

IZA DP No. 7860

**The RUMiC Longitudinal Survey:  
Fostering Research on Labor Markets in China**

Mehtap Akgüç  
Corrado Giulietti  
Klaus F. Zimmermann

December 2013

# The RUMiC Longitudinal Survey: Fostering Research on Labor Markets in China

**Mehtap Akgüç**

*IZA and Toulouse School of Economics*

**Corrado Giuliatti**

*IZA*

**Klaus F. Zimmermann**

*IZA and University of Bonn*

Discussion Paper No. 7860  
December 2013

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **The RUMiC Longitudinal Survey: Fostering Research on Labor Markets in China**\*

This paper describes the Longitudinal Survey on Rural Urban Migration in China (RUMiC), a unique data source in terms of spatial coverage and panel dimension for research on labor markets in China. The survey is a collaboration project between the Australian National University, Beijing Normal University and the Institute for the Study of Labor (IZA), which makes data publicly available to the scientific community by producing Scientific Use Files. The paper illustrates the structure, sampling frame and tracking method of the survey, and provides an overview of the topics covered by the dataset, and a review of the existing studies based on RUMiC data.

JEL Classification: C81, J01, P36, R23

Keywords: household survey, China, migration, labor markets

Corresponding author:

Corrado Giuliatti  
IZA  
P.O. Box 7240  
53072 Bonn  
Germany  
E-mail: [giuliatti@iza.org](mailto:giuliatti@iza.org)

---

\* 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. This paper is part of the duties IZA has taken to document the data and make it available to the scientific community. We thank Xin Meng and Shi Li for their trust and support in this project.

## 1 Introduction

China has witnessed rapid demographic and socio-economic changes during the last three decades. Starting from the end of the 1970s and accompanied by the rise in foreign investment, the economic reforms introduced by the government led to a sudden expansion in the demand for unskilled labor in urban areas. Excess labor force was generated in many rural areas thanks to productivity growth in the primary sector. The household registration system that was used by the government as a policy to control and restrict internal movement of labor – the hukou – was progressively relaxed, albeit not eliminated. These rapid transformations set the background of the largest movement of labor force within a country, referred to as the Great Migration in China throughout this paper. According to the National Bureau of Statistics of China, more than 260 million Chinese left their hometown for at least 6 months, including more than 160 million migrant workers moving from rural to urban areas (NBS, 2013). This large-scale movement of labor is one of the driving forces of economic growth in China.

Mass internal migration in China has important reverberations for the global economy. On the one hand, a substantial part of the global demand has been indirectly supported by the shift of labor from the primary sector to export-oriented industries and services located in urban areas. On the other hand, the growing influence in the world trade and the massive upsurge in domestic consumption have led China to affect the world prices of many commodities, including food and energy.

The Great Migration brought along unprecedented levels of urbanization. To put it in perspective, while the rate of urbanization in Europe increased from 30% to 50% in the first half of the 20<sup>th</sup> century – at its fastest pace - urbanization in China is at least twice as fast (Frijters and Meng, 2009). Nowadays, there are over 700 million individuals living in urban areas (more than double that of just twenty years ago). Projections indicate that such number could be as large as 900 million by 2030 (Kamal-Chaoui et al, 2009), with most of this increase expected to be fuelled by rural-to-urban migration.

The persistence of the hukou system means that migrants generally cannot permanently settle in cities, making most migrations temporary. Nonetheless, the frequency, circularity and volume of the flows are such that rural-to-urban migrants hold a constant presence in urban areas, while many villages face an increasing lack in of the working-age population. As a consequence,

besides affecting migrants themselves, migration also has socio-economic repercussions on family members left behind in rural villages, as well as urban residents.

With the goal of understanding the relationship between the Great Migration and changing labor markets in China, a group of international researchers has established the Rural-Urban Migration in China (RUMiC) project. The project's main output has been the design and implementation of a large scale longitudinal household survey covering individuals in rural and urban areas, as well as temporary migrants working in cities.

This paper outlines documents the structure, sampling frame and tracking method of the first two waves of the RUMiC survey, and provides an overview of the topics covered, and a review of the existing studies based on these data.

## **2 The RUMiC Longitudinal Survey**

### **2.1 Overview**

The Longitudinal Survey on Rural Urban Migration in China (RUMiC) consists of three independent surveys: the Urban Household Survey (UHS), the Rural Household Survey (RHS) and the Migrant Household Survey (MHS). It was initiated by a group of researchers at the Australian National University, the University of Queensland and the Beijing Normal University and supported by IZA, which provides the Scientific Use Files through IDSC, its data bank center. 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.

The fieldwork started in 2008, and since then four waves of the UHS and RHS and five waves of the MHS have been collected. The RHS and UHS have been conducted in collaboration with the National Bureau of Statistics of China (NBS), while the MHS has been conducted in partnership with a professional survey company. The project is designed to track households as long as they remain in the surveyed cities and villages. A systematic tracking strategy – especially relevant for migrant households – is adopted to follow individuals over the project's lifespan.

Its large scale, in-depth topics and longitudinal aspect make RUMiC a unique tool to explore migration and labor markets in China. The RHS comprises around 8,000 households, while the UHS and MHS each involve around 5,000 households. Urban (rural) residents are individuals who possess urban (rural) hukou. A migrant is defined as an individual who has rural hukou, but is living in a city at the time of the survey. The availability of three surveys allows multiple “control groups” to investigate the effects of rural-urban migration (Kong, 2010). For instance, when analyzing the impact of migration on rural areas, non-migrant individuals serve as a control group for those who are currently migrating in cities or those who have returned. Similarly, urban residents can be used as a control group when investigating the economic situation of migrant workers.

