employment choices. They control for the endogeneity between return migration and occupational choices. They find that returnees are more likely than non-migrants to be self-employed. Finally, Wahba and Zenou (2012) focus on the channels through which migration influences entrepreneurship. They distinguish between the role played by overseas savings and experience and the trade-off implied by the loss of social networks in the sending areas. Unlike the studies above, we focus entirely on individuals with no migration experience. In particular, we explore whether having current migrants or return migrants in the household affects the probability of self-employment of these individuals.

A number of studies have focused on the impact of migration and remittances on those left behind. Woodruff and Zenteno (2007) find that migration is associated with a significantly higher

rate of microenterprise investment. They document that migration networks help to overcome capital constraints in Mexico. Using a survey of self-employed workers and small firm owners with access to remittance flows, they estimate the impact of attachment to migration networks on the level of capital investment, the capital-output ratio, sales and microenterprises' profits. Other studies have examined the impact of migration on the income and poverty levels of household members left behind. For example, Gibson et al. (2011) examine the consequences of migration on a number of outcomes, including labor supply, income, and health. They exploit the randomisation provided by an immigration ballot under the Pacific Access Category (PAC) of New Zealand's immigration policy. They survey applicants to the 2002-05 PAC ballots in Tonga and compare outcomes for the remaining household members of emigrants with those for members of similar households who were unsuccessful in the ballot. They find that, due to the loss of labour earnings after some of their household members migrated, left behind individuals are worse off in terms of per capita household resources, despite receiving remittances. The impact of migration on the labour market participation of women left behind has been studied by Mendola and Carletto (2009). Another positive externality of migration is related to the brain gain. A few studies have shown that migration prospects raise the expected return to human capital and thus foster education investment at origin (see, e.g., Beine et al., 2001). In our paper, we contribute to this literature by examining another spill-over effect of migration on those who have not migrated, namely entrepreneurship.

### 3 Empirical Strategy

In order to explore the self-employment choices of individuals who have not migrated, we adopt an empirical approach that builds upon the literature on return migration. The standard approach consists of estimating a specification where the self-employment choice (henceforth *SE*) is a function of a set of standard covariates and an indicator of whether the individual is a return migrant or non-migrant. By acknowledging the existence of endogeneity/simultaneity between the employment and migration decisions, several studies have employed a bivariate Probit specification to estimate the parameters of interest (see e.g., Wahba and Zenou, 2012; Démurger and Xu, 2011).

We depart from this approach by only considering the self-employment choice among individuals

in rural areas, differentiating between individuals who live in households with return migrants (*HRM*), households with current migrants in the city (*HCM*), and households where no one is a migrant (*HNM*).<sup>3</sup> Our aim is to answer questions such as: are left behind individuals more likely to choose self-employment if a household member had not migrated? Does the presence of returnees in the household create positive spill-overs for starting a business?<sup>4</sup>

The type of household where individuals live is likely to be co-determined with the self-employment choice, which raises concerns of potential endogeneity between the variables of interest. Accordingly, this leads us to explore our research question using the following model (subscripts for the observational units are suppressed):

$$\left\{ \begin{array}{ll} SE^* & = \beta_1 X + \gamma HRM + \delta HCM + \varepsilon_1 & \text{with } SE & = 1 \text{ if } SE^* > 0; \quad 0 \text{ otherwise} \\ HRM^* & = \beta_2 X + \lambda_1 z_1 + \varepsilon_2 & \text{with } HRM & = 1 \text{ if } HRM^* > 0; \quad 0 \text{ otherwise} \\ HCM^* & = \beta_3 X + \lambda_2 z_2 + \varepsilon_3 & \text{with } HCM & = 1 \text{ if } HCM^* > 0; \quad 0 \text{ otherwise} \end{array} \right.$$

The *SE* indicator corresponds to the self-employment state (1 =self-employed; 0 =wage or farm work, or other states); the matrix *X* contains standard socio-demographic and economic covariates. The system is estimated using a recursive trivariate Probit model. The parameters of interest are  $\gamma$  and  $\delta$ , which capture conditional differences in the probability of being self-employed between left-behind individuals and individuals in non-migrant households, and between individuals in households with returnees and individuals in non-migrant households, respectively.

The terms  $z_1$  and  $z_2$  are variables representing exclusion restrictions that do not appear in the *SE* equation. We construct these variables using information on the economic conditions in the potential migration destinations of rural individuals. Economic conditions in the areas of destinations are thought to be unrelated to the outcomes of interest, in our case self-employment (see Yang, 2008 and Antman, 2011, for applications using destination-based variables to identify the exogenous impact of migration).

Our exclusion restrictions are derived by combining the past thirty years' urban employment

---

<sup>3</sup>Given that we are only focusing on individuals in rural areas, current and return migrants are not part of our analysis. Information about their characteristics is used to define the *HCM* and *HRM* states.

<sup>4</sup>While our focus is not on return migrants themselves, ancillary regressions that we conducted confirm the results of previous studies, namely that return migrants are more likely to be self-employed. Results of these analyses are available upon request.

growth in migrant destinations with the share of migrants coming from a certain province of origin. The growth of urban employment measures the economic development that started in migrant destination areas after the economic reforms at the end of the 1970s and can be considered as a “pull” factor for internal migration. Given that they are measured in the area of destinations, both  $z_1$  and  $z_2$  are unlikely to be linked to the economic conditions in sending areas, and thus should exclusively influence self-employment through the *HRM* and *HCM* equations, respectively.

The first step in deriving  $z_1$  and  $z_2$  involves constructing the variables of urban employment growth in province  $j$ , which we indicate as  $G_j$ . In the case of  $z_1$ , we use the urban employment growth in the period 1978-2000, while for  $z_2$  we use the urban employment growth in the period 2000-2007. The rationale is that  $G^{(1978-2000)}$  is likely to represent a strong pull factor for return migration, given that more than two-thirds of return migrants had their first migration during the period 1978-2000, as shown in Figure A1 in the Appendix. Consequently, this variable will also explain the probability of a given individual to have return migrants in the household (*HRM*). Similarly, the variable  $G^{(2000-2007)}$  is likely to be correlated with patterns of current migration, since as many as 70% of current migrants had their first migration during this period (see Figure A1). Hence, it will also correlate with the probability of being left behind (*HCM*). At the same time, the two growth patterns  $G^{(1978-2000)}$  and  $G^{(2000-2007)}$  are found to be essentially uncorrelated with each other, as shown in Figure A2. The reason for such different patterns between the two periods relates to the profound political and economic changes that took place in China around the end of the 1990s.<sup>5</sup> While this is not a formal test of independence, it suggests that  $z_1$  (constructed using  $G^{(1978-2000)}$ ) could be excluded from the *HRM* equation and  $z_2$  (constructed using  $G^{(2000-2007)}$ ) from the *HCM* equation.

