Zentrum für Entwicklungsforschung

# Impacts of Rural-urban Migration on Demographic Transition, Human Capital Investment and Agricultural Productivity in China's Rural Areas

Inaugural-Dissertation

zur

Erlangung des Grades

Doktor der Agrarwissenschaften

(Dr. agr)

der

Landwirtschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

Vorgelegt von

Yi Zhang

aus

Jiangxi, China

Referent: Professor Dr. Joachim von Braun
Korreferent: Professor Dr. Thomas Heckelei
Tag der mündlichen Prüfung: 10.07.2015
Erscheinungsjahr: 2015

## Abstract

This dissertation focuses on the impacts of rural-urban migration on the rural areas in China. It consists of five chapters. The first chapter introduces the research problems and presents the framework for studying three selected impacts of migration. These impacts are on the demographic change of the rural population, human capital investment, and agricultural productivity, and are all respectively discussed in the middle three chapters. The last chapter is a case study of a typical Chinese village with massive rural-urban migration.

The second chapter first estimates the scale and age-structure of the rural-urban migrants, and then separates the effects of migration on the rural demography by performing simulations with the Cohort Component Method and using data from China's latest censuses in 2000 and 2010. In addition, it uses household data to confirm the huge effect of rural-urban migration on the demographic structure.

The third chapter develops a theoretical framework to investigate the relationship between migration and education. Empirical research reveals a robustly positive effect of migration on educational attainment among the stayers by proposing a novel instrument of the availability of local train stations to deal with the endogeneity.

The fourth chapter sets up a theoretical model to study the impacts of migration on agricultural productivity and empirically employs a Simultaneous Equations Model estimated by two-step-least-square method. Empirical results show that migration of the labor force reduces agricultural productivity and households with migration do not invest more in agriculture unless the land size reaches an optimal level. Migration along with land transfer can improve agricultural productivity.

The fifth chapter presents a case study of a Chinese village which is transforming its labor-intensive agriculture into a capital-intensive one based on changes in relative scarcity of production factors triggered by the rural-urban migration. It indicates that migration as an external force has broken equilibrium of the traditional agriculture and leads modern agriculture to take off by inducing capital to substitute for labor in agriculture.

## Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit den Auswirkungen von Land-Stadt-Migration auf ländliche Gebiete in China. Die Arbeit besteht aus fünf Kapiteln. Das erste Kapitel beschreibt den Forschungsgegenstand und das Bezugssystem für die Analyse von drei ausgewählten Auswirkungen von Migration. Diese drei ausgewählten Auswirkungen sind der demographische Wandel auf dem Land, die Investitionen in Humankapital und landwirtschaftliche Produktivität. Diese drei Aspekte werden jeweils in den drei mittleren Kapiteln behandelt. Das letzte Kapitel ist eine Fallstudie eines typischen chinesischen Dorfes mit massiver Land-Stadt-Migration.

Das zweite Kapitel schätzt zunächst die Ausmaße der Land-Stadt-Migration und die Altersstruktur der Migranten und separiert dann die Effekte der Migration auf die ländliche Demographie mithilfe von Simulationen mit der Kohorten-Komponenten-Methode und Daten der letzten chinesischen Volkszählungen in 2000 und 2010. Zusätzlich werden Haushaltsdaten genutzt, um die enormen Effekte der Land-Stadt-Migration auf die demographische Struktur zu erfassen.

Im dritten Kapitel wird ein theoretischer Rahmen entwickelt, in dem die Beziehung zwischen Migration und Ausbildung untersucht werden. Empirische Forschung mithilfe einer IV Schätzung mit fixen Effekten, die das Vorhandensein von lokalen Bahnhöfen als neues Instrument nutzt, zeigt einen robusten positiven Effekt von Migration auf den Ausbildungsstand derer, die auf dem Land bleiben.

Im vierten Kapitel wird zunächst ein theoretisches Modell zur Untersuchung der Auswirkungen von Migration auf die landwirtschaftliche Produktivität aufgestellt. Empirisch wird mithilfe der 2SLS-Methode ein Modelsimultan geschätzt. Die empirischen Ergebnisse zeigen, dass Arbeitskräftemigration die landwirtschaftliche Produktivität reduziert und Haushalte mit Migration nicht mehr in die Landwirtschaft investieren bis ihr Landbesitz eine optimale Größe erreicht. Migration kombiniert mit Landtransfers kann jedoch die Produktivität verbessern.

Das fünfte Kapitel präsentiert eine Fallstudie eines chinesischen Dorfes, das aufgrund relativer Knappheit von Produktionsfaktoren, die durch Land-Stadt-Migration ausgelöst wurde, seine arbeitsintensive Landwirtschaft in eine kapitalintensive umwandelt. Die Fallstudie zeigt, dass Migration als externer Einflussfaktor das Gleichgewicht von traditioneller Landwirtschaft gebrochen hat und nun dazu führt, dass durch die Substitution von Arbeit durch Kapital die moderne Landwirtschaft Einzug hält.

## Acknowledgments

First of all, I would like to express my very great appreciation to my supervisor, Professor Joachim von Braun. From the research proposal to the final dissertation, Professor von Braun gave me his guidance and support whenever I needed it. I am also very thankful to my two tutors, Dr. Danial Tsegai who helped me draft my research proposal from 2010 to 2012 and Dr. Julia Anna Matz who gave me invaluable advice and comments during my thesis writing from 2012 to 2014. Discussions with them always provided me with inspiration. Without their invaluable help I could not have finished this dissertation.

I would like to thank the academic and administrative staff at ZEF, particularly, Dr. Gunther Manske, Mrs. Rosemarie Zaber, and Mrs. Maike Reta-Amin. Not only I, but also every doctoral student has been impressed by their kind and patient help in our daily life. In addition, I am grateful to Mr. Guido Luchters for his patience and help with data analysis. I wish to acknowledge Mrs. Christine Husmann's help in my German general abstract. I would like to thank all the staff and my colleagues at ZEF for their help in offering me the resources during my study.

I am particularly grateful for the support given by Professor Jingbei Hu. As a specialist in the field of the Chinese economy, he gave me a lot of comments and suggestions on my research problem. I would like to offer my special thanks to Mrs. Susanne Reiter. She often discussed the social problems of China with me and offered me enormous help in many aspects of my life in Germany.

I wish to acknowledge the China Scholarship Council supporting me financially between 2010 and 2013 and the Foundation Fiat Panis, which financed my field survey in China in 2011. I also thank the Carolina Population Center at the University of North Carolina at Chapel Hill for supplying me with the data set of China Health and Nutrition Survey.

Finally, my sincere gratitude goes to my dear family, my parents and older brother, who have always supported me throughout my life and education. I surely would not have been able to go so far without your understanding and encouragement.

## Table of Contents

## List of Tables

## List of Figures

## List of Abbreviations

| | |
|---|---|
| CHNS | China Health and Nutrition Survey |
| HRS | Household Responsibility System |
| IMR | Infant Mortality Rate |
| NBSC | National Bureau of Statistics of China |
| NELM | New Economics of Labor Migration |
| RLCL | Rural Land Contracting Law |
| TFR | Total Fertility Rate |

# 1 Introduction and overview

## 1.1 Research objectives and questions

Joseph Stiglitz, a Nobel Prize laureate, said that technological innovation in America and urbanization in China would be "two keys" to mankind's development in the 21st century (The Economist 2014a). China's policymakers share the same opinion. In 2013, China's new leaders emphasized urbanization as a key "gigantic engine" to the country's economic growth and set an ambitious target for urbanization in China. The Chinese "National New-type Urbanization Plan (2014-2020)" sets very clear targets: the urban population will reach 60% of its total population in 2020 compared to 53.7% in 2014.[1] To promote urbanization, the Chinese central government has enacted the change of the *hukou* system and completely relaxed the household registration in township and small cities since July, 2014. The ability to reach this target of 60% of urbanization by 2020 rests not only on whether the urban areas can absorb the large flow of rural-urban migrants, but also on whether the rural areas can supply urban development with enough surplus labor force and match the food demands resulting from urban expansion while the rural labor migrates out in massive scale and also if the target is economically viable. This dissertation has a more specific focus, and will address the impacts of rural-urban migration on the rural areas and the dynamic responses of the rural areas to migration.

The impacts of migration on the rural areas in China have not been studied exhaustively. Regarding relevant studies on China, the available literature tends to focus on the impacts of migration on urban areas (Bhattacharyya and Parker 1999; Zhang and Song 2003; Au and Henderson 2006; Lu and Song 2006; Fu and Gabriel 2012; Zhong et al. 2013), or on migrants per se (Wang and Zuo 1999; Zhao 1999; Tuan et al. 2000; Carrillo 2004; Li 2010; Knight and Gunatilaka 2010; Zhang and Luo 2012). Less attention has been paid to the rural areas. It is certain that the urban expansion occurring due to inflows of millions of migrants has corresponding impacts in the rural areas. Rural-urban migration can significantly influence the welfare of the populations in the

[1] The full content of "National New-type Urbanization Plan (2014-2020)" can be found on http://news.xinhuanet.com/english/china/2014-03/16/c_133190495.htm (access on January 26, 2015)

rural areas, e.g. by increasing household income and decreasing income inequality (Zhang et al. 2007).

This dissertation will focus on rural areas. The key objectives of this study are twofold. The first objective is to contribute to closing the literature gap pertaining to the impacts of rural-urban migration on the demography and human capital in rural areas. Second, this study seeks to provide theoretical and empirical evidence on the effects of rural-urban migration on agricultural productivity.

There is no doubt that the decrease in the rural population was the most prominent change in China's demographic structure over the past three decades. As shown in Figure 1.1, China's total population grew continuously between 1949 and 2012 while its rural population began to decrease in 1995 from 859 million to 642 million in 2012. China's rural population has been falling by an average of 12 million people a year since 1995, an unprecedented scale in the history of human urbanization so far.[2] However, the rural population is still in surplus according to the agricultural stock (Mi 2008).

The starting point of this dissertation is the demographic change to the rural population in China due to rural-urban migration, in particular the change of age structure. This study will start with an estimation of rural-urban migration because the official statistics and surveys in China do not offer detailed demographic information on rural-urban migration at a national level (Zhong et al. 2013). My study will cover the following four research questions which are also the topics of the four chapters of this dissertation.

[2] Rural-urban migration is not the only contributor to urbanization. Urban population growth can be broken into three components: natural increase, net migration , and urban administrative reclassification. Chan (2012) found that in China urban natural increase, net migration, and urban reclassification accounted for about 15, 43, and 42 percent of the urban growth, respectively. Administrative reclassification could contribute a lot to the increase in the total number of urban population. But net migration could not only influence total number but also the structure of the population. This is why net migration is the most important factor to population changes.

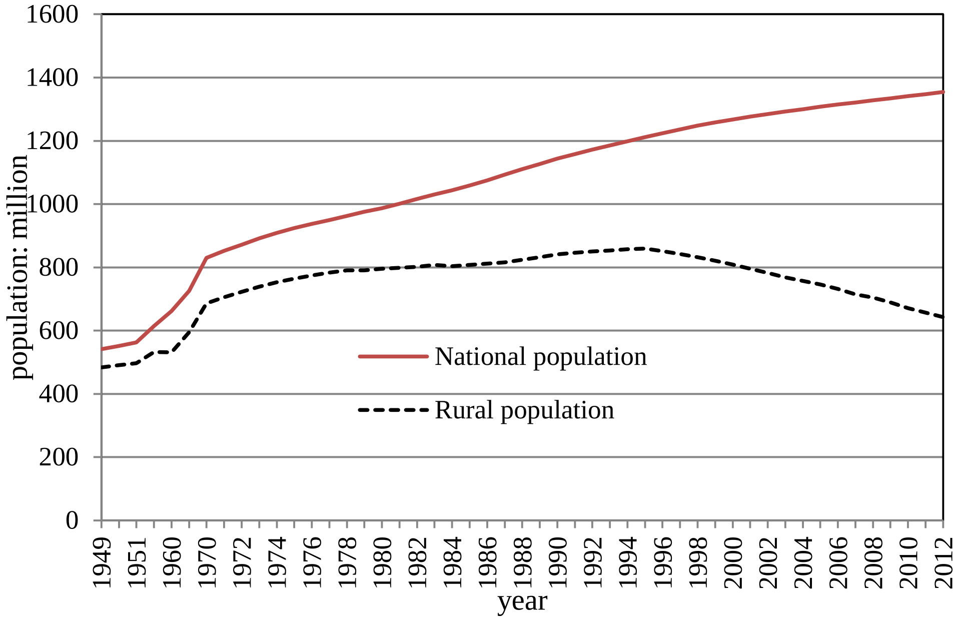



Figure 1.1 Change of national and rural population in China during 1949-2012
Source: NBSC, China Statistics Yearbook 2013

### 1.1.1 Demographic impacts of rural-urban migration on the rural population

Rural-urban migration changes not only the number of people, but also the population composition in the rural areas. In general, the age structure of the rural population differs significantly from the national population due to rural-urban migration. Figure 1.2 shows the national and the rural population pyramids in 2010. We can see that the rural population pyramid demonstrates a more aging population than the national one given that most rural-urban migrants form part of the young population (Taylor et al. 2003). It is obvious that of the large quantity of young people migrating to urban areas results in a huge gap in this age group between the national and the rural populations. Therefore, the first research question is: what are the net impacts of rural-urban migration on age structure of the rural population?

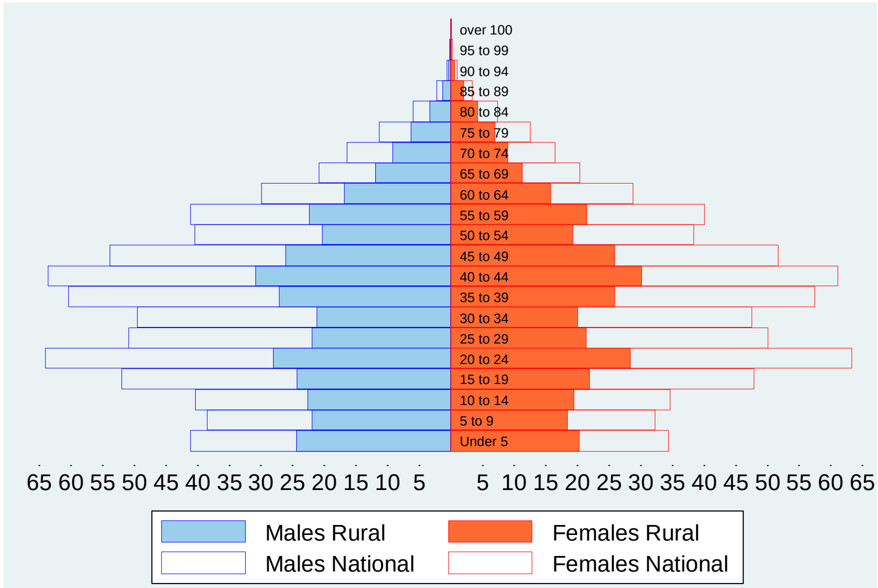



Figure 1.2 China's rural population and national pyramid in 2010
Source: National Bureau of Statistics of China, Census 2010, population in millions

### 1.1.2 Impact of rural-urban migration on human capital investment in rural areas

The decrease of quantity of the rural population let us focus on the productivity of the rural population. The reduction and ageing of the agricultural labor force due to rural-urban migration calls for more human capital investment. Human capital has proved to be one of the most important inputs in farm production which can increase income and alleviate poverty in rural areas (Huffman 2001, 3-5). As such, it is worthwhile to focus on human capital investment in the young left behind in rural areas when I consider the impacts of rural-urban migration in the long run. To that end, the second research question is: what are the impacts of rural-urban migration on human capital investment in rural areas?

### 1.1.3 Impact of migration on agricultural productivity

An important question is whether the massive and continuous decrease of the agricultural labor force has harmed Chinese agriculture. Fortunately, the real agricultural output has been increasing over the past three decades with the reduced labor force input. As shown in Figure 1.3, it shows that the agricultural labor force was less than 300 million in 1980, peaked in 1990 to almost 400 million, then declined to 350 million in 1996 and more or less continued to decline. Since 2003, the agricultural labor force has been on a steady decline. By contrast, real agricultural output has increased fivefold from 100 billion yuan to 550 billion yuan between 1980 and 2012, which triggers us to consider the question: What are the impacts of rural-urban migration on agricultural productivity?

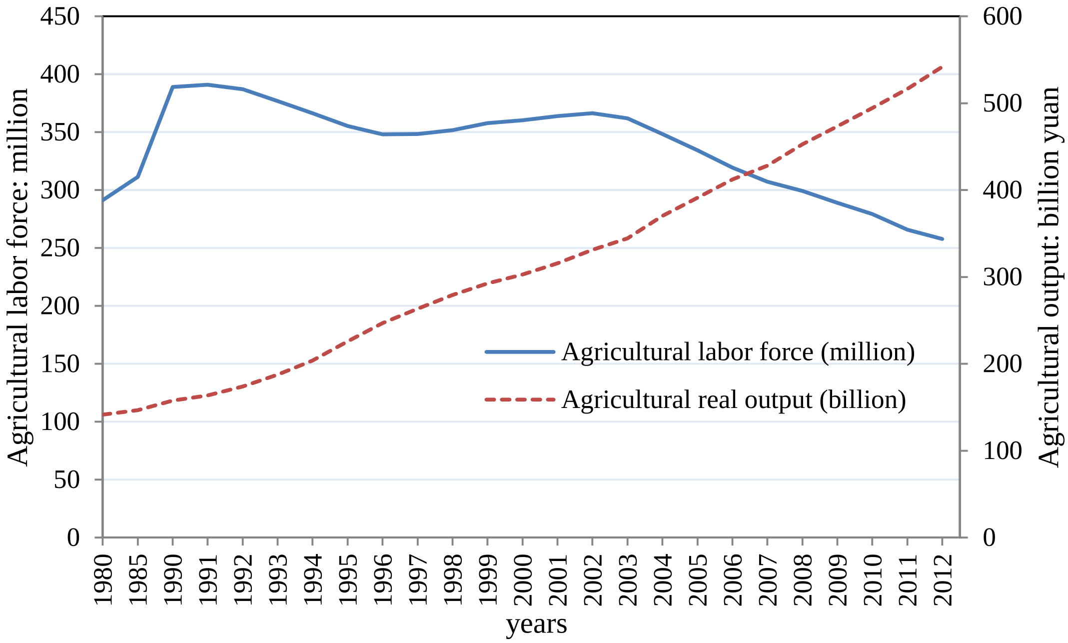



Figure 1.3 Change of China's agricultural labor force and China's agriculture production
Source: NBSC, China Statistical Yearbook in 2013
Note: output is deflated to 2012 prices (yuan)

### 1.1.4 Rural-urban migration and transformation of agricultural villages in China

A village, the most basic administrative unit of the nation, has experienced huge challenges due to massive rural-urban migration in China since 1980s. It is more intriguing for us to consider the impact of massive rural-urban migration on an independent social unit, a village. However, there are few studies on rural-urban migration at a village level (Wu and Yao 2010). This study will investigate the last question: what are the impacts of rural-urban migration on agricultural production in a rice village in China?

The above four research questions can be illustrated in Figure 1.4 where the starting point of analysis is the demographic change of population in Chapter 2. Then the research extends in three directions. One direction is the remaining young adults in the village.

They will become the future labor force in the rural areas. I will investigate whether they invest more in human capital or not in Chapter 3. The other direction is the remaining agricultural labor force. In Chapter 4, I will study whether the remaining agricultural labor force has the capacity to improve agricultural productivity. Finally, in Chapter 5, I will employ a case study to reflect how a traditional village adopts their agriculture production for loss of labor due to rural-urban migration.

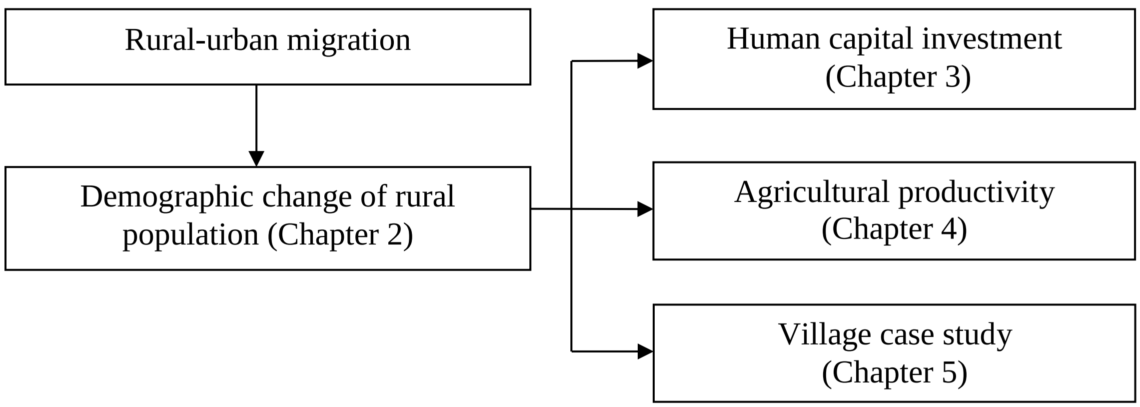



Figure 1.4 Framework of analyzing the impact of rural-urban migration on source communities

### 1.2 Some typical features of rural-urban migration in China

Compared to other developing countries, China has three important distinguishing characteristics regarding rural-urban migration.

First, the quantity of migrants is tremendous, according to the latest two censuses in 2000 and 2010, the number of internal temporary migrants has increased from about 144 million to 261 million (Qiao and Huang 2013). This number of internal temporary migrants accounted for 20% of the total population in 2010.

Second, migration has lasted over the past three decades and will continue to do so for a long time. Some scholars predict that China may need another 35 years before its agricultural labor share declines to 10%, the percentage that the agricultural labor shares of Japan, South Korea, and Taiwan have stabilized (Holz 2008).

The last characteristic is that most migrants keep close touch with their rural families because the household registration policy denies migrants permanent residence status in urban areas. Most migrants leave their children and the elderly at the rural home and send remittances to help their family with agricultural production and the children's education.

## 1.3 Definitions of terms

The definitions of urban and rural populations in this study are in accordance with those official definitions by the central government of China. According to the official definitions[3]: Urban population refers to the following population: the total population of districts under the jurisdiction of a city with district establishment, the population of street committees under the jurisdiction of a city without district establishment, population of resident-committees of towns under the jurisdiction of a city without district establishment, and the population of resident-committees of towns under the jurisdiction of a county. Rural population refers to total population except urban population defined above.

The definition of rural-urban migrant is twofold. In the household survey, a rural-urban migrant is defined for the purpose of this study as the rural population that does not live in the rural household, but migrates to urban areas for employment and does not participate in the farm production at the rural home. It excludes other forms of migration, for instance, migration for marriage, study, military service and so on. The key words of this definition are migration for employment and lack of involvement in farm production.

In the national aggregate data, this above definition of rural-urban migrant does not hold true. For instance, the rural-urban migrants in Chapter 2 are estimated through calculation of the changed rural population from Census data in 2000 and 2010 where the definition of urban population also includes the changed statistics of areas that are declared “urban” and were rural before. Such reclassification of population would be counted as rural-urban migration if I use the change of urban population or rural population to estimate the migrants. However, the other chapters used the data from household survey (CHNS) do not have the problem of administrative reclassification.

[3] The definitions can be found in the category of classifications & methods in the official website of national bureau of statistics of China on http://www.stats.gov.cn (access on January 20, 2015).

## 1.4 Institutional backgrounds

Before I introduce the history and characteristics of rural-urban migration in China, it is important to clarify two fundamental institutions associated with rural-urban migration in China.

### 1.4.1 *Hukou* system

The *hukou* system instituted in 1958 is a kind of household registration system which classify households by location and occupation (Chan and Zhang 1996). It is known as the dual household registry classification. The first classification classifies all Chinese citizens by residential location. Social security for every citizen is covered by the local government in the city where the citizen registers. The second classification classifies all residents by economic activity agriculture-oriented or non-agriculture oriented in every residential location. This classification is based on the 1950s occupational division. The population classified under agriculture *hukou* is entitled to obtain a certain amount of collectively-owned farm land in the local community. The population classified as non-agriculture *hukou* is granted subsidies in health care, public education, unemployment insurance, and guaranteed a minimum income and basic welfare support by the local urban government.

To change one's *hukou* in China is extremely difficult according to regulations before 2014. The local governments don't want to accept new migrated populations to register in their *hukou* system permanently because it will increase their financial burden to include these new migrants in their social security system (Bao et al. 2011). Therefore, the majority of rural-urban migration is informal or temporary migrants in China. Based on the latest Chinese population census in 2010, about 260 million Chinese were internal migrants, of which 220 million were temporary internal migrants (Peng 2011).

The two most important reforms of the *hukou* system occurred in the 1980s and in 2014. The first was to allow the rural population to work in cities starting in the 1980s. This has stimulated substantial growth of rural-urban migration (Chan 2010). But most of these rural-urban migrants are not registered in the local *hukou* system. The second was to relax the restriction of the *hukou* registration in urban areas according to the new policy published on 30 July, 2014. The ministry of Public Security introduced a guideline for upcoming reform of China's *hukou* system. Migrant workers are encouraged to register in towns and small cities. The restrictions of *hukou* registration in medium-sized cities will be gradually relaxed. The difference between agriculture *hukou* and non-agriculture *hukou* will be eliminated, and the discriminatory policies in education, healthcare, employment and other social welfare systems against migrant workers will be abolished in future.

### 1.4.2 Household Responsibility System (HRS)

The Household Responsibility System (HRS) was first adopted in the agricultural sector in 1979. It is a farming institution that leased collectively-owned land to every rural household with agriculture *hukou.* The ownership still belongs to the village collective but households can make agricultural operating decisions independently (Lin 1989). A household's obligations were to fulfill state procurement quotas and pay various forms of local taxes during the 1980s and 1990s (Crook 1985). However, a household could retain surplus above the stated obligations. In late 2002, the ministry of finance began to subsidize farming households. In 2006, the Chinese government abolished all state procurement quotas and agricultural taxes. Nowadays, there are four Chinese agricultural subsidies including a grain subsidy, an input subsidy, a quality seed subsidy, and an agricultural machinery subsidy (Huang et al. 2013).

Normally, land is distributed within the village. The principle of egalitarianism is the guiding law in the distribution of land leases (Lin 1989). The length of the period to redistribute land can be between 5 and 10 years. The land distribution process is as follows. Land is first classified into several different grades and then every household obtains a parcel from each grade. In 1988, to prevent land being left idle when a household migrates, the Chinese government issued a series of policies to promote agricultural land transfers. In 2002, to strengthen the contract of land use-rights, the Chinese government extended the duration of land-use rights from 15 to 30 years in the Rural Land Contracting Law.

### 1.4.3 The evolution of rural-urban migration policy in China

The evolution of rural-urban migration policy in China since the Chinese reform and open policy in 1979 can be summarized as follows:

1979-1983: The Chinese government prohibited rural labor migration. Under the no-migration policy, the rural population was prohibited to leave their home area without official permission.

1984-1988: The Chinese government relaxed the no-migration policy. The government started to allow rural laborers to enter cities for employment as long as their food was self-supplied.

1989-1991: The Chinese government restricted rural labor migration to fight high inflation. The boom of township enterprises reduced rural-urban migration.

1992-2000: The government allowed, and to some extent encouraged, rural-urban migration due to the huge labor demand of highly labor-intense industrial sectors.

2000-2014: The government encouraged temporary migration. There is no limit to temporary rural-urban migration. However, the *hukou* system is still an institutional block for the migrants' access to the urban social security system.

July, 2014: The government relaxed the *hukou* system for temporary migration. The Chinese central government relaxed the registration of the *hukou* system in township and small cities. The migrants registered in those cities can now enjoy the same social welfare as the local residents. The *hukou* system in medium and mega cities will gradually be open to migrants but under some conditions formulated by local government.

## 1.5 Dataset

Three main data sources will be used in this dissertation. The statistical data from yearbooks and censuses published by National Bureau of Statistics of China will be used in Chapter 2. The statistical data of yearbooks from 1996 to 2014 and censuses in 2000 and 2010 can be downloaded from the official website of the National Bureau of Statistics for the People's Republic of China.[4] A village survey collected by the author

[4] The website of statistical data published by National Bureau of Statistics of the People's Republic of China is http://www.stats.gov.cn/tjsj/ (access on January 19, 2015)

will be used in Chapter 5. These two datasets will be introduced in the correspondent chapters. The main dataset of this dissertation comes from the China Health and Nutrition Survey (CHNS) which will be used in parts of Chapter 2, Chapter 3 and Chapter 4. To avoid reiteration, the dataset (CHNS) is introduced here. The data of China Health and Nutrition Survey (CHNS) can be downloaded from the Carolina Population Center website at the University of North Carolina at Chapel Hill.[5]

The China Health and Nutrition Survey (CHNS) dataset is longitudinal data collected between 1989 and 2009 by the Carolina Population Center of the University of North Carolina at Chapel Hill.[6] The survey applies a multistage cluster sample design where the first layer is made up of nine densely populated provinces that account for 56% of the country's population.[7] Figure 1.5 displays the provinces surveyed. Counties of low, middle, and high average income levels are randomly chosen from each province and three villages are randomly selected from each county. The sample covers approximately 4,000 households with 26,000 individuals every survey round and these participants are partly followed over time. I restrict my analysis to the rural sample (approximately 2,700 households and 11,000 individuals per round). Data consists of information on the individual, household and community levels for every round. Furthermore, I am not able to use the early data of the survey due to questions on migration only being included from 1997 onwards and thus restricting the dataset to the latest five rounds that took place in 1997, 2000, 2004, 2006, and 2009. The survey includes questions on demography, education, employment, housing conditions, income, agricultural practices, time use,

[5] The China Health and Nutrition Survey website of the Carolina Population Center at the University of North Carolina at Chapel Hill is http://www.cpc.unc.edu/projects/china (access on January 19, 2015)

[6] I thank the National Institute of Nutrition and Food Safety, China Center for Disease Control and Prevention; the Carolina Population Center, University of North Carolina at Chapel Hill; the National Institutes of Health (NIH; R01-HD30880, DK056350, and R01-HD38700); and the Fogarty International Center, NIH, for financial support for the collection and preparation of the China Health and Nutrition Survey data. Furthermore, I am grateful to the China-Japan Friendship Hospital and the Chinese Ministry of Health for their support while working with the survey data.

[7] These nine provinces are Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong. Heilongjiang was added in 1997 when Liaoning quit the survey, but Liaoning has returned to the project since the 2000 survey.

commune facilities, and health and nutritional measures. CHNS has regarded as one of the most used panel datasets on China's household level (Chen and Zhang 2009).[8]


Figure 1.5 Map of the survey provinces
Source: available on http://www.cpc.unc.edu/projects/china/about/proj_desc/chinamap
(access on January 19, 2015)

### 1.6 Methodologies, limitations and recommendations for the future research

In this section I will briefly introduce the methodology used for analysis, identify the limitations of each chapter and give some recommendations for the future research.