Each of the three surveys include comprehensive information on household and personal characteristics, detailed health status, employment, income, training and education of adults and children, social networks, family and social relationships, life events, and mental health measures of the individuals. The MHS additionally includes questions related to migration history.

In addition to its rich set of information at such a large scale, one of the key features of RUMiC lies in its longitudinal structure, which allows researchers to incorporate dynamic aspects in their empirical analyses. Each of the three surveys can be utilized as repeated cross-section – which is particularly useful when analyzing trends in economics outcomes of certain populations of interest – or as panel data – allowing the use of fixed effect estimators.

RUMiC complements existing surveys that have recently been conducted in China, such as the Chinese Household Income Project (CHIP) survey, which collects data on rural and urban surveys yet largely excludes migrant workers, and the China Health and Retirement Longitudinal Study (CHARLS), a biennial survey which aims to support research on the elderly, and thus only collects data about individuals aged 45 and older.<sup>12</sup> RUMiC also complements other surveys that IZA has made or is planning to make publicly available to the research community, such as the IZA Evaluation Dataset (see Caliendo et al, 2011; Arni et al, 2013) and the Ukrainian Longitudinal Monitoring Survey for Ukraine (see Lehmann et al, 2012). IZA currently provides the scientific use files for the first two RUMiC waves, and the IZA team is currently processing the remaining waves, which will be made available soon.

## **2.2 Listing and Sampling Frame**

The RHS and UHS were conducted using random samples from the annual household income and expenditure surveys carried out in cities and rural villages (Kong, 2010). During the sampling process, efforts were made to cover rural and urban areas with representative income and population levels.

Due to the mobile and temporary nature of internal migration in China, there was no existing sampling frame to conduct the MHS. Migrants are typically clustered in dormitories near factories and construction sites, often without a registered address. Therefore, the sampling frame from the UHS would not be representative for migrant workers. To cope with this issue and ensure reasonable representativeness of the migrant worker population, the RUMiC team devised an innovative sampling frame.<sup>3</sup> The first step involved creating a listing of migrants and workplaces to estimate the total size of the migrant population in each city. This was based on a pre-survey census collected across randomly selected blocks in which the cities were divided into (see Kong, 2009, for details). Using the listing data, a sampling frame based on workplaces (rather than residences) was created. Furthermore, all businesses – including street vendors – in randomly selected enumeration areas within defined city boundaries were included. For each city, a sample of migrant workers was randomly selected within each workplace, based on their birth month. The enumerators subsequently conducted face-to-face interviews with migrant workers and their families.

## **2.3 Coverage**

The RUMiC survey covers principal migrant sending and receiving regions. The RHS was conducted in villages across nine provinces, while the UHS and MHS were carried out in nineteen and fifteen cities, respectively. Table 1 represents the spatial coverage of RUMiC. Maps showing the sample size of each wave in the three surveys are reported in the Appendix. The 2000 Census shows that two-thirds of all migrant workers in China are located in the provinces of Shanghai, Guangdong, Jiangsu and Zhejiang, while 47% of all migrants emigrate from Sichuan, Chongqing, Anhui, Hubei and Henan provinces.

**Table 1 – RUMiC: Spatial Coverage**

	<b>RHS</b>	<b>UHS</b>		<b>MHS</b>	
<i>Spatial unit</i>	Villages	City		City	
<i>Coverage (provinces / cities)</i>	Anhui	Anyang	Leshan	Bengbu	Luoyang
	Chongqing	Bengbu	Luoyang	Chengdu	Nanjing
	Guangdong	Chengdu	Mianyang	Chongqing	Ningbo
	Hebei	Chongqing	Nanjing	Dongguan	Shanghai
	Henan	Dongguan	Ningbo	Guangzhou	Shenzhen
	Hubei	Guangzhou	Shanghai	Hefei	Wuhan
	Jiangsu	Hangzhou	Shenzhen	Hangzhou	Wuxi
	Sichuan	Hefei	Wuhan		Zhengzhou
	Zhejiang	Jiande	Wuxi		
			Zhengzhou		

Source: RUMiC Wave 1 and 2.

## 2.4 Tracking Method

A systematic tracking strategy is adopted to follow individuals over the lifespan of the project, as long as they remain in the surveyed cities and villages. For the RHS and UHS, enumerators track individuals using their permanent addresses. For the MHS, due to the higher mobility of migrants and the fact that they usually do not have a permanent address in cities, a more complex tracking strategy was adopted.<sup>4</sup> The survey team recorded the individuals' work and home addresses, as well as other contact details in both cities and home villages. They also recorded the phone numbers of three close relatives or friends to be contacted in the case that households moved (Kong, 2010). As an incentive for improving tracking, the team designed three lotteries for each year, with prizes assigned to survey participants ranging from 50 to 2000 Yuan, as well as a yearly dispatch of small presents before the Chinese New Year (Kong et al, 2009).

## 2.5 Sample Size and Panel Attrition

Table 2 provides the sample size for each survey and wave. For wave 2, the table includes figures concerning households and individuals who are tracked over time, the attrition rate, the additional household members who were not surveyed in wave 1, and for the MHS, the size of the new sample that was collected. The panel attrition was essentially inexistent for individuals in the RHS (0.4%) and rather low for those in the UHS (5.8%). In contrast, the attrition rate was rather

significant for the MHS, despite substantial efforts to track individuals over time (58.4%). This is partly due to the mobile nature of migrant workers, as well as the consequences of the global financial crisis that also hit China’s economy in 2009 – and especially export-oriented sectors in which many migrant workers are concentrated.