The second step involves deriving a matrix of weights based on interprovincial migration, i.e.,  $M_{ij}$ , where  $i$  is the province of origin and  $j$  the province of destination. To obtain these weights, we use data from the 1990 Census of China, which identifies the province of residence of individuals in 1990 and the one in which they lived in 1985. For each province of destination, we calculate

---

<sup>5</sup>Figure A2 shows that coastal provinces such as Shanghai, Jiangsu and Zhejiang had very low or moderate employment growth during the period 1978-2000, and much higher during 2000-2007. This change is attributable to public investments and reforms of state-owned enterprises promoted by the central government at the end of 1990s on targeted areas, such as east coast cities along the Yangzi and Yellow rivers. During the same period, other provinces such as Fujian, Shandong and Henan experienced a slowdown in urban employment growth with respect to the rather fast growth recorded in the previous twenty years.



the share of migrants coming from all origins, i.e.,  $M_{ij}/\sum_i M_{ij}$ . These destination-based weights represent the “linkage” between the origin and destination and inform how economic conditions in each destination are transmitted to the provinces of origin. Using weights that refer to an earlier time period implies that we are capturing how economic conditions would be transferred had the linkages not changed.<sup>6</sup> Table A1 in the Appendix reports the weights used for the nine sending provinces in our sample.

The final step is to combine the weights with the urban employment growth measures. We aggregate over each destination to obtain a weighted average effect for each sending area i.e.,  $z = \sum_j G_j \times (M_{ij}/\sum_i M_{ij})$ . This measure constitutes an arguably exogenous pull factor of migration for each province of origin.

While the exclusion restrictions are exogenous with respect to the conditions of the sending areas, our identification strategy assumes that there is no selection of return migrants, e.g., that only a self-selected group returns from the city to the village. This seems a plausible assumption since, due to the existence of hukou regulations, most of rural-to-urban migrations in China are of temporary nature. In section 5 we conduct tests to confirm that selection of return migrants is not affecting our results.

## 4 Data and Summary Statistics

Our analysis is based on the 2008 wave of a large scale household survey conducted in China within the Rural Urban Migration in China (RUMiC) project.<sup>7</sup> Each wave is composed by three parts: 1) a rural household survey (RHS); 2) a rural to urban migrant household survey (MHS); and 3) a urban household survey (UHS). This paper focuses on the RHS, which collects

---

<sup>6</sup>To ensure that weights are fully “exogenous”, we would ideally construct them using information on migration that dates back as far as possible, possibly before the economic reform in 1978 (for a discussion of an “historical” instrument, see Hanson and Woodruff, 2003). However, 1990 is the first year for which interprovincial migration data is available. It must be noted nonetheless that since the province of origin refers to residence in 1985, it is unlikely that in the short period after the reform (1978-1985) economic conditions influenced substantially the origin-destination patterns. As a check, we have derived the growth in urban employment between 1978 and 1985 ( $G^{(1978-1985)}$ ) and correlate it with the total migration in each urban destination i.e.,  $M_j$ ; we found a very small correlation, 0.051. Furthermore, we calculated the correlation between the weights  $M_{ij}/\sum_i M_{ij}$  and interprovincial differences in growth patterns, i.e.,  $G_i^{(1978-1985)} - G_j^{(1978-1985)}$ , finding again a very small value (lower than 0.01). These tests suggest that migration – at least until 1985 – was not substantially influenced by urban growth patterns, but driven by other causes, such as distance or origin-related, push factors (e.g., increased agricultural productivity in rural areas).

<sup>7</sup>Data have been obtained from IZA IDSC (<http://idsc.iza.org?page=27&id=58>) - accessed on 15/11/2011.

data in nine of the largest sending provinces of migrants, and includes detailed information about socio-demographic characteristics, labour market outcomes and family arrangements.<sup>8</sup> Importantly, it is possible to ascertain whether there are individuals in the household who are currently living in urban areas or are return migrants.<sup>9</sup>

Our sample is restricted to non-migrant individuals aged between 16 and 64. We use a rather conservative definition of non-migrants: individuals who have no migration history and indicate that they are not planning to migrate in the near future. Consequently, this means that our main analysis excludes individuals who are current migrants (defined as rural residents temporarily living in urban areas) and individuals who are return migrants (defined as those who have returned before 2007 and indicate that they do not want to migrate again).<sup>10,11</sup>

The sample largely comprises (98%) of individuals who work as self-employed, in a salaried job or farm work. We also include a few cases of unemployed individuals and family workers without pay.<sup>12</sup> The final sample consists of 8,188 individuals.

An inspection of the self-employed individuals in our data (Table 1) shows that the typical self-employment activity consists of a small business, mostly on own account. Furthermore, there is indication that less than half of these businesses are run within a self-employment household (e.g., family business).

The majority of self-employed work is in wholesale/retail trade or in services, while only a few are involved in agriculture-related businesses. Furthermore, our data shows that individuals chose to be self-employed because they want to make greater profits or engage in a more flexible/independent type of job. Less than 10% of the individuals report being self-employed due to difficulty in finding employment.

We distinguish three groups of individuals: (i) those who live in households where there is at least a return migrant (*HRM*), (ii) those who live in households where there is at least a migrant currently living in the city (*HCM*), and (iii) those who live in non-migrants house-

---

<sup>8</sup>See Kong (2010) for a description of the RUMiC survey and its implementation.

<sup>9</sup>One of the features of RUMiC data is that most of the interviews in rural areas are conducted during the period of the Chinese New Year, when many temporary migrants briefly return to their homes for the celebrations. Information on current migrants not present at the time of the interview is provided by other household members (generally the household head or his/her spouse).

<sup>10</sup>For completeness and since we will use some of the information pertinent to both current and return migrants, we report summary statistics of their characteristics in Table A2 in the Appendix.

<sup>11</sup>The sample also excludes all individuals who do not report information about their migration history, i.e., whether they have ever migrated.

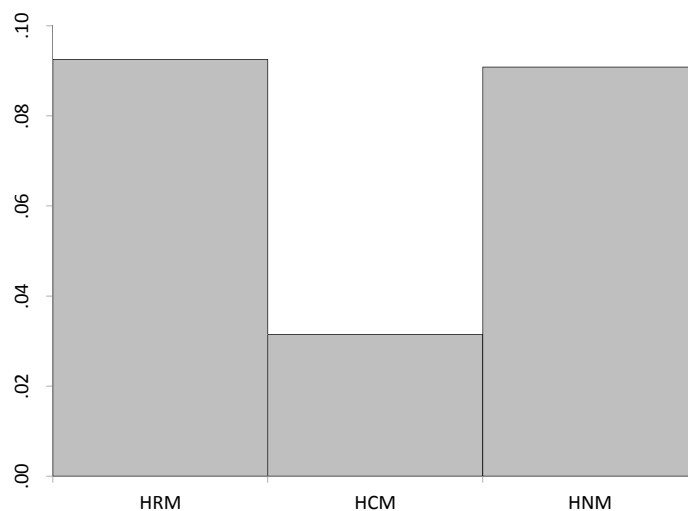
<sup>12</sup>The exclusion of these cases does not change the results of our analysis.