I first analyzed the demographic impact of rural-urban migrants in China during 2000-10 in Chapter 2. Specially, I measured the net effect of rural-urban migration on the age structure of the population in rural China. The study revealed that rural-urban migration is the most significant factor to change the age structure of rural population. The method I used is to simulate a closed rural population based on Census data of 2000 and 2010 using the Cohort Component Method and compare the simulated rural population with the actual population. The difference was assumed to be rural-urban migrants. An assumption of this method is the change of rural population consisting of natural change and net migration. But in a fast developing country like China the

[8] In China, there are five major household level panel datasets including RCRE (Research Center for Rural Economy), RHS (Rural Household Survey), CHNS (China Health and Nutrition Survey), CHIPS (Chinese Household Income Project Survey), and CAPM (Chinese Academy of Preventative Medicine).

definition of “urban” and “rural” changed several times up until 2006 (Kamal-Chaoui et al. 2009). Therefore, the third source of change to the rural population is the administrative reclassification which meant that in some rural areas the rural population decreased as a result of parts of these areas becoming reclassified as urban. However, data is limited to measure the population change due to the administrative reclassification. A possible avenue for future research is to calculate the number of people affected by the Chinese government’s reclassification.

Chapter 3 focuses on whether rural-urban migration could promote human capital investments in rural areas. Meanwhile, educational attainment is an important contributor to migration decisions. Therefore, migration decision and educational attainment have the relationship of reverse causality. I proposed a novel instrument of the availability of train station to deal with the endogeneity due to the reverse causality. I also employed a fixed effect regression to deal with the fixed effect of local community. The limitations of this method cannot be neglected. Although the statistical tests for the instrument proved that the availability of train station is a strong and valid instrumental variable for migration, the complicated relationship between migration and a train station is still not very clear. For instance, a train station supplies convenient transportation for the local people who want to migration. On the other hand, it can promote the local job market which could stop the local people from migrating. Future research should focus on revealing the clear relationship between access to a train station and migration. Actually, the relationship between the local labor market, migration and railway station is a pertinent topic, so much so that China’s central government has invested on railway tremendously in recent years (The Economist 2014b).

Future demographic changes to the rural population will also influence rural agricultural productivity. I studied the impact of migration on agricultural productivity in Chapter 4. However, for a farm household, the decision to migrate and agricultural production are simultaneously determined. The simultaneity variable bias in the functions of migration and agricultural production would result in the problem of endogeneity. To account for this, I constructed a system of three equations to determine agricultural productivity, migration decision, and capital investment in agriculture. Considering the constraint of land size for farmers to invest in agriculture, I included the effect of interaction between migration and farm size in the empirical analysis. I proposed an instrument, percentage of land collectively plowed in village for household's capital investment in agriculture. Due to limited information about this instrument in the survey, the relationship between the instrument and capital investment in agriculture is not very clear despite the fact that the relevant statistical tests confirmed the validity of the instrument. As such, the effect of migration on agriculture is debatable (Yang et al. 2014). Future research should try to separate the different effects of migration on agricultural production.

Chapter 5 was a case study which sought to examine the demographic changes of the rural population, educational attainment and agricultural production in an actual village. I described these topics in a typical village which has long experienced the effects of rural-urban migration. The case study sought to observe the actual transformation of traditional agriculture due to rural-urban migration. Rural-urban migration has changed the relative price of agricultural inputs, namely labor, land and physical capital, which trigger this village to transform from labor-intensive agriculture to capital-intensive agriculture. In particular, machinery operation has become an unavoidable trend in agriculture in this village. A limitation to this case study is its representativeness. This case study can represent the traditional agriculture villages in southern China. Furthermore, the economic principle of agricultural transformation reflected in this case study is very meaningful for the villages experiencing rural-urban migration. Possible future research will be to study this village over a very long-term. It will illustrate the process of how the induced innovation happens in agriculture in this village through the change of relative prices of inputs due to migration.

# 2 The demographic impact of rural-urban migration on the rural population

## 2.1 Introduction

In 1995, the rural population in China peaked at 859 million and accounted for 71% of the total population in China. From then, it stopped growing and began a stable downward trajectory. Only 15 years later, in 2010, the rural population had decreased by 21% to 671 million people, which made up just 50% of the total population in China. The following year, the rural population in China had decreased to less than its urban population for the first time in the country's history. The Chinese government expects that the proportion of China's rural population will decrease to 30% of the total Chinese population in another 15 years later, namely by 2025.[9]

The rapid decline of China's rural population has been the most striking demographic change for China over the last decades. Over the last two decades, given that the rural population has had a positive natural growth rate, the urban population has stably increased, and that China has been a rather closed society towards international migration, the primary reason for the rapid decline of China's rural population must be the increase in rural-urban migration within the country. Therefore, rural-urban migration is the most important contributor to changes in China's rural demography, especially in age composition of its rural population (Cai 2010).

The objective of this chapter is to evaluate the effects of rural-urban migration on the age-structure of the rural population in China. It should be noted that the impacts of rural-urban migration on changes to the rural demographic have rarely been studied in and outside of China (Peng 2011). One of the main reasons for this may be the absence of relevant data on migration. The population censuses of China did not include any information on migration prior to the 1990s (Liang 2001). There are no data on rural-urban migration in regularly published statistics in China to date. To overcome these statistical shortages, this chapter will first estimate the magnitude of rural-urban

[9] In June 2013, China's central government sets a target urbanization rate of 70% by 2025, which equates to roughly 900 million people living in cities. To realize this target implies that the rural population will continue to decline from about 50% to 30% by 2025, which means that more than 250 million rural residents (20% of the rural population) will migrate to the cities by 2025 (The New York Times 2013).

migration in China by performing simulations of the rural population during 2000-10, since such quantitative estimations are indispensable for studying the impacts of migration on the rural demographics. This method of simulation is widely used in the international demographic literature and by key international organizations that estimate demographic trends. However, it has not been applied to research on China's rural-urban migration so far. Thus, the present study may be the first to apply this method as a means of inquiry into the demographic impacts of rural-urban migration in China.

The rest of this chapter is organized as follows. Section 2.2 reviews the literature related to the methodology used in this study. Section 2.3 introduces the demographic changes of the rural population in the long and short run. Section 2.4 illustrates the methodology of the cohort component model and assesses the simulated results. Section 2.5 confirms evidence of demographic characteristics of rural-urban migrants by household data. And finally Section 2.6 summarizes my findings in this chapter and draws conclusions.

## 2.2 Literature review

Most of the literature on the demographic impacts of rural-urban migration on the rural population in China is descriptive. The literature argues that both the elderly and the young are more likely to be left behind in the rural areas. For example, using the Chinese 1990 population census, Liang et al (2002) found that non-migrants in rural areas were five years older than migrants. Similarly, Biao (2007) summarized that the population left behind in rural China included mostly the elderly and children because migrants were mostly men and young people. For instance, Hare (1999) found that young people between the ages 16–25 and 26–35 were most likely to migrate. Yang and Guo (1999) claimed that men had a higher propensity to migrate.

Rural-urban migration has become the dominant factor changing the rural population. Cai (2013) found that the population of out-migrants from rural communities in central and western China had surpassed the population of natural growth between 2000-2010 and induced negative population growth in these provinces. Not only has the size of the rural population been affected by rural-urban migration but the structure of the rural population has also changed. The rural elderly who are left behind face the precarious situation of a lack of young adults to support them (Giles et al. 2010). The old-age dependency ratio was 14% in rural China in 2010. This ratio is estimated to increase to 23% by 2020 and 30% by 2030 due to rural-urban migration (Cai et al. 2012). Zhong and Xiang (2012) analyzed the demographic structure of the rural population based on a survey of five provinces in 2010-2011 and their results showed that the rural population would be seriously aging due to out-migration of the young, and that aging of the agricultural labor force is speeding up.

However, as stated in the introduction, to date there is no national statistics available on rural-urban migrants.[10] To deal with the data shortage, several different methods have been developed to estimate migrants. How to choose them depends on the diversity of need. Among them, the Cohort Component Method has been widely applied because it can include some useful demographic characteristics including age, gender, and fertility rate (Smith et al. 2001). For example, the U.S Census Bureau has used the Cohort Component Method to project national population. The United Nations (UN) has employed this method to estimate global and national population since the 1950s (O'Neill et al. 2001).

Although the demographic impact of rural-urban migration on the rural population has been rarely studied in China, some western researchers have assessed the impact of international migration on the national populations of western countries. They used methods of simulation to evaluate the effect of migration on population size and age composition in western countries (Kusek and Rist 2004; Khandker et al. 2010; Philipov and Schuster 2010). Philipov and Schuster (2010) simulated a closed population without migration and then compared it with the observed population in Europe during 1960-2007. The difference between the simulated and observed population was net migration. They found that in-migration increased the population of western European countries while decreased that of southern European countries. Le Bras (1991) employed a similar method to analyze the demographic impact of international migration in OECD countries between the first censuses after World War II and in early 1980s. His results found that the effects of migration had a significant impact on population growth rather than on the

[10] The latest three population censuses have included migration information but they only published the total number of migrants according to their *hukou* registered places.

age structure of the population. Coleman (2008) evaluated the impact of international migration on European countries and argued that migration was the dominant factor determining the size and composition of most European countries. Loh and George (2007) examined the effect of net international migration on the population size and age structure of Canada by comparing two projected population scenarios for the next 50 years. They found that the effect of international migration contributed significantly to an increase in population, but was limited in its effectiveness at ameliorating population aging in Canada.

The other method to study the demographic impacts of migration is to integrate demographic analyses within econometric models. Zimmermann et al. (2006) reviewed different approaches to estimate structural change using Econometric and Simulation models in which cohort analyses could be employed to study the relationship between demographic and economic factors when the population dynamics and life cycle table were available. Von Braun (1979, 48-60) employed a demographic cohort model to study the effects of a decline of the agricultural labor force on economic growth in West-Germany during the 1960s and 1970s. He separated the off-farm migration from other demographic factors such as ageing, death, disability, and retirement. However, similar studies in China cannot be done due to a lack of aggregate demographic data on the agricultural labor force, for instance, ageing, death, and retirement of labor force. These data have not been collected by the China's statistical authority. There are also few demographic data on the rural population published so far. Even the fertility rate and mortality rate of the rural population are not published regularly in China except in censuses.

## 2.3 Demographic change of China's rural population

### 2.3.1 Change of China's rural population in the long run

Until the late 1980s, the central government strictly regulated rural-urban migration in China through the household registration system of the *hukou* policy. The *hukou* policy functioned so effectively that until 1985 the average rural-urban migration rate was close to zero (Zhao 2000). China launched the economic reform and open policy in 1978, and then promoted the institutional reform of household responsibility system (HRS) in agriculture. Since then, the demand for labor in urban areas and pressure of surplus labor in rural areas has induced the largest scale of rural-urban migration in the world (Taylor 2001). In the last three decades rural-urban migration has been the most prominent factor to influence the rural population.

Figure 2.1 illustrates the evolution of China's rural population between 1949 and 2010. From 1949 until 1995, the rural population increased from 484.02 million to its peak with 859.47 million. Then the rural population fell into a fast downward track after its historical peak in 1995 (Chi et al. 2012). In 2010, the rural population had decreased to the same amount as the urban population, namely 671.13 million.

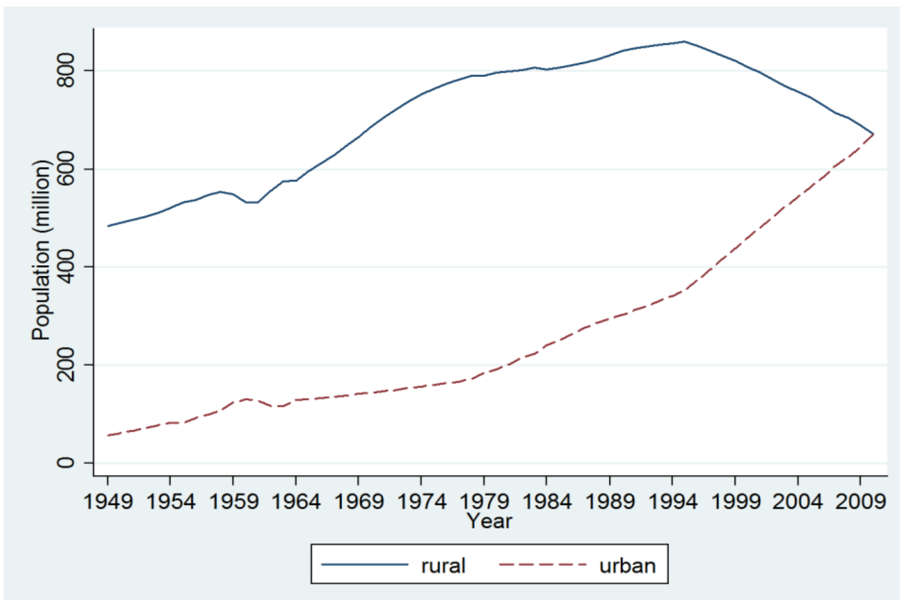



Figure 2.1 Evolution of urban and rural population in China 1949-2010
Source: NBSC, China Statistic Yearbook 2011

The Chinese national statistics bureau doesn't publish the data on rural birth rates and death rates each year. However, the latest research shows that during 1950-2000 the differences between the national birth rate and the rural birth rate are small (Feng 2013). Therefore, it is acceptable to use the national birth and death rates to illustrate the changes in rural birth and death rates in the long run. Figure 2.2 shows the changes in national population growth, birth and death rates during 1949-2010. As shown in Figure 2.2, the demographic transition started in the 1960s in China when the birth and death rates began to fall continuously (Hussain 2002; Mason and Feng 2005; Wang and Mason

2008). In 1964, the population had a presumably bunching of births due to a rebound from the famine in 1959-1963.[11] The death rate remained at a steady low level from 1960s onward while the birth rate continuously declined during 1970s (Hussain 2002). In 1979, China implemented the "one-child family planning policy" to control the birth rate (Hess 2013). The population growth steadily decreased from 1980 to 2010 except for some years of fluctuation in 1980s. The natural growth rate remained under 15 per thousand which is the threshold between high and low growth by international standards (Leete and Alam 1993). After 2000, the natural growth rate fell close to the lower level of five per thousand. According to estimates China's low birth rate will remain low in the long term (Zhang and Zhao 2006; Cai 2008).

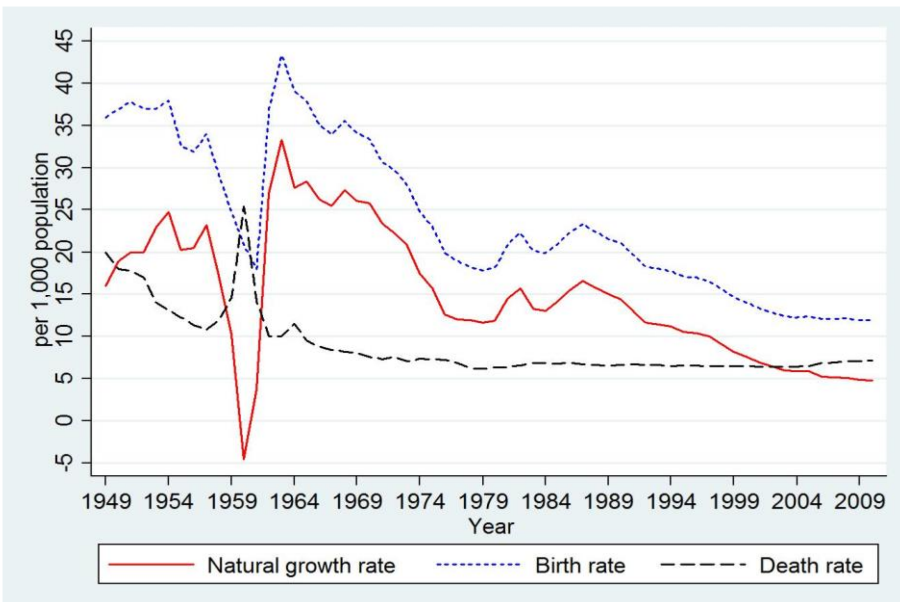



Figure 2.2 Population growth, birth rate, and death rate 1949-2010
Source: NBSC, China Statistic Yearbook 2011

### 2.3.2 Change of China's rural population in the short run

The latest censuses of 2000 and 2010 offered the most comprehensive data on the Chinese population so far. Specifically, they included the demographic information on rural areas. Figure 2.3 compares the rural population pyramids in 2000 and 2010. It

[11] The precipitous variation between 1958 and 1961 attributed to the large death toll from the famine during a period of unprecedented political and social disaster known as "Great Leap Forward" (Drèze and Sen 1989). The detailed explanation can be found in Ashton et al. (1984).

shows that the young rural population has fallen a lot because they are more likely to migrate (Akay et al. 2012). The out-migrants are mainly young people aged 10-35. For instance, the group of age 10-14 had more than 45 million people in 2000. But in 2010, there were less than 30 million people in the age group 20-24, which indicated that at least 15 million people had migrated into urban areas without considering the population change of the age group 10-14 due to mortality. A similar situation happened in the age groups 25-29 and 30-34.

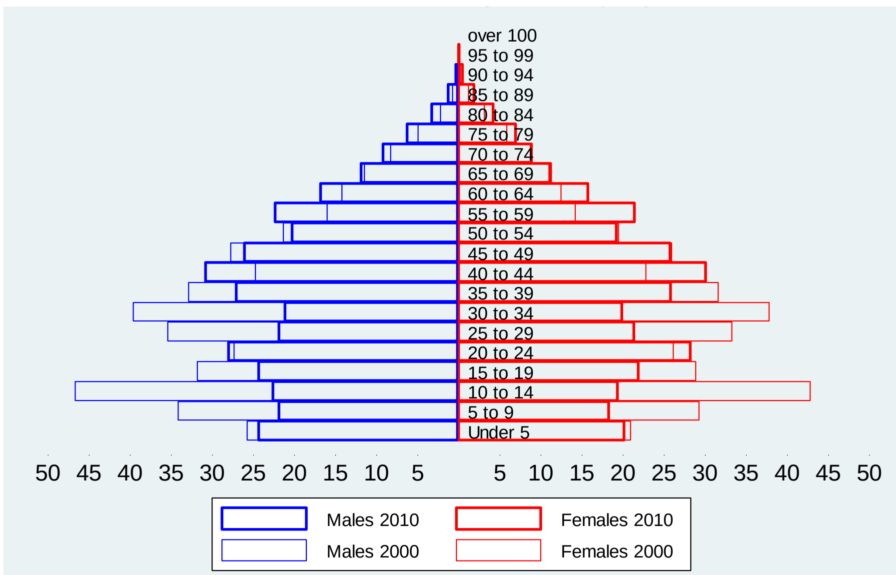



Figure 2.3 China's rural population pyramid in 2000 and 2010
Source: Chinese population censuses in 2000 and 2010. Population in millions

The massive young out-migrants from rural areas, which must change the age structure of rural population. The total dependency ratio is an indicator of changes to the age structure.[12] Figure 2.4 shows the total dependency ratios in Chinese cities, townships and villages in 2000 and 2010. The total dependency ratio was highest in villages both in 2000 and in 2010, which suggested that the dependent population, people of age 0-14 and over 64, represented a higher percentage of the total population in villages than in townships and in cities. This obvious difference in age structure is the consequence of

[12] Dependency ratio is a ratio of the population of different age group. For instance, the population is divided into three age groups, namely aged 0-14, aged 15-64 and aged 65 and over. Total dependency ratio, according to China's statistics, is defined as the population of aged 0-14 and aged 65 and over divided by population of aged 15-64.

rural-urban migration. The rural aging problem could be a challenge for social security and the public health system in rural areas (Joseph and Phillips 1999; Giles et al. 2010).

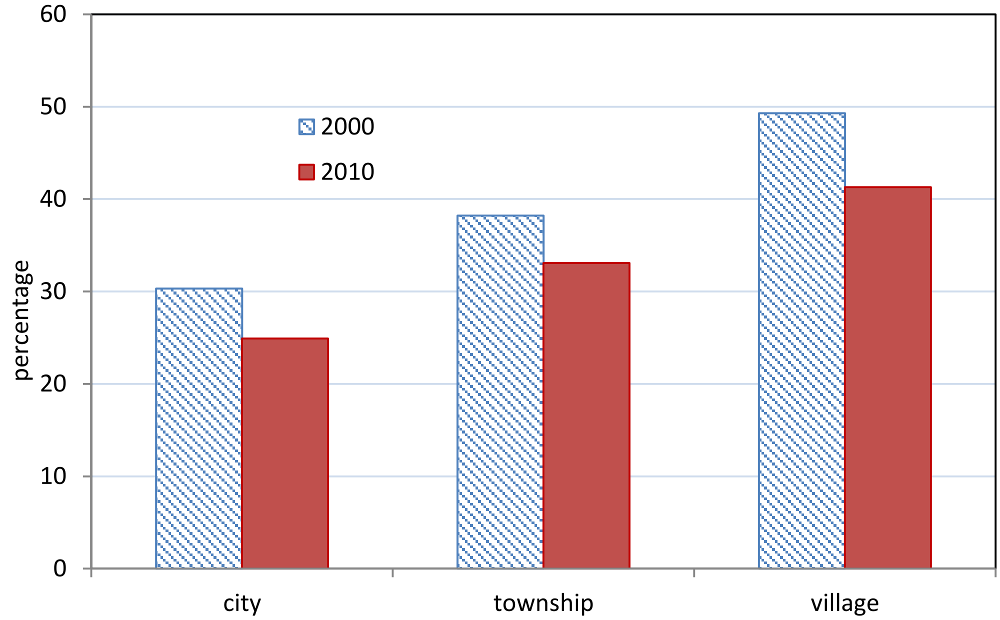



Figure 2.4 Total dependency ratios in cities, townships and villages in 2000 and 2010
Source: Chinese population censuses in 2000 and 2010

## 2.4 Evaluation of the net impact of rural-urban migration on the rural population

### 2.4.1 Cohort Component Method to simulate population growth

In the Cohort Component Method, the basic components of population change include births, deaths, immigration and outmigration. Births and immigration add to the population while deaths and emigration are subtracted from it. At any given time interval, the population can change due to any of these four components. The net plus of the birth and death rates is referred to natural growth. The difference between immigration and outmigration is net migration. If the population is closed without migration, e.g., the global population, the population growth depends entirely on the natural growth. However, if population is open, the growth of an open population consists of natural growth and net migration.

The mechanism of Cohort Component Method is as follows, each cohort of the population in this period is computed by the survival population of the last period plus net migrants. So at $t$ period, the population of age group $i$ evolves from the age group

$i-1$ at $t-1$ period, as stated in the following Equation (2.1) and (2.2), $i$ denotes the age group, $i = 0,1,2,3 \ldots 95$, here the number zero represents the new births, $t$ represents period, the rural-urban migration, the birth and death rates are flow variables measured over an interval of time while the population is a stock variable measured at a point of time.

#### 2.4.1.1 The survival component and migration component

The non-new-birth population ($i > 0$) at $t$ period can be calculated by the survival rate and net migration of population age $i-1$ during period $t-1$ and $t$.

$$Population_i^t = Population_{i-1}^{t-1} * Survival\ Rate_{i-1}^{t-1,t} + Net\ Migrants\ _{i-1}^{t-1,t}\ . \quad (2.1)$$

Rewritten in the concise form,

$$P_i^t = P_{i-1}^{t-1} * S_{i-1}^{t-1,t} + M_{i-1}^{t-1,t}\ , \quad (2.2)$$

where $P_{i-1}^{t-1}, and\ S_{i-1}^{t-1,t}$ stand for population of age group $i-1$ at period $t-1$, survival rate of age group $i-1$ during an interval from period $t-1$ to period $t$, and $M_{i-1}^{t-1,t}$ denotes net migration of age group $i-1$ during an interval from period $t-1$ to period $t$ respectively. All the demographic rates in this study are calculated as per thousand persons per year.

#### 2.4.1.2 The fertility component

As shown in Equations (2.3) and (2.4), the newborn population denoted as age zero ($i = 0$) is calculated from all the new births from women at childbearing ages. According to China's official population statistics, the age is 15-49.

$$Population_0^t = \sum_{i=15}^{49} Population_i^{t-1} * Women\ ratio_i^{t-1} * Fertility\ Rate_i^{t-1,t}\ . \quad (2.3)$$

Replacing the variables with symbols gives the following equation.

$$P_0^t = \sum_{i=15}^{49} P_i^{t-1} W_i^{t-1} * F_i^{t-1,t}\ , \quad (2.4)$$

where $W_i^{t-1}$ denotes the ratio of women at childbearing age $i$, $F_i^{t-1,t}$ represents the fertility rate of the women in age group $i$ during period $t-1$ and $t$.

#### 2.4.1.3 Putting the survival, migration and fertility components together

A general matrix equation is constructed to estimate the population. If I know the population's initial age and sex ratio, survival rate, birth rate, and net number of migrants by age and sex, I can simulate this process in the given arguments. The age groups are divided into every age span of one year. The interval of period $t$ is set as one year. This setting increases the accuracy of the simulation because it can be used to estimate the people of age $t+1$ based on the people of age $t$.[13] The total population at period $t$ can be calculated from period $t-1$ as the Matrix Equation (2.5).

$$\begin{bmatrix} P_0^t \\ P_1^t \\ P_2^t \\ \vdots \\ P_{95}^t \end{bmatrix} = \begin{bmatrix} W_0^{t-1}F_0^{t-1} & W_1^{t-1}F_1^{t-1} & \dots & W_a^{t-1}F_a^{t-1} \\ S_1^{t-1} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & S_2^{t-1} & \dots & 0 \\ 0 & 0 & \dots & S_a^{t-1} \end{bmatrix} \begin{bmatrix} P_0^{t-1} \\ P_1^{t-1} \\ P_2^{t-1} \\ \vdots \\ P_{95}^{t-1} \end{bmatrix} - \begin{bmatrix} M_0^{t-1,t} \\ M_1^{t-1,t} \\ M_2^{t-1,t} \\ \vdots \\ M_{95}^{t-1,t} \end{bmatrix}. \tag{2.5}$$

The populations in periods $t$ and $t-1$ are normally known, therefore I can utilize the Matrix Equation (2.5) to calculate the migrant populations. Net migrants can be calculated in each age group through the simulation with the Matrix Equation (2.6), where the survival rate $S$ is replaced by $1-D$. $D$ denotes the mortality rate.

$$\begin{bmatrix} M_0^{t-1,t} \\ M_1^{t-1,t} \\ M_2^{t-1,t} \\ \vdots \\ M_{95}^{t-1,t} \end{bmatrix} = \begin{bmatrix} W_0^{t-1}F_0^{t-1} & W_1^{t-1}F_1^{t-1} & \dots & W_a^{t-1}F_a^{t-1} \\ 1-D_1^{t-1} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1-D_2^{t-1} & \dots & 0 \\ 0 & 0 & \dots & 1-D_{95}^{t-1} \end{bmatrix} \begin{bmatrix} P_0^{t-1} \\ P_1^{t-1} \\ P_2^{t-1} \\ \vdots \\ P_{95}^{t-1} \end{bmatrix} - \begin{bmatrix} P_0^t \\ P_1^t \\ P_2^t \\ \vdots \\ P_{95}^t \end{bmatrix}. \tag{2.6}$$

With the help of Matrix Equation (2.6), I can simulate the closed population starting with a scenario without migration. Then net migrants are the difference between the simulated closed population and the actual closed population.

[13] Traditionally, demographers estimate the people of age $t+5$ from the people of age $t$. But this method is not as accurate for my study. E.g. in my estimate, I need to calculate 5 times if I want to estimate the population of age $t+5$, namely, estimate the population of age $t$+1 from $t$, then $t+2$, $t+3$, $t+4$, and finally $t+5$. The traditional method employs a one-time calculation to estimate the population of age $t+5$ from age $t$. In other words, my estimation is like compounds interest yearly while the traditional estimation uses an interest rate for five years.

### 2.4.2 Simulate the closed population without migration

My final objective is to estimate net rural-urban migrants. The accuracy of the estimated migrant population depends on the accuracy of the closed population simulation. Thus I need to test the accuracy of the closed population. A straightforward test is to simulate the total population of China, assuming it is a closed population, and then compare the simulated population of China with the actual one. The differences between them reflect the error of the simulation.

The assumption that the Chinese population is closed is reasonable. According to statistical data from the United Nations population division,[14] China's net international migration rate was only -0.35 per thousand persons in 2000-2005 and -0.28 per thousand persons in 2005-2010. The total net number of international migrants in China was about 4 million in the last decade (2000-2010). As I discussed before, there are two components that shape the population, one is net migration, and the other is natural growth. Chinese statistics show that the total population increased by 90 million people during 2000-2010, which indicated that natural growth accounted for about 95% of the total population growth. Accordingly, Chinese population growth depends mostly on natural growth.

#### 2.4.2.1 Parameters for simulation of the national population

In the following simulation, I employ statistical data from the 2000 and 2010 Chinese population censuses. The data used in the simulation of the population in 2010 includes population by age and sex in 2000, the mean mortality and fertility rates by age and sex in 2000 and 2010, and the sex ratio of new birth in 2000 and 2010. Two assumptions are made due to data limitations. First, I assume the invariable fertility and mortality rates used in the simulation to be the average value of fertility and mortality rates between 2000 and 2010. Second, I assume that the sex ratio of new birth during 2000-10 is the same as it was in 2000.

The fertility rates are shown in Figure 2.5. According to the standard of the National Bureau of Statistics, the female fertility age is 15-49. The peak fertility rate is at age 24 with about 145 births per thousand women in 2000 and 100 births per thousand women in

[14] These data are available on the United Nations population division website. Available on: http://esa.un.org/unpd/wpp/Excel-Data/migration.htm (access on August, 16, 2014)

2010. The fertility rate fell as a whole and the average fertility age increased in 2010. Total fertility rate (TFR) in 2000 was 1.2 children per woman.[15] This TFR is generally recognized as underestimated by authorities and academics. The Chinese statistics department acknowledged that it had been underestimated and raised the TFR to 1.8 after the population census of 2000 (Tyers et al. 2006). However, some scholars don't agree with this amendment. Zhang and Zhao (2006) provided an extensive survey of the literature on fertility rates and concluded that the TFR should be 1.6 rather than 1.8. I will consider these two options of fertility rate, 1.6 and 1.8, in the following simulation.