To counter the substantial attrition of the MHS and maintain the original sample size, a re-sampling based on the pre-survey census was conducted (see Kong et al, 2009 and Meng, 2013, for details). This implies that, starting from the second wave, the MHS consists of two separate samples: the “old sample”, which tracks migrants from the first wave, and the “new sample”, a fresh randomly drawn sample that is followed in the subsequent waves. The representativeness of the new sample is discussed in detail in Kong et al (2009).<sup>5</sup>

**Table 2 – Attrition Rates**

	RHS		UHS		MHS	
	Households	Individuals	Households	Individuals	Households	Individuals
Wave 1	8,000	31,791	5,005	14,695	5,007	8,446
Wave 2	7,992	32,171	4,735	14,859		
Tracked	7,992	31,652	4,735	13,841	1,821	3,512
(Attr. Rate)	0.1%	0.4%	5.4%	5.8%	63.6%	58.4%
Add. members		519		1,018		409
New sample					3,422	5,426

Source: RUMiC Wave 1 and 2.

## 2.6 Questionnaire Modules and Variable Content

The RUMiC survey provides a rich set of variables. The questionnaires cover detailed standard demographic and socioeconomic characteristics of household heads and members, and include also questions on physical and mental health status, life events, social networks, household consumption, assets and expenditure. This information offers a significant opportunity to investigate interesting yet under-researched topics concerning migration and labor markets in China.

Table 3 provides an overview of the questionnaire modules for each wave. All three surveys include questions about personal characteristics such as age, marital status, children, hukou status, as well as migration status and experience. Additional questions are asked about social

insurance coverage (medical, unemployment, etc.), health and health status (recent illness, disabilities, medical expenditures, etc.) as well as a set of subjective well-being questions are asked. The three surveys include comparable information on years of schooling, highest level of completed education, training, and scores at the national university entrance exam.

The modules about employment provide information on the type of employment (e.g. wage-work, self-employment, domestic work without pay), hours of work, earnings, job search, firm ownership, as well as occupation and industry codes (which have been harmonized across the two waves). Several questions are asked about entrepreneurship and self-employment, including aspects related to borrowing and the existence of credit constraints to starting a business. The MHS includes additional questions about the workplace conditions (e.g. catering, accommodation) and retrospective queries about the first job after migration.

Several modules include questions about family members such as the spouse, the parents and siblings of the household head, as well as the young and adult children living in the household. The module on social and family network includes information on the number of greetings sent during Chinese New Year (a proxy for network size), as well as detailed characteristics concerning up to five closest contacts who have provided help in the previous twelve months, including the number of gifts exchanged.

All surveys include questions about life events that occurred in the previous twelve months (marriage, childbirth, house purchase, death of household member), as well as information on whether such events were expected. Questions about household income, consumption, assets, savings, remittances and expenditures are also included. Furthermore, the MHS contains information on housing conditions, durable goods and the rural hometown of origin.

While the majority of the questions are included in both waves, a few are only present in one wave. Most of these refer to new modules introduced in the second wave, e.g. the supplementary survey on health status, questions on risk attitudes, and the comparison of satisfaction level and income. Questions only asked in the first wave yet not in the second usually reflect characteristics and traits that are arguably stable over time, at least in the short run.

**Table 3 – Questionnaire Modules**

	RHS		UHS		MHS		
	Wave 1	Wave 2	Wave 1	Wave 1	Wave 1	Wave 2 Old S.	Wave 2 New S.
Personal characteristics	✓	✓	✓	✓	✓	✓	✓
Social insurance	✓	✓	✓	✓	✓	✓	✓
Health status	✓	✓	✓	✓	✓	✓	✓
Supplementary health condition questions				✓		✓	✓
Subjective well-being	✓	✓	✓	✓	✓	✓	✓
Risk and preferences		✓		✓		✓	✓
Education and training background	✓	✓	✓	✓	✓	✓	✓
Employment situation	✓	✓	✓	✓	✓	✓	✓
Information on young children (≤16 years old)	✓	✓	✓	✓	✓	✓	✓
Information on adult children (>16 years old)	✓		✓		✓	✓	
Information on spouse living apart					✓	✓	✓
Information on parents <sup>†</sup>	✓		✓		✓	✓	✓
Information on siblings <sup>†</sup>		✓		✓			
Social network <sup>†</sup>	✓	✓	✓	✓	✓	✓	✓
Information on farm land <sup>†</sup>		✓				✓	✓
Life events <sup>†</sup>	✓	✓	✓	✓	✓	✓	✓
Comparison of satisfaction level and income <sup>†</sup>		✓		✓		✓	✓
Household income, expenditure, assets <sup>†</sup>		✓	✓	✓	✓	✓	✓
Durable goods listing <sup>†</sup>					✓	✓	✓
Present housing and living conditions <sup>†</sup>			✓		✓	✓	✓
Information on rural hometown <sup>†</sup>					✓	✓	✓

Source: RUMiC Wave 1 and 2. <sup>†</sup> questions asked to household head or spouse only.

## 2.7 Summary Statistics

We provide a simple effective picture of RUMiC data in Table 4, reporting statistics for selected variables including age, gender, marital status, number of children, work status, and an indicator for self-employment. The general health status is a self-reported categorical variable indicating “very poor”, “poor”, “average”, “good” or “excellent” health. Happiness refers to a question asking whether individuals are happy when considering all aspects of their life, with responses including “much less than usual”, “less so than usual”, “same as usual” and “more so than usual”.