Table 1: Characteristics of the self-employed

Characteristics of the business	Mean (s.d.)	N	Industry	Mean (s.d.)	N
Number of persons hired (excluding family)	1.286 (5.613)	598	Agriculture & Mining	0.110 (0.313)	608
Business within a self-employment household (share)	0.443 (0.497)	609	Manufacturing	0.163 (0.37)	608
Amount invested in starting the business (RMB)	33,590 (75100)	582	Construction	0.039 (0.195)	608
Amount borrowed for starting the business (RMB)	10,340 (35412)	555	Transport, Storage and Post	0.138 (0.345)	608
Share borrowed from private lenders	0.264 (0.434)	555	Wholesale and Retail Trade	0.280 (0.449)	608
Previous job was self-employment (share)	0.226 (0.419)	593	Hotel and Catering Services	0.059 (0.236)	608
Monthly net income from the business (RMB)	1,843 (4576)	601	Other services	0.211 (0.408)	608
Number of hours worked in a week	49.921 (20.542)	606			
Years since the start of the business	10.390 (8.098)	608	<b>Reasons for choosing self-employment</b>		
Started business with help of family and friends (%)	0.508 (0.500)	608	Make profits	0.451 (0.498)	597
Has either job/injury/pension insurance	0.268 (0.443)	609	Want to be a boss	0.079 (0.27)	597
			Flexibility	0.382 (0.486)	597
			Cannot be employed / other	0.089 (0.285)	597

Source: RUMiC 2008, own elaborations. Figures refer to the sample of 609 self-employed rural residents with no migration history, aged 16-64. N refers to the number of cases with valid responses for the pertinent variable.

Figure 1: Self-employment rates by group



Source: RUMiC 2008, own elaborations. The sample is composed by rural residents with no migration history aged 16-64. *HRM* = individuals in households with return migrants; *HCM* = individuals in households with current migrants; *HNM* = individuals in households with no migrants. See text for the definition of self-employment.

holds (*HNM*).<sup>13,14</sup> Raw statistics (see Figure 1) show that individuals who live in households with return migrants exhibit self-employment rates that are similar to individuals living in non-migrants households, while self-employment is particularly low for individuals who are left behind.

Table 2 provides summary statistics of our sample, distinguishing by *HRM*, *HCM* and *HNM* states. The table contains the set of variables used in the analyses. These include individual and household level socio-demographic characteristics, and some variables capturing economic conditions of the village and the geographic location.

As seen above, self-employment rates are strikingly low for *HCM* individuals. Indeed, this group has observable characteristics that are different from both *HNM* and *HRM*. For example, they are slightly older, report fewer years of education and live in relatively larger households. Our analysis in the next section will shed light on whether the observed gaps in self-employment persist after controlling for observable and unobservable characteristics by jointly modelling the migration states and self-employment decisions.

## 5 Results

### 5.1 Results from Univariate Probit Regression

As a preliminary step, we estimate univariate Probit models of self-employment including the relevant indicators for *HRM* and *HCM*. Table 3 shows the results: column I contains the bivariate correlations between *SE* and *HRM* and *SE* and *HCM*; in column II, we add individual level covariates; in column III, we introduce household characteristics; and in column IV, we add variables capturing the economic development of the village and a dummy for whether the province of residence is in the coastal area. Across the various specifications, the estimate for the *HRM* indicator indicates a positive association between living in a household with return migrants and self-employment which, after controlling for observables, yields a substantially

---

<sup>13</sup>In classifying current and return rural-to-urban migrants, we exclude those migrants who move for non-working reasons and those who moved to other rural areas. One potential limitation in our data is the lack of information on households who might have migrated to the cities in their entirety. However, given the presence of the hukou regulations, such cases are relatively limited. For the same reasons, the large majority of migrations are temporary. Very few rural residents can obtain a city hukou (e.g., by marrying a urban resident) and therefore permanently live in urban areas.

<sup>14</sup>There are 122 individuals who belong to households which have both current and return migrants. These have been classified under both *HCM* and *HRM*. The trivariate Probit approach allows us to model groups which are not mutually exclusive.

Table 2: Summary statistics by group

	<i>HRM</i>	<i>HCM</i>	<i>HNM</i>
<b>Individual</b>			
Self-employed (%)	0.093 (0.290)	0.031 (0.174)	0.091 (0.287)
Age	45.732 (10.902)	49.495 (8.835)	44.358 (11.413)
Age sq/100	22.101 (9.442)	25.278 (8.111)	20.978 (9.818)
Male	0.336 (0.473)	0.467 (0.499)	0.523 (0.500)
Married	0.920 (0.272)	0.924 (0.265)	0.882 (0.322)
Years of education	6.988 (2.637)	6.667 (2.42)	7.564 (2.576)
Has local urban Hukou	0.044 (0.206)	0.023 (0.151)	0.077 (0.266)
<b>Household</b>			
Household size	4.408 (1.372)	4.641 (1.366)	3.867 (1.277)
N. elderly (>70) in the household	0.131 (0.403)	0.097 (0.334)	0.110 (0.346)
N. children (<6) in the household	0.256 (0.485)	0.242 (0.490)	0.165 (0.399)
Value of the house (1,000 RMB)	51.226 (59.412)	34.227 (32.822)	53.197 (66.294)
<b>Village</b>			
Labour employed by local enterprises (% pop)	0.043 (0.078)	0.024 (0.066)	0.070 (0.125)
Arable land (log Mu)	7.774 (1.157)	7.949 (0.916)	7.692 (1.243)
Average monthly income (RMB)	412.06 (227.621)	294.94 (124.825)	423.34 (262.771)
N	497	2,387	5,426

Source: RUMiC 2008, own elaborations. The sample is composed by rural residents with no migration history aged 16-64. *HRM* = individuals in households with return migrants; *HCM* = individuals in households with current migrants; *HNM* = individuals in households with no migrants.

higher marginal effect than the bivariate correlation (0.024 vs 0.015). The estimate for *HCM* is negative, with a marginal effect of -0.035 in the preferred specification. Despite being smaller than the observed raw difference in Figure 1, this estimate is statistically significant and suggests a persistent strong negative association between being left behind and self-employment. These results suggest that return migrants are not the only individuals whose self-employment choices are influenced by migration. There is also an effect on individuals who are indirectly involved in the migration process, e.g., those left behind and those who live in households with returnees.