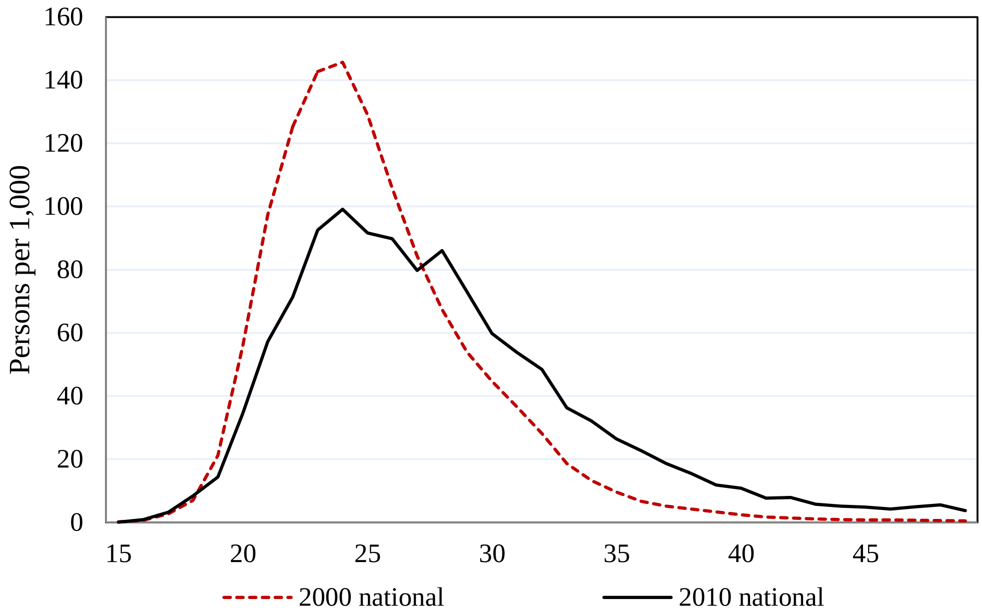



Figure 2.5 National fertility rates by age in China in 2000 and 2010
Source: 2000 and 2010 Chinese population censuses

The mortality rate is illustrated in Figure 2.6. The mortality rate remained almost the same between 2000 and 2010 for the population below age 50 with the exception of the infant mortality rate (IMR).[16] The age is limited between zero and ninety-five to avoid death analysis. If I had not limited an upper age, I should have had to analyze the age of death. It would have complicated the analysis. It is impossible to perform a death analysis since the census data does not include such death information. Furthermore, in my

[15] This abbreviation of TFR is the same as the one adopted by United Nations, Population division on http://esa.un.org/wpp/Excel-Data/fertility.htm (access on January 19, 2015).
[16] This abbreviation of IMR is the same as the one adopted by United Nations, Population division on http://esa.un.org/wpp/Excel-Data/mortality.htm (access on January 19, 2015).

analysis, the age of death is not my concern and does not influence the conclusion of this study.

The total mortality rate decreased between 2000 and 2010. It is noteworthy that the IMR in 2010 was only 3.8 per thousand compared to 26.9 per thousand in 2000. The IMR for 2010 was widely regarded to be underestimated. The Chinese National Bureau of Statistics has since adjusted it to 13.9 in China's statistics yearbook.[17] I will consider this adjustment in the later parameter setting.

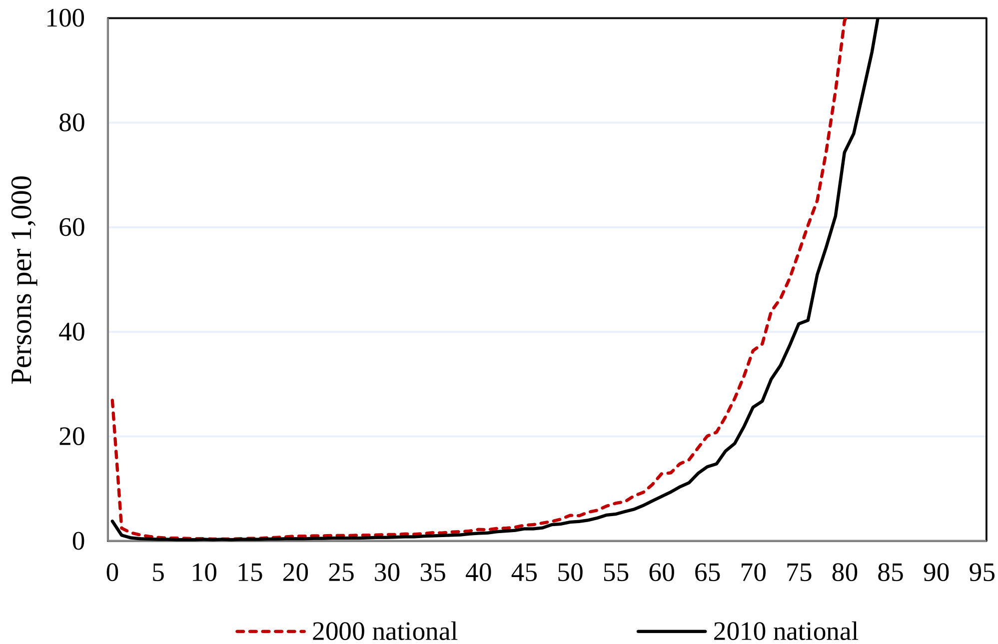



Figure 2.6 Mortality rates by age in China in 2000 and 2010
Source: 2000 and 2010 Chinese censuses

**2.4.2.2 Results of national population simulation**

The national population in 2010 can be simulated from the population in 2000. The gaps between the simulated and real population are defined as simulation errors. First of all, in Figure 2.7 and 2.8, I do not adjust the underestimated fertility and infant mortality rates, the curves of simulation and real population by sex are illustrated in Figure 2.7 and Figure 2.8.

[17] These data can be found in the official dataset of National Bureau of Statistics of China, available on http://data.stats.gov.cn/workspace/index?m=hgnd (access on January 19, 2015).

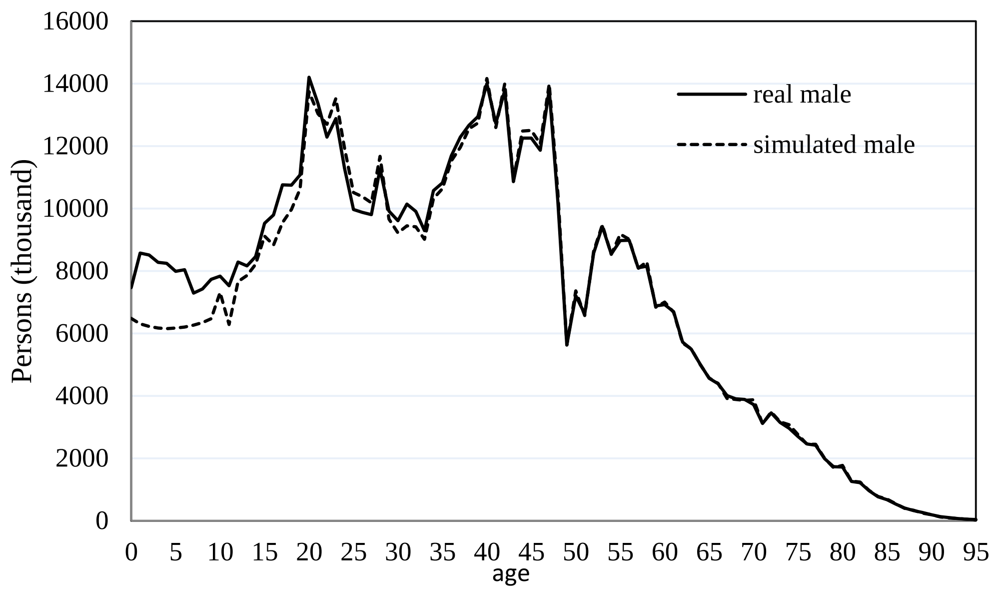



Figure 2.7 Simulated and real male population in China in 2010

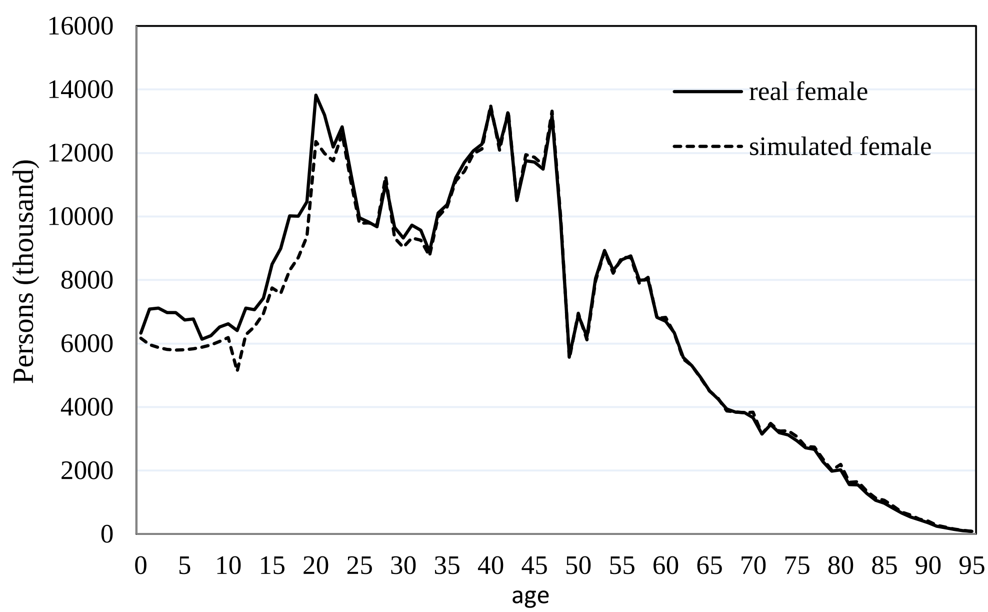



Figure 2.8 Simulated and real female population in China in 2010

The comparisons of simulated and real populations in China show that the model can simulate population very well except for the age group 0-10.[18] To measure the simulated error, I calculated the ratio of the difference between real and simulated to the real

[18] In this chapter, I tried different adjusted rates to estimate population of ages 0-95. But the population under age ten is not the focus in this chapter. Instead, the main purpose is to estimate the migrated labor forces. As shown in next two chapters, this study focuses on human capital and agricultural productivity in which the population above the age of 10.

population in each age group in Table 2.1. In general, the total simulated error rate is about -3%, which means the simulated population is 3% underestimated (the final row in Table 2.1). The age group with the highest simulated error rate is age 0-9 with -21% for males and -12% for females (the first row in Table 2.1). Generally, except for the three age groups 0-9, 10-19 and 90-95, other age groups have relatively low error rates with about 2% error.

Table 2.1 Measurement of simulated error by age group

| | Male (thousand) | | | Female(thousand) | | |
|---|---|---|---|---|---|---|
| Age group | real | simulated | error rate | real | simulated | error rate |
| 0-9 | 79,527 | 62,799 | -21% | 66,887 | 59,144 | -12% |
| 10-19 | 92,172 | 85,396 | -7% | 82,625 | 72,793 | -12% |
| 20-29 | 114,846 | 117,334 | 2% | 113,581 | 109,744 | -3% |
| 30-39 | 109,913 | 106,809 | -3% | 105,251 | 103,304 | -2% |
| 40-49 | 117,385 | 118,838 | 1% | 112,963 | 113,649 | 1% |
| 50-59 | 81,446 | 81,963 | 1% | 78,619 | 78,337 | 0% |
| 60-69 | 50,583 | 50,501 | 0% | 49,198 | 49,181 | 0% |
| 70-79 | 27,682 | 28,036 | 1% | 29,142 | 29,902 | 3% |
| 80-89 | 8,117 | 8,222 | 1% | 10,888 | 11,653 | 7% |
| 90-95 | 567 | 549 | -3% | 1,129 | 1,259 | 12% |
| Total | 682,239 | 660,448 | -3% | 650,284 | 628,965 | -3% |

**2.4.2.3 Parameters adjustment**

It is possible that the underestimated population in the age group 0-9 is due to the underestimated fertility rate in 2000. To adjust the simulated error of underestimated fertility rate in the age group 0-9, I consider two optional scenarios of fertility rates as suggested in section 2.4.2.1. The first is to multiply the original fertility rate by 1.3 which would change the total fertility rate (TFR) from 1.22 to 1.56, close to the 1.6 rate suggested by scholars Zhang and Zhao (2006). The second is to multiply the original fertility rate by 1.5, which would change the TFR to 1.8, the value adjusted by the Chinese statistics department. The simulation results can be found in Figures A2.1-A2.4 (TFR=1.56) and Tables A2.1-A2.2 (TFR=1.8) in the Appendix A2. The simulated error rates in age group 0-9 change from -21% to 8% for the male population and from -12% to 9% for the female population with the adjusted TFR of 1.56. The total simulated error falls to zero for the male population and -1% for the female. By contrast, the simulated error rates in age group 0-9 are still as high as 25% for the male and female populations with the adjusted TFR of 1.8. Obviously, among the three options, the adjusted fertility rate suggested by Zhang and Zhao (2006) had the better simulation in the age group 0-9. Accordingly, the total fertility rate should be adjusted to the level of 1.56 based on the Zhang and Zhao's study.

To adjust the underestimated infant mortality rate (IMR), I adjusted it to 13.8 based on the data published by the Chinese National Bureau of Statistics and used the average birth sex ratios between 2000 and 2010 instead of the birth sex ratio in 2000. The simulated results are shown in Figures A2.5-A2.6 and Tables A2.3. The results are similar to those shown in Figures A2.1-A2.2 whereby the infant mortality rate and birth sex ratio are not adjusted. The total simulated errors remain the same as prior to the adjustment of IMR. The simulated error of the age group 0-9 falls to 8%. These results indicate that the underestimated fertility rate is the biggest source of the estimated error in this simulation. The adjustments of IMR and birth sex ratio are not necessary because the simulated errors do not decrease.

However, three sources of error cannot be controlled in the simulation due to data limitations. First, the assumption that the average fertility and mortality rates are invariable could be problematic in the long run. But in a decade the changes to the fertility rate and the death rate are assumed to be linear. Even if the rates had changed throughout the decade, the average values of the rates are still valid as long as the change is linear. Second, although international migration has little effect on changes to the population, I cannot eliminate the error from it because the demographic data on international migration is not available. The least controlled factor in this study are the statistical errors in the censuses, for instance, the population in 2000 is regarded to be underestimated in the census (Liang and Ma 2004). If the population in 2000 is underestimated, the simulated population in 2010 is also underestimated. However, I can do little to deal with this part of the statistic error.

### 2.4.3 Simulate the closed rural population without migration

As the parameter settings discussed above, the fertility rate, infant mortality rate and birth sex ratio need to be adjusted. This process of parameter settings of national data can be applied in the rural population data as long as I assume that the rural population data is consistent with the national population data. The following section 2.4.3.1 will illustrate that the changes to rural population data are quite consistent with the national data.

As shown in Figures 2.9 and 2.10, the trends related to changes in the fertility rate and the mortality rate are similar for both the rural and national populations, but the parameters for the rural population are bigger than that of the national population.

The parameters for the rural population need to be adjusted as do the parameters for the national population because they show similar trends in the fertility and mortality rates. As I discussed in section 2.4.2.2, the fertility rates have been inflated by 1.3 in order to adjust the total fertility rate (TFR) to 1.56.

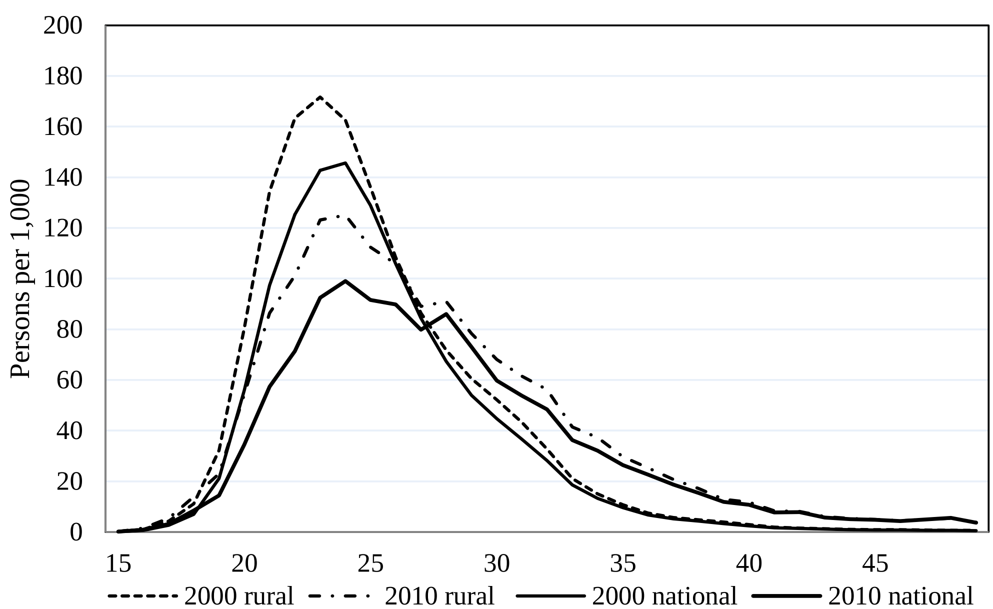



Figure 2.9 Fertility rates by age in China in 2000 and 2010
Source: 2000 and 2010 Chinese population censuses

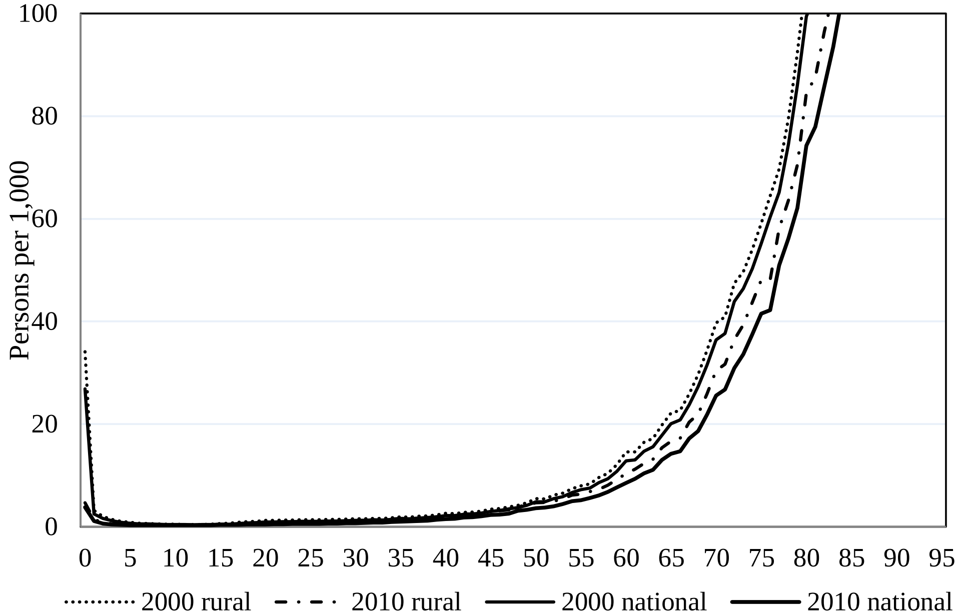



Figure 2.10 Mortality rates by age in China in 2000 and 2010
Source: 2000 and 2010 Chinese population censuses

Unlike the national population, the rural population cannot be assumed as a closed population because net rural-urban migration is a dominant factor shaping the rural population. I can simulate the scenario of the rural population without rural-urban migration by employing the data of rural population by sex and age in 2000, and the adjusted rural fertility and mortality rates between 2000 and 2010. The gaps between simulated rural population and real rural population theoretically equal the number of net rural-urban migrants.

Figures 2.11 and 2.12 compare the real and simulated rural populations in 2010. As shown in the gaps between the two curves, which represent net rural-urban migrants, most rural-urban migrants are around age 15-45. There are more male migrants than female migrants.

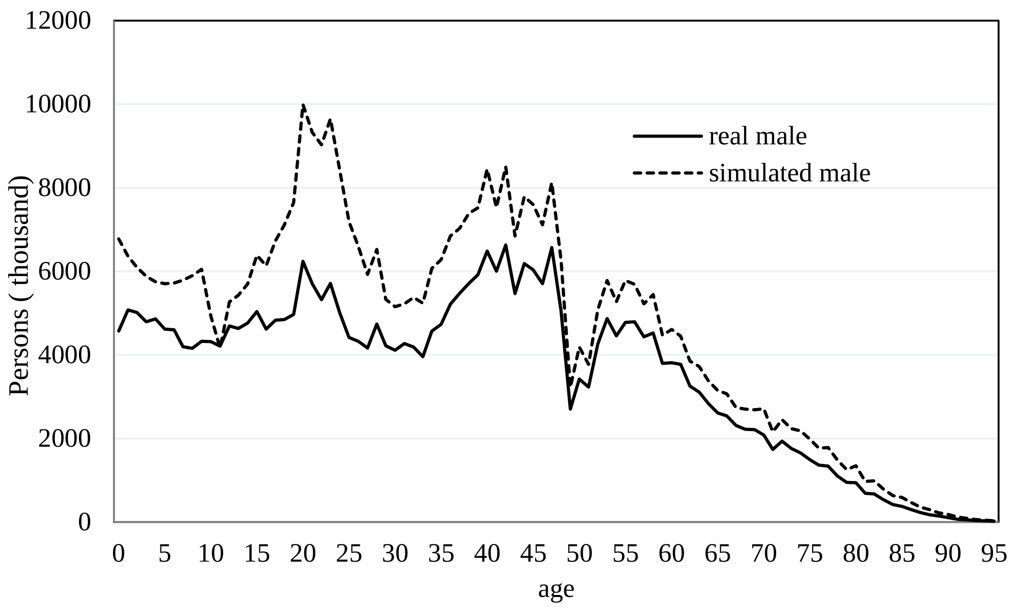



Figure 2.11 Simulated and real male population of China's rural areas in 2010

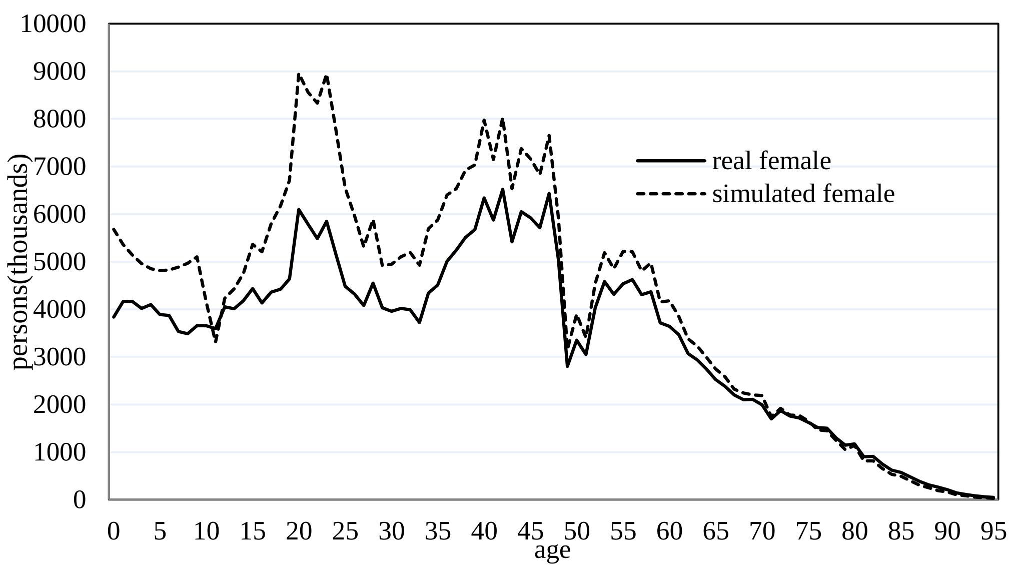



Figure 2.12 Simulated and real female population of China's rural areas in 2010

Table 2.2 illustrates the migrants of each age group. In total there were 177 million rural-urban migrants between 2000 and 2010. Most migrants are between the ages of 20 and 29 and younger than 50.

Table 2.2 Simulated rural-urban migrants by age group between 2000 and 2010

| | Male (thousand) | | | Female(thousand) | | |
|---|---|---|---|---|---|---|
| Age group | real | simulated | migrants | real | simulated | migrants |
| 0-9 | 46,197 | 60,005 | 13,809 | 38,715 | 50,590 | 11,875 |
| 10-19 | 46,904 | 59,472 | 12,569 | 41,497 | 50,153 | 8,657 |
| 20-29 | 49,830 | 77,976 | 28,146 | 49,858 | 71,214 | 21,356 |
| 30-39 | 48,132 | 62,104 | 13,973 | 45,996 | 58,639 | 12,643 |
| 40-49 | 56,816 | 71,405 | 14,589 | 56,116 | 67,797 | 11,681 |
| 50-59 | 42,553 | 50,680 | 8,128 | 40,889 | 46,252 | 5,363 |
| 60-69 | 28,649 | 34,340 | 5,690 | 27,184 | 29,730 | 2,546 |
| 70-79 | 15,392 | 19,980 | 4,588 | 16,124 | 16,271 | 147 |
| 80-89 | 4,470 | 6,638 | 2,169 | 6,380 | 5,588 | -792 |
| 90-95 | 303 | 515 | 212 | 651 | 471 | -180 |
| Total | 339,245 | 443,116 | 103,871 | 323,410 | 396,706 | 73,296 |

Source: author simulations based on 2000 and 2010 China Censuses data

The purpose of this study is to estimate the impact of migration on the labor force, so I will focus on the age groups between 15 and 69, both of which have very low estimated errors (around 2%). Figure 2.13 reports rural-urban migrants calculated from the difference between the real and simulated populations between the ages of 15 to 69. I only consider the population aged between 15 and 69 because the simulated error rates for ages 0-9 and 10-19 are relatively high and I am more interested in the labor force. Figure 2.13 shows the distribution of rural-urban migrants. Most migrants are in the age group 20-24 with 32 million migrants, followed by the age groups 15-19 and 25-29 with around 16 million migrants, and then the age groups 35-39 and 40-44 with about 15 million migrants. The pyramid shows that there is selectivity of gender for the migrants. There are 30 million more male migrants than female migrants, i.e., the male population is more likely to migrate. This conclusion coincides with that of Tan and Ma's analysis (2007) based on Chinese Census data from 2000.

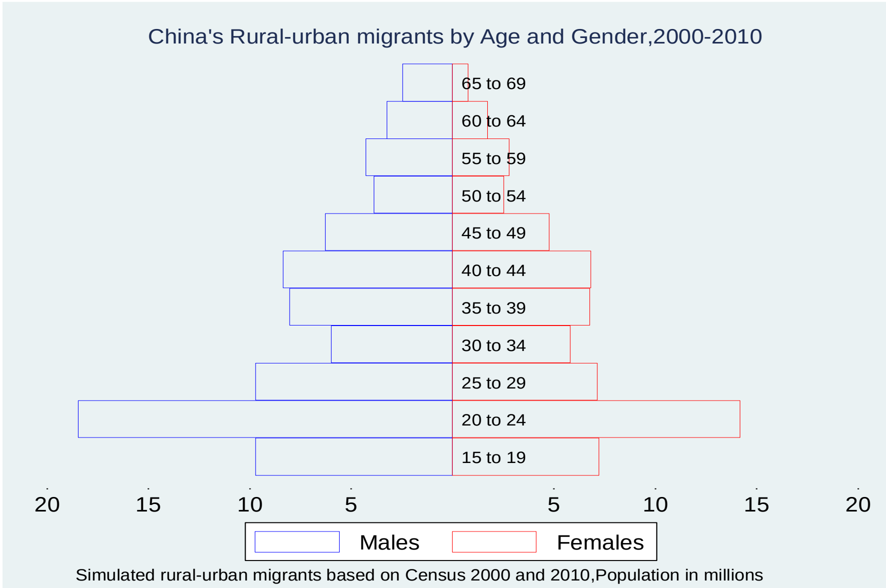



Figure 2.13 Simulated rural-urban migrants by age and sex (millions)

## 2.5 Demographic selectivity of rural-urban migration based on household data

In Section 2.4, the aggregate data from the national statistics and censuses illustrated the demographic changes of the rural population due to rural-urban migration. The present section employs household data in China to present more details about the demographic characteristics of rural-urban migration. The objective of this section is to present evidence from household data to confirm demographic characteristics of rural-urban migrants found in the last simulated model from the national censuses data.

### 2.5.1 Data Description

I restricted the data of China Health and Nutrition Survey (CHNS) to the last five rounds of surveys, namely 1997, 2000, 2004, 2006 and 2009. The data covers nine provinces. Three eastern provinces are more developed, and two western provinces belong to the least developed regions in China. The remaining four provinces are central provinces with intermediate development. Since the empirical test of variance shows that region is a significant influence of the migrant rate, migrants come from the central or western regions are more than migrants from the eastern region. Therefore, it is necessary to analyze the characteristics of rural-urban migrants based on different geographical regions. In Table 2.3 I categorized the 9 provinces into 3 regions, namely, Eastern, Western and Central Regions based on China's official classification standards (Peng 2011).

Table 2.3 Region classification of survey provinces

| Region | Province | Observations |
|---|---|---|
| Eastern | Liaoning Jiangsu Shandong | 14,144 |
| Central | Heilongji Henan Hubei Hunan | 21,399 |
| Western | Guangxi Guizhou | 13,272 |
| Total | | 48,815 |

Source: selected sample from CHNS

Table A2.4 in the Appendix A2 illustrates summary statistics for individuals, households, and communities in three regions. In Table A2.4 the western rural region had a higher migration rate than the eastern and central regions, namely 23% in the western region, 18% in the central region, and 17% in the eastern region respectively. The average age of the eastern population was 39 years old compared to 36 years old in the western and eastern regions. The population in the eastern region had a higher educational attainment (6.99 years) than the western (6.31 years) and central regions (6.79 years) because the eastern region is more developed compared to the central and western regions.

The household variables reflect significant differences between the three regions. Household size in the western region is bigger than in other regions with 5.55 members compared to around 4.50 members in other regions, respectively. Households in the western region have more children (2.18) than in the eastern (1.54) and central regions (1.77). The educational attainment of the household head was 7.42 years in the western areas compared to 8.3 years in the eastern and central regions. Households in the central region invested more in agricultural machinery with 1,204 yuan while only 946 yuan in the eastern region and 291 yuan in the western region. Annual household income in the eastern region was the highest at 27,125 yuan, about 7,000 yuan more than in other regions.

The community variables indicate that the agricultural labor force accounts for 49% of the total labor force of the communities in the central region, higher than that in western and eastern regions. Nevertheless, agriculture is still the primary income source in the rural areas in all three regions. The communities in the western areas live in more remote villages rather than close to a township. Only half of the villages are near bus stations in the central region while more than 65% of the villages in the eastern and western regions are close to bus stations. Communities with railway stations can be found more in the western region (19%) compared to the central region (15%) and eastern region (13%).

### 2.5.2 Model specification

As stated in the new economics of labor migration (NELM) (Stark and Bloom 1985; Stark 1991; Banister and Taylor 1989; Taylor 1999; Taylor and Martin 2001), the decision to migrate depends on the context of the individual, the household and the community. At the individual level, the most important factors affecting rural-urban migration are demographic variables, namely age and education known as the fundamental variables to human capital. At the household level, household size can influence how a household allocates its limited labor force. The education level of the household head and his/her partner, the income level of the household, and the fixed agricultural assets could influence a household's decision to migrate (Stark and Wang 2002). In the estimate, I controlled some of the community variables which include the transportation options, i.e., whether the village has a bus station, or has a railway station, as bus and train are the main modes of transport for migrants, the market conditions, i.e., whether the community is located close to a township, and economic conditions, i.e., the share of agriculture labor force in the village.