The average age is about 37 years for individuals living in rural areas and about 40 for urban residents. Migrant workers are relatively younger, with an average age of about 29 years. The share of males is around 0.5 in both the RHS and UHS, but is somewhat higher in the MHS, reflecting that males are more likely to migrate. The share of married individuals varies across the three surveys, and is highest in the UHS (between 0.68 and 0.69) and lowest in the MHS (between 0.54 and 0.60). Within the MHS, the share of married migrants is relatively higher in the old sample (second wave), reflecting the higher attrition characterizing unmarried migrants, who are generally more mobile.

Individuals in rural areas report a number of children slightly above 2; in contrast, the number for urban residents is slightly above 1. This difference is a consequence of the one-child policy, which was strictly enforced in cities but was less binding in rural areas, with a second child allowed in many provinces if the first was a girl. All individuals report good levels of health and happiness.

With regard to employment, there is much variation across the three surveys. Only 40% of individuals in rural areas are employed in non-farm work, reflecting the continuing importance of the agricultural sector in villages. About 60% of individuals in urban areas are in wage work, while 5% engage in self-employment activities. Approximately two-thirds of migrants in the first wave were working as employees; however, as many as 20% were in self-employment, reflecting that many migrants decide to start their own activity once in the city – often in the informal sector. One remarkable aspect is the sharp decrease in the share of wage workers in the tracked sample of migrants, and the increase in the share of individuals engaging in self-employment. One likely explanation behind such a change is that the economic crisis had a relatively larger impact on export oriented sectors, inducing migrants to change city or return to their home village and thus drop out from the survey.

**Table 4 – Selected Descriptive Statistics**

Variables	RHS		UHS		MHS		
	Wave 1	Wave 2	Wave 1	Wave 2	Wave 1	Wave 2 Old S.    New S.	
Age	36.290 (19.065)	37.028 (19.213)	40.139 (19.214)	40.808 (19.307)	28.790 (12.695)	29.311 (13.499)	29.433 (12.760)
Male (%)	0.518 (0.500)	0.516 (0.500)	0.495 (0.500)	0.496 (0.500)	0.568 (0.495)	0.545 (0.498)	0.568 (0.495)
Married (%)	0.620 (0.485)	0.631 (0.483)	0.684 (0.465)	0.691 (0.462)	0.557 (0.497)	0.608 (0.488)	0.539 (0.499)
N. of children	2.069 (1.195)	2.060 (1.200)	1.349 (0.891)	1.387 (0.996)	0.795 (0.921)	1.397 (0.812)	1.457 (0.883)
Health status	1.975 (0.816)	2.041 (0.830)	2.182 (0.789)	2.289 (0.761)	1.765 (0.744)	2.013 (0.759)	1.822 (0.740)
Happiness *	1.758 (0.610)	1.795 (0.592)	1.749 (0.603)	1.764 (0.578)	1.778 (0.646)	1.888 (0.624)	1.843 (0.619)
Wage work **	0.347 (0.476)	0.354 (0.478)	0.631 (0.483)	0.619 (0.486)	0.682 (0.466)	0.538 (0.499)	0.711 (0.453)
Self-employment **	0.059 (0.235)	0.062 (0.242)	0.050 (0.218)	0.055 (0.229)	0.213 (0.409)	0.339 (0.473)	0.210 (0.407)

Source: RUMiC Waves 1 and 2. Notes: standard deviation in parentheses. \* statistics refer to individuals older than 16 present during the survey. \*\* statistics refer to individuals 16-64. Health status is measured on a scale 1 to 5, where 1 is “Excellent” and 5 “Very poor”. Happiness is measured on a scale 1 to 4, where 1 is “Happy more than usual” and 4 is “Happy much less than usual”.

## 2.8 Access to RUMiC

To date, IZA has made the scientific use files for the first two waves publicly available. Access is granted for scientific purposes to researchers from universities and research institutes. Data application forms can be downloaded from the International Data Service Center (IDSC) of IZA website (<http://idsc.iza.org/rumic>) and should be completed with a description of the research project and submitted to [idsc@iza.org](mailto:idsc@iza.org). Data files are provided in formats that are compatible with standard statistical software.

### 3 A Review of Literature Based on RUMiC

Despite being relatively recent, a significant amount of research has already been produced using RUMiC data. In this section, we provide an overview of selected studies. Table 5 lists the research topics and main findings of these papers.

**Table 5 Selected Papers Based on RUMiC**

<b>Paper</b>	<b>Research topics</b>	<b>Survey</b>	<b>Key findings</b>
Akay et al (2012)	Subjective well-being of migrants; positional concerns	RHS, UHS, MHS	Economic success of other migrants and rural workers reduces subjective well-being, while comparison with urban residents increases it.
Akay et al (2013)	Subjective well-being; remittances	MHS	Migrants experience welfare gains by sending remittances; both altruistic and contractual motivations are at work.
Biavaschi et al (2013)	Left-behind; human capital; sibling effects	RHS	Sibling influence on schooling performance is stronger among left-behind children, compensating potential adverse effects of migration.
Cui et al (2013)	Returns to education; entrepreneurship; wage distribution	MHS	Emergence of a dual labor market with rising returns to education for urban residents and disadvantaged migrant workers.
Démurger and Li (2013)	Remittances; left-behind; occupational choice	RHS	Own migration experience leads to subsequent off-farm work. Migration of a family member induces more farming work.
Démurger et al (2012)	Wage inequality; entrepreneurship;	UHS	Earning gaps across firm ownership decrease over time.
Frijters et al (2010)	Wage inequality; labor market segmentation	UHS, MHS	Migrants have lower wages and welfare insurance contributions compared to urban residents.
Frijters et al (2011)	Entrepreneurship; credit constraints	UHS, MHS	Discrimination in salaried jobs leads to migrants choosing self-employment, which offsets the negative effects of credit constraints on self-employment.
Ge and Lehmann (2013)	Worker displacement; labor market segmentation	UHS, MHS	Re-employment outcomes after displacement differ between migrants and urban workers.
Giulietti et al (2012)	Self-employment; wage differentials	RHS, MHS	Wage differential is a key determinant in the choice between self-employment and wage work.
Giulietti et al (2013)	Entrepreneurship; left-behind; return migration	RHS	Return migration promotes self-employment among non-migrants, while current migration reduces it.
Qu and Zhao (2013)	Wage inequality	UHS, MHS	During the 2000s, wage inequality decreased among migrant workers while it increased among urban workers.
Zhang and Zhao (2011)	Entrepreneurship; social networks	MHS	Migrants with larger social-family network are more likely to choose self-employment.
Zhang and Zhao (2013)	Migration decision	MHS	Rural-urban migrants who move further away need to be compensated with larger income gains.