A possible interpretation for the positive effect on *HRM* individuals is that return migration generates positive spill-overs on other household members' self-employment choice, which could be attributable to non-migrants benefitting of the human capital (e.g., entrepreneurial skills)

Table 3: Probability of self-employment - Probit regression

	<i>Prob(SE = 1)</i>							
	I		II		III		IV	
Household w/ return migrants ( <i>HRM</i> )	0.109		0.209	**	0.208	**	0.187	*
	(0.092)		(0.095)		(0.097)		(0.097)	
-> Marginal effect	<b>0.015</b>		<b>0.029</b>		<b>0.028</b>		<b>0.024</b>	
Household w/ current migrants ( <i>HCM</i> )	-0.531	***	-0.465	***	-0.414	***	-0.341	***
	(0.065)		(0.068)		(0.070)		(0.073)	
-> Marginal effect	<b>-0.060</b>		<b>-0.048</b>		<b>-0.042</b>		<b>-0.035</b>	
Age			0.098	***	0.101	***	0.103	***
			(0.019)		(0.020)		(0.020)	
Age sq./100			-0.122	***	-0.125	***	-0.127	***
			(0.022)		(0.022)		(0.023)	
Male			0.441	***	0.451	***	0.452	***
			(0.038)		(0.039)		(0.039)	
Married			0.203	**	0.195	*	0.189	*
			(0.103)		(0.106)		(0.105)	
Years education			0.029	***	0.021	**	0.017	
			(0.010)		(0.010)		(0.011)	
Has local urban Hukou			0.307	***	0.192	**	0.158	*
			(0.091)		(0.096)		(0.094)	
Household size					-0.018		-0.010	
					(0.025)		(0.026)	
N. elderly					-0.171	**	-0.174	**
					(0.086)		(0.087)	
N. children					0.017		0.008	
					(0.072)		(0.074)	
Value of the house					0.002	***	0.001	***
					(0.000)		(0.000)	
Village local labour							-0.110	
							(0.230)	
Village land							-0.019	
							(0.023)	
Village average monthly income							0.001	***
							(0.000)	
Coastal province							0.030	
							(0.066)	
Constant	-1.336	***	-3.863	***	-3.878	***	-4.039	***
	(0.029)		(0.384)		(0.407)		(0.451)	
Pseudo-R2	0.03		0.07		0.09		0.10	
N	8,188		8,188		8,188		8,188	

Source: RUMiC 2008. \*/\*\*/\*\* indicate significance at the 10%/5%/1% level. Robust standard error clustered at the household level in parentheses. The dependent variable is probability of self-employment. *HRM* is an indicator which equals to 1 if the individual lives in a household where there is at least one return migrant and 0 otherwise. *HCM* is an indicator which equals to 1 if the individual lives in a household where there is at least one migrant currently living in the city and 0 otherwise. The marginal effect for *HRM* is calculated as  $P(SE = 1|HRM = 1) - P(SE = 1|HRM = 0)$ . The marginal effect for *HCM* is calculated in similar fashion.

and physical capital (e.g., savings) accumulated by the return migrants during their migration. Similarly, the negative effect for *HCM* individuals indicates that the absence of the migrant from the rural household generates disincentives to engage in business activities. Accordingly, the benefits of having a current migrant in the city (e.g., through receiving remittances) do not appear to be sufficient in compensating for the costs determined by their absence, for example caused by the foregone earnings and temporary shortage of entrepreneurial skills. While identifying the exact channels that create these positive and negative spill-overs is arduous, we

will attempt to provide some evidence of them in sub-section 5.3.

The estimates of other covariates in the model are in line with previous studies on rural self-employment in China (such as Zhang et al., 2006 and Démurger and Xu, 2011). For example, there is a concave relationship between age and self-employment; males are more likely to be self-employed than females; marital status and education are positively associated with self-employment. There is indication that individuals in larger households are less likely to engage in self-employment, yet this estimate is not statistically significant. On the other hand, the number of elderly in the household negatively correlates with self-employment, while the value of the housing is positively associated with it. Regarding village variables, individuals who live in villages with more available land are more likely to engage in activities other than self-employment (although the estimate is statistically insignificant at conventional levels). On the other hand, self-employment is more likely to occur in villages with relatively higher per capita income.

## 5.2 Results from trivariate Probit Regression

As discussed in the previous section, *HRM* and *HCM* states could be co-determined with the self-employment choice. Left-behind individuals and those who live with returnees might possess characteristics that are unobservable, yet (positively or negatively) correlated with the self-employment choice.

We address this issue by estimating a recursive trivariate Probit model where the *HRM* and *HCM* equations are jointly model with self-employment.<sup>15</sup> The parameters of interest are identified by using exclusion restrictions that do not appear in the *SE* equation. Results in Table 4 show that the estimate of *HCM* is still negative, and around the same size of the preferred specification in the Probit model (the marginal effect is identical). Meanwhile, the estimate of *HRM* is positive and much larger than in the Probit model.<sup>16</sup> This could be attributable to the moderate correlation between the *HCM* and *HRM* equations, which is negative and statistically significant at conventional levels. On the other hand, the correlations

---

<sup>15</sup>Our estimations are based on the Stata routine MVPROBIT developed by Cappellari and Jenkins (2006).

<sup>16</sup>We have also estimated a model where we interact the *HRM* and *HCM* variables in the *SE* equation. This interaction captures the cross-effects for the 122 individuals who have both return and current migrants in the household. The estimate is essentially zero (0.001 s.e. 0.237), suggesting that "positive" and "negative" spill-over effects neutralise for individuals living in this type of household.

between the *SE* and *HRM* and between the *SE* and *HCM* equations are negligible in size and statistically insignificant.

Our estimates suggest a “premium” for individuals living with returnees of about 4.7 points and a “penalty” for individuals left behind of about 3.5 points. These estimates are not directly comparable with those of previous studies, since we look at spill-over effects of migration and not at the effect on migrants themselves. However, they seem in line with results from a recent study conducted in one county in China (Démurger and Xu, 2011). In their case, the marginal effect of returnees versus non-migrants was 0.099 - although the average self-employment rate for returnees in that province was much higher (44%).