The specification for the empirical estimation is based on the variables described above. A reduced form of migration decision for individual $i$ from household $h$ in year $t$ is expressed as Equation (2.7). Considering the binary decision to migrate, I employed the Logit model in the migration decision model as follows (Greene 2012, 760-770). I compare the Probit and Logit model with the same specification of Equation (2.7) because the Logit and Probit models are nested. Alaike's information criterion (AIC) and Schwarz's Bayesian information criterion (BIC) are two standard measures to compare the Probit and Logit models. AIC and BIC in Logit models are smaller than in Probit models. The log likelihood of the model has a higher value in Logit models than in Probit models. These standards suggest that Logit models are better in this study than Probit models.

$$M_{ihjt} = \boldsymbol{D}'_{it}\beta_1 + \beta_2 I^r_{ht} + \beta_3 I^u_{ht} + \boldsymbol{V}'_{jt}\beta_4 + \boldsymbol{X}'_{ht}\beta_5 + v_t + e_{ihjt}\,, \qquad (2.7)$$

where $M_{ihjt}$ is a binary variable to indicate whether individual $i$ from household $h$ in the village $j$ in year $t$ is a migrant. If the individual migrates, then $M_{ihjt}$ equals 1, otherwise it equals 0.

$\boldsymbol{D}'_{it}$ is a vector of demographic characteristics which include sex, age group, educational attainment. Age is categorized into the six groups, namely, 0-5, 5-14, 15-25, 26-35, 36-45, 46-60, and above 60. Educational attainment is categorized as illiterate, primary school, middle school, high school, and higher education.

$I^r_{ht}$ is a variable to indicate the rural household income at $t$ period. This variable of income at home reflects the opportunity cost of migration. $I^u_{ht}$ is the median urban household income in the same province at $t$ period. This can reflect the opportunity costs of not migrating. The survey does not supply information on the income at destination, but it is reasonable to assume that median urban household income in the province as the expected income. For the rural households, the information on income in the same province should be easy to access. I chose median value instead of mean value because the urban household income is highly unequal. The median value can represent the income of a larger portion of the population than mean value.

$\boldsymbol{V}'_{jt}$ is the vector of village characteristics that influence the migration in the village $j$ in year $t$, which include whether the village has a bus station, whether it is located close to a township, and the share of the agriculture labor force in the village.

$\boldsymbol{X}'_{ht}$ is a vector for household characteristics, including educational attainment of the head and his/her partner of household, number of children, household size, and the value of agricultural machinery. $v_t$ is a dummy variable for year and $e_{ihjt}$ is the error term.

Some variables may have the problem of multicollinearity in this specification. Notice the educational attainment of the head of household and of individual are both involved in the regression because they represent different levels of control variables: the household level and the individual level. They have a correlation coefficient of 0.38 which does not result in serious multicollinearity. Similarly, number of children and household size may have high multicollinearity. They have a correlation coefficient of 0.56. It is a little high. However, it is still at an acceptable range. Considering the possibility of multicollinearity between bus station and railway station, I calculated the covariance coefficient of them and found that these two variables are weakly correlated with 0.18.

It is necessary to use robust standard errors because the error term is heterosecedastic as the samples from different regions greatly differ in economic, geographic, and demographic conditions (Bertrand et al. 2004).

**2.5.3 Coefficients and marginal effects of Logit models**

Marginal effects are more informative than coefficients because the Logit model is a nonlinear modal (Cameron 2009, 334-35). The marginal effect at mean of the estimated Equation 2.7 is reported in Table 2.4. All of the Logit model coefficients can be found in Table A2.5.

As shown in Table 2.4. Gender is not a typical characteristic for the migrants in eastern and central areas. However, for the western population, females are 2.3% more likely to migrate than males. Overall, females are 0.6% more likely to migrate. Although the magnitude of positive effect is very small, this result is different from the simulation analysis in Section 2.4.3 where the simulated migrants from census data indicate that females are less likely to migrate. However, because this data set is not national data, migrant population does not show the selectivity of male as the census data does. It does not conclude that the simulation in Section 4 is not correct.

The selectivity of age in migration has been confirmed in the household data. The young of ages 16-25, 26-35, and 36-45 have high possibilities to migrate compared with other age groups. Specifically, the young of ages 16-25, and 26-35 have about 25% higher likelihood to migrate than the other age groups.

The selectivity of age in migration shows differences in three regions. The migrants from the western and central regions migrate younger than their counterparts from the eastern region. In the eastern region, the age most likely to migrate is between 26 and 35. In the central and western regions, the young of ages 16-25 are most likely to migrate. In all regions, the young of ages 26-35 and ages 16-25 have the same likelihood to migrate with 29.7%.

The elderly above age 60 are more likely to stay in the rural areas. In summary, the selectivity of migrants demonstrates that the young of age 16-45 are more likely to migrate while the elderly above age 60 are more likely to stay in the rural areas.

Table 2.4 Marginal effects of demographic factors at means

| | Total | Eastern | Central | Western |
|---|---|---|---|---|
| Dependent | migrant | migrant | migrant | migrant |
| Female | 0.006** | -0.002 | 0.004 | 0.023*** |
| | (0.002) | (0.004) | (0.003) | (0.052) |
| 6-15 | -0.014 | -0.015 | -0.473 | -0.083* |
| | (0.009) | (0.060) | (0.035) | (0.464) |
| 16-25 | 0.297*** | 0.231*** | 0.212*** | 0.260*** |
| | (0.011) | (0.057) | (0.032) | (0.419) |
| 26-35 | 0.297*** | 0.252*** | 0.204*** | 0.254*** |
| | (0.010) | (0.057) | (0.032) | (0.041) |
| 36-45 | 0.112*** | 0.164*** | 0.103*** | 0.163*** |
| | (0.009) | (0.057) | (0.033) | (0.417) |
| 46-60 | 0.005 | 0.059 | -0.0032 | -0.029 |
| | (0.008) | (0.057) | (0.032) | (0.426) |
| above 60 | -0.017* | -0.016 | -0.071* | -0.103* |
| | (0.008) | (0.058) | (0.033) | (0.043) |
| Wald Chi-square | 7,093 | 1,946 | 3,184 | 2,076 |
| Pseudo R-square | 0.28 | 0.26 | 0.29 | 0.31 |
| N | 48,815 | 14,144 | 21,399 | 13,272 |

Note: Standard errors in parentheses, * p<0.05 ** p<0.01*** p<0.001
Source: selected samples from CHNS

As the Logit models have shown, the econometric analysis illustrated that rural-urban migrants had significant demographic characteristics based on household data from 9 provinces in China. Generally, the young are more likely to migrate. Specifically, the most likely cohort of migration is the age group 26-35 in the eastern region and age 16-25 in the central and western regions, which is consistent with the features of the simulated rural-urban migrants in Section 2.4.

Besides selectivity of demographic characteristics migrants have other significant features as shown in Table A2.5. Education has significant positive effects on migration. Independent of the region, higher education individuals have a higher probability to migrate. Furthermore, larger households are more likely to migrate. The households with less income are more likely to migrate. The expected income in urban areas can encourage rural households to migrate. Whether there is a bus station or a railway station is negatively related to migration because bus station and railway station reflect not only the good transportation conditions but also good labor markets for this village. It is reasonable to assume that a bus station or a railway station can promote the local labor market. This could stop the local residents from migrating to find a job. I will do a sensitive test to prove this assumption for explaining the negative relationship between train station and migration in Chapter 3. The households investing more in agricultural machinery are less likely to migrate. The relationship between migration and agricultural investment will be further explained in Chapter 4. Finally, the people living in villages are more likely to migrate than those living near a township. It is not difficult to imagine that people living in villages have fewer opportunities to find jobs near their village while people living close to townships have more opportunities. People from villages have to migrate due to lack of local job opportunities.

## 2.6 Conclusions

In this chapter, using the latest population censuses from 2000 and 2010, I first estimated rural-urban migrants between 2000 and 2010 through the Cohort Component Method and separated the impact of migration from natural growth on size and age structure of the rural population. The simulation results suggested the dominant effect of rural-urban migration in changing the rural population structure. In order to confirm the demographic effect of rural-urban migration, I employed the household data of CHNS to test the demographic features of rural-urban migrants. The evidence from household data confirmed that the young are more likely to migrate from the rural areas while the elderly are more likely to stay in the rural areas. The most important demographic consequence of rural-urban migration is that the age structure of rural population has been experiencing a process of accelerated aging for the last decade.

Some implications arise to dealing with the changes to the structure of the rural population. First, as the rural labor force continuously decreases, the change of age structure in rural areas accelerates population ageing. It will be necessary to improve the quality of the rural labor force left behind through investment in human capital. Although the young, who have a higher level of education in rural areas are more likely to migrate to urban areas, this should not lead to underestimate the function of public investment in education for the rural population. Secondly, the government should pay more attention to agricultural production in the rural areas given the ageing problem of the labor force. It is an open issue if the aging labor force is less productive in agriculture, or actually more productive as a consequence of self-selection and resource endowment changes. The nation-wide economic net benefits of migration are a more complex issue which has not been addressed here. However, preventing the rural young people from migrating to urban areas is generally not a good policy.

## Appendix A2

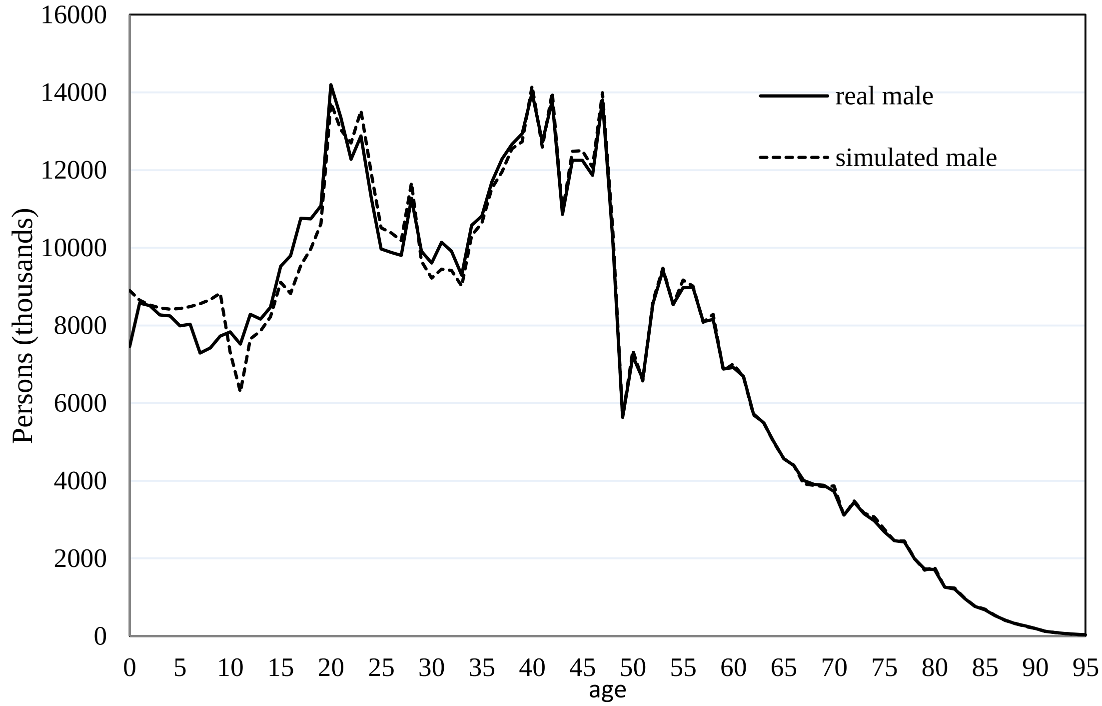



Figure A2.1 Simulated and real male population of China in 2010 (TFR=1.56)

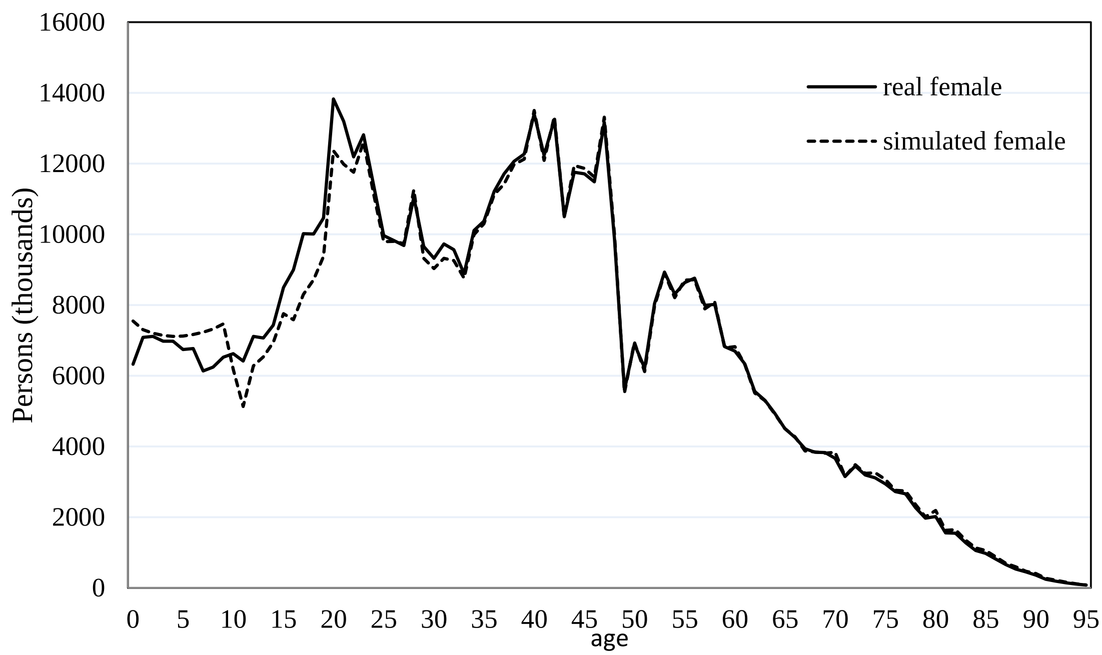



Figure A2.2 Simulated and real female population of China in 2010 (TFR=1.56)

Table A2.1 Measurement of simulated error by age group (TFR=1.56)

| | Male (thousand) | | | Female(thousand) | | |
|---|---|---|---|---|---|---|
| Age group | real | simulated | Error rate | real | simulated | Error rate |
| 0-9 | 79,527 | 85,913 | 8% | 66,887 | 72,631 | 9% |
| 10-19 | 92,172 | 85,396 | -7% | 82,625 | 72,793 | -12% |
| 20-29 | 114,846 | 117,334 | 2% | 113,581 | 109,744 | -3% |
| 30-39 | 109,913 | 106,809 | -3% | 105,251 | 103,304 | -2% |
| 40-49 | 117,385 | 118,838 | 1% | 112,963 | 113,649 | 1% |
| 50-59 | 81,446 | 81,963 | 1% | 78,619 | 78,337 | 0% |
| 60-69 | 50,583 | 50,501 | 0% | 49,198 | 49,181 | 0% |
| 70-79 | 27,682 | 28,036 | 1% | 29,142 | 29,902 | 3% |
| 80-89 | 8,117 | 8,222 | 1% | 10,888 | 11,653 | 7% |
| 90-95 | 567 | 549 | -3% | 1,129 | 1,259 | 12% |
| Total | 682,239 | 683,561 | 0% | 650,284 | 642,453 | -1% |

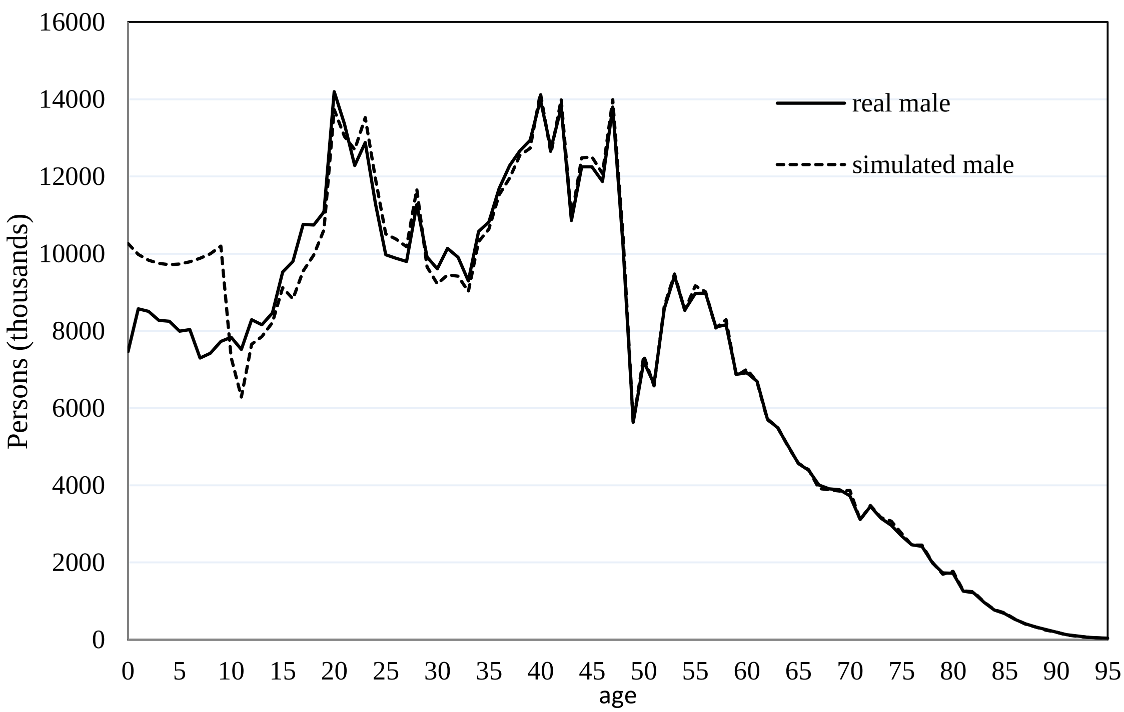



Figure A2.3 Simulated and real male population of China in 2010 (TFR=1.80)

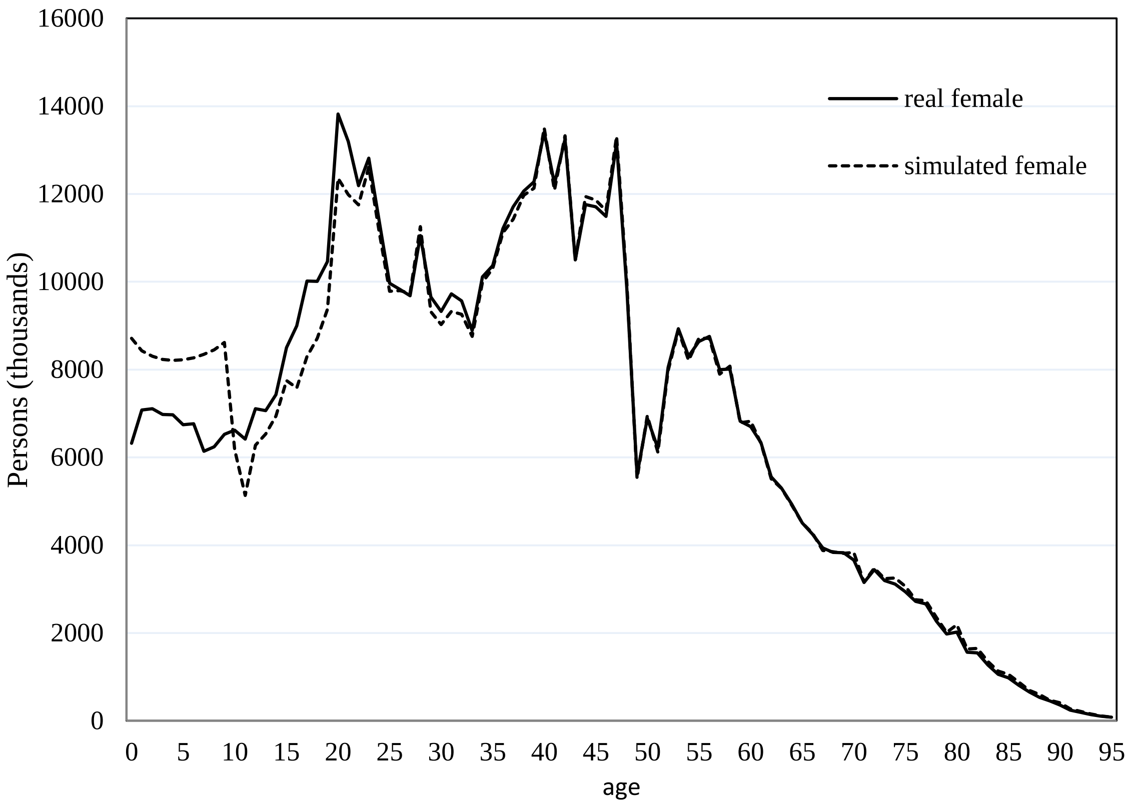



Figure A2.4 Simulated and real female population of China in 2010 (TFR=1.80)

Table A2.2 Measurement of simulated error by age group (TFR=1.80)

| | Male (thousand) | | | Female(thousand) | | |
|---|---|---|---|---|---|---|
| Age group | real | simulated | Error rate | real | simulated | Error rate |
| 0-9 | 79,527 | 99,130 | 25% | 66,887 | 83,805 | 25% |
| 10-19 | 92,172 | 85,396 | -7% | 82,625 | 72,793 | -12% |
| 20-29 | 114,846 | 117,334 | 2% | 113,581 | 109,744 | -3% |
| 30-39 | 109,913 | 106,809 | -3% | 105,251 | 103,304 | -2% |
| 40-49 | 117,385 | 118,838 | 1% | 112,963 | 113,649 | 1% |
| 50-59 | 81,446 | 81,963 | 1% | 78,619 | 78,337 | 0% |
| 60-69 | 50,583 | 50,501 | 0% | 49,198 | 49,181 | 0% |
| 70-79 | 27,682 | 28,036 | 1% | 29,142 | 29,902 | 3% |
| 80-89 | 8,117 | 8,222 | 1% | 10,888 | 11,653 | 7% |
| 90-95 | 567 | 549 | -3% | 1,129 | 1,259 | 12% |
| Total | 682,239 | 696,778 | 2% | 650,284 | 653,626 | 1% |

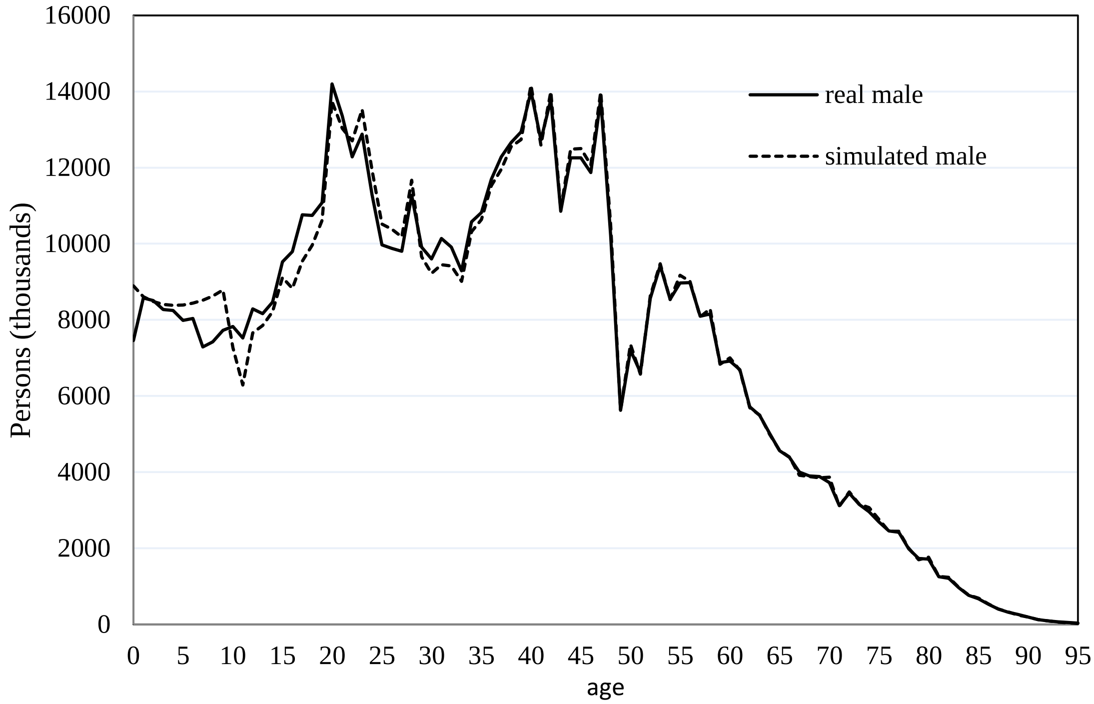



Figure A2.5 Simulated and real male population of China in 2010 (TFR=1.56, IMR=13.8)

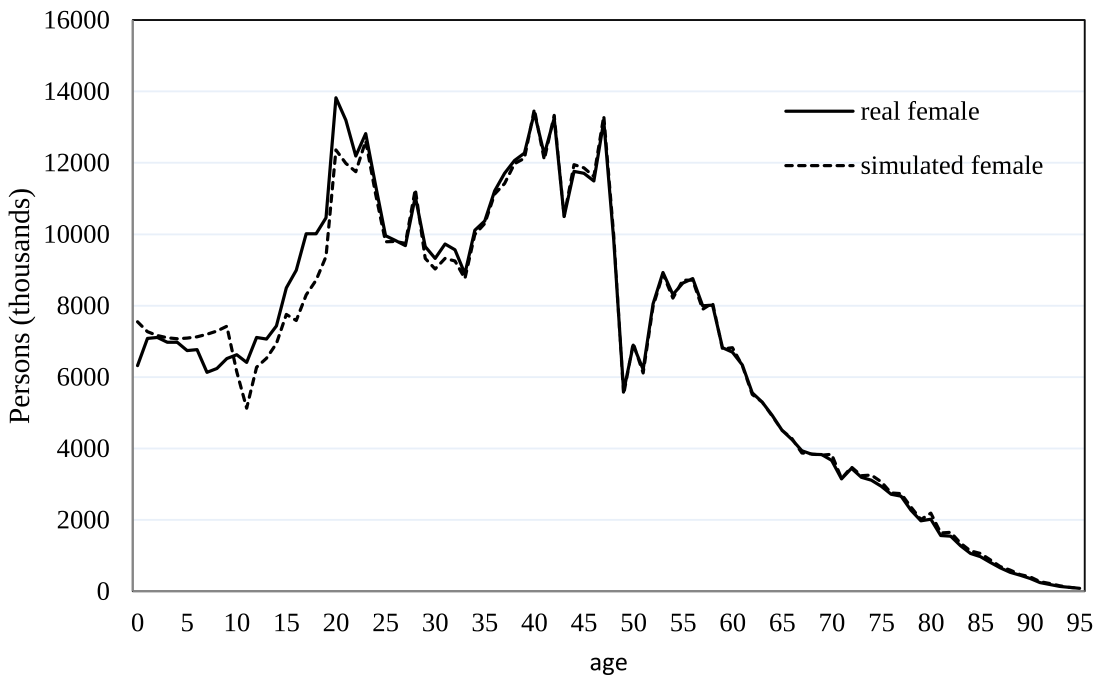



Figure A2.6 Simulated and real female population of China in 2010 (TFR=1.56, IMR=13.8)

Table A2.3 Measurement of simulated error by age group (TFR=1.56, IMR=13.8)

| | Male (thousand) | | | Female(thousand) | | |
|---|---|---|---|---|---|---|
| Age group | real | simulated | Error rate | real | simulated | Error rate |
| 0-9 | 79,527 | 85,528 | 8% | 66,887 | 72,287 | 8% |
| 10-19 | 92,172 | 85,360 | -7% | 82,625 | 72,760 | -12% |
| 20-29 | 114,846 | 117,334 | 2% | 113,581 | 109,744 | -3% |
| 30-39 | 109,913 | 106,809 | -3% | 105,251 | 103,304 | -2% |
| 40-49 | 117,385 | 118,838 | 1% | 112,963 | 113,649 | 1% |
| 50-59 | 81,446 | 81,963 | 1% | 78,619 | 78,337 | 0% |
| 60-69 | 50,583 | 50,501 | 0% | 49,198 | 49,181 | 0% |
| 70-79 | 27,682 | 28,036 | 1% | 29,142 | 29,902 | 3% |
| 80-89 | 8,117 | 8,222 | 1% | 10,888 | 11,653 | 7% |
| 90-95 | 567 | 549 | -3% | 1,129 | 1,259 | 12% |
| Total | 682,239 | 683,140 | 0% | 650,284 | 642,076 | -1% |

Table A2.4 Description of rural samples from different regions

| | Total | | Eastern | | Central | | Western | |
|---|---|---|---|---|---|---|---|---|
| Variable | Mean | Std. Dev | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. |
| Individual Variable | | | | | | | | |
| Migrant | 0.19 | 0.40 | 0.17 | 0.38 | 0.18 | 0.39 | 0.23 | 0.42 |
| Age | 37.13 | 19.02 | 39.01 | 18.37 | 36.24 | 18.80 | 36.57 | 19.89 |
| Education (Year) | 6.72 | 3.86 | 6.99 | 3.97 | 6.79 | 3.78 | 6.31 | 3.85 |
| Female | 0.50 | 0.50 | 0.51 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| Household Variable | | | | | | | | |
| Children Number | 1.81 | 1.09 | 1.54 | 0.96 | 1.77 | 1.02 | 2.18 | 1.24 |
| Head's Education (years) | 8.04 | 3.27 | 8.30 | 3.27 | 8.26 | 3.15 | 7.42 | 3.37 |
| Household Size | 4.85 | 2.13 | 4.41 | 1.83 | 4.69 | 1.97 | 5.55 | 2.47 |
| Machinery Value (Yuan) | 881.52 | 4,473.17 | 946.34 | 4,431.30 | 1,204.36 | 5,563.75 | 291.90 | 1,496.70 |
| Annual household income (Yuan) | 22,155.84 | 30,113.40 | 27,125.38 | 35,169.15 | 19,806.77 | 28,836.20 | 20,647.29 | 25,280.78 |
| Expected annual urban income | 22,802.24 | 9,098.62 | 29,266.09 | 11,576.75 | 21,732.24 | 6,034.19 | 17,638.91 | 5,470.69 |
| Community Variable | | | | | | | | |
| Bus Station | 0.59 | 0.49 | 0.65 | 0.48 | 0.50 | 0.50 | 0.67 | 0.47 |
| Railway station | 0.15 | 0.36 | 0.13 | 0.34 | 0.15 | 0.35 | 0.19 | 0.39 |
| Village | 0.79 | 0.41 | 0.79 | 0.41 | 0.78 | 0.42 | 0.80 | 0.40 |
| Percentage of agricultural labor force | 46.38 | 30.28 | 40.87 | 30.36 | 49.77 | 31.08 | 46.77 | 27.96 |
| Observation | 48,815 | | 14,144 | | 21,399 | | 13,272 | |