It is possible to group the studies in Table 5 into the main topics covered. These include occupational choice and entrepreneurship, subjective well-being, wage inequality and labor market segmentation, and the determinants and consequences of migration.

### **3.1 Entrepreneurship and occupational choice**

A strand of the literature focuses on the determinants of entrepreneurship among migrant workers. Frijters et al (2011) analyze the link between self-employment and credit constraints. In order to explore the effect of credit constraints, they use information from the second wave of the UHS and MHS, classifying workers into self-employed, involuntary self-employed, want-to-be self-employed and happy-to-be-salaried. Their descriptive analysis shows that being discriminated in salaried jobs leads to migrants becoming self-employed, thereby offsetting the negative effects of credit constraints on entrepreneurship. Giulietti et al (2012) model migrant workers' choice between wage work and self-employment using the first wave of the MHS. Their key explanatory variable is the wage differential between the two sectors, estimated through an endogenous switching model. Migration selectivity is taken into account using information from the RHS. The authors document that self-employed migrants have higher earnings than wage workers, and that they are positively selected in terms of unobservable characteristics. Zhang and Zhao (2011) investigate the link between the self-employment choice and social networks among migrant workers using the first wave of the MHS. They use the number of greetings sent during the Chinese New Year as a proxy for the size of the social network. To deal with the endogeneity between social networks and the self-employment choice, they instrument the network variable using the distance between the province of origin and the city to which migrants moved. Their instrumental variable estimates show that migrants with larger social networks are more likely to be entrepreneurs.

A few other papers study the occupational choices of individuals left behind in rural areas. Using the first wave of the RHS, Démurger and Li (2013) investigate the impact of remittances on the occupational choices. The authors use a switching probit to estimate the impact of belonging to a migrant household on individuals' occupational choice. They find that an individual's own migration experience leads to subsequent off-farm work, while the migration experience of a family member increases the likelihood of left-behind individuals being in farm work. They interpret this last result as the consequence of an increase in the reservation wage induced by remittances. A complementary study by Giulietti et al (2013) uses a recursive trivariate probit model to model three simultaneous states: self-employment, living in a household with return migrants, and living in a household with migrants currently in the city. They show that return migration promotes self-employment among household members who have not yet migrated; however, left-behind individuals are less likely to become self-employed compared to those living in non-migrant households.

### **3.2 Subjective well-being**

Another line of research focuses on the determinants of migrants' subjective well-being (SWB). Akay et al (2012) study the impact of relative income on rural-to-urban migrants' SWB, arguing that this depends on different reference groups such as peer migrants, urban residents and individuals in the rural areas of origin. Using the first wave of the MHS, the authors aggregate responses from all SWB questions into one synthetic measure of well-being and regress it on a rich set of personal characteristics, including the income level of the reference group. Their estimates document that the economic success of other migrants and rural workers reduces migrants' subjective well-being. In contrast, it is positively correlated with the income level of urban residents, which is interpreted as being determined by better income prospects for migrants in urban areas. A study by Akay et al (forthcoming), also based on the first wave of the MHS, relates the well-being of migrants to the remittances sent to the families left behind. Their SWB measure is regressed on the amount of remittances sent home and other socio-demographic characteristics, with the results indicating that migrants experience welfare gains by sending remittances. Additional specifications in which remittances are interacted with selected characteristics document the importance of family arrangements, e.g. whether migrants are married, and whether the spouse and/or kids are left behind in villages. In particular, migrants with family responsibilities are not as satisfied by sending remittances as those with no such obligations. Furthermore, using SWB data, the authors investigate the motivations behind sending remittances, finding that both altruistic and contractual motivations are at work among their sample.

### **3.3 Wage inequality and labor market segmentation**

Another group of studies explores wage inequality and labor market segmentation in urban areas. Frijters et al (2010) use the first wave of the UHS and MHS to estimate a wage equation, and include the score of the university entrance exam as a proxy for individual ability. They show that returns to ability and education are lower for migrants compared to urban residents. The authors argue that migrants' lower earnings are likely related to their shorter labor market experience, caused by the temporary nature of migrations. Cui et al (2013) use the 1995 and 2002 waves of the CHIP together with the first wave of the MHS to estimate a wage equation for migrants and urban residents. In contrast to rising returns to education for urban residents, the authors find no significant changes in the returns for

migrants. According to their analysis, the wage increases are mostly attributable to the type of job, industry and the firm ownership, but not to the level of education. Their results highlight the emergence of a dual labor market, with outcomes of migrant workers worsening over time. Démurger et al (2012) investigate the earning differentials by firm ownership among urban workers. The authors conduct a regression analysis by ownership type and a decomposition of the earnings differentials using the 2002 wave of CHIP together with the first wave of the UHS. Regression results show that the earning gap across types of firm ownership decreases over time. Their decomposition analysis documents the increasing importance of differences in endowments in determining earning gaps by ownership type. Qu and Zhao (2013) investigate the evolution of wage inequality over time using the 2002 CHIP data and the first wave of the UHS and the MHS. Applying decomposition methods, the authors find that wage inequality within migrants decreased, while wage inequality within urban residents increased.