It is insightful to also discuss the determinants of the *HRM* and *HCM* equations, which are reported in the second and third columns of Table 4. As could be expected, what determines the probability of living in a household with current and return migrants works in an antithetical way to what would predict the action of migrating. For example, the probability of *HCM* increases with age, while males and married people are less likely to be left behind. As for household characteristics, the probability of being left behind increases with household size yet decreases with the number of elderly and children. This suggests that migration is seen as a channel for sustaining the more vulnerable members of the household. The probability of being left behind is also negatively correlated with assets available at the household level (housing). Finally, individuals living in villages that are more economically developed (e.g., larger land availability and higher per capita income) are less likely to be left-behind indicating smaller incentives to migrate in relatively well-off villages. Fewer estimates are statistically significant for the *HRM* equation, most likely due to the smaller sample size of this group. Most of these estimates are similar in terms of sign to those of the *HCM* equation, albeit in this case the income of the village increases the probability of living with a return migrant.

Finally, the estimates of the two exclusion restrictions are statistically significant for the pertinent equations. The urban employment growth between 1978 and 2000 is a strong predictor of living in a household with return migrants; similarly, the urban employment growth between 2000 and 2007 is a strong determinant of being left behind.<sup>17</sup>

---

<sup>17</sup>We have conducted a series of robustness tests to support our identification strategy. Our results could be biased if the assumption that there is no selection on return migration does not hold. Although every migrant eventually returns home due to hukou regulations, the time of return might determine some self-selection pattern if, for example less able migrants are those who leave the cities first. However, in our robustness checks we found no evidence of this. We have performed a regression on the sample of current and return migrants using the data in Table A2. We model the odds of being a return migrant (vs



Table 4: Probability of self-employment - recursive trivariate Probit

	<i>Prob(SE = 1)</i>	<i>Prob(HRM = 1)</i>	<i>Prob(HCM = 1)</i>
<i>HRM</i>	0.309 ** (0.144)		
-> <b>Marginal effect</b>	<b>0.047</b>		
<i>HCM</i>	-0.301 *** (0.096)		
-> <b>Marginal effect</b>	<b>-0.035</b>		
Age	0.102 *** (0.020)	-0.005 (0.017)	0.121 *** (0.016)
Age sq./100	-0.127 *** (0.023)	0.008 (0.019)	-0.102 *** (0.017)
Male	0.459 *** (0.039)	-0.343 *** (0.044)	-0.193 *** (0.025)
Married	0.189 * (0.105)	0.089 (0.092)	-0.170 ** (0.073)
Years education	0.017 (0.011)	-0.012 (0.011)	-0.009 (0.008)
Has local urban Hukou	0.162 * (0.094)	-0.168 (0.129)	-0.285 *** (0.108)
Household size	-0.015 (0.026)	0.075 *** (0.023)	0.368 *** (0.021)
N. elderly	-0.171 * (0.087)	0.049 (0.079)	-0.375 *** (0.066)
N. children	0.010 (0.074)	0.054 (0.072)	-0.246 *** (0.060)
Value of the house	0.001 *** (0.000)	0.000 (0.000)	-0.002 *** (0.001)
Village local labour	-0.086 (0.231)	-1.153 *** (0.358)	-1.893 *** (0.341)
Village land	-0.019 (0.023)	-0.010 (0.027)	0.084 *** (0.024)
Village average monthly income	0.001 *** (0.000)	0.001 *** (0.000)	-0.001 *** (0.000)
Coastal province	0.032 (0.066)		
Urban empl. growth 1978-2000		0.236 *** (0.089)	
Urban empl. growth 2000-2007			0.289 *** (0.074)
Constant	-4.015 *** (0.453)	-1.949 *** (0.445)	-5.186 *** (0.410)
N		8,188	
	Rho12	Rho13	Rho23
Corr. between equations	-0.058 (0.051)	-0.023 (0.039)	-0.129 *** (0.034)

Source: RUMiC 2008. \*/\*\*/\*\* indicate significance at the 10%/5%/1% level. Robust standard errors clustered at the household level in parentheses. The dependent variables are probability of self-employment (col 1), probability of living in a household where there is at least one migrant currently in the city (col 2) and probability of living in a household where there is at least one return migrant (col 3). The marginal effect for *HRM* is calculated as  $P(SE = 1|HRM = 1) - P(SE = 1|HRM = 0)$ . The marginal effect for *HCM* is calculated in similar fashion.

current migrant) and found that the significant explanatory determinants are demographic characteristics and village attributes. The estimate for the years of education is small and statistically insignificant. We also added additional measures of observed ability such as having attempted the university entrance exam and having had good test scores while in school: none of the estimates are neither economically nor statistically significant. We obtained the same results when we restrict the sample to more comparable

### 5.3 Exploring the Spill-Over Effects

The analysis in the previous sub-section suggests that individuals living in households with return migrants are more likely to engage in self-employment, and also that the opposite is true for individuals left behind. Given that these are interesting findings requiring further investigation, we attempt in this sub-section to explore the potential channels that might be at work.

As previously discussed, there are several ways through which migration impacts the entrepreneurship of the individuals who have not migrated. For *HRM* individuals, having a returnee in the household means being able to “share” skills and savings accumulated by the migrant while in the city. Meanwhile, for *HCM* individuals, an absent member of the family might entail the loss of labour and entrepreneurial skill embedded in the migrant, leading to a lower probability of becoming self-employed. Furthermore, migration might lead to receiving a stream of remittances that can enable left-behind individuals to invest in a new business, but also to stay home and not work. Moreover, migration might also weaken social connections in rural areas yet can also lead to the formation of a new network in urban areas.

It is important to emphasise the difficulty in achieving a clear disentangling of these channels, which requires rich information concerning the processes of skill acquisition whilst in the city, the amount and frequency of remittances, the fraction of savings attributable to migrants, and the social networks before, during and after migration. While all such information is unfortunately not available in our data, we are still able to provide evidence on whether, for example, the “premium” for *HRM* is linked to the entrepreneurial skills of the return migrant, or whether the observed “penalty” for *HCM* is related to remittances.

To explore the various channels at work, we estimate a trivariate Probit model where the self-employment equation is represented by:

$$SE^* = \beta_1 X + \gamma_1 HRM \times (1 - D) + \gamma_2 HRM \times D + \delta_1 HCM \times (1 - D) + \delta_2 HCM \times D + \varepsilon$$

Here,  $D$  is an indicator representing certain characteristics of either the current migrant or the returnee. In practice, we are interested in exploring whether the effects of *HRM* and *HCM* vary depending on the characteristics of current and return migrants, which we achieve by comparing

---

individuals, i.e., individuals whose first migration was between 2006 and 2007. This further suggests that, at a given point in time, (observed) ability is not a strong determinant of who returns and who stays in the city.

the estimates of  $\gamma_1$  with  $\gamma_2$  for the case of *HRM*, and of  $\delta_1$  with  $\delta_2$  in the case of *HCM*.<sup>18</sup>

We use several characteristics of the current/return migrants as crude proxies for the main channels through which migration could have spill-overs on the self-employment of non-migrants: human capital and entrepreneurial skills, physical capital and social networks. We consider whether the current/return migrant: a) is self-employed; b) has a formal education above 9 years (upper decile of the distribution); c) has received training/apprenticeship in a firm; d) has attempted the national entrance exam for university (a measure of “cognitive ability”); e) earns labour income above the sample mean; f) sends a high level of remittances (above 13000 RMB/year, i.e. in the upper quartile - this variable is only defined for *HCM*); and g) has obtained the current job through friends’ network. The summary statistics for these characteristics are reported in Table A2 in the Appendix.