Table A2.5 Logit models of rural-urban migration from 1997 till 2009

| | Total | Eastern | Central | Western |
|---|---|---|---|---|
| Dependent | migrant | migrant | migrant | migrant |
| Female | 0.0719** | -0.0275 | 0.0475 | 0.226*** |
| | (0.0273) | (0.0525) | (0.0422) | (0.0504) |
| Age group: Base=0-5 | | | | |
| 6-15 | -0.538 | -0.189 | -0.568 | -0.796 |
| | (0.279) | (0.760) | (0.425) | (0.442) |
| 16-25 | 2.618*** | 2.917*** | 2.556*** | 2.476*** |
| | (0.258) | (0.722) | (0.395) | (0.403) |
| 26-35 | 2.617*** | 3.173*** | 2.449*** | 2.424*** |
| | (0.258) | (0.721) | (0.395) | (0.402) |
| 36-45 | 1.562*** | 2.070** | 1.240** | 1.559*** |
| | (0.258) | (0.720) | (0.395) | (0.401) |
| 46-60 | 0.140 | 0.754 | -0.0381 | -0.279 |
| | (0.258) | (0.721) | (0.392) | (0.406) |
| above 60 | -0.739** | -0.203 | -0.855* | -0.988* |
| | (0.267) | (0.739) | (0.406) | (0.413) |
| Education: Base=No schooling | | | | |
| Primary School | 0.631*** | 0.608*** | 0.541*** | 0.843*** |
| | (0.0872) | (0.146) | (0.149) | (0.166) |
| Middle School | 0.990*** | 0.953*** | 0.844*** | 1.276*** |
| | (0.0881) | (0.147) | (0.153) | (0.167) |
| High School | 1.036*** | 0.943*** | 1.016*** | 1.284*** |
| | (0.0941) | (0.163) | (0.159) | (0.180) |
| Higher Education | 1.321*** | 1.343*** | 1.321*** | 1.395*** |
| | (0.121) | (0.200) | (0.208) | (0.237) |
| Head‘s Education | -0.09*** | -0.07*** | -0.09*** | -0.091*** |
| | (0.0048) | (0.0094) | (0.0077) | (0.0085) |
| Children Number | 0.112*** | 0.101** | 0.127*** | 0.107*** |
| | (0.0162) | (0.0349) | (0.0269) | (0.0255) |
| Household Size | 0.131*** | 0.158*** | 0.136*** | 0.118*** |
| | (0.0088) | (0.0177) | (0.0142) | (0.0141) |
| Machinery Value | -0.04*** | -0.00967 | -0.06*** | -0.0226* |

Table A2.5(continued) Logit models of rural-urban migration from 1997 till 2009

| | Total | Eastern | Central | Western |
|---|---|---|---|---|
| | (0.0044) | (0.0079) | (0.0065) | (0.0100) |
| Household annual Income | -0.24*** | -0.30*** | -0.21*** | -0.306*** |
| | (0.0143) | (0.0273) | (0.0216) | (0.0287) |
| Expected income in urban areas | 0.112 | 0.580*** | -1.60*** | 1.425*** |
| | (0.0734) | (0.0943) | (0.167) | (0.384) |
| Bus station | -0.09*** | 0.0576 | -0.0545 | -0.160** |
| | (0.0285) | (0.0583) | (0.0431) | (0.0599) |
| Train station | -0.0360 | -0.213** | -0.135* | 0.114 |
| | (0.0385) | (0.0818) | (0.0618) | (0.0683) |
| Location: Base=Township | | | | |
| Village | 0.407*** | 0.933*** | 0.463*** | -0.232* |
| | (0.0467) | (0.0966) | (0.0671) | (0.0954) |
| Percentage of agricultural labor force | 0.000307 | -0.01*** | 0.000394 | 0.0071*** |
| | (0.0006) | (0.0011) | (0.0008) | (0.0013) |
| Constant | -3.98*** | -9.04*** | 12.22*** | -15.61*** |
| | (0.772) | (1.203) | (1.666) | (3.664) |
| Central region | -0.079* | | | |
| | (0.037) | | | |
| Western region | 0.058 | | | |
| | (0.050) | | | |
| Control year | Yes | Yes | Yes | Yes |
| Wald chi2 | 7,093.4 | 1,946.6 | 3,184.7 | 2,076.2 |
| R-square | 0.27 | 0.25 | 0.28 | 0.31 |
| N | 48,815 | 14,144 | 21,399 | 13,272 |

# 3 The impact of rural-urban migration on human capital investment in rural China*

## 3.1 Introduction

China has witnessed immense internal migration from the rural to the urban areas since the 1980s, which relates to an issue that economists have been highly interested in: the effect of migration on the accumulation of human capital (Frisbie 1975; Zhao 1997; Coleman 2008; Lu 2012). The relationship is especially interesting as migration may exacerbate existing inequalities in the investment in this type of capital between rural areas with, on average, lower educational attainment and urban areas with, on average, higher levels of human capital. The implications of migration for the destination region are the topic of a large body of research with contradictory findings (e.g. Card and Krueger 1990; Borjas 2003) but the question whether migration is harmful or beneficial for the source region, i.e. the discussion of whether migration leads to a 'brain drain' or 'brain grain', respectively, is even more controversial and has been the topic of a lot of literature in the field of migration (Stark et al. 1997; Fan and Stark 2007; Beine et al. 2008; Beine et al. 2011; Marchiori et al. 2013). This chapter aims to add evidence to the effect of migration on the investment in human capital with the help of panel data from China that permits investigating changes within, rather than between villages, using the availability of train stations as a proxy for migration flows. Controlling for unobserved differences across villages appears crucial as I find a negative effect of migration on educational outcomes of stayers when I do not control for this and a robust positive effect if I do.

There are theoretical arguments for both positive and negative impacts of migration on educational outcomes, which one of the two dominates is an empirical question that is difficult to answer due to the nature of the relationship. Specifically, while the prospect of migration is proposed to have an effect on the level of human capital accumulation by changing investment incentives (Rapoport and Docquier 2006; Dustmann and Glitz 2011),

---

* This chapter heavily relies on an unpublished manuscript titled "Rural-urban migration, train stations, and education in rural China" which is coauthored with Dr. Julia Anna Matz. It was submitted to the conference "China after 35 Years of Economic Transition" in London on May 8-9, 2014. Furthermore, I am grateful to participants of various seminars and workshops for their valuable feedback on this paper. However, all errors in this chapter are mine.

education levels may in turn also impact on the likelihood of migration, which makes it difficult to isolate either effect in an empirical fashion. The most promising strategy to identify the causal effect of migratory movements on the accumulation of human capital is the use of instrumental variables. While a number have been proposed (Hanson and Woodruff 2003; Hildebrandt et al. 2005; Mishra 2007; McKenzie and Rapoport 2011), most studies do not account for the unobserved differences between villages. This chapter adds to the literature by suggesting a novel instrument for internal migration: the availability of train services in the village, and by focusing on changes within villages over time in response to a new train station being available. In contrast to de Brauw and Giles (2008), who also use of a large panel dataset from China and find a robust negative relationship between the opportunity to migrate and high school enrollment, my main empirical findings suggest that the exposure to migration encourages the accumulation of human capital measured by educational attainment, which may be due to my study controlling for the unobserved heterogeneity across villages. I support my results in sensitivity checks using different measures of the exposure to migration and varying specifications.

Rural-urban migration may not only transfer existing human capital from the countryside to cities, it may also impact on investment in human capital in the source region. To begin with the possible positive channels for the effect of migration on educational attainment, migration of household members may relax credit constraints associated with the education of children due to remittances being sent (Edwards and Ureta 2003; Acosta 2006). Yang (2008), for example, studies how sudden shocks in exchange rates affected child schooling and educational expenditure in the Philippines through their effect on remittances during the Asian financial crisis in 1997 and finds positive effects of remittances on human capital accumulation, and López-Córdova (2006) shows that municipalities in Mexico that receive relatively high remittances have higher literacy and school attendance rates among children aged 6 to 14. Furthermore, while the majority of the skilled labor force may leave the source region for destinations where the return to their education is higher, i.e. a ‘brain drain’ may take place (Marchiori et al. 2013), the prospect of migrating to an urban area (or abroad) in which high-skilled jobs are more prevalent than in the rural areas may increase the possible payoff of education and thus, educational attainment. This would imply migration encouraging the formation of human capital (Mountford 1997), thereby leading to a 'brain gain' if some of these prospective migrants end up staying (Stark et al. 1997; Stark et al. 1998; Stark 2005). Empirically, Beine et al. (2010) investigate the possibility of a 'beneficial brain drain' using both cross-sectional and panel data for a large set of developing countries and find evidence for higher emigration rates being positively associated with the accumulation of human capital. Similarly, Batista et al. (2007, 2012) argue that the existing figures on the brain drain are too high and that significant gains from migration are possible for the source country if out- and return migration are allowed.

Conversely, there are channels through which migration may negatively impact on the formation of human capital. For example, besides the possible direct negative effect of migration through highly-skilled individuals leaving and the average level of education arithmetically decreasing, there may also be indirect effects. Firstly, there are possible labor market effects. For example, when the educated leave a rural area, local wages for highly-skilled jobs increase due to a shortage of skilled labor, which would, in turn, increase rural wages for unskilled labor, thereby increasing the cost of migration and possibly lowering the investment in human capital due to lower relative returns (Zhang et al. 2011). Furthermore, migration of parents may have adverse effects on the educational involvement of their children (Djajić 2003). Hanson and Woodruff (2003), for example, argue that parental migration leads to a lower intensity of parental supervision, resulting in a reduction of study for children and Antman (2011) argues that children may have to increase work hours and sacrifice study time to make up for the migrated parent's lost work input. Similarly, Zhao et al (2012) find a negative relationship between parental migration and the performance of students with respect to test scores. McKenzie and Rapoport (2011) state that children in migrant families are less likely to attend school than children in non-migrant households.

The remainder of this chapter is structured as follows: Section 3.2 presents a simple theoretical framework on the relationship between migration and human capital accumulation. Section 3.3 introduces the data and presents descriptive statistics. Section 3.4 outlines my empirical strategy to identify the causal effect of migration on the investment in education. Section 3.5 discusses the results of this exercise, Section 6 concludes.

### 3.2 A framework on the relationship between migration and rural education

In this section I develop a simple theoretical framework on the relationship between rural-urban migration and the level of investment in human capital formation in rural areas, based on the work of Stark et al. (Stark et al. 1997, 1998; Stark and Dorn 2013). By understanding the opportunity cost of education as the investment in human capital, I reduce my framework to two key parameters: migration and years of schooling.

I assume that a rural worker lives for two periods, out of which decisions on human capital formation are made in the first one. A rural worker spends the first period of his life in the rural area and may invest in education or work in the agricultural sector. In the second period, the rural worker can choose to stay and continue working in agriculture or to migrate to the urban area, where he has a chance of a non-agricultural job whose wage is higher than the income from agriculture in the rural area.

The rural worker faces a credit constraint in the first period as he is endowed solely with one unit of labor and needs both time and capital to enhance his level of human capital. He can decide which proportion of his labor time $l$ to allocate to human capital formation in the first period. The function of human capital investment $H$ in the first period can be expressed as the continuously differentiable function

$$H=H(l), \text{ where } H_l' > 0, H_{ll}'' < 0 \ \forall l \in (0,1), \lim_{l\to 0} H_l' = \infty, \text{ and } \lim_{l\to 1} H_l' = 0 .$$

Rural (denoted by subscript $r$) farm production $Y$ in the first period (denoted by subscript *1*) is described by $Y_{1r} = f(1-l)$, where $1-l$ is the worker's input of time into agriculture. The opportunity cost of devoting human capital to agriculture in the first period is $Pf(l)$, where $P$ is the price of farm product. As mentioned in the beginning of this section, here I use the same income function as in Stark's model where the cost of labor, e.g., wages, is not considered (Stark et al. 1997, 1998).

The total income from agriculture in the first period is:

$$I_{1r} = Pf(1-l). \tag{3.1}$$

In the second period, the laborer has accumulated the amount of human capital based on the amount of time he chose to invest, $H=H(l)$, and the return to human capital depends on whether the worker migrates or not. If the worker stays in the rural area, farm production is equal to $Y_{2r} = f(H)$, and income from agriculture equal to

$$I_{2r} = Pf(H). \quad (3.2)$$

Non-agricultural jobs in the urban area (denoted by subscript *u*), on the other hand, offer a competitive wage $w_u$ , leading to an income in the urban area of $I_{2u} = w_u H.$

The utility of a worker, derived from income over the two periods of his life, can be expressed as $U(I)$, where $U(0) = 0$, $U'(I) > 0$, $U''(I) < 0$, $\forall I \geq 0$. The objective of the household is to maximize inter-temporal utility:

$$\max_{l,\theta}\{U(I_{1r}) + \varphi[U(\theta I_{2u}) + U((1-\theta)I_{2r})]\}, \quad (3.3)$$

subject to $l \in (0,1)$ and $\theta \in (0,1)$, where $0 < \varphi \leq 1$ is the subjective time discount rate and $\theta$ denotes the probability of migration with $0 \leq \theta \leq 1$.

Using the properties developed above and substituting into equation (3.3), the first derivative with respect to *l* is:

$$\frac{U'(I_{1r})(Pf(1-l)'_l)}{\psi[\theta U'(I_{2u})w_u + (1-\theta)U'(I_{2r})(Pf(H)'_l)]} = H'_l . \quad (3.4)$$

The denominator of the left-hand side of equation (3.4) can be rewritten as

$$Z(l,\theta) = H'_l. \quad (3.5)$$

The second derivatives of equation (3.5) with respect to $l$ and $\theta$ are:

$$\frac{d(Z(l,\theta))}{dl} = H''_{ll}, \quad (3.6)$$

$$\frac{d(Z(l,\theta))}{d\theta} = H''_{l\theta}, \quad (3.7)$$

which, according to the chain rural of differentiation, can be expressed as

$$\frac{dl}{d\theta} = -\frac{H''_{l\theta}}{H''_{ll}}, \quad (3.8)$$

where $H''_{ll} < 0$ due to diminishing returns of time invested into human capital formation. Thus, the sign of the relationship of human capital investment and probability of migration described by equation (3.8) depends on

$$H''_{l\theta} = \frac{-U'(I_{1r})(Pf'_l)[U'(I_{2u})w_u - U'(I_{2r})(Pf'_H)]}{\psi[\theta U'(I_{2u})w_u + (1-\theta)U'(I_{2r})(Pf'_H)]^2}. \quad (3.9)$$

Only if $H''_{l\theta} > 0$, equation (3.8) is positive, which would imply that the probability of migration is positively related to the formation of human capital. If $H''_{l\theta} < 0$, equation

(3.8) is negative, which would mean that the probability of migration reduces educational attainment. The denominator of equation (3.9) is positive because $[\theta U'(I_{2u})w_u + (1-\theta)U'(I_{2r})(Pf'_H)]^2 > 0$ and $\psi > 0$, so the sign of $H''_{l\theta}$ depends on the sign of numerator of equation (3.9). As I has assumed that $U'(I_{1r}) > 0$ and $(Pf'_l) > 0$, the sign of $H''_{l\theta}$ depends on

$$[U'(I_{2u})w_u - U'(I_{2r})(Pf'_H)], \quad (3.10)$$

Let us denote equation (3.10) as $\boldsymbol{\chi} = U'(I_{2u})w_u - U'(I_{2r})(Pf'_H)$, which is a utility function representing the relative risk aversion (Stark et al. 1998) .

It follows that, if $\boldsymbol{\chi}$ >0, then $H''_{l\theta} < 0$, hence $\frac{dl}{d\theta} > 0$, which indicates a positive relationship between human capital investment and the probability of migration. On the other hand, if $\boldsymbol{\chi}$ <0, then $H''_{l\theta} > 0$, hence $\frac{dl}{d\theta} < 0$, which indicates a negative relationship between human capital investment and the probability of migration.

The factors which, also intuitively, influence the decision to invest in human capital follow easily:

$$H_{l,\theta} = H(Pf'_l, Pf'_H, w_u, U', \theta), \quad (3.11)$$

where educational attainment $H_{l,\theta}$ is a function of the marginal returns to agricultural labor and human capital, $Pf'$ and $Pf'_H$, respectively, non-agricultural wages in the urban area $w_u$, properties of the household's utility function $U'$, and the probability of migration $\theta$. The aggregate effect of migration on human capital accumulation is, therefore, to be settled empirically as it depends on the specific form of the utility function that is difficult to assess, both theoretically and empirically. In the empirical part of the paper, I will account for it by including individual, household, and village characteristics.

## 3.3 Data description

Due to questions on migration of household members only being included from 1997 onwards I restrict the dataset of CHNS to the latest five rounds that took place in 1997, 2000, 2004, 2006, and 2009. Besides information on the migration and education of household members, the survey includes questions on demography, education, employment, housing conditions, income, agricultural practices, time use, community facilities, and health and nutritional measures.

### 3.3.1 Variables of migration

As the definition in Chapter 1, I define a migrant in this survey as a member of a rural household who does not currently live in the rural household but has migrated (to an urban area) for the purpose of finding a job.[19] As I am specifically interested in the relationship between migration and education with respect to informed decisions about payoffs of human capital accumulation in the labor force, I, thus, ignore individuals who have left the household for reasons related to marriage, education, military service, or other reasons. Figure 3.1 displays the percentage of migrants according to the definition by years of age between 1997 and 2009. The highest share of migrants is found for individuals of just above 20 years of age. I am specifically interested in the effect of exposure to migration and define my key measure of migration as the relative number of migrants in the village, i.e. the ratio of the number of migrants from a village to the total population of stayers in the village, both as measured with the household survey data available. As can be seen in Table 3.1, this variable takes an average value of 8% and ranges from 0 to 36%. It is also shown that villages have a mean number of migrants of almost six as measured by my survey data.

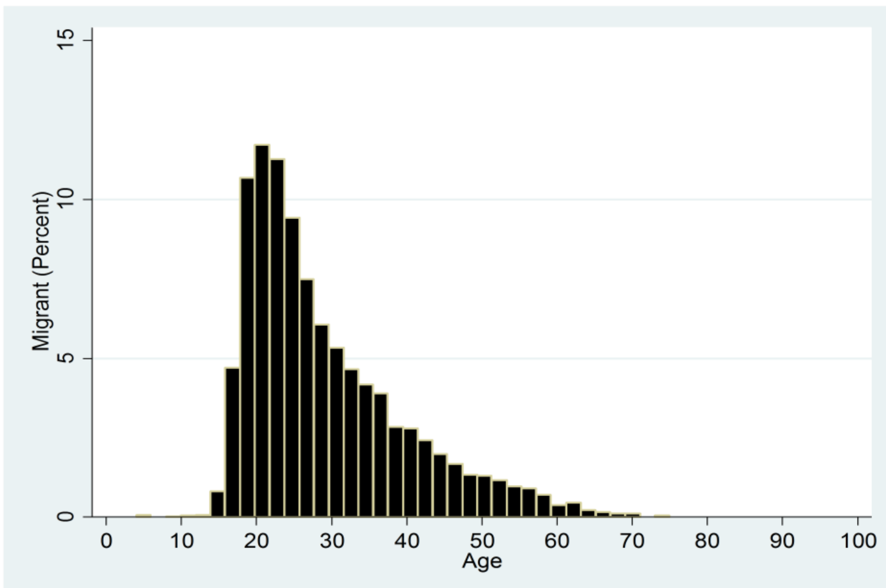
### 3.3.2 Measuring the investment in human capital

With respect to education, the other one of the two main concepts I am interested in, it should first be noted that educational attainment of women and men, migrants and non-migrants has increased by about one year on average over the time span of my data, and especially strongly between 1997 and 2000. Figure 3.2 displays mean years of schooling of the rural sample by gender and migration status in relation to age. It is easy to see that women generally complete less years of education than men and that migrants who leave at a young age (possibly for unskilled jobs in the urban area) are generally less educated than the ones that stay behind (to invest in their human capital). This is supported when

[19] While this is not clearly spelled out in the English translation of the questionnaire, the option of a member having left the household to search for employment in the original version is understood as migrating to an urban area to find employment.

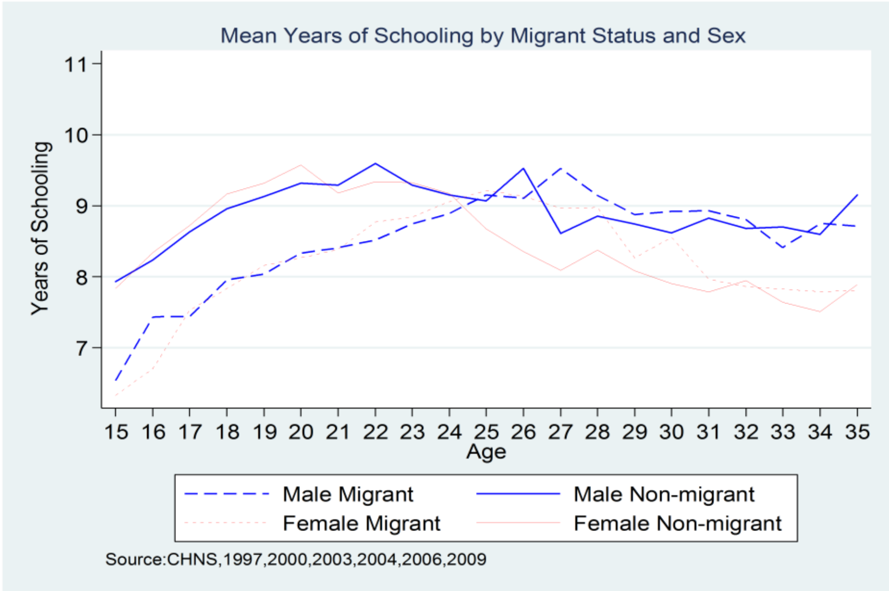
taking into account that at age 16, the youngest legal age workers are allowed to have in China. Migrants have completed seven years of schooling on average, while non-migrants possess over eight years of education. And, furthermore, while non-migrants reach the compulsory minimum level of schooling of nine years at 18 years of age, migrants on average possess only eight years of schooling at this point in their lives. The picture changes and from an age of about 25, however, migrants have completed more years of education than non-migrants on average.

Figure 3.1 Share of migrants by age group




The fact that China's educational system is strongly regulated in terms of the age at which children may start school is beneficial to my analysis. Specifically, children are legally required to start school at age 6 (Brown and Park 2002), which means that they normally finish primary school at age 12, middle school at an age of 15 years, and high school at age 18. Due to the compulsory minimum education of nine years in China (Connelly and Zheng 2003), primary and middle schools are highly subsidized by the government and parents only faced with relatively small monetary costs of sending children to school (Tsang 1996). Attending high school, on the other hand, is not compulsory and associated with tuition fees that may amount to a large fraction of annual household income so many young adults, when facing the opportunity cost of continued education, decide to drop out and look for employment (Glewwe and Jacoby 2004).

Figure 3.2 Mean years of schooling by migration status and gender




Another factor that aids my analysis is the *hukou* system. Even though the discrepancy in possibilities with respect to education between children in rural and urban areas is unique in China and public educational facilities are, on average, better in urban areas, children from rural areas are usually not able to attend them due to the difficulty of becoming officially registered in the respective urban district. This difficulty is rooted in the relatively high financial burden that temporarily enrolled children put on the municipality. For this reason, migrating parents usually leave their children in the source region unless they are able and willing to pay tuition fees for a private school. While this is surely not ideal in itself, it mitigates concerns of selection that would be apparent if children of migrants could easily migrate with their parents. If such a family was exceptionally able, for example, and the children of this household would have done relatively well and stayed in education relatively long also in the rural area, this form of selection would lead to a downward bias in my estimate of the effect of migration on educational attainment.

Table 3.1 Descriptive statistics

| | Mean | Min. | Median | Max. | Std. Dev. | N |
|---|---|---|---|---|---|---|
| **Village level** | | | | | | |
| ratio of migrants | 0.08 | 0 | 0.06 | 0.36 | 0.08 | 147 |
| number of migrants | 5.79 | 0 | 4 | 33 | 5.68 | 147 |
| size of the population | 3,394 | 162 | 1,894 | 74,501 | 5,830 | 147 |
| primary school | 0.80 | 0 | 1 | 1 | 0.40 | 147 |
| middle school | 0.29 | 0 | 0 | 1 | 0.45 | 147 |
| high school | 0.13 | 0 | 0 | 1 | 0.34 | 147 |
| rural town | 0.20 | 0 | 0 | 1 | 0.40 | 147 |
| near trade area | 0.34 | 0 | 0 | 1 | 0.47 | 147 |
| labor share of agriculture | 48.41 | 0 | 50 | 100 | 30.33 | 147 |
| **Province level** | | | | | | |
| median urban household income | 20,900 | 12,405 | 18,658 | 55,686 | 8,283 | 9 |
| **Household level** | | | | | | |
| education of the mother | 5.97 | 0 | 6 | 17 | 3.83 | 1305 |
| education of the father | 8.10 | 0 | 9 | 18 | 2.95 | 1305 |
| number of siblings | 2 | 0 | 2 | 6 | 0.98 | 1305 |
| household size | 4.96 | 2 | 5 | 29 | 1.89 | 1305 |
| household income | 24,229 | 45 | 16,560 | 855,270 | 31,792 | 1305 |
| **Individual level** | | | | | | |
| years of schooling | 9.04 | 0 | 9 | 18 | 2.56 | 2463 |
| age | 21.67 | 18 | 22 | 25 | 2.34 | 2463 |
| female | 0.47 | 0 | 0 | 1 | 0.50 | 2463 |

As I am interested in the effect of exposure to migration on the (post-compulsory) education choices of young adults who stay in the rural area, I restrict samples to individuals aged between 18 and 25 living in a rural area, which yields a sample size of 1,962 individuals with 2,463 observations from 147 rural villages. The lower cutoff relates to the age at which individuals graduate from high school if they decide to stay in secondary education, which means that I investigate a sample that has most likely finished their educational career and the associated decisions are not being made anymore, which would falsify my approach and findings. While I am in accordance with the existing literature (Chiquiar and Hanson 2002; Hanson and Woodruff 2003; McKenzie and Rapoport 2011) in restricting my sample by age and assuming that age is a good predictor of the amount of schooling, the measure may be inaccurate due to delayed enrollment or shorter primary schooling in some regions, grades being skipped or repeated. Unfortunately, I am unable to infer more specific information from my data but believe that the mentioned reasons for inaccuracy relate to unusual cases and should not strongly influence my results. To ameliorate this concern, however, I demonstrate that my conclusions are robust to varying restrictions on the sample with respect to age.

As shown in Table 3.1, the mean years of education is just above nine years; Table A3.1 in the Appendix A3 presents a detailed picture of the distribution of years of schooling in my sample. It is obvious that very few individuals have less than complete primary education and that approximately 75 percent of individuals have at least the compulsory nine years of education, which is also the median level of schooling. It follows that considering the effect of migration on post-compulsory education exclusively is sensible as this is where variation in decisions related to human capital investment exists.

### 3.3.3 Control variables

As shown in Table 3.1, 80% of the villages from which I have data have a primary school, 29% have their own middle school, and 13% have a high school, which impacts on the likelihood of children attending further education, in particular in rural areas where public transportation is often problematic. It should be pointed out that many rural schools provide dormitories for the students, which could influence attendance. Unfortunately, this survey does not supply available data on dormitories. I cannot control for this influence.

Twenty percent of villages are classified as rural towns and about a third of the villages from which I have data are close to special trade areas that provide relatively good employment opportunities.[20] The mean labor share in agriculture takes a value of 48%.

A little less than half of the sample is composed of women with a mean age of 21.7 years. The education of mothers takes an average value of just under six years, while fathers have received over two years more on average. Households have a mean size of almost 5 members and individuals an average of two siblings, which is not unusual, even considering the Chinese ‘one child-policy’, which was applied in a less strict fashion in rural areas. Average real annual household income is equal to 24,229 yuan, the mean of

[20] The relevant question yielding the latter variable is: “Is there an open trade area, an open city, or a special economic zone near this village/neighborhood (within two hours by bus)?” (Question O40 in the community questionnaires 1997 to 2009).

median urban household income within the province takes a value of 20,900 yuan.[21] It is surprising that mean income is higher in the rural area. This may either be driven by large income disparities in the urban areas or possibly also by outliers in rural household income as indicated by the large maximum value relative to the mean and median.[22]

Remittances may be one of the key channels through which migration influences decisions related to education as mentioned above. Unfortunately, my data show severe shortcomings in terms of missing data and possible misreporting. I am unable to account for this factor. However, remittances are a consequence of migration and the two concepts, therefore, inevitably intertwined (McKenzie and Sasin 2007). I am in accordance with a lot of the literature that does not explicitly consider the effect of remittances when studying the comprehensive impact of migration (Brauw et al. 2008; Mckenzie and Rapoport 2007).

Finally, it should be noted that attrition with respect to entire villages or households is not a serious concern in my study. To be specific, 37 % of villages are included in the sample in all rounds and 35% are included in four rounds. This means that over two thirds of the villages are represented in at least four of the five rounds. Households, however, are not as continuously represented in the data I use due to the restrictions imposed according to the age of the individuals being studied and them "growing out" of the sample. Looking at the whole survey, over one third of households are surveyed in all five periods, and about 60% are included in at least four out of the five survey rounds, which is not unusual in surveys covering such a long time frame, unfortunately.

## 3.4 Empirical analyses

In this section I outline the empirical strategy with which I aim to assess the causal impact of migration on the accumulation of human capital. I discuss the difficulties in estimating the causal relationship, outline a widely adopted remedy, the approach of using instrumental variables, and discuss why the instrument I propose, the availability of a local train station, is sensible.

[21] Values for annual household income are inflated to 2009. For reasons of comparison, one US-dollar corresponded to 6.831 Yuan in 2009 according to official exchange rates available from China's Statistical Yearbook 2012.

[22] Note that I pay attention to outliers in my empirical approach.

### 3.4.1 The problem of endogeneity

To begin with, I outline a naïve and reduced form equation (3.12) for the impact of the exposure to migration on educational attainment similar to (de Brauw and Giles 2008) :

$$years\ of\ schooling_{ihjpt} = \beta_1 share\ of\ migrants_{jpt} + \boldsymbol{Z}'_{ihjpt}\boldsymbol{\beta_2} + \boldsymbol{X}'_{hjpt}\boldsymbol{\beta_3} + \boldsymbol{V_{jpt}}'\boldsymbol{\beta_4} + \beta_5 \ln\left(I^u_{pt}\right) + v_t + e_{ihjpt}, \quad (3.12)$$

where the dependent variable is the years of schooling of individual $i$ from household $h$ in village $j$ and province $p$ at time $t$. The ratio of migrants to the total population of the village is my key variable of interest measuring the exposure to migration. The median urban household income $I^u$ within the province acts as a measure of expected income if migrating. $Z$ is a vector of individual characteristics such as age and gender, while $X$ represents household level control variables like the education of the mother and the father, the number of siblings, the logarithmic value of household income and household size. Furthermore, I control for whether the village an individual resides in has a middle or high school, whether it is near a special trade area, and for the share of employment in the village being in agriculture. Survey round indicators are included with the help of $v$, $e$ is a stochastic error term. I initially estimate equation (3.12) using Ordinary Least Squares (OLS) with heteroskedasticity-robust standard errors.