In a recent contribution, Ge and Lehmann (2013) investigate the cost of job loss and the consequences of worker displacement in urban labor markets, using the second wave of the UHS and MHS. They first distinguish between displaced workers, quitters and stayers, before subsequently analyzing these groups upon re-employment, assessing the length of unemployment spell, earnings, happiness and health. Their results indicate that displaced migrant workers do not face relatively long unemployment spells or wage penalties, unlike displaced urban workers. The authors interpret these results as evidence of segmented labor markets in urban areas.

### **3.4 Determinants and consequences of migration**

Zhang and Zhao (2013) investigate the determinants of migration choice by focusing on the role of distance. They develop a framework in which the distance between the home village and destination city is included in the utility function. For the empirical exercise, the authors first calculate a variable measuring individuals' expected income in each potential destination, which is included in their OLS and discrete choice models analyses as an explanatory variable. Their results based on the first wave of the MHS suggest that rural-to-urban migrants do not find it convenient to move far away from their home villages. In particular, their estimates imply that in order to induce migrants to move 10% further away from home, their income has to increase by at least 15%.

A recent paper by Biavaschi et al (2013) focuses on the effects of parental migration on the educational outcomes of children left behind, highlighting the importance of sibling interactions in such a context. Exploiting the panel dimension of the RHS, the authors estimate the relationship between older and younger siblings' school performance using OLS and fixed effects methods. Their main findings suggest that sibling influence is stronger among left-behind children. In particular, older sisters primarily have a positive influence on their younger siblings. The authors interpret their results as evidence that sibling effects are a mechanism to shape children's educational outcomes and that adjustments within the family left behind can reduce the hardship determined by parental migration.

#### **4 Conclusions**

This paper provides a brief description of the RUMiC Survey, which has been created with the aim to foster research on migration and labor markets in China. The survey comprises the Rural Household Survey, the Urban Household Survey and the Migrant Household Survey. Scientific Use Files for the first two waves are made publicly available by IDSC, the data bank center of IZA. Further waves will be made available in the near future.

The large scale, in-depth topics and longitudinal aspect are the core features of RUMiC. The surveys follow more than 18,000 households over time and include comprehensive information on household and personal characteristics. With such features, RUMiC data allow studying the consequences of the Great Migration for rural and urban areas in China, as well as the migrant themselves. Furthermore, it is possible to better understand the evolution of labor markets over the past few years, including the consequences of the financial crisis.

In the paper, we have surveyed a few studies based on RUMiC data, highlighting topical questions concerning entrepreneurship, happiness and wage inequality. Much more research is expected to be produced now that Scientific Use Files for the second wave have been made publicly available, which allow exploiting the panel aspect of the survey. Future lines of research could include a more rigorous study of the consequences of migration for left-behind individuals (especially children and the elderly) and the in-depth analysis of internal migration patterns over time, devoting particular attention to modeling circular and return migration.

## Endnotes

---

<sup>1</sup> More information about CHIP project can be obtained from <http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/21741?archive=ICPSR&q=China+Household+Income+Project>.

Information about CHARLS data can be obtained from <http://charls.ccer.edu.cn/en>.

<sup>2</sup> See Gong et al (2008) for an overview of other household surveys implemented in China.

<sup>3</sup> It is beyond the scope of this paper to describe all the details of the census listing and sampling; we refer to Kong (2010) and Gong et al (2008) for further details regarding the listing scheme and random sampling procedures of the survey.

<sup>4</sup> The RUMiC survey does not track migrants who return to their villages.

<sup>5</sup> For example, as mentioned by Knog et al (2009) using the pre-survey census provides information on whether workplaces had shut down or changed. However, it does not inform about the magnitude of downsizing within each workplace.