It is important to once again underscore that we are not able to identify the “causal effects” of each of these channels, given that some of the characteristics of current and return migrants are likely to be endogenous or co-determined with the self-employment of the non-migrant. Consequently, estimates of  $\gamma_2$  and  $\delta_2$  should be interpreted as correlations.

The results in Table 5 indicate that the positive effect of *HRM* is stronger when the return migrant is self-employed. This might suggest that the entrepreneurial skills accumulated by the return migrant are transferred to or help non-migrant household members.<sup>19</sup> A stronger effect, as indicated by the marginal effects, is also found when the return migrant has comparatively more “human capital” – proxied by having more education, having received training and attempted the university entrance exam – although the estimates are not always statistically significant.<sup>20</sup> Finally, a larger effect is estimated when the return migrant has found a job through friends; this provides an indication of the role of social skills and connections that the

---

<sup>18</sup>Note that the main effect of a certain characteristic is absorbed in both interaction terms and hence estimates in Table 5 are not directly comparable to those in Table 4.

<sup>19</sup>The estimate indicates a rather large marginal effect. One necessary remark is that this interaction is estimated with a relatively small number of observations. As emphasised, it is also possible that the self-employment states of return migrants and non-migrants are endogenously determined, which might lead to a positive bias of  $\gamma_2$ . However, our robustness checks suggest otherwise. An inspection of the data suggests that only very few self-employed who live in a household with current/return migrants started their business before the current/return migrant did. Not surprisingly, excluding these cases from the analysis does not change our results.

<sup>20</sup>Since we do not know when and where the training/apprenticeship was acquired, it is possible that – especially in the case of individuals living with return migrants – the training of the returnees and the self-employment of non-migrants are endogenously determined, which could generate a biased estimate of  $\gamma_2$ .

Table 5: Probability of self-employment - recursive trivariate Probit with interactions

	Self-employed	Education	Training	National exam	Labour income	Remittances	Friends network
$HRM \times (D = 0)$	0.130 (0.142)	0.247 * (0.139)	0.236 * (0.138)	0.257 * (0.136)	0.237 * (0.141)		0.207 (0.147)
-> Marginal effect	<b>0.015</b>	<b>0.032</b>	<b>0.030</b>	<b>0.034</b>	<b>0.030</b>	0.305 ** (0.134)	<b>0.025</b>
$HRM \times (D = 1)$	0.848 *** (0.207)	0.295 (0.202)	0.391 * (0.216)	0.389 (0.245)	0.235 (0.192)	<b>0.041</b>	0.294 * (0.176)
-> Marginal effect	<b>0.164</b>	<b>0.040</b>	<b>0.056</b>	<b>0.056</b>	<b>0.030</b>		<b>0.039</b>
$HCM \times (D = 0)$	-0.319 *** (0.09)	-0.361 *** (0.095)	-0.350 *** (0.093)	-0.351 *** (0.092)	-0.247 ** (0.098)	-0.320 *** (0.093)	-0.237 ** (0.101)
-> Marginal effect	<b>-0.026</b>	<b>-0.029</b>	<b>-0.027</b>	<b>-0.028</b>	<b>-0.021</b>	<b>-0.026</b>	<b>-0.020</b>
$HCM \times (D = 1)$	0.092 (0.264)	0.157 (0.112)	0.189 (0.132)	0.259 * (0.142)	-0.160 (0.106)	0.015 (0.139)	-0.152 (0.104)
-> Marginal effect	<b>0.010</b>	<b>0.019</b>	<b>0.023</b>	<b>0.034</b>	<b>-0.015</b>	<b>0.002</b>	<b>-0.014</b>

Source: RUMiC 2008. \*/\*\*/\*\* indicate significance at the 10%/5%/1% level. Robust standard errors clustered at the household level in parentheses. The dependent variable is probability of self-employment.  $HRM$  is an indicator which equals to 1 if the individual lives in a household where there is at least one return migrant and 0 otherwise.  $HCM$  is an indicator which equals to 1 if the individual lives in a household where there is at least one migrant currently living in the city and 0 otherwise.  $D$  is an indicator for the status of the current/return migrants, and is equal to 1 if a given state - indicated in the column heading - is realised in at least one current/return migrant and 0 otherwise. "Self-employment" refers to self-employment; "Education" refers to having years of education in the upper decile (above 9 years of education); "Training" refers to having received training or apprenticeship in businesses/enterprises; "National exam" refers to having attempted the national university entrance exam; "Labour income" refers to monthly income above the mean level (1,500 RMB); "Remittances" refers to whether the current migrant sends an amount of yearly remittances that falls in the upper quartile (above 13,000 RMB); "Friends network" refers to having found a job with the help of friends. The marginal effect for  $HRM \times (D = 0)$  is calculated as  $P(SE = 1|HRM = 1, D = 0) - P(SE = 1|HRM = 0, D = 0)$ ; the marginal effect for  $HRM \times (D = 1)$  is calculated as  $P(SE = 1|HRM = 1, D = 1) - P(SE = 1|HRM = 0, D = 1)$ . The marginal effects for  $HCM \times (D = 0)$  and  $HCM \times (D = 1)$  are calculated in similar fashion. In the model "Remittances", there is no interaction between living in a household with returnees and remittances, hence the effect of  $HRM$  alone is estimated, similarly to Table 4.

migrant might have developed in urban areas.<sup>21</sup>

Results for *HCM* indicate a very similar pattern. Although most of the interactions  $HCM \times (D = 1)$  are estimated with relatively large standard errors, they suggest a positive (or less negative) effect. Of particular interest is the case when a large amount of remittances is sent. While remittances are important to compensate for the absence of the migrant from the household, they are not sufficient to counterbalance the observed self-employment “penalty”.

## 6 Conclusions

This paper contributes to the growing literature on the effects of migration on the sending regions by exploring the entrepreneurship decision of individuals left behind in rural China. Using a unique dataset that provides information on both current and return migration, and differing from the common approach adopted in the literature, we compare individuals who live in households with return migrants, with individuals who live in households with migrants currently in cities, and individuals who live in non-migrant households. This approach allows us to control for the endogeneity between the self-employment and migration decisions.