As briefly touched upon above, the difficulty in estimating the causal effect of migration on education lies in the fact that there may be reverse causality between the two, and a simultaneity or omitted variable bias, all of these implying endogeneity in the presence of which OLS is unable to produce unbiased estimates (Greene 2012, 219-233). Specifically, it is also likely that a relatively high level of education is beneficial for migration due to higher expected incomes in the destination and lower costs of migration due to easier access to employment (for example, Huffman 1980; Zhao 1999; Wu and Yao 2003; Rong et al. 2012), and thus, that causality does not exclusively run from migration to education. Alternatively, it may be that a factor that is not included in equation (3.12) drives both the decisions to migrate and how much time to invest in education. Think of motivation or ambition on part of the parents, for example, that could lead to both part of the family migrating and children being urged to stay in education for a relatively long time.

**3.4.2 The instrument**

In accordance with much of the literature, I adopt an instrumental variables technique to estimate the causal effect of migration on educational attainment to circumvent the problems outlined above (Hanson and Woodruff 2003; Antman 2011; McKenzie and Rapoport 2011). A relatively large number of instruments for this specific question have been proposed that can be broadly categorized as either relating to migrant networks that facilitate the migratory process and have been the topic of a large body of research themselves (e.g. Rozelle et al. 1999; de Brauw and Giles 2008; Zhang and Zhao 2011; Zhao et al. 2012; Giulietti et al. 2013) or to directly lowering the costs associated with migration. As examples of instruments in the first category, some scholars (Hanson and Woodruff 2003; Hildebrandt et al. 2005; Acosta 2006) use historical migration rates to proxy current migration, de Brauw and Giles (2008) use the time of the initial distribution of identity cards for rural residents of China to measure migrant networks. With respect to instruments related to a change in migration costs, the distance to urban areas ( McKenzie and Sasin 2007; McKenzie and Rapoport 2011) and the occurrence of natural (Munshi 2003) or economic shocks (Yang 2008) have been applied.

I propose a novel instrument, the availability of a local train station, and argue that it is both relevant to the endogenous explanatory variable, the exposure to migration, and uncorrelated with the error term, i.e. that it causally impacts on education exclusively through migration. While I do not know the exact timing a train station was built or opened, the community questionnaire asks: "Is this village near a train station?"[23] so I end up with nine possibilities for when a local train station became available: before 1989, between 1989 and 1991, between 1991 and 1993, between 1993 and 1997, between 1997 and 2000, between 2000 and 2004, between 2004 and 2006, between 2006 and 2009, and no train station until 2009. The distribution of these is displayed in Table 3.2. I generate a binary variable taking a value of one if a train station is available in a period, and zero otherwise. The distribution of new train stations being available is presented in Table 3.2: almost 60% of villages state that a local train station was not yet available as of data

[23] This question is number 13 (O35) and can be found on page 12 in section 9 (Other Facilities and Services) of the 2009 commune questionnaire (the page number may differ in the other rounds).

collection in 2009, while a large number of over 12% of villages received a local train station between 1989 and 1991, for example.

Table 3.2 Local train station openings for villages over time

| Railway station | Number of Villages | Percent | Cumulative |
|---|---|---|---|
| None yet | 88 | 59.86 | 59.86 |
| Until 1989 | 7 | 4.76 | 64.63 |
| 1989-1991 | 18 | 12.24 | 76.87 |
| 1991-1993 | 3 | 2.04 | 78.91 |
| 1993-1997 | 9 | 6.12 | 85.03 |
| 1997-2000 | 3 | 2.04 | 87.07 |
| 2000-2004 | 11 | 7.48 | 94.56 |
| 2004-2006 | 7 | 4.76 | 99.32 |
| 2006-2009 | 1 | 0.68 | 100 |
| Total | 147 | 100 | |

The study by de Brauw and Giles (2008) is closely related to my paper and empirical approach. They investigated how the opportunity to migrate influenced high school enrollment in rural China between 1986 and 2003 based on data from four provinces, two of them also being investigated here and two neighboring provinces. They used the time of the initial distribution of national identity cards in villages as their instrument for migration by arguing that ID cards reduce the costs associated with migration. They found a negative relationship between the opportunity to migrate and high school enrollment (de Brauw and Giles 2008).

Furthermore, I argue that my instrument may work along similar lines as historical migration rates of Mexicans to the United States of America from the 1920s as used by Mckenzie and Rapoport (2007, 2011), which are argued to be the result of largely historical demand-side factors combined with the spatial pattern of Mexican railroad availability. Specifically, immigrant workers were needed for the booming American economy at the time and recruited mainly from regions along the railroads going from the U.S. south into Mexico. This led to the West-Central Mexican states generally having relatively high migration rates compared to the rest of the country and even current

Mexican migrant networks being associated with the historical railroad system (Mckenzie and Rapoport 2007).

While historical migration rates are virtually equal to zero for a large part of Chinese history, railroad services may play an important role in the decision to migrate either by reducing migration costs or by facilitating local employment and commuting as an alternative to migration. With respect to direct costs of migration, travelling by train is cheaper and more comfortable than any other mode of transportation in China, thus, highly demanded by migrants when planning to return to their village of origin for important holidays, for example.[24] Furthermore, a local train station may be associated with a strong network of migrants from the “home region” in the destination, which lowers the costs associated with migration. On the other hand, train services may provide local employment through more firms locating there or, for rural communities that are relatively close to special trade areas, train services may also offer a daily mode of transport to work, hence making migration for employment purposes unnecessary. While all of these are possibly at play in China, the empirical question is whether one of them outweighs the other.

While the connection between the availability of a local train station and migration flows is relatively straightforward, I now outline why I believe the former is a sensible instrument for the latter in my situation, i.e. why there is no relationship between the availability of a local train station and educational attainment of young adults in the region other than through its effect on migration. A train station may affect the educational attainment of young adults because schools are so far that they need to take a train to get there. If schools could provide dormitories, these students could avoid relying on trains. However, taking the train to school is a very rare situation for Chinese rural students. Due to a lack of information on schools with dormitories, I cannot check this possibility.

[24] Another mode of transport is provided by buses. Bus stations, however, may not be used as an instrument as the exclusion restriction does not hold. To be specific, the local government that strongly influences whether bus stations are being erected is also the one making decisions related to educational facilities so there may be a relationship between bus stations and educational attainment of stayers other than through migration flows. Note that approximately 90% of the villages in my sample have a bus station and that I do not find any apparent relationship between the availability of bus and train stations in villages; the correlation coefficient is only 0.2.

The first possibility for the exclusion restriction being invalid is that both may be driven by wealth of a village, i.e. richer villages may be able to build local train stations earlier and to provide better educational facilities. However, in the setting of my study, China, this is not a valid concern as decisions to build new schools or an additional train station are made by different levels of government. Specifically, while the local government usually decides on investments related to education, it is the central and provincial governments that decide on and provide the financing for additional train stations (Li 2013). As local layers of government are unlikely to have an influence on decision processes within the central or provincial governments, the opening of a local train station can be understood as an exogenous shock leading to a change in migration rates, thereby satisfying the criteria that need to be fulfilled for instruments to be valid (Angrist 2001).

If the assignment of a local train station is exogenous and not related to characteristics of the village such as wealth, I should not be able to detect any differences between the characteristics of villages that have and those that do not possess a local train station. Table 3.3 presents basic summary statistics and mean comparison tests for the average value of the migration ratio, average years of education in the village, size in terms of population, classification as a rural town, proximity to a special trade area, the percentage of the labor force working in agriculture, median urban income within the province, household income and indicators of whether the village has a middle or high school in the survey round of 2009. It is reassuring to see that most differences are not statistically significant. Rural communities with a train station are statistically significantly more likely to be classified as a rural town, however. The migration ratio is slightly higher in villages with a railway station but not statistically significant, which is unexpected. Average household income is higher in villages with a train station, which are also more likely to have a middle school but not a high school. None of these differences are statistically significant, however, which supports my choice of instrument.

The fact that rural communities with a train station exhibit statistically significantly higher average educational attainment may be a preview of my results. On the other hand, if trains were also used to commute to schools, the exclusion restriction would break down. While I do not believe this is prevalent, I address this concern as part of the robustness checks.

Table 3.3 Characteristics of villages with and without train stations in 2009

| | with station | | | without station | | | with-without |
|---|---|---|---|---|---|---|---|
| | N | Mean | Std. Dev. | N | Mean | Std. Dev. | Difference |
| ratio of migrants | 59 | 0.11 | 0.01 | 88 | 0.12 | 0.01 | -0.01 |
| average years of education | 59 | 7.54 | 0.21 | 88 | 6.74 | 0.13 | 0.86*** |
| size of the population | 59 | 4,224 | 786.23 | 88 | 3,301.95 | 356.52 | 922.05 |
| labor share of agriculture | 59 | 37.29 | 4.29 | 88 | 44.10 | 2.88 | -6.81 |
| median urban HH income | 59 | 30,485 | 1,042 | 88 | 31,052 | 1,058 | -567 |
| household income | 59 | 22,913 | 1,301 | 88 | 22,32 | 1,194 | 589.26 |
| primary school | 59 | 0.64 | 0.06 | 88 | 0.61 | 0.05 | 0.03 |
| middle school | 59 | 0.29 | 0.06 | 88 | 0.3 | 0.05 | -0.01 |
| high school | 59 | 0.12 | 0.04 | 88 | 0.1 | 0.03 | 0.02 |
| rural town | 59 | 0.35 | 0.06 | 88 | 0.2 | 0.04 | 0.15** |
| near trade Area | 59 | 0.36 | 0.06 | 88 | 0.3 | 0.05 | 0.06 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Two-sample t-tests for unpaired data with unequal variances in all cases.

In addition, it may be that a newly built train station has an influence on characteristics of villages, which in turn also affect educational attainment, e.g. households and villages may become richer with a train station due to better possibilities of trade. I divide the sample of villages almost equally by separating those that received the train station before the survey round of 1997, and those that have received a new train station in and after the survey round of 1997 and display basic summary statistics and mean comparison tests in Table 3A.2 in the Appendix 3A. First of all, the absence of a statistically significant difference in migration rates is surprising at first but may be rooted in longer histories of railroad transportation being associated with a higher ratio of migrants, and of local employment or commuting acting as a substitute for migration and the two outweighing each other. Furthermore, household income is slightly higher in villages that received a train station early, which are also more likely to be classified as a rural town but less likely to be close to a special trade area and to have a middle school. Only the latter difference is statistically significant, which is not a source of big concern as middle school is compulsory and I investigate the effect of exposure to migration on post-compulsory education decisions. It is, thus, reassuring that the difference in high schools is not statistically significant.

When putting my instrumental variable into practice, I employ a standard two stage least squares approach (IV- 2SLS) and a village fixed effects instrumental variables estimator for panel data (IV- FE) in different specifications. Controlling for unobserved heterogeneity across villages is crucial, I aim to see the effect of a village receiving a new train station over time rather than seeing the average effect across villages, which may be driven by factors I am unable to observe. Therefore, the fixed effects instrumental variables estimator is my main specification.

## 3.5 Evidence on the relationship between migration and education

In this section I present the empirical results. I start by discussing the first stage, i.e. the effect of having a train station on migration and move on to discussing the results of the instrumental variables approach for the effect of migration on educational attainment. Finally, I present several robustness and sensitivity checks.

### 3.5.1 The first stage

Table 3.4 presents the key estimation result of the first stage, i.e. where migration is the dependent variable and my instrument, the binary variable for the existence of a local train station, is the key explanatory variable. I use main measure of migration, the relative number of migrants in the village, in columns (1) and (2) and the ratio of migrants above age 25 in relation to the size of the population above age 25 in column (3). In column (4), migration is simply measured as the number of migrants as not only the relative, but also the absolute size of migration may be important. I find statistically significant associations of the variable for the existence of a train station in all columns and achieve values of the F-statistic for weak identification well above the conventional threshold of 10 so the instrument appears valid. Furthermore, the Durbin-Wu-Hausman test rejects the hypothesis of migration being exogenous with a p-value of 0.0004 in the two-stage least squares specification. This is not a surprise, however, as apparent from my discussion of the literature. The sign of the relationship is negative, which is in line with the findings of de Brauw and Giles (2008).

As touched upon above, train services could enhance migration by reducing migration costs but may also reduce migration by facilitating local employment or commuting, thereby reducing migratory movements for the purpose of finding employment.

However, this relationship may depend on the size and remoteness of the community. I will check these factors in the robustness checks. In this setting, this is what appears to be driving the results of the first stage. I split the sample into rural communities that are close to a special trade area (defined as within two hours of reach by bus) and those that are not and I find evidence in support of the explanation being mainly driven by commuting.

As shown in Table A3.3, the effect is stronger in villages that are close to special trade areas, i.e. those for which train stations are likely to make commuting to employment easier. This is also supported by the fact that villages with a train station have a lower share of the labor force being active in agriculture as presented in Table 3.4, even if this difference is not statistically significant.

Table 3.4 The first stage– Migration flows and train stations

| | ratio of migrants | | (age >25) | number of migrants |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | IV-2SLS | IV-FE | IV-FE | IV-FE |
| train station | -0.017*** | -0.025*** | -0.043*** | -2.611*** |
| | (0.003) | (0.005) | (0.005) | (0.40) |
| Individual controls | Yes | Yes | Yes | Yes |
| Household controls | Yes | Yes | Yes | Yes |
| Village controls | Yes | Yes | Yes | Yes |
| Village fixed effects | No | Yes | Yes | Yes |
| Year indicators | Yes | Yes | Yes | Yes |
| Number of observations | 2463 | 2463 | 2463 | 2463 |
| F- test statistic | 62.17 | 24.03 | 19.38 | 25.45 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Robust standard errors are presented in parentheses. Individual, household and village controls include all those discussed in relation to equation (1). Instrumental variables estimation with village fixed effects except in column (1). The dependent variable is the ratio of migrants in a village in relation to its population size in columns (1) and (2), the ratio of migrants above age 25 in relation to the population size above age 25 in column (3) and the number of migrants in column (4).

### 3.5.2 The main results

Table 3.5 and Table A3.4 present main results with the results of the OLS specification presented in column (1), the instrumental variables two stage least squares ones in column (2), and the ones for instrumental variables with village fixed effects in columns (3) through (5). While I use the ratio of migrants to the population of the village in the first three columns, I use the relative number of migrants above age 25 in relation to the population of the village above age 25 in column (4) to ensure that my findings are not simply the result of a an arithmetic exercise combined with a selection effect. If, for example, individuals with a low level of education were the ones most likely to leave and the measure of migration did not exclude individuals of the same age span as those in the sample used for estimation, the effect of migration on education should be positive by construction. In column (5) I do not use the ratio of migrants to population but simply the number of migrants originating from a village. In the last three columns, I actually employ a fixed effect model to control the fixed effects of the community, e.g., size, remoteness. For reasons of space, the main results are split with the key explanatory variables presented in Table 3.5 and the remainder in Table A3.4 in the Appendix A3.

With respect to my coefficient of interest, the different specifications show opposite signs of the coefficient of interest: I find a negative and statistically significant effect of the exposure to migration on educational attainment in the OLS specification in column (1) and in the simple two-stage least squares specifications in column (2). As soon as I control for the unobserved heterogeneity across rural communities, however, the effect of migration is consistently statistically significant and positive across the different measures of migration. To be specific, the OLS and simple instrumental variables results indicate that an increase in the ratio of migrants in the village by ten percent is associated with a reduction in the years of schooling of young adults by 0.2 or 2 years, respectively, while the same increase in the ratio of migrants to the population leads to an increase in the years of education by 3.8 years. The latter sounds like a very big effect, but if the relative number of migrant increases by 1%, the average educational attainment increases by 0.4 years, which is still big but not unrealistic.

Table 3.5 The impact of migrant ratio on educational attainment

| | years of schooling | | | | |
|---|---|---|---|---|---|
| | (1) OLS | (2) IV-2SLS | (3) IV-FE | (4) IV-FE | (5) IV-FE |
| ratio of migrants | -2.108*** | -21.99*** | 38.52*** | | |
| | (0.659) | (6.412) | (13.62) | | |
| ratio of migrants (>25) | | | | 22.59*** | |
| | | | | (6.915) | |
| number of migrants | | | | | 0.374*** |
| | | | | | (0.120) |
| median urban HH income | 0.660*** | 0.478** | 2.041** | 1.493** | 1.577** |
| | (0.195) | (0.234) | (0.794) | (0.590) | (0.630) |
| household income | 0.288*** | 0.367*** | 0.193*** | 0.208*** | 0.201*** |
| | (0.0523) | (0.0629) | (0.0612) | (0.0531) | (0.0554) |
| female | -0.0740 | -0.134 | -0.0728 | -0.0648 | -0.0903 |
| | (0.0918) | (0.107) | (0.112) | (0.0973) | (0.102) |
| age | -0.0087 | -0.00005 | -0.0385 | -0.0093 | -0.0234 |
| | (0.0199) | (0.0231) | (0.0275) | (0.0213) | (0.0231) |
| education of the mother | 0.0598*** | 0.0346* | 0.0197 | 0.0198 | 0.0217 |
| | (0.0136) | (0.0178) | (0.0192) | (0.0166) | (0.0173) |
| education of the father | 0.144*** | 0.134*** | 0.106*** | 0.124*** | 0.113*** |
| | (0.0188) | (0.0217) | (0.0243) | (0.0199) | (0.0213) |
| middle school | 0.443*** | 0.408*** | 0.352 | 0.420** | 0.509** |
| | (0.113) | (0.132) | (0.218) | (0.197) | (0.220) |
| high school | -0.176 | -0.027 | -1.12*** | -1.31*** | -1.42*** |
| | (0.161) | (0.178) | (0.338) | (0.321) | (0.356) |
| Village fixed effects | No | No | Yes | Yes | Yes |
| Year indicators | Yes | Yes | Yes | Yes | Yes |
| N | 2463 | 2463 | 2463 | 2463 | 2463 |
| Adjusted R-squared | 0.22 | | | | |
| Wald Chi-squared | | 491.06 | 28484.15 | 37974.95 | 34790.96 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Robust standard errors are presented in parentheses. Columns (4) and (5) are identical to column (3) with the exception of the endogenous regressor here being the ratio of migrants above age 25 to the size of the population above age 25, and the number of migrants, respectively.

When investigating the control variables, I find, in most cases as expected, that both household and median urban income within the province are positively associated with educational attainment. Parental education, especially of the father, is positively associated with years of schooling. Having a middle school in the village yields statistically significant positive coefficients, while those for having a high school in the village are statistically significant and negative, which is surprising.

The findings of my main specification are in contrast to other empirical research which suggests that low returns to high school education are a likely explanation for the negative relationship between exposure to migration and educational attainment (de Brauw et al. 2002; Zhang et al. 2002; Cai et al. 2008; Chi et al. 2012). It should be noted, however, that most existing research uses cross-sectional data or panel data without exploiting one of the most valuable characteristics of panel data: the possibility of controlling for unobserved heterogeneity. Doing exactly this and addressing the endogeneity in the relationship by employing an instrumental variables approach, I obtain results that contradict those of OLS and many existing studies. Controlling for the unobserved heterogeneity across villages, I find a positive effect of migration on the investment in human capital, which is in line with the literature on the "brain gain" ( Stark et al. 1997, 1998, 2005), possibly by changing investment incentives as suggested by Beine et al. (2001, 2008, 2011).

**3.5.3 Robustness checks**

I perform a series of robustness and sensitivity checks to support my main findings and present and discuss them in this section.

**3.5.3.1 Proximity to a special trade area**

As the effect of the existence of a train station impacts on migration differently depending on whether a village is close to a special trade area or not as discussed above and presented in Table A3.4, I also replicate the main results in these sub-samples. The results for villages close to a special trade area are displayed in column (1) of Table 3.6 and in column (2) of the same table for villages where this is not the case. Interestingly, migration does not yield a statistically significant coefficient in villages that are close to a special trade area, although it should be kept in mind that the sample size is diminished. The main results are supported in villages that are not close to a special trade area, however. This suggests that a possible facilitation of commuting may be responsible for the negative relationship between the availability of train services but does not impact on the incentives to invest in human capital. In villages that do not have the opportunity of commuting because train services become available, the impact of migration is statistically significant, positive, and comparable to the main results in terms of magnitude. This also mitigates the concern of train services facilitating commuting to educational facilities driving the result, which would have violated the exclusion restriction as mentioned above.

**3.5.3.2 The effect of other types of migration**

While I specifically investigate the effect of an exposure to migration for the purposes of finding a job for my main results, I now investigate whether the effect depends on migration being for this specific reason or whether it is driven by migration in general. To be specific, I investigate the effect of the relative number of migrants in the village that migrated specifically for the purpose of higher education and those that left for any type of migration, irrespective of their motivation. The results of the former are presented in column (3) of Table 3.6, the latter in column (4) of the same table. The effect of migration for higher education is much stronger than the one of migration for the purpose of finding employment in the main results. Together with the one for migration in general in column (4) being smaller than the one reported in the main findings in Table 3.6, this suggests that it may be the expected payoff from migration for a highly-skilled job that acts as an incentive to invest in the formation of human capital.

Table 3.6 Proximity to a special trade area and different definitions of migrants

| | close to a trade area<br>(1)<br>IV-FE | far from a trade area<br>(2)<br>IV-FE | migration for higher education<br>(3)<br>IV-FE | all forms of migration<br>(4)<br>IV-FE |
|---|---|---|---|---|
| ratio of migrants | -27.24 | 37.81** | | |
| | (20.72) | (16.18) | | |
| ratio of migrants (education) | | | 102.7** | |
| | | | (48.43) | |
| ratio of migrants (all) | | | | 28.00*** |
| | | | | (9.660) |
| Number of observations | 840 | 1623 | 2463 | 2463 |
| Wald Chi-squared | 14014.71 | 19333.36 | 16007.99 | 29908.86 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Robust standard errors are presented in parentheses. All columns include individual, household, and village controls, village fixed effects and time indicators as in the main results. Column (1) restricts the sample to villages that are close to a trade area, column (2) to those that are not. The denominator of the ratio of migrants in column (3) includes only migrants that have left for the purpose of higher education; the one in column (4) includes all migrants, irrespective of their motivation.

#### 3.5.3.3 Different effects by demographic characteristics

With respect to demographic characteristics, I first split the sample into sub-samples according to age and replicate the main results reported in column (3) of Table 3.7. The key results for the sub-sample of individuals aged 18-21 are reported in column (1) of Table 3.7, while those for the subsample of individuals aged 22 to 25 are presented in column (2) of the same table. The coefficients on the ratio of migrants in the village are statistically significant and positive in both columns but the one in column (1) is larger, which indicates that the effect is stronger for younger individuals.

It may be that gender plays a role as well so I split the sample: male stayers are investigated in column (3) Table 3.7, while the results for female ones are reported in column (4) of the same table. It is interesting to see that the effect is exclusively apparent for male individuals in the rural community and absent for female ones, which may be rooted in male individuals being more active in the labor market, especially in high-skilled employment.

Table 3.7 A binary measure of education and in sub-samples according to age

| | Age 18-21 | Age 22-25 | male stayers | female stayers |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | IV-FE | IV-FE | IV-FE | IV-FE |
| ratio of migrants | 38.88** | 32.92* | 31.23* | 35.03 |
| | (16.51) | (19.32) | (15.19) | (22.20) |
| Number of observations | 1143 | 1320 | 1309 | 1154 |
| Wald Chi-squared | 13873.49 | 16638.50 | 18204.93 | 14036.07 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Robust standard errors are presented in parentheses. All columns include individual, household, and village controls, village fixed effects and time indicators as in the main results. Column (1) restricts the sample to individuals aged 18-21, column (2) to those aged 22-25. Male individuals are included in column (3), female ones in column (4).

#### 3.5.3.4 Other sensitivity checks

It is apparent from the summary statistics that the majority of villages do not have a train station by 2009, which is when my data end. For this reason, I verify the main results in the sub-sample of villages that have a train station by 2009. The results are presented in column (1) of Table A3.5 in Appendix A3 and support my main findings, also with respect to the magnitude of the effect.

Furthermore, outliers in household income may be an issue as touched upon above. I, therefore, replicate the main results without individuals who live in households that report incomes in the top or bottom fifth percentile and present the results in column (2) of Table A3.5 in Appendix A3 and see that the results are not sensitive to this measure.

Finally, the size and remoteness of a community could influence the labor market opportunity for the local residents, which plays an important role in migration decisions. Remote and small villages have limited labor opportunities in the local market. People from remote and small villages are more likely to migrate. It is possible that the relationship between migration and education is more significant. I checked these two sensitivity checks in Table A3.6. I divided villages into remote villages and township village according to whether the village is located in the township in column (1) and (2). In the last two columns, (3) and (4), I categorized the village into halves by median population of villages. Villages which are smaller than the median population are defined as small villages. Those bigger than the median population are considered as big villages. The results of the tests confirm my analysis that migration in remote and small villages has a positive effect on education while township and big villages showed insignificant results.

### 3.6 Conclusions

My study investigates the relationship between migration and educational attainment, which is the heart of a large body of literature due to its relevance, particularly in developing economies and due to the difficulty of clearly estimating the causal effect in applied studies. I first developed a simple framework in which the sign of the effect may be negative or positive depending on the returns to education and form of the utility function. Empirically, estimating the effect of migration on education is difficult due to reverse causality, i.e. the prospect of migration may influence the investment in human capital, but different levels of education may also impact on the likelihood of migrating.

I follow a large number of other scholars in their approach to provide causal evidence for the impact of migration on education by using an instrumental variable approach and take advantage of the fact that rural communities in China have not been connected to the railroad system at a uniform point in time. This allows me to propose a novel instrument for migratory flows: the availability of a local train station in Chinese villages. While it is relatively straightforward that the possibility of using railway services enhances migration, the relationship may also work in the other direction through the facilitation of local employment or commuting, thereby making migration unnecessary. I find evidence for the latter outweighing the former as the existence of a train station is negatively associated with out-migration in my data. Furthermore, I argue at length that there is no direct link from a village having a train station to the educational attainment of young local adults, a critical criterion of a valid instrument.

I use the ratio of migrants to the total population in a village as my main measure of the exposure to migration. My results add to the literature by demonstrating a negative effect when using ordinary least squares or a basic two stage least squares procedure. I argue that accounting for the unobserved heterogeneity across villages and investigating the effect within villages over time rather than average effects across villages is crucial, however. When the instrumental variables approach is employed with village fixed effects, i.e., when I investigate changes within villages over time rather than across villages, I find a positive effect of migration on educational attainment among the stayers that is robust to using different definitions of my measure of the exposure to migration and in additional sensitivity checks. The findings suggest that internal migration should not be discouraged when enhancing educational attainment is also a topic on the policy-making agenda.

## Appendix A3

Table A3.1 Distribution of the years of schooling

| Years of schooling | Frequency | Percent | Cumulative |
|---|---|---|---|
| 0 | 22 | 0.89 | 0.89 |
| 1 | 5 | 0.20 | 1.10 |
| 2 | 9 | 0.37 | 1.46 |
| 3 | 26 | 1.06 | 2.52 |
| 4 | 34 | 1.38 | 3.90 |
| 5 | 145 | 5.89 | 9.78 |
| 6 | 160 | 6.50 | 16.28 |
| 7 | 80 | 3.25 | 19.53 |
| 8 | 134 | 5.44 | 24.97 |
| 9 | 1,170 | 47.50 | 72.47 |
| 10 | 51 | 2.07 | 74.54 |
| 11 | 140 | 5.68 | 80.23 |
| 12 | 391 | 15.87 | 96.10 |
| 13 | 9 | 0.37 | 96.47 |
| 14 | 13 | 0.53 | 97.00 |
| 15 | 54 | 2.19 | 99.19 |
| 16 | 16 | 0.65 | 99.84 |
| 17 | 3 | 0.12 | 99.96 |
| 18 | 1 | 0.04 | 100.00 |
| Total | 2,463 | 100 | |

Table A3.2 Characteristics of villages with train stations until and after the 1997 survey

| | Until 1997 | | | From 1997 | | | Until-From 1997 |
|---|---|---|---|---|---|---|---|
| | N | Mean | Std. Dev. | N | Mean | Std. Dev. | Difference |
| ratio of migrants | 31 | 0.11 | 0.02 | 28 | 0.10 | 0.02 | 0.01 |
| average years of education | 31 | 7.55 | 0.33 | 28 | 7.53 | 0.27 | 0.02 |
| size of the population | 31 | 3216 | 632.08 | 28 | 5339 | 1,489.58 | -2,123 |
| labor share in agriculture | 31 | 36.16 | 6.32 | 28 | 38.53 | 5.83 | -2.37 |
| median urban HH income | 31 | 31,718 | 1,662 | 28 | 29,119 | 1,178.71 | 2,598.6 |
| household income | 31 | 23,210 | 1,827 | 28 | 22,585 | 1,883.19 | 625.29 |
| primary school | 31 | 0.71 | 0.08 | 28 | 0.58 | 0.09 | 0.13 |
| middle school | 31 | 0.16 | 0.07 | 28 | 0.42 | 0.95 | -0.26** |
| high school | 31 | 0.06 | 0.04 | 28 | 0.18 | 0.07 | -0.12 |
| rural town | 31 | 0.32 | 0.08 | 28 | 0.39 | 0.09 | -0.07 |
| near trade area | 31 | 0.26 | 0.08 | 28 | 0.46 | 0.96 | -0.2 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Two-sample t-tests for unpaired data with unequal variances in all cases.

Table A3.3 The first stage split by proximity to a special trade area

| | Close to a special trade area | | | Not close to a special trade area | | |
|---|---|---|---|---|---|---|
| | ratio of migrants | | number of migrants | ratio of migrants | | number of migrants |
| | (1) IV-2SLS | (2) IV-FE | (3) IV-FE | (4) IV-2SLS | (5) IV-FE | (6) IV-FE |
| train station | -0.01*** | -0.05*** | -4.07*** | -0.01*** | -0.02*** | -2.65*** |
| | (0.003) | (0.012) | (0.959) | (0.003) | (0.006) | (0.465) |
| Individual controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Household controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Village controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Village fixed effects | No | Yes | Yes | No | Yes | Yes |
| Year indicators | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of observations | 840 | 840 | 840 | 1623 | 1623 | 1623 |
| R-adjust | 0.44 | | | 0.21 | | |
| F- test statistic | 49.76 | 76.73 | 63.23 | 35.67 | 57.44 | 55.71 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Robust standard errors are presented in parentheses. Individual, household and village controls include all those discussed in relation to equation (1).