## References

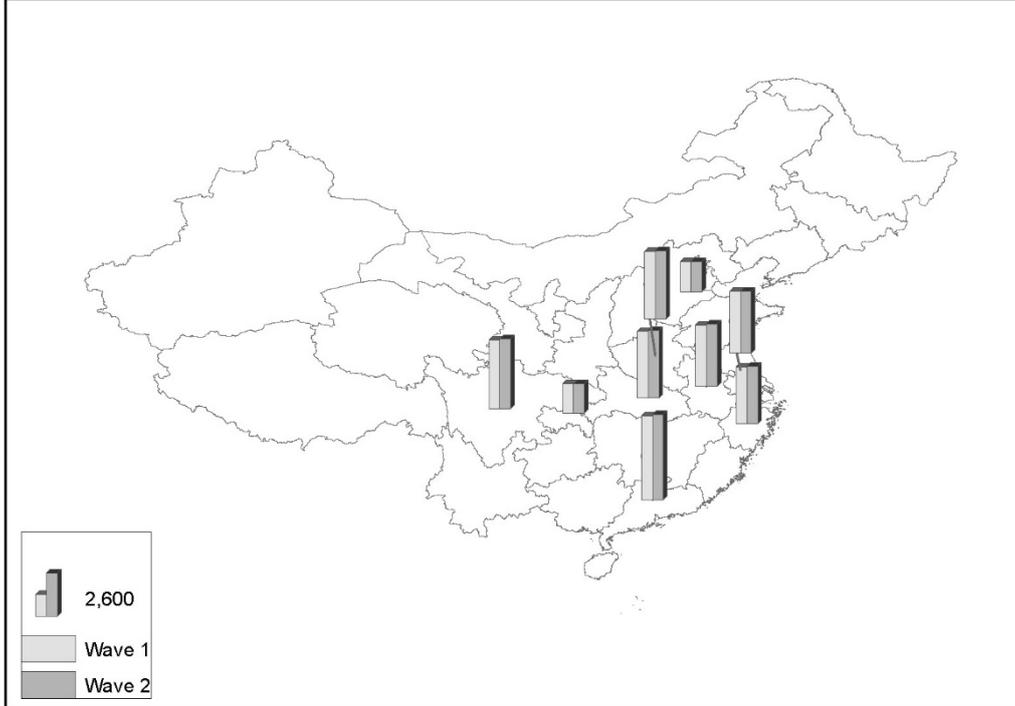
- Akay, A., Bargain, O., & Zimmermann, K. F. (2012). Relative concerns of rural-to-urban migrants in China. *Journal of Economic Behavior and Organization*, 421-441.
- Akay, A., Giuliatti, C., Robalino, J. D., & Zimmermann, K. F. (forthcoming). Remittances and well-being among rural-to-urban migrants in China. *Review of Economics of the Household*.
- Arni, P., Caliendo, M., Künn, S., & Zimmermann, K. F. (2013). The IZA Evaluation Dataset Survey: A Scientific Use File. IZA Discussion Paper, forthcoming.
- Biavaschi, C., Giuliatti, C., & Zimmermann, K. F. (2013). Sibling Influence on the Human Capital of the Left Behind. IZA Discussion Paper No. 7859.
- Caliendo, M., Falk, A., Kaiser, L. C., Schneider, H., Uhlendorff, A., van der Berg, G., et al (2011). The IZA Evaluation Dataset: towards evidence-based labor policy making. *International Journal of Manpower*, 32(7), 731-752.
- Cui, Y., Nahm, D., & Tani, M. (2013). Earnings Differentials and Returns to Education in China, 1995-2008. IZA Discussion Paper No. 7349.
- Démurger, S., & Li, S. (2013). Migration, Remittances, and Rural Employment Patterns: Evidence from China. *Research in Labor Economics*, 37, 31-63.
- Démurger, S., Li, S., & Yang, J. (2012). Earnings differentials between the public and private sectors in China: Exploring changes for urban locals residents in the 2000s. *China Economic Review*, 23, 138-153.
- Frijters, P., & Meng, X. (2009). Rural to Urban Migration in China: an Overall View. Australian National University Working Paper.
- Frijters, P., Kong, T., & Meng, X. (2011). Migrant Entrepreneurs and Credit Constraints under Labour Market Discrimination. IZA Discussion Paper No. 5967.
- Frijters, P., Lee, L., & Meng, X. (2010). Jobs, Working Hours and Remuneration Packages. In X. Meng, C. Manning, L. Shi, & T. N. Effendi (Eds.), *The Great Migration: Rural-Urban Migration in China and Indonesia* (pp. 47-73). Edward Elgar Publishing.
- Ge, Y., & Lehmann, H. (2013). The costs of worker displacement in urban labor markets of China. *IZA Journal of Labor and Development*, 2(4).
- Giuliatti, C., Ning, G., Zimmermann, & F, K. (2012). Self-employment of rural-to-urban migrants in China. *International Journal of Manpower*, 33, 96-117.
- Giuliatti, C., Wahba, J., & Zimmermann, K. F. (2013). Entrepreneurship of the left-behind. *Research in Labor Economics*, 37, 65-92.

- Gong, X., Kong, S. T., Li, S., & Meng, X. (2008). Rural-urban migrants. In L. Song, & W. T. Woo (Eds.), *China's Dilemma* (pp. 110-152). Asia Pacific Press.
- Kong, S. T. (2010). Rural-Urban Migration in China: Survey Design and Implementation. In X. Meng, C. Manning, L. Shi, & T. N. Effendi (Eds.), *The Great Migration: Rural-Urban Migration in China and Indonesia* (pp. 135-150). Edward Elgar Publishing.
- Kong, S. T., Meng, X., & Zhang, D. (2009). Impact of economic slowdown on migrant workers. In R. Garnaut, L. Song, & W. T. Woo (Eds.), *China's New Place in a World in Crisis: Economic, Geopolitical, and Environmental Dimensions* (pp. 233-260). The Australian National University E Press.
- Kamal-Chaoui, L., Leman, E. and Zhang, R. (2009). Urban Trends and Policy in China. OECD Regional Development Working Papers 2009/1.
- Lehmann, H., Muravyev, A., & Zimmermann, K. F. (2012). The Ukrainian longitudinal monitoring survey: Towards a better understanding of labor markets in transition. *IZA Journal of Labor and Development*, 1(9).
- Meng, X. (2012). Labor Market Outcomes and Reforms in China. *Journal of Economic Perspectives*, 26, 75-101.
- Meng, X. (2013). Rural-Urban Migration: Trend and Policy Implications (2008-2012). In R. Garnaut, & L. Song (Eds.), *China: A New Model for Development and Growth* (pp. 179-197). The Australian National University E Press.
- Meng, X. M., Shi, L., & Effendi, T. N. (2010). *The Great Migration: Rural-Urban Migration in China and Indonesia*. Edward Elgar Publishing.
- National Bureau of Statistics of China (NBS). (2012). *China Statistical Yearbook 2012*. China Statistics Press, Beijing.
- Qu, Z., & Zhao, Z. (2013). Wage inequality of Chinese rural-urban migrants between 2002 and 2007. PEP Working Paper 2013-04.
- Zhang, J., & Zhao, Z. (2011). Social-Family Network and Self-Employment: Evidence from Temporary Rural-Urban Migrants in China. *IZA Discussion Paper No. 5446*.
- Zhang, J., & Zhao, Z. (2013). Measuring the Income-Distance Tradeoff for Rural-Urban Migrants in China. *IZA Discussion Paper No. 7160*.