Our results show that return migration promotes self-employment among household members that have not migrated. However, left-behind individuals are less likely to be self-employed when compared to those living in non-migrant households. Overall, our findings provide evidence of spill-over effects of migration that have not previously been studied. On the one hand, our paper highlights an additional gain from return migration, which extends to other members of the household benefitting from the migration experience of the return migrant in setting-up businesses. On the other hand, we emphasise the need for a more thorough understanding of the complex impact of migration on individuals left behind, who might face a loss of labour supply and entrepreneurial skills due to the absence of the migrant, which are not compensated for by receiving remittances.

---

<sup>21</sup>To estimate a causal impact of the return migrant’s network on the non-migrant employment, one would need to observe the network of the returnee before and after migration, which is not possible in our data.

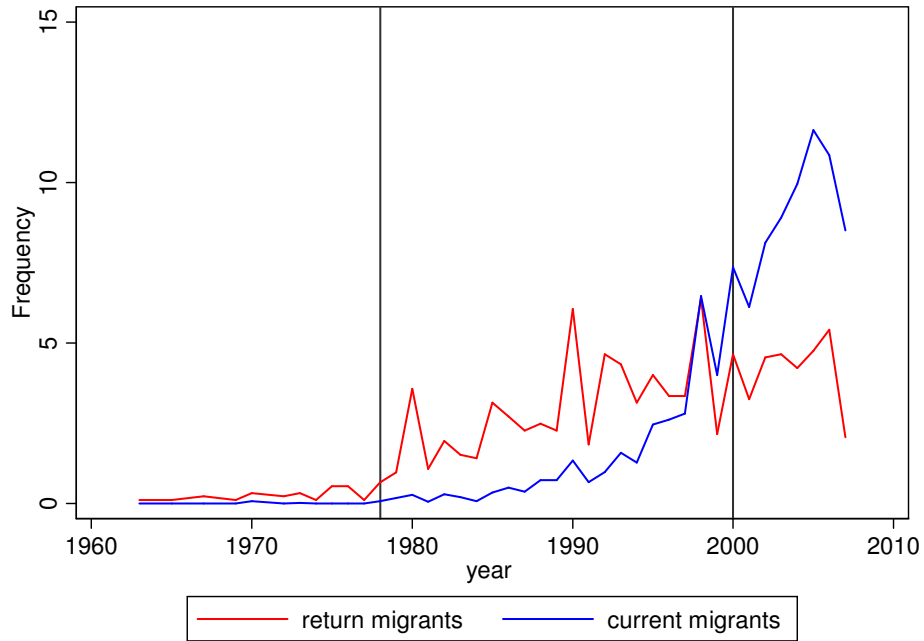
## References

- Antman, F. M. (2011). The Intergenerational Effects of Paternal Migration on Schooling and Work: What can we Learn from Children's time Allocations? *Journal of Development Economics* 96(2), 200–208.
- Antman, F. M. (2013). The Impact of Migration on Family Left Behind. In A.F. Constant and K.F. Zimmermann (Ed.), *International Handbook on the Economics of Migration*. Edward Elgar Publishing, forthcoming.
- Beine, M., F. Docquier, and H. Rapoport (2001). Brain Drain and Economic Growth: Theory and Evidence. *Journal of Development Economics* 64(1), 275–289.
- Cappellari, L. and S. P. Jenkins (2006). MVPROBIT: Stata Module to Calculate Multivariate Probit Regression Using Simulated Maximum Likelihood. *The Stata Journal* 3(3), 278–294.
- Démurger, S. and H. Xu (2011). Return Migrants: The Rise of new Entrepreneurs in Rural China. *World Development* 39(10), 1847–1861.
- Dustmann, C. and O. Kirchkamp (2002). The Optimal Migration Duration and Activity Choice After re-Migration. *Journal of Development Economics* 67(2), 351–372.
- Gibson, J., D. McKenzie, and S. Stillman (2011). The Impacts of International Migration on Remaining Household Members: Omnibus Results from a Migration Lottery Program. *Review of Economics and Statistics* 93(4), 1297–1318.
- Giulietti, C., G. Ning, and K.F. Zimmermann (2012). Self-Employment of Rural-to-Urban Migrants in China. *International Journal of Manpower* 33(1), 96–117.
- Hanson, G. H. and C. Woodruff (2003). Emigration and Educational Attainment in Mexico. *University of California San Diego, mimeo*.
- Ilahi, N. (2002). Return Migration and Occupational Change. *Review of Development Economics* 3(2), 170–186.
- Kong, T. (2010). Rural-Urban Migration in China: Survey Design and Implementation. In X. Meng and C. Manning (Ed.), *The Great Migration: Rural-Urban Migration in China and Indonesia*, pp. 135–150. Edward Elgar Publishing.

- McCormick, B. and J. Wahba (2001). Overseas Work Experience, Savings and Entrepreneurship Amongst Return Migrants to LDCs. *The Scottish Journal of Political Economy* 48(2), 164–178.
- Mendola, M. and C. Carletto (2009). International Migration and Gender Differentials in the Home Labor Market: Evidence from Albania. *World Bank Policy Research Working Paper Series* 4900.
- Mesnard, A. (2004). Temporary Migration and Capital Market Imperfections. *Oxford Economic Papers* 56(2), 242–262.
- NBS China (2011). *China Statistical Yearbook 2010*. National Bureau of Statistics of China. China Statistics Press.
- Piracha, M. and F. Vadean (2010). Return Migration and Occupational Choice: Evidence from Albania. *World Development* 38(8), 1141–1155.
- Wahba, J. and Y. Zenou (2012). Out of Sight, Out of Mind: Migration, Entrepreneurship and Social Capital. *Regional Science and Urban Economics* 42(5), 890–903.
- Woodruff, C. and R. Zenteno (2007). Migration Networks and Microenterprises in Mexico. *Journal of Development Economics* 82(2), 509–528.
- Yang, D. (2008). International Migration, Remittances and Household Investment: Evidence from Philippine Migrants Exchange Rate Shocks. *The Economic Journal* 118(528), 591–630.
- Zhang, J., L. Zhang, S. Rozelle, and S. Boucher (2006). Self-employment with Chinese Characteristics: The Forgotten Engine of Rural China’s Growth. *Contemporary Economic Policy* 24(3), 446–458.