Table A3.4 The Impact of migration ratio on educational attainment – Control variables

| | years of schooling | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | OLS | IV-2SLS | IV-FE | IV-FE | IV-FE |
| number of siblings | -0.0361 | 0.0473 | -0.104 | -0.0738 | -0.0892 |
| | (0.0585) | (0.0714) | (0.0751) | (0.0626) | (0.0665) |
| household size | -0.0555** | 0.00988 | -0.0666* | -0.0479 | -0.0622* |
| | (0.0271) | (0.0364) | (0.0371) | (0.0315) | (0.0333) |
| near trade area | -0.0462 | 0.169 | -0.346 | 0.0169 | -0.395* |
| | (0.0996) | (0.132) | (0.223) | (0.145) | (0.213) |
| labor share in agriculture | -0.0171*** | -0.00725** | -0.00272 | 0.000713 | 0.00177 |
| | (0.00160) | (0.00366) | (0.00358) | (0.00328) | (0.00355) |
| Village fixed effects | No | No | Yes | Yes | Yes |
| Year indicators | Yes | Yes | Yes | Yes | Yes |
| Number of observations | 2463 | 2463 | 2463 | 2463 | 2463 |
| Adjusted R-squared | 0.22 | | | | |
| Wald Chi-squared | | 491.06 | 28484.15 | 37974.95 | 34790.96 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Robust standard errors are presented in parentheses. Columns (4) and (5) are identical to column (3) with the exception of the endogenous regressor here being the ratio of migrants above age 25 to the size of the population above age 25, and the number of migrants, respectively.

Table A3.5 Sub-samples according to availability of the train station and excluding outliers in household income

| | train station in 2009 | exclude the top and bottom 5% |
|---|---|---|
| | (1) | (2) |
| | IV-FE | IV-FE |
| ratio of migrants | 32.44** | 37.09*** |
| | (13.55) | (12.21) |
| Number of observations | 1062 | 2216 |
| Wald Chi-squared | 14825.95 | 27159.49 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Robust standard errors are presented in parentheses. All columns include individual, household, and village controls, village fixed effects and time indicators as in the main results. The sample in column (1) includes individuals in villages that have a train station in 2009, column (2) excludes individuals from households with incomes in the top and bottom 5th percentile.

Table A3.6 Sub-samples according to remoteness and size of village

| | Remote villages | Township villages | Small villages | Big villages |
|---|---|---|---|---|
| | (1) | (3) | (4) | (5) |
| | IV-FE | IV-FE | IV-FE | IV-FE |
| ratio of migrants | 39.44** | 24.04 | 95.66* | 8.43 |
| | (13.55) | (21.37) | (55.43) | (8.32) |
| Number of observations | 1970 | 493 | 1228 | 1231 |
| Wald Chi-squared | 9820 | 9820 | 5146 | 22616 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Robust standard errors are presented in parentheses. All columns include individual, household, and village controls, village fixed effects and time indicators as in the main results.

# 4 The impact of rural-urban migration on agricultural productivity in China

## 4.1 Introduction

In China, the average annual growth rate of agricultural productivity has remained at a high level of over 3% in both labor and land productivity over the past three decades. The effects of migration on the process of agriculture modernization are still not very clear. Migration of the agricultural labor force is argued to be an important cause to explain the high growth rate of agricultural productivity (Shen et al. 1996).[25] On the other hand, migration has been found to decrease agricultural productivity (Schmook and Radel 2008) and has been shown not to influence the technical efficiency of agricultural production (Yang et al. 2014). In fact, because migration, agriculture productivity and agricultural capital investments are interrelated variables, it is difficult to identify the net effects of migration on agricultural productivity.

These empirical studies are not the only ones with contradictory results; Theoretical analyses also cannot give a conclusive result. The New Economics of Labor Migration (NELM) argues that migration of the labor force is a strategy to overcome imperfect capital markets in agriculture. The argument of NELM is that a migrant can bring remittances to a rural household to increase investments in agriculture. On the other hand, migrants reduce the labor input in agriculture. Therefore the impact of migration has the mixed effect of reducing labor inputs and increasing capital investments. In China, Rozelle et al (1999) argued that remittances could compensate the loss of output resulted from reduced labor by increasing agricultural yields. However, is it true that the migrant labor force has a negative effect on the labor input in production and a positive effect on capital investment? This will be the key question to discuss in this chapter. If the agricultural labor force is in surplus, then the migrant labor force from agriculture does not influence the output due to the zero marginal return of surplus labor forces. Furthermore, even if migrant households received remittances from migrant workers, they would not increase their capital investments in agriculture since they prefer consumption (De Brauw and Rozelle 2008). Neither of the situations discussed above has

[25] The other very important contributor to growth of agricultural productivity is the institutional reform of the household responsibility system (HRS) (Runsheng 2010).

a clear conclusion on whether rural-urban migration influences agricultural production negatively or positively.

Chinese agriculture is characterized by small farms, therefore capital investments, especially mechanization, rely on the farm size. This study investigates the impacts of rural-urban migration on agricultural productivity with regards to the farm size. Using recent data in CHNS from 2004, 2006, and 2009, this chapter examines the role of rural-urban migration on the agricultural productivity in China. In the empirical analysis, a system of equations is constructed to investigate the relationship between agricultural productivity, rural-urban migrants, farm size and capital investment. After controlling endogeneity in the empirical analysis, the results suggest that the effect of migration on agricultural productivity has two dimensions, the negative effect of loss labor force and positive effect of capital investment. The net effect would depend on the household's farm size.

The rest of this chapter is organized as follows. Section 4.2 will review the literature on the relationship of migration and agricultural productivity. Section 4.3 will construct a simple theoretical model to reveal the relationship between agricultural productivity and rural-urban migration. Section 4.4 will describe the data set. Section 4.5 will discuss the empirical analysis including specification, identification, and the results. Section 4.6 concludes.

## 4.2 Literature review

The relationship between migration and agriculture has been investigated by many economists. The earliest reference to this topic is David Ricardo's The Principles of Political Economy and Taxation from 1817. He formulated the concept of labor surplus and believed that migrated surplus laborers did not reduce agriculture production. In the first dualistic economic model, Lewis (1954) proposed that the marginal product of surplus labor was near zero in the agricultural sector because of unlimited labor force in agriculture.

Based on the Lewis's work, Ranis and Fei (1961) developed the hypothesis of perfectly elastic labor supply. Labor supply from the agricultural sector was not unlimited anymore. To achieve a balanced growth between the industrial sector and the agricultural sector, they emphasized that investments in the agricultural sector should be adequate to increase agricultural productivity sufficiently to adapt to the reduced agricultural labor force input.

Harris and Todaro's model (Todaro 1969; Harris and Todaro 1970) assumed that both the agricultural and industrial sectors were neoclassical while in the Lewis, and Ranis and Fei two-sector models only the industrial sector was assumed to be neoclassical. According to Harris and Todaro's model, the transfer of labor out of agriculture would reduce agricultural output. If the agricultural sector wants to maintain its agricultural output, agricultural productivity would have to increase enough to compensate for the reduction in output. However, Harris and Todaro's model does not pay special attention to agricultural productivity.

The new economics of labor migration presents a joint household model in which households allocate members to be migrants to overcome the barriers of an imperfect capital market in agricultural production ( Stark and Bloom 1985; Stark and Lucas 1988; Stark 1991; Taylor 1999;). Agricultural production benefits from the added financial capital input of remittances from migration. The NELM model redefined the analytical unit from the individual to the household. As Rozelle et al (1999) stated, the negative effect of migration on farm production is in part compensated through remittances. Additionally, rural-urban migration relaxes land constraints through land transfer in rural areas and then promote the agricultural productivity (Mullan et al. 2011; Deininger et al. 2012). Although the channel of influence may differ, migrant households increase the agricultural productivity in response to the change of factor inputs.

Out-migration of the labor force is a main cause of technical change and innovation in agriculture development, which can continuously increase agricultural productivity. In the 1990s, as the labor force in China shifted off the farm and arable land stagnated, agricultural output continued to grow with TFP growth at a rate of around 2% per year (Jin et al 2010). Mendola (2008) found that households who engaged in international migration were more likely to adapt to modern farming technology, thereby achieving higher productivity. Many empirical studies have confirmed a positive relationship between agricultural productivity growth and agricultural out-migration in China. These empirical studies can be divided into two categories based on the data sets. The first category was based on panel data of aggregated statistics in China (Fan and Pardey 1997; Carter et al. 2003; Ito 2010; Jin et al. 2010). The second category used cross-sectional data from household surveys (Rozelle et al. 1999; Taylor and López-Feldman 2007; Mendola 2008; Ito 2010).

## 4.3 Theoretical framework

In this section, I construct a simple theoretical model to illustrate the relationship between rural-urban migration and agricultural yields. The key purpose of this theoretical model is to discuss how migration influences agricultural investments and productivity. This model examines the combined effect of loss labor force and agricultural investments made as a result of remittances from migration to explain how a migrant household might influence agricultural productivity.

Agricultural productivity is a general concept. It measures the amount of output per unit of a kind of input. For instance, labor productivity refers to output per active person, land productivity measures output per unit of land. Generally agricultural productivity is known as labor productivity unless specified otherwise. However, labor productivity is not a good indicator of technical progress in agriculture. Labor productivity increases as long as the labor force decreases while the real output remains the same. This change of labor productivity does not reflect any technical progress or efficient improvement. Labor productivity can increase as long as the labor force decreases through outmigration from agriculture. By contrast, land productivity is a better indicator (Zepeda 2001). If the labor forces migrate out of agriculture, land productivity does not change as a result of the reduced labor forces. Only when the total output per unit of land increases, can the land productivity increase. Therefore, in this study, agricultural productivity refers to land productivity instead of labor productivity. Rozelle et al. (1999) adopted the same indicator to measure agricultural productivity. Agricultural productivity and land productivity (yield) are used interchangeably in the following sections.

To transform the output into yield (land productivity), I divided both sides of a production function by land area.

$$\frac{\text{Output}}{Land} = F\left(\frac{Labor}{Land}, \frac{Capital}{Land}, \frac{Land}{Land}\right). \quad (4.1)$$

The left side of the equation is yield denoted as $y$. On the right side, input of labor per unit land is $l$, capital per unit land is $k$. The function of agricultural yield for this household can be expressed as

$$\mathrm{y} = f(l, k). \quad (4.2)$$

This production function is an implicit function. However I assume that the production function has some common properties including positive marginal products of inputs and nonnegative cross derivatives.

### 4.3.1 Migration and agricultural investment

Assuming that a rural household initially has $l$ units of labor per unit of land and $k$ units of capital per unit of land for the agricultural production function $f(l, k)$, the household can send a portion of labor $m$ $(0 < m < l)$ out as a migrant worker in the non-agricultural sector. The wage income of the migrant worker is $lw$. Accordingly, the household loses the portion of labor $l$ input in agricultural production, but the migrant worker can remit a proportion $\phi$ $(0 < \phi < 1)$ of their wage income $\phi mw$ back to the rural household in agricultural production. So the migrant household's total income can be expressed as farm income $Pf(l - m, k + \phi mw)$ plus wage income$(1 - \phi)mw$. The household's total income is

$$Pf(l - m, k + \phi mw) + (1 - \phi)mw, \quad (4.3)$$

where $L = l - m, and\ K = k + \phi mw, then\ f'_K > 0, f''_{KK} < 0, f'_L > 0, f''_{LL} < 0.$

The objective function of the household is to maximize total income and can be written as,

$$\max_{l>m>0, 1>\phi>0} \{Pf(l - m, k + \phi mw) + (1 - \phi)mw\}, \quad (4.4)$$

where $P$ is the price of farm produce.

The objective function (4.4) is denoted as $R$. The Kuhn-Tucker first order conditions of optimization problems are

$$\frac{dR}{dm} = \phi wPf'_K - Pf'_L + (1 - \phi)w = 0, \quad (4.5)$$

$$\frac{dR}{d\phi} = mwPf'_K - mw = 0. \tag{4.6}$$

I can get the reduced form of (4.5) and (4.6), which are the labor input and capital input conditions of the maximized household income.

$$Pf'_L - w = 0, \tag{4.7}$$

$$Pf'_K - 1 = 0. \tag{4.8}$$

Equation (4.7) states that marginal returns to labor force are equal between agricultural production and non-agricultural production. The rural labor force will leave agriculture only when marginal returns to labor force in the non-agricultural industry are no less than that in the agricultural production, which means $w \geq Pf'_L$. In reality, the wages in the non-agricultural industry are always higher than the marginal returns to labor force in agriculture.

Equation (4.8) indicates that marginal returns to physical capital in agricultural production are equal to one, which means that a household would add one unit of capital when the marginal return of this one unit of capital equals the interest rate. It is easy to understand when I rewrite Equation (4.8) as $f'_K = \frac{1}{P}$. The capital input is not enough in production when the marginal returns to capital are more than the interest rate that is $Pf'_K > 1$, the household should increase capital input because it can get higher returns to capital than the interest rate. Otherwise, the household should decrease its use of capital when $Pf'_K < 1$, since the returns to capital are less than interest rate. Thus, a household will use capital to the scale determined by the equation of $Pf'_K = 1$. But households often cannot achieve equilibrium in use of capital because of obvious scarcity in capital and financial constraints in agriculture. In most cases of rural households in developing countries, there is not enough capital available to invest in agriculture and the returns to capital remains with $Pf'_K > 1$. Therefore, I will make an assumption of $Pf'_K > 1$ for following analysis.

Using (4.5) divided by (4.6), I can get the relationship between the proportion of migrant labor $m$ and proportion of investment in farm production $\phi$.

$$\frac{d\phi}{dm} = \frac{\phi w(Pf'_K - 1) + (w - Pf'_L)}{mw(Pf'_K - 1)}. \tag{4.9}$$

As discussed before, migrant wages are normally higher than the return to labor force in agriculture and the returns on one unit of capital are higher than one, namely $w - Pf_L' > 0$ and $Pf_K' - 1 > 0$. Therefore, the Equation (4.9) is positive. This suggests that migration and investment in agriculture has a positive relationship only if the returns to capital and labor satisfy certain conditions.

### 4.3.2 Migration and agricultural yield

As stated in NELM theory (Rozelle et al. 1999), it is assumed that the rural household can compensate the loss of labor input through investment in physical capital in agricultural productivity. To check this hypothesis, I considered the relationship between agricultural yield and migrated labor force. I calculated the derivation of migrant labor $m$ for agriculture production $y = Pf(l - m, k + \phi mw)$.

$$\frac{dy}{dm} = P(\phi w f_K' - f_L'). \tag{4.10}$$

Equation (4.10) illustrates that the effects of migrant labor are mixed. The first component in the right bracket indicates a positive effect of increased investment, $\phi wPf_K'$, while the second component in the right bracket represents a negative effect of the loss of labor, $-Pf_L'$, However, the comprehensive effect of migration depends on which effect is dominant.

Assuming the physical capital investment can compensate the loss of labor input, which means

$$\phi wPf_K' - Pf_L' > 0. \tag{4.11}$$

The economic meaning of inequality (4.11) is that returns to capital invested in agriculture should be higher than returns to labor due to migration.

This conclusion indicates that the negative or positive effect of migration on agricultural productivity depends on the comparison of the returns to labor and capital in agriculture and returns to labor in non-agricultural activities.

As discussed in the theoretical framework, rural-urban migration can influence agricultural productivity in two ways, namely, labor input and capital input. If migrant households have extra financial capital from a migrant worker, they have the capacity to invest more in agricultural production. On the other hand, migrant households lose some of their labor force in the agricultural production due to migration. However, whether the reduction in labor force decreases the yield is not clear since this reduced labor force may be surplus in production. The empirical analysis will focus on the effects of these two factors, migrated labor force and financial capital investment in the household.

## 4.4 Data description

I employ the longitudinal dataset of the China Health and Nutrition Survey (CHNS). Considering the completeness of variables on agriculture the sample is limited to households involved in agriculture in the 2004, 2006, and 2009 surveys. The basic observation is a farm household instead of an individual because the farm production is arranged by the unit of a household. The sample includes 2756 households in 102 villages, from 9 provinces. The data used in this study includes detailed information on agricultural production and the characteristics of each household and village. The key variables are described as follows.

### 4.4.1 Crop yield

In order to combine different crops into the aggregate output, the total crop output is measured by a household's total income from crops.[26] The total income from crops incudes not only the revenue from sales and estimated value of crops in stock,[27] but also crops consumed by the household in the survey year. As discussed in the theoretical framework, agricultural productivity refers to land productivity or yield in this study. The yield unit is yuan per mu. Yuan is the Chinese currency unit. Mu is a Chinese official area unit which equals 1/15 hectare. To remove the effect of inflation, all data measured in currency are deflated to the price level in 2009 through the CPI deflated index in the dataset. As shown in the first row of household level in Table 4.1, the mean household yield is 1,673 yuan per mu.

[26] Animal production is not considered here because the relevant information on land, labor and physical capital input in animal production is not available. I focus on the crop production in this study.

[27] This includes grains, tobacco, and greenhouse flowers. The relevant information can be found in question E14a in the household questionnaire from 2004 to 2009.

### 4.4.2 Migrant labor force

Migration has become a common strategy for farm households to maximize income. Most farm households send some of their labor force out as migrant workers to work in the urban manufacturing or service industries. Figure 4.1 illustrates the distribution of number of migrants in the household. The percentages of households with one, two and three migrants are 26%, 20% and 10%. Non-migrant households account for only 37% of all farm households. Migrant households form the majority of the rural households. The average number of migrants is 1.28 persons which can be found in the household level in Table 4.1.

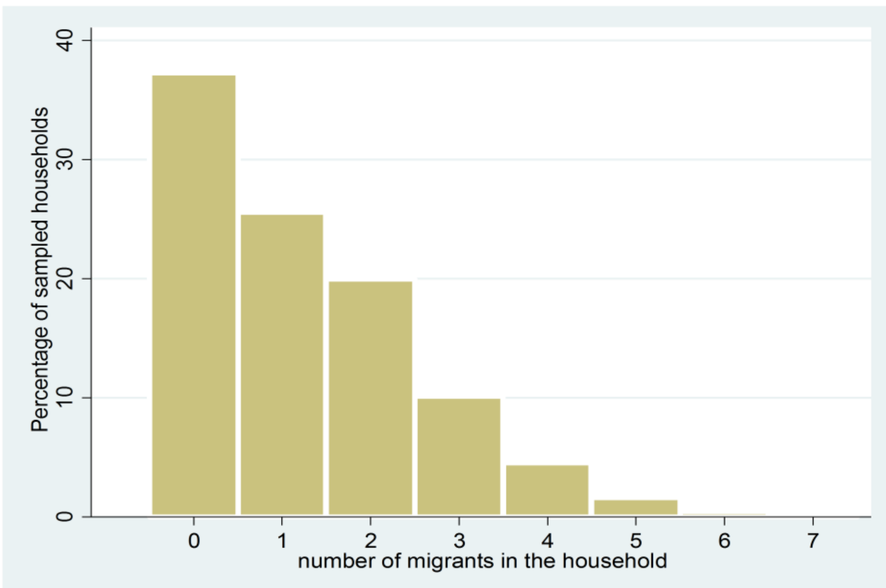



Figure 4.1 Histogram of number of migrants in the household
Source: selected samples from CHNS 2004, 2006, 2009

### 4.4.3 Capital investment in farm production

Capital investment in farm production is the sum of variable costs and fixed costs in production. Variable costs include purchasing seedlings, fertilizers, tools, insecticides, and hiring labor for the crops. The fixed costs refer to farm machinery including tractors, irrigation equipment, power thresher, water pump and others.[28] All capital investment is

[28] Detailed information on variable costs and farm machinery costs in the survey can be found in part 5 and part 14 of the CHNS household questionnaire in 2004, 2006 and 2009.

divided by area of farm land. The mean capital investment is 630 yuan per mu. As shown in the household level in Table 4.1.

The capital investment is associated with the size of the farm, namely the area of farm land. Chinese agriculture is dominated by small farms of less than 1 hectare of land each. In 1997 eighty-three percent of all Chinese household farms were less than 1 hectare while only 0.24% of farms were bigger than 6.6 hectares (Fan and Chan-Kang 2005). If the farm size is too small, it is impossible to promote modern agriculture, especially mechanized production. So farm size can influence the capital investment in agriculture.

### 4.4.4 Control variables

In addition to the above three key variables, crop yield, migrant labor force and capital investment in farm production, other characteristics of household and village should be controlled for the analysis. At the village level, a train station is available for one third of the villages. In some villages, parts of farm land have been plowed collectively. On average, 11% of the farm land in the villages is plowed by the collective. In general, most of the labor force in these villages (54%) is engaged in agriculture. One third of these villages are near the trade area.

At the household level, the age of farmers is an important influence on agricultural production. I calculated the average farmer age in a household because there is normally more than one farmer in a household. The average farmer age is 39 years old. The heads of households have two more years of education than their partner, namely 7.8 years compared to 5.6 years. There are 1.8 children in these households with an average of 4.6 people. The average number of migrants per household is 1.28 persons. The average annual rural household income is 23,155 yuan which is less than the median urban household income of 27.629 yuan. The rural and urban household incomes are important to represent the opportunity cost of rural-urban migration and of non-rural-urban migration. Migrants have to give up their rural income when they migrate and expected urban income when they do not migrate. The median urban household income in the province is assumed to be the expected income for the migrant household because the data of real migrant wages is not available. This assumption is plausible because rural households can easily access information on the income of urban households in their provinces. The average farm field is 1.673 yuan/mu while the farm capital investment is 629 yuan/mu. Half of the farm capital investment comes from the machinery investment (316 yuan/mu). The average land size for a household is 6.41 mu.

Considering the potential problem of multicollinearity between rural household income and farm income, education of household head and his partner, household size and number of children, I calculated the correlation coefficient of each pair of variables and found they were 0.18, 0.42 and 0.58, which show no serious problem of multicollinearity.

Table 4.1 Summary statistics of variables

| | Mean | Min. | Median | Max. | Std. Dev. | N |
|---|---|---|---|---|---|---|
| Village level | | | | | | |
| train station | 0.31 | 0 | 0 | 1 | 0.46 | 102 |
| collectively plowed land (%) | 10.56 | 0 | 0 | 100 | 28.58 | 102 |
| labor share in agriculture (%) | 54.34 | 0 | 50 | 100 | 24.09 | 102 |
| trade area | 0.30 | 0 | 0 | 1 | 0.46 | 102 |
| Province level | | | | | | |
| median urban household income | 27,629 | 12,404 | 26,900 | 55,686 | 10,777 | 9 |
| Household level | | | | | | |
| Average age of farmer | 39.35 | 20 | 37.7 | 82.5 | 9.86 | 2,756 |
| Farm yield (yuan/mu) | 1,673 | 2.27 | 1,539 | 195,472 | 1,064 | 2,756 |
| number of migrants (person) | 1.28 | 0 | 1 | 7 | 1.35 | 2,756 |
| machinery investment (yuan/mu) | 316 | 0 | 0 | 34,986 | 1,253 | 2,756 |
| capital investment (yuan/mu) | 629 | 4.33 | 366.8 | 35,335 | 1,291 | 2,756 |
| land size (mu) | 6.41 | 1 | 4 | 120 | 9.31 | 2,756 |
| education of head of household | 7.81 | 0 | 9 | 17 | 2.82 | 2,756 |
| education of head's partner | 5.56 | 0 | 6 | 15 | 3.49 | 2,756 |
| number of children | 1.77 | 0 | 2 | 7 | 1.04 | 2,756 |
| household size | 4.60 | 1 | 4 | 26 | 2.13 | 2,756 |
| rural household income (yuan) | 23,155 | 249.29 | 16,633 | 337,697 | 24,354 | 2,756 |

Source: selected samples from CHNS 2004, 2006, 2009

Note: all values measured in price are deflated to 2009 prices, including farm yield, capital investment, and rural household income.

## 4.5 Empirical analyses

To address whether migration can influence agricultural productivity, I constructed a system of three different functions to simultaneously decide the migrant members, agricultural productivity, and agricultural investment. The system of functions is adapted from a similar study by Rozelle et al. (1999) and Mendola (2008). Unlike their functions my system of functions considered the interaction of migration and land size as a function of capital investment.

### 4.5.1 Specification

The specification is expressed as a system of equations as follows:

$$\ln(Y) = \alpha_0 + \alpha_1 M + \alpha_2 \ln(K) + \alpha_3 \boldsymbol{X_Y} + \varepsilon_Y, \qquad (4.13)$$

$$M = \beta_0 + \beta_1 \boldsymbol{X_M} + \beta_2 \boldsymbol{Z_M} + \varepsilon_M , \qquad (4.14)$$

$$\ln(K) = \gamma_0 + \gamma_1 M * S + \gamma_2 \boldsymbol{X_K} + \gamma_3 \boldsymbol{Z_K} + \varepsilon_K, \qquad (4.15)$$

where $Y$ is the farm yield, $M$ is the number of migrants in the household, $K$ denotes calculated capital investment. $S$ is the area of farm land. Function (4.13), (4.14), and (4.15) are used to respectively determine agricultural productivity (farm yield), number of migrant members in the household, and capital investment in farm production. As Wooldridge (2002, 67-70) and Greene (2012, 699-701) have suggested, I set an interaction term in the function of capital investment. $M * S$ represents the interaction term of migrant members in the household and farm size because the capital investment is jointly influenced by both of them.

$\boldsymbol{X_Y}$, $\boldsymbol{X_M}$ and $\boldsymbol{X_K}$ are vectors of household and village characteristics influencing the farm yield, migration, and capital investment. These household characteristics include *Average age of farmers in the household*, *Educational attainment of the head of the household*, *Education of the head of household's partner*, *Number of Siblings,* and *Household Size, Farm size*. Furthermore, the villages characteristics consist of whether the village is near a special trade area, and the share of labor force in agriculture, whether the village has access to a train station, and percentage of land plowed collectively.

$\boldsymbol{Z_M}$ and $\boldsymbol{Z_K}$ are exogenous variables to be used as instruments for the endogenous migration variable and the capital investment variable. The instrument for the migrant member $\boldsymbol{Z_M}$ is whether the village access to the train station which has been stated as in Chapter 3. The instrument for capital investment $\boldsymbol{Z_K}$ is the area of collectively plowed land in the village. Both instruments for the migration decision and capital investment at the household level come from the village level. This design is to guarantee the instruments exogenous to the household farm production. However, I will argue the validity of these instruments in next section of the identification strategy.

$\varepsilon_Y$, $\varepsilon_Y$, and $\varepsilon_Y$ are stochastic variables of the correspondent estimated equations. They are assumed to be normally and independently distributed with variance in each estimated equation.

Actually, the farm yield, migration decision and capital investment are simultaneously determined. For instance, a household can make the migration decision by comparing average farm income with the wage of a migrant worker; meanwhile, their migration decision can directly influence the labor and capital input in agricultural production.

To reduce the heteroskedasticity of the monetary variables, I transform the variables measured in monetary units into logarithm form. The logarithmic transformation can fit the non-linear relationship better and make the homoscedastic normal assumption more plausible (Wooldridge 2002, 520). The testable hypothesis in the empirical analysis is whether rural-urban migration decreases or increases farm yield.

### 4.5.2 Identification strategy

In the system of functions (4.13), (4.14) and (4.15), it should be noted that migrant members $M$ and capital investment $K$ are explanatory variables in function (4.13), but explained variables in function (4.14) and (4.15) respectively. This would result to the problem of endogeneity in functions (4.14) and (4.15). To identify the number of migrants and capital investment in the system of simultaneous equations, I introduce two instruments to deal with the problem of endogeneity in the estimates.

In order to guarantee the system of equations to be identified, I should find at least two instrumental variables to identify the endogenous variables of migrants and capital inputs in farm production. These two instrumental variables should solely appear in the identified equation of migration and capital investment. This restriction imposed in a system of equation is known as an exclusion restriction because it can exclude certain endogenous and exogenous variables from each equation in the system (Wooldridge 2002, 183-191).

Specially, the first instrumental variable for number of migrants is whether a train station is available in the village. The second instrumental variable for capital investment in agricultural production is the percentage of collectively plowed land in the village. Both of these variables are valid instrumental variables because they are associated with endogenous variables but isolated with the disturbance in the function of household's farm productivity as shown in Equation (4.14).

A train station has both positive and negative effect on migration. It can reduce the cost of transportation for migration. The relationship between access to a train station and the number of migrants in the village should be positive. On the other hand, a train station can contribute to the development of the local labor market which can decrease the possibility of migration for the local population. The effect of a train station is mixed for migration. But I argue that a train station can hardly influence local agricultural productivity. A concern of using train station as an instrument is that these villages with train stations have higher agricultural productivity than villages that do not have a train station. I compared the characteristics of villages with and without a station to discuss whether a train station has a significant influence on the characteristics of agricultural production in the village.

In Table 4.2, I compare the villages with and without train stations until 2009 by the characteristics of agricultural production including farm income, farm capital investment, and farm size, and other characteristics of the village consisting of household income, labor share of agriculture and proximity to a trade area. It presents summary statistics and means comparison tests. The results show that none of the differences between the villages with and without train stations are statistically significant. Therefore, the instrumental variable, availability of train station, is independent from the agricultural production but associated with the migration of the household in the village.

Table 4.2 Characteristics of villages with and without train stations until 2009

| variables | with station | | | without station | | | with-without |
|---|---|---|---|---|---|---|---|
| | N | Mean | Std. Dev. | N | Mean | Std. Dev. | Difference |
| farm income | 35 | 7,821.67 | 4,382.94 | 67 | 7,820.51 | 6,078.75 | 1.16 |
| capital invested/mu | 35 | 716.08 | 569.23 | 67 | 568.40 | 476.60 | 147.68 |
| household size | 35 | 5.14 | 3.31 | 67 | 4.97 | 1.96 | 0.17 |
| farm size | 35 | 5.87 | 4.14 | 67 | 7.61 | 9.69 | -1.74 |
| household income | 35 | 22,472 | 8,719 | 67 | 23,713 | 14,260 | -1,240 |
| labor share of agriculture | 35 | 57.89 | 26.17 | 67 | 50.70 | 21.24 | 7.18 |
| near trade area | 35 | 0.34 | 0.48 | 67 | 0.30 | 0.46 | 0.04 |

Notes: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Two-sample t-tests for unpaired data with unequal variances in all cases.

The second instrumental variable, the percentage of collectively plowed land, has a direct effect on the capital investment in agriculture for the households in the village because there are some necessary conditions to plow farm land collectively. First, the size of land which is possible to be plowed collectively has to be big enough. Otherwise, the small and fragmented farm land cannot be collectively plowed. The average land size in the villages with collectively plowed land may be bigger than in the villages without collectively plowed land. Second, the collective plowed land needs machinery operation in mass scale. The machinery cost in these villages with the collective plowed land may be bigger than in the villagers without collective plowed land. However, there is no clear evidence that higher costs of machinery can increase or decrease land productivity. Therefore, except for influencing agricultural capital investment (i.e. machinery investment), the percentage of collectively plowed land cannot directly influence the yields.