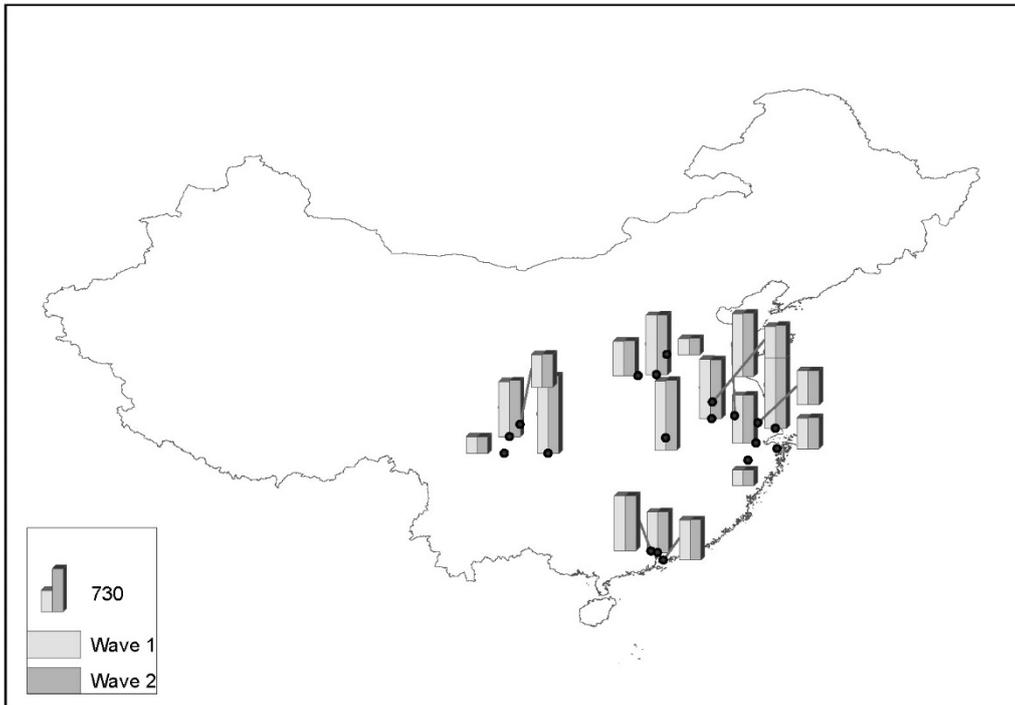
Appendix

Figure 1 – RUMiC Spatial Coverage, Wave 1 and 2

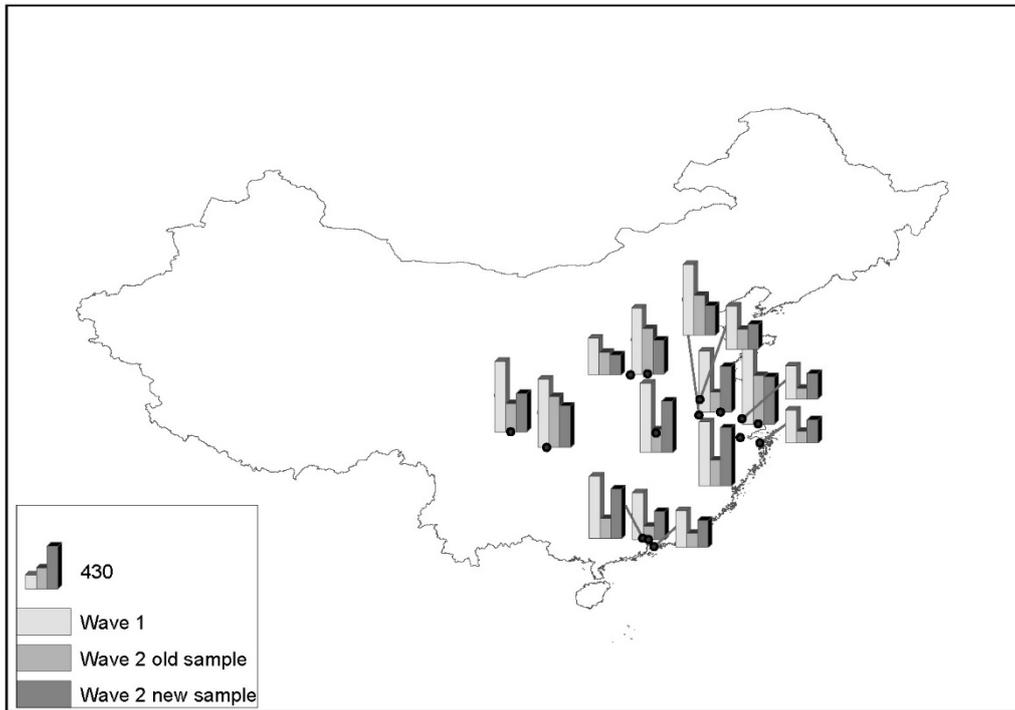
RHS



UHS



# MHS



Source: RUMiC Wave 1 and 2. The height of bins refers to sample size as indicated in the legend.