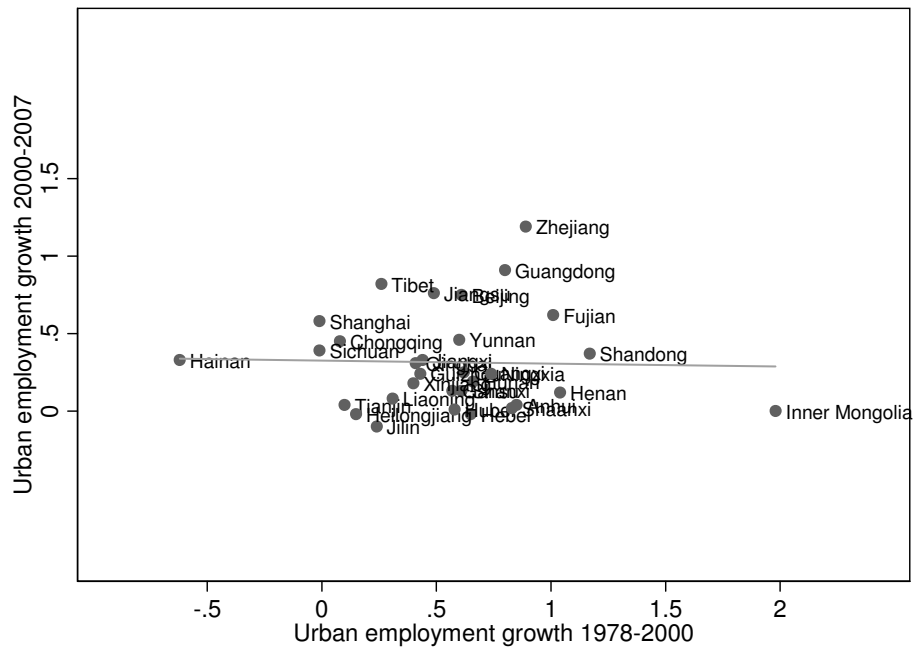
# Appendix

Figure A1: Year of first migration - frequency by group



Source: RUMiC 2008.

Figure A2: Scatterplot of urban growth in the two periods



Source: National Bureau of Statistics of China <http://www.stats.gov.cn/english/statisticaldata/yearlydata/>



Table A1: Destination-based migration weights

Destin./Orig.	Hebei	Jiangsu	Zhejiang	Anhui	Henan	Hubei	Guangdong	Chongqing	Sichuan
Beijing	0.233	0.038	0.022	0.040	0.066	0.025	0.007	0.006	0.035
Tianjin	0.263	0.018	0.011	0.022	0.030	0.013	0.004	0.003	0.019
Hebei	0.646	0.008	0.007	0.008	0.030	0.010	0.002	0.002	0.038
Shanxi	0.071	0.008	0.008	0.008	0.052	0.008	0.001	0.002	0.030
Inner Mongolia	0.087	0.006	0.003	0.004	0.012	0.002	0.001	0.001	0.012
Liaoning	0.039	0.009	0.005	0.007	0.012	0.004	0.002	0.001	0.012
Jilin	0.031	0.008	0.003	0.004	0.008	0.004	0.001	0.000	0.005
Heilongjiang	0.036	0.013	0.003	0.013	0.012	0.003	0.001	0.000	0.007
Shanghai	0.005	0.209	0.104	0.134	0.019	0.014	0.006	0.005	0.031
Jiangsu	0.004	0.640	0.025	0.115	0.019	0.013	0.002	0.005	0.038
Zhejiang	0.003	0.037	0.503	0.089	0.022	0.029	0.002	0.017	0.062
Anhui	0.005	0.079	0.015	0.730	0.035	0.009	0.002	0.002	0.026
Fujian	0.002	0.012	0.026	0.026	0.010	0.023	0.016	0.022	0.084
Jiangxi	0.004	0.024	0.053	0.029	0.008	0.025	0.023	0.002	0.012
Shandong	0.019	0.024	0.009	0.013	0.021	0.007	0.001	0.002	0.016
Henan	0.019	0.012	0.007	0.021	0.758	0.023	0.003	0.002	0.025
Hubei	0.010	0.016	0.011	0.011	0.066	0.690	0.006	0.015	0.042
Hunan	0.005	0.006	0.007	0.004	0.007	0.031	0.011	0.003	0.014
Guangdong	0.002	0.007	0.006	0.014	0.039	0.057	0.395	0.018	0.097
Guangxi	0.003	0.005	0.007	0.004	0.005	0.007	0.033	0.002	0.011
Hainan	0.003	0.008	0.006	0.011	0.015	0.032	0.139	0.009	0.045
Chongqing	0.005	0.008	0.008	0.004	0.007	0.021	0.005	0.627	0.215
Sichuan	0.006	0.007	0.006	0.003	0.007	0.007	0.003	0.060	0.808
Guizhou	0.006	0.010	0.009	0.005	0.011	0.008	0.006	0.023	0.112
Yunnan	0.004	0.007	0.014	0.004	0.008	0.010	0.008	0.026	0.109
Tibet	0.005	0.007	0.009	0.007	0.028	0.008	0.001	0.033	0.314
Shaanxi	0.022	0.020	0.011	0.013	0.100	0.016	0.003	0.003	0.039
Gansu	0.024	0.021	0.012	0.009	0.052	0.007	0.002	0.002	0.027
Qinghai	0.034	0.028	0.012	0.014	0.086	0.007	0.002	0.003	0.047
Ningxia	0.037	0.022	0.015	0.020	0.061	0.003	0.001	0.002	0.020
Xinjiang	0.017	0.050	0.010	0.031	0.153	0.021	0.002	0.015	0.165

Source: Own calculations based on China Census 1990. National Bureau of Statistics of China.

Table A2: Characteristics of return and current migrants

	PRM	PCM
<b>Individual characteristics</b>		
Age	42.613 (10.851)	29.257 (8.934)
Age sq/100	19.335 (9.217)	9.358 (6.065)
Male (=1)	0.642 (0.480)	0.623 (0.485)
Married (=1)	0.919 (0.273)	0.543 (0.498)
Years of education	7.745 (2.382)	8.610 (2.168)
Has local urban Hukou (=1)	0.064 (0.246)	0.039 (0.195)
<b>Characteristics used in regression Table 5 (in %)</b>		
Self-employed	0.151 (0.330)	0.057 (0.231)
Years of education in the upper decile (>9)	0.124 (0.330)	0.166 (0.372)
Has received training/apprenticeship in firm	0.106 (0.307)	0.119 (0.324)
Attempted national university entrance exam	0.058 (0.234)	0.064 (0.245)
Has monthly labour income above the mean (>1,500 RMB)	0.167 (0.373)	0.411 (0.492)
Sends annual remittances in upper quartile (>13,000 RMB)		0.220 (0.414)
Has found a job through the social network	0.706 (0.456)	0.868 (0.339)
N	1,072	4,440

Source: RUMiC 2008, own elaborations. The sample is composed by rural residents who are return migrants (*PRM*) and by temporary migrants currently living in the city (*PCM*). The percentages in the second panel refer to whether the named characteristic is observed in at least one current/return migrant.