I compared the differences of agricultural characteristics of villages with and without collectively plowed land to support the above analysis on the validity of the instrumental variable. In Table 4.3, I report the summary statistics and mean comparison tests. The results suggest that only household size and household income significantly differ. However, some variables differ obviously but not significantly because the relevant variances are very big. For instance, as I explained before, the machinery cost per mu in the villages with collectively plowed land is 314 yuan more than in the villages without collectively plowed land. Capital investment per mu in the villages with collectively plowed land is 358 yuan more than those without collectively plowed land. Farm income in the villages with collectively plowed land is 2,219 more than in the villages without it. The farm size in the villages with collectively plowed land is twice that of the villages without collectively plowed land. However, their differences are not statistically significant. Besides the relatively huge variances of the above variables, the other possible reason may come from the small sample of villages with collectively plowed lands (13) compared to big sample villages without collectively plowed land (89). However, the instrument is plausible because there is no significant evidence to prove that the instrument is a factor to influence the agricultural productivity between the villages with and without collectively plowed lands.

Table 4.3 Characteristics of villages with and without collectively plowed lands

| variables | with collectively plowed land | | | without collectively plowed land | | | with-without |
|---|---|---|---|---|---|---|---|
| | N | Mean | Std. Dev. | N | Mean | Std. Dev. | Difference |
| farm income | 13 | 9,757.37 | 6,427.42 | 89 | 7,538.06 | 5,372.83 | 2,219.32 |
| machinery cost/mu | 13 | 594.65 | 243.98 | 89 | 281.10 | 37.87 | 313.55 |
| capital invested/mu | 13 | 931.13 | 880.60 | 89 | 573.49 | 423.10 | 357.64 |
| household size | 13 | 3.77 | 1.74 | 89 | 5.21 | 2.54 | -1.44** |
| farm size | 13 | 12.21 | 13.97 | 89 | 6.26 | 6.83 | 5.96 |
| household income | 13 | 29,867 | 11,521 | 89 | 22,326 | 12,524 | 7,540** |
| agricultural labor share | 13 | 55.85 | 26.83 | 89 | 52.78 | 22.74 | 3.07 |
| near trade area | 13 | 0.46 | 0.52 | 89 | 0.29 | 0.46 | 0.17 |

Notes: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Two-sample t-tests for unpaired data with unequal variances in all cases.

### 4.5.3 Main results

I started from the OLS regression estimate of the impact of migration on agricultural productivity. In Table 4.4, Column (1), (2) and (3) are estimates of specification (4.13), (4.14) and (4.15) respectively. Although I controlled some household and village characteristics, the number of migrants has no significant effect on farm yield and capital investment. However, capital investments per mu have a positive effect on farm yield.

The reduced regression between capital investment per mu and the percentage of collectively plowed land is significant and positive. The monetary unit variables have been transformed into logarithmic form because they can make the homoscedastic normal assumption of error term more plausible to satisfy (Wooldridge 2002, 520). The results of OLS must be biased without considering the endogeneity in the system of the equation. The stochastic disturbances in the equations are correlated because the agricultural yield, number of migrants and capital investment in agriculture are more likely to be determined simultaneously. OLS estimates are not suitable here. The results of OLS regression are shown in Table 4.4. Coefficients of control variables are reported in Table A4.1.

Table 4.4 OLS regression estimate of the impact of rural-urban migration on agricultural productivity

| | Dependent variables | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Independent variables | ln(yield) | number of migrants | ln(capital investment/mu) |
| Migration effects: | | | |
| Number of migrants | -0.00519 | | -0.0165 |
| | (0.00949) | | (0.0202) |
| (number of migrants*farm size) | | | 0.00219 |
| | | | (0.0019) |
| Farm size (mu) | -0.0162*** | | -0.015*** |
| | (0.00132) | | (0.0028) |
| ln(capital investment/mu) | 0.204*** | | |
| | (0.0106) | | |
| Instruments | | | |
| Near train station | | 0.0286 | |
| | | (0.0455) | |
| Collectively plowed land (%) | | | 0.00201*** |
| | | | (0.0007) |
| Control variables | Yes | Yes | Yes |
| Time indicators | Yes | Yes | Yes |
| Province indicators | Yes | Yes | Yes |
| N | 2,756 | 2,756 | 2,756 |
| Adj. R-sq | 0.368 | 0.430 | 0.193 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Standard errors are presented in parentheses. The fixed effects of province and time are included in all equations but not reported in this table. The variables associated with income are calculated by logarithms.

As discussed in 4.5.2 Identification strategy, in order to overcome the endogeneity problem which cannot be solved by OLS estimates, I estimated the system of Equations consisting of (4.13), (4.14) and (4.15) using a two-stage least squares estimator (2SLS) and a three-stage least squares estimator (3SLS). 3SLS estimate applies to the procedures for instrument variables (IV) to generalize the consistent estimates like the 2SLS estimate. However, it also takes advantage of the least squares to account for the associated disturbances terms across equations in the system (Mendola 2008), which 2SLS does not. The advantage of the 3SLS is that the 3SLS estimator is generally more efficient if the system extension of the homoscedasticity assumption holds true, as stated by Wooldridge (2002, 198-99). However, if the assumption does not hold, the 3SLS estimator is inconsistent. In this case, 2SLS is a better option to choose because it can generate consistent estimates regardless of whether the system extension of homoscedasticity assumption holds true or not. The results of 3SLS are reported in Table A4.2. Some of estimated coefficients from 3SLS are much bigger than those from 2SLS.

When choosing an estimation method for this simultaneous equations system, it is important to remember the advantages and disadvantages of 3SLS (Wooldridge 2002, 222). If all equations are correctly specified, 3SLS are asymptotically more efficient than 2SLS. But 2SLS is more robust as I am particularly interested in the equation 4.13, 2SLS is consistent and asymptotically normal if equation 4.13 is correctly specified and the instruments are exogenous. To avoid the inconsistent estimates of 3SLS resulting from misspecified equations in the system, I chose the 2SLS as the main estimate in this study because it can get more robust estimates.

The main results of the empirical analysis come from the 2SLS estimation of model. As shown in Table 4.5, I performed some statistical test to check the validity of the two instruments, i.e. endogeneity and weak instruments. I first carried out the Durbin and Wu-Hausman tests to test a subset of the endogenous regressors in the model, and found that the number of migrants and capital investment in the model are endogenous. The first instrument variable, near a train station, is not significant in the second Column. But the second one, the percentage of collectively plowed land, is significant. The significance of each instrument cannot decide whether this set of instruments is strong or weak. I performed Stock and Yogo’s test (2005) to test whether these combined instruments are weak. The Stock and Yogo's test suggests that I can reject the null hypothesis of a set of weak instrument. The set of instruments is strong.

Both instruments show a positive sign in the regression which is expected in the instrument selection analysis. Compared to the negative sign of train station in Chapter 3, the positive sign in this chapter may derive from the limited samples of households involved in agriculture, which makes the negative effect of train station on local labor market small. The positive relationship between the percentage of land plowed collectively and capital investment in agricultural production has been discussed before. Collectively plowed land requires more agricultural investment, e.g. machinery operation.

The main results in Table 4.5 support two conclusions about the impact of rural-urban migration on agricultural productivity. First, as shown in Column (1), the number of migrants has a negative impact on agricultural productivity, but the effect is very weak. One migrant could decrease 6.84% of crop yields, which indicates that the labor force in agriculture is still sufficient. The effect of reducing the labor force on agricultural productivity is small.

Second, One percent of capital investment per mu increases crops yields by 1.09%. The results are consistent with the findings of Rozelle et al (1999). Their study on migration, remittances and agricultural productivity in China found that one additional migrant decreased maize yield by 230 kg per mu and an additional yuan remitted increased yield by 0.22 yuan per mu. Many scholars confirm that capital investment in crop production can improve productivity (Chen et al. 2010). However, migrant households do not increase capital investment in agriculture in the beginning. Instead they decrease their productive investment. Other scholars found the same investment behavior for the migrant households (de Brauw and Rozelle 2008). The negative effect changed with the interaction of number of migrants and farm size. The marginal effect of migration on capital investment per mu depends on the interaction term of migrants and farm size. According to the coefficients of these variables in column (3) of Table 4.5, the marginal effect of migration on capital investment can be expressed as follows,

$$\frac{dK}{dM} = -1.744 + 0.091\, S \qquad (4.16)$$

Where $\frac{dK}{dM}$ denotes the marginal effect of migration on capital investment per mu in crop production, and S is farm size. $\frac{dK}{dM} \geq 0$ if $S \geq 19.16$ mu. This is the minimum farm size where marginal effect of migration changes from negative to positive.

The economic explanation of marginal effect is that migrant households with more than 19.16 mu (1.28 hectares) would increase investments in crop production while the other households with less than 19.16 mu would decrease investments in crop production. The increased capital investment is mainly in the form of machinery which requires a large size of farm land. This estimation of 1.28 hectares is consistent with correspondent research by Van den Berg et al. (2007) who found that rice cultivation was only mechanized with at least 1.8 hectares of farm land and 2.4 hectares of mixed farm. Their findings of the relationship between farm size and mechanization are close to my estimations.

Some of the other explanatory variables are also noteworthy. The area of land has a negative relationship to agricultural productivity. When there is one mu land increase, crop yields decrease by 0.6%. This indicates that small farms are more productive in China, which has been found by the many studies based on the panel data of household survey covering 8 provinces (Gao and Zhang 2006) and 1 province ( Li et al. 2009) in China. The age of farmers has a positive influence on crops yield. The effect of educational attainment of the head of household and their partners on agricultural productivity is negative. A plausible explanation to the above relationship is that aging farmers with more experience but less education are more productive in agriculture.

The educational attainment of household head has a negative effect on the number of migrants. The household size and number of children has a positive effect on agricultural productivity and the number of migrants while farm income and rural household income have a negative effect on migration. Household size and the number of children could promote the number of migrants. Because farm income and rural household income are the opportunity cost for migration, the incidence of migration decreases with the increase of opportunity cost.

Table 4.5 Estimation of the impact of rural-urban migration on yields using 2SLS estimate

| | Dependent variable | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Independent variable | ln(yield) | number of migrants | ln(capital investment/mu) |
| number of migrants | -0.0684** | | -1.744** |
| | (0.0333) | | (0.681) |
| number of migrants*farm size | | | 0.0911*** |
| | | | (0.0352) |
| Farm size | -0.00647** | | -0.0838*** |
| | (0.00259) | | (0.0276) |
| ln(capital investment/mu) | 1.092*** | | |
| | (0.0631) | | |
| Near train station | | 0.0286 | |
| | | (0.0455) | |
| Collectively plowed land (%) | | | 0.00316** |
| | | | (0.00133) |
| Control variables | | | |
| Average age of farmers | 0.0122*** | | 0.0154 |
| | (0.00251) | | (0.0112) |
| education of head | -0.00690 | -0.0205*** | -0.0196 |
| | (0.00728) | (0.00772) | (0.0160) |
| education of head's partner | -0.000440 | -0.0276*** | -0.0104 |
| | (0.00625) | (0.00647) | (0.0117) |
| number of children | 0.0741** | 0.562*** | 0.794** |
| | (0.0328) | (0.0239) | (0.319) |
| household size | 0.0142 | 0.147*** | 0.189** |
| | (0.0117) | (0.0116) | (0.0794) |
| ln(farm income) | | 0.00699 | |
| | | (0.0289) | |
| ln(urban household income) | | 0.153 | |
| | | (0.206) | |
| ln(rural household income) | | -0.152*** | -0.0392 |
| | | (0.0253) | (0.0825) |
| near trade area | 0.136*** | -0.0836* | |
| | (0.0468) | (0.0492) | |
| labor share in agriculture | -0.00867*** | -0.0000781 | 0.00571*** |
| | (0.000996) | (0.000947) | (0.00181) |
| N | 2756 | 2756 | 2756 |
| F-statistics | 32.86 | 104.86 | 9.13 |
| P-value | 0.000 | 0.000 | 0.000 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Standard errors are presented in parentheses. The fix effects of province and time are included in all equations but not reported in this table. The variables associated with income are calculated by logarithm.

In summary, households with migration decrease investment in farm production unless land size reaches an optimal level. This finding is different from Rozelle et al's (1999) and Cao and Birthenall (2013). They found that migration can increase agricultural productivity investments. My argument is that migrant households do not increase investments in agricultural production. Productive investments are not rural household priorities even if their incomes increase through migrant employment. This argument is supported by other relevant studies. de Brauw and Rozelle (2008) argued that there was no evidence of a relationship between migration and productive investments. Rural households prefer to spend their additional income on housing and other durable goods. The reasons why they do not invest in agriculture are varied, but one of the most important is the insecurity of land tenure (Gao et al. 2011).

## 4.6 Conclusions

To illustrate the impact of migration on agricultural productivity in this chapter, I discussed a theoretical model and tested for empirical evidence. The theoretical model showed that the migrant labor force could reduce agricultural productivity while capital from migrant remittances can be invested in agriculture. However, the net effect of migration is mixed. It depends on the comparison of returns to labor force and returns to capital investments. The net effect relies on empirical evidence.  The empirical analysis shows that the effect of reduced labor force due to migration is estimated to reduce 6.8% of agricultural productivity. On the other hand, the effect of migration on capital investments is negative at first but changes to positive when the farm size reaches an optimal area. It indicates that there is a minimum farm size for households to increase investments in agriculture.

Chinese agriculture is experiencing an important transition. Rural-urban migration can provide capital investments to change from labor intensive agriculture to capital intensive agriculture. However, the effect of migration has not been completely recognized so far. This study has a significant implication for policymakers. The motivation of capital investment depends on farm size.  Encouraging land transfer or joint service in rural areas can accelerate agricultural mechanization. The negative effect of migration has been overstated in the process of policy making. As proved in the history of developed countries, agricultural productivity has significantly improved with the labor force leaving agriculture. Policy should permit land transfer to facilitate the labor force more into the non-agricultural industry.

## Appendix A4

Table A4.1 Estimation of the impact of rural-urban migration on yields using OLS estimates (control variables)

| Household controls | ln(yield) | number of migrants | ln(capital investment) |
|---|---|---|---|
| Average age of farmers | 0.000034 | | -0.0110*** |
| | (0.00123) | | (0.00219) |
| education of head (year) | 0.00178 | -0.0205*** | 0.00322 |
| | (0.00382) | (0.00772) | (0.00690) |
| education of head's partner | 0.000907 | -0.0276*** | -0.00252 |
| | (0.00329) | (0.00647) | (0.00589) |
| number of children | 0.0187 | 0.562*** | -0.00553 |
| | (0.0143) | (0.0239) | (0.0256) |
| household size (person) | 0.00867 | 0.147*** | -0.00636 |
| | (0.00580) | (0.0116) | (0.0106) |
| farm income (yuan) | | 0.00699 | |
| | | (0.0289) | |
| urban household income(yuan) | | 0.153 | |
| | | (0.206) | |
| rural household income | | -0.152*** | 0.142*** |
| | | (0.0253) | (0.0215) |
| village controls: | | | |
| near trade area | 0.000934 | -0.0836* | |
| | (0.0241) | (0.0492) | |
| labor share in agriculture | -0.0018*** | -0.000078 | 0.00796*** |
| | (0.000465) | (0.000947) | (0.000823) |
| Time indicators | Yes | Yes | Yes |
| Province indicators | Yes | Yes | Yes |
| N | 2,756 | 2,756 | 2,756 |
| Adj. R-sq | 0.3676 | 0.4299 | 0.1933 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Standard errors are presented in parentheses. The fix effects of province and time are included in all equations but not reported in this table. The variables associated with income are calculated by logarithm.

Table A4.2 3SLS estimates of the impact of rural-urban migration on agricultural productivity

| | Dependent variable | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Independent variable | ln(yield) | number of migrants | ln(capital investment) |
| number of migrants | -0.139*** | | -2.806*** |
| | (0.0331) | | (0.629) |
| number of migrants*farm size | | | 0.0832*** |
| | | | (0.0326) |
| farm size | -0.00543*** | | -0.0850*** |
| | (0.00256) | | (0.0250) |
| ln(capital investment) | 1.267*** | | |
| | (0.0592) | | |
| Near train station | | 0.0639** | |
| | | (0.0284) | |
| collectively plowed land (%) | | | 0.00216*** |
| | | | (0.000858) |
| Control variables | | | |
| Average age of farmers | 0.0142*** | | 0.0136 |
| | (0.00248) | | (0.00998) |
| education of head of household | -0.0104 | -0.0195** | -0.0397** |
| | (0.00725) | (0.00769) | (0.0155) |
| education of head's partner | -0.00218 | -0.0299*** | -0.0437*** |
| | (0.00622) | (0.00643) | (0.0114) |
| number of children | 0.116*** | 0.575*** | 1.411*** |
| | (0.0326) | (0.0237) | (0.294) |
| household size | 0.0232** | 0.150*** | 0.353*** |
| | (0.0116) | (0.0116) | (0.0736) |
| ln(farm income) | | -0.168*** | |
| | | (0.0217) | |
| ln(urban household income) | | -0.438** | |
| | | (0.191) | |
| ln(rural household income) | | -0.0914*** | -0.220*** |
| near trade area | 0.107** | -0.0178 | |
| | (0.0435) | (0.0347) | |
| labor share in agriculture | -0.0099*** | 0.00069 | 0.00605*** |
| | (0.00098) | (0.00094) | (0.00177) |
| N | 2756 | 2756 | 2756 |
| Wald chi2 | 850.65 | 2187.30 | 1154.68 |
| P-value | 0.000 | 0.000 | 0.000 |

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Standard errors are presented in parentheses. The fix effects of province and time are included in all equations but not reported in this table. The variables associated with income are calculated by logarithm.

# 5 Rural-urban migration and the transformation of traditional agriculture: A case study of a rice village in southeast China[29]

## 5.1 Introduction

In the chapters above, I described the rural-urban migration occurring in China and analyzed its effects on population size and structure, education, and agricultural productivity. The present chapter will highlight my earlier description and analysis through a case study of a typical rural village exposed to rural-urban migration. This case study will emphasize the transformation of traditional agriculture (labor-intensive agriculture) into modern agriculture (capital-intensive agriculture) as a result of the changes in relative scarcity of production factors triggered by labor loss through rural-urban migration.

Schultz (1964, 176-77) stated that farmers, in an equilibrium of traditional agriculture, do not search for modern factors given the normal preferences and motives. His argument is based on the assumption that the traditional farmers do not change their normal preferences and motives. His fundamental solution is to invest in farmers' human capital, particularly in education (Schultz 1964, 176). However, according to Hayami's induced innovation theory (Hayami and Ruttan 1971; Hayami and Godo 2005, 16-20), I argue that instead of human capital investment, the traditional farmers could also change their normal preferences and motives as the relative price of factors change. Traditional farmers would search especially for a cheaper factor to substitute for it. They would also acquire a modern or new factor, as long as the relative prices of factors change substantially enough.

As Chinese agriculture has experienced modernization in the past few decades, more and more modern factors, particularly agricultural machinery, have been introduced into agriculture. This chapter will illustrate how this process of agricultural mechanization happened in a typical Chinese village with the process of out-migration of human capital in agriculture. This case study argued that a long-term decrease of the agricultural labor

[29] This village survey was supported by a grant from the Dr. Hermann Eiselen Doctoral Program of the Foundation Fiat Panis in 2011 and the Doctoral Scholarship from the Center for development research (ZEF) at university of Bonn in 2014. The author wish to thank the Foundation Fiat Panis and ZEF for their finance supports.

force due to rural-urban migration can trigger the transformation of traditional agriculture, including mechanization, consolidation and intensification

## 5.2 Theory of transforming traditional agriculture

The Mansholt Plan has put forward a controversial proposal to reform European agriculture in 1968, labor had been steadily migrating out of farming and there was a lot of concern regarding the aging of farmers. The median age of a farmer was over 57 years at that time. The key suggestion was to modernize production methods and enlarge the size of small farms (Bouman 2014). Rice farming in China also needs a Mansholt Plan, because rural-urban migration has accelerated the decrease in the young labor force and increased aging in agriculture. A fundamental solution to modernize Chinese agriculture is to decrease the number of farmers and let the remaining farmers cultivate more land.

To date, the process of agriculture modernization has not been given enough attention by academics and policy makers. According to Schultz (1964, 30), there are three conditions that maintain traditional agriculture in equilibrium. First, the state of the art remains constant. Second, the state of preferences and motives for holding and acquiring sources of income remain constant, Third, both of these states remain constant long enough for marginal preferences and motives for acquiring new agricultural factors of production approach zero. In this case, the traditional agricultural equilibrium prevents the introduction of new production factors into agriculture. If the price of an input changes, farmers can adjust their production accordingly to a new efficient method based on the inputs at their disposal. But this change of factor input should be permanent and big enough. Otherwise, even if productive factors change, e.g. a new irrigation facility, or a reduction in the cost of any agricultural input, the factor inputs can be temporarily adjusted in production. Over time, agriculture will return to the particular equilibrium of traditional agriculture (Schultz 1964, 33). This is why a new factor is so difficult to introduce into agricultural production. Schultz's solution to break this equilibrium and transfer traditional agriculture into a modern one is investment in human capital in rural areas.

In this section, I will argue that large-scale outmigration will break the traditional agricultural equilibrium and lead famers to modernize their agricultural production. It is well known that prices of productive factors are substantially determined based on scarcity. If relative scarcity clearly changes, relative prices of agricultural inputs, that is, capital, labor and land, must change correspondingly. As long as farmers are price-sensitive, they will change their production inputs in the directions the prices indicate. If input prices go up, they will reduce this input and vice versa. In the framework of Schultz's theory, farmers are regarded as rational, that is, price sensitive. A large-scale outmigration of labor will fundamentally change the relative scarcity of essential agricultural inputs, particularly the man-land ratio. The area of land per capita for the remaining labor force will clearly be enhanced with outmigration of labor. Prices for agricultural labor will increase while those for land and capital will relatively fall. Therefore, less and less labor will be used while more and more capital will be allocated into agriculture to substitute for labor, In this manner, labor-intensive traditional agricultural will be gradually transformed into a more capital-intensive one. In the case of my village study in the following analysis, it was transforming towards mechanized agriculture at the time of my survey. The transition from traditional agriculture described above can be properly explained by Hayami and Ruttan's induced technological and institutional innovation theory (Hayami and Ruttan 1971; Hayami and Godo 2005, 16-24). The theory of "induced technological innovation" stems from Hicks (1932) where he proposed that such kind of change in technology is induced by reducing production costs through substituting scarcer resources with abundant ones. Regardless of whether there is a competitive market, as assumed by neoclassical economists, this theory can be used to explain the change of production technology as long as the producers can recognize the relative factor scarcities, e.g. through shadow prices in non-markets or subsistence markets (Hayami and Godo 2005, 16).

Under the basic mechanism of induced innovation, the process of agricultural mechanization can be illustrated in Hayami and Ruttan's model in Figure 5.1. It depicts the process of a farm household changing from labor-intensive production to capital (machinery)-intensive production. The left Land-Labor (A-L) quadrant includes two basic productive factor inputs, land and labor, U is a isoquant of the meta-production function, which is the envelope of less elastic isoquants, such as $u_1$ and $u_2$, corresponding to two different types of machinery and technology. Isoquant U is assumed to be an existing concave innovation possibility curve. $P_1$ and $P_2$ are two different relative price ratios (or shadow price ratios) of land and labor that reflect the relative scarcity of these two factor inputs. Correspondently, $AL_1$ and $AL_2$ represent points of tangency of different isoquants and factor price ratios. The line to the right Land-Machinery Power (A-M) quadrant reflects a complementary relationship between land and a third factor input; machinery power. The points, $AM_1$ and $AM_2$ , are two kinds of different combinations of land and power.

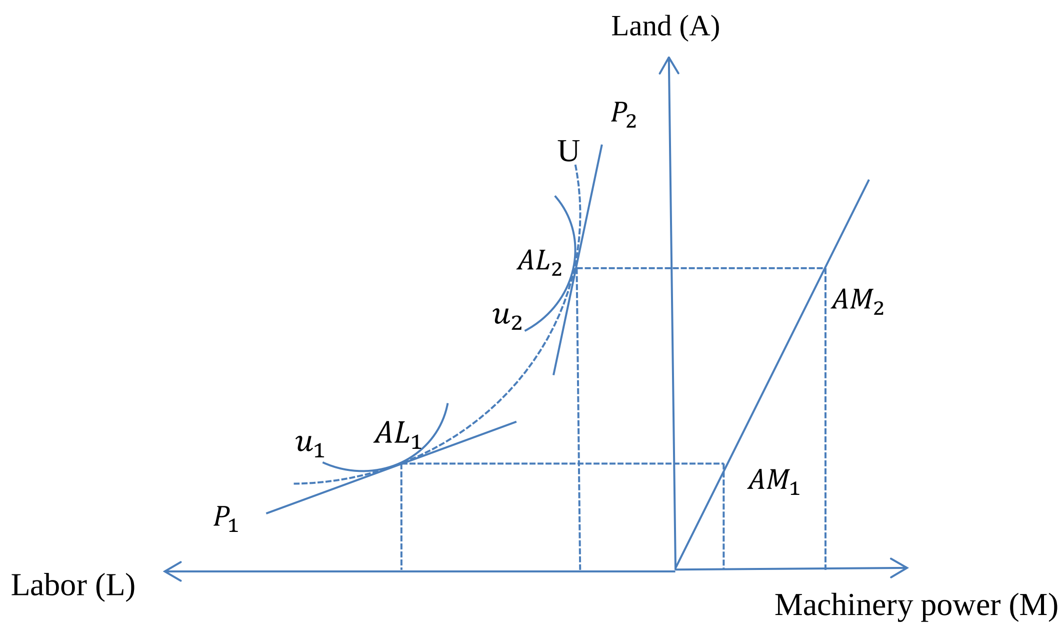



Figure 5.1 A model of induced agriculture mechanization for a farm household
Source: Hayami and Ruttan 1971, 17, Figure 3

With a decrease of labor force and an increase of land consolidation in the rural areas due to rural-urban migration, a household faces a decreased labor force and an increased cultivated area of arable land. To adapt to this change of scarcity of input factors, agriculture production changes from labor-intensive production $AL_1$ to labor-save production $AL_2$. It is noteworthy that it is not only the change of combination of input factors during the process of this change from $AL_1$ to $AL_2$. In the meantime, innovation is induced by this change of input factors. The larger area of land and less labor force in production demand for larger scale mechanized production. Therefore, as labor decrease from $AL_1$ to $AL_2$, the machinery power that complemented with area of land changes from $AM_1$ to $AM_2$.

In summary, most of farm households in the village can cultivate more areas of arable land by renting it from other rural-urban migrant households. The wage of farm labor rises with a decrease in the agricultural labor supply. Farmers start to use machinery to substitute the labor force. Finally, agricultural production changes from labor-intensive and a small scale system $AL_1$to a labor-save and relatively massive scale one $AL_2$. The whole process can be illustrated in Figure 5.1 and summarized as the substitution of a combination of land and machinery power for labor in response to a change to the ratio of labor and land (Hayami and Ruttan 1971). In what follows I will illustrate how rural-urban migration in the village I studied functions as the method to launch the substitution of a labor for a combination of land and machinery power.

## 5.3 Methodologies

The study of the Chinese village has not received enough attention. However, Hayami et al.'s research (Hayami et al. 1978; Hayami and Kikuchi 2000) established a good model of a village study, which will act as a reference in this village study. The methodology applied in the village study is mainly the survey through questionnaires. In addition, similar to Schultz's empirical analysis (1964, 44-52), The methodologies of anthropology will also be employed in this study. I will also carry out some in-depth interviews to investigate the economic motivation behind the behavior of the people questioned from the anthropological point of view (Fei 2001; Fei and Zhang 2006).

### 5.3.1 Criteria for village selection

The village in my study represents the main characteristics of a rural and agricultural community with out-migration in the area of southeastern China. According to the definitions proposed by Hayami and Kikuchi (2000, xvii) and Wolf (1957), a typical rural and agriculture village should meet the following criteria:

(1) Be modest in size, for instance, one hundred households (Hayami and Kikuchi 2000, xvii). But this criterion is not essential.

(2) Be clearly demarcated from other villages, particularly not exposed too much to urban activities, as Wolf (1957) pointed out, it should have clear geographic boundaries and be a "closed corporate community".
(3) Its main economic activity should be agriculture. There should be no or almost no industrial factories in the village or in the immediate neighborhood.[30]
(4) It should be a residential and agricultural community in which the most important social and cultural institutions are family, lineage, and neighborhood.[31]

Given that the objective of this research is focused on rural-urban migration and the transformation of traditional agriculture, there are two other criteria emphasized in the study in addition to the four criteria mentioned above.

(5) It has experienced massive rural-urban migration in the past few decades.

(6) It has a single-product (in particular, rice) agriculture. Because the purpose of this study is to study the relationship between agriculture and migration, single-product agriculture could help to clarify the problem and avoid other irrelevant influences.

**5.3.2 Selection of the village**

The village I chose, Yangyi, is located in Ningdu County, Jiangxi province, southeastern China (Point A in Figure 5.2). Jiangxi is close to three coastal provinces: Zhejiang, Fujian and Guangdong. These provinces were the first ones in China open up to foreign investors after 1978 and since then have received most migrant workers from inland provinces. Because of its geographical advantage, Jiangxi is one of the provinces with the highest out-migration rates in China.[32]

Yangyi village meets the essential criteria in section 5.3.1. First, it is "clearly demarcated from other villages" based on its unique geographical location. As shown in Figure 5.3 and Figure 5.4, Yangyi is located in a flood plain surrounded on three sides by

---

[30] Most of Chinese villages are in remote areas and difficult to develop local industries (Croll and Ping 1997).

[31] Scott (1976) stated a village should be built on this traditional premise called "moral economy", a perspective regarding the behavior of peasants

[32] Some researchers have done similar case studies on rural-urban migration in this province before because of this. Murphy (2002) studied the return migration in 1997 and 1998 in this province. Kuiper (2005) analyzed the village model by combining the general equilibrium and household modeling in a village in the same province.

a meandering river. The river serves as a natural boundary for the village as shown in Figure 5.3. Most rice paddies are located either on the plains or on the other side of the river that surrounds the village. The village, the meandering river, and the rice paddies are embraced by huge crescent mountains, forming a basin where the village and its rice-paddies are located. A four-kilometer-long and three-meter-wide paved road connects the village with a provincial road as illustrated in Figure 5.3. The village is six kilometers away from the nearest township and fifty kilometers away from the nearest city. There are no regular transportation means, e.g. bus lines or railway, connecting the village with the nearest township. As such, motorbikes and bicycles are popular transportation tools used by villagers to go to local markets. Accordingly, the village is not immediately exposed to urban activities due to its remote location and poor transportation.


Figure 5.2 Location of Yangyi in China (point A)

Source: https://maps.google.com (access on April 7, 2014)

It should be emphasized in this study that Yangyi is a natural village rather than the official administrative village of Yangyi defined by the local government.[33] The

[33] The administrative village is also called Yangyi. The administrative structure of this village is as follows. The official governance structure of the village consists of a general secretary of the communist party in the village committee, a village head, an accountant. The general secretary owns the real power in the village and is appointed by the party committee of the township. Although the general secretary normally has a three-year term, generally they serve a very long term, e.g. the general secretary of Yangyi village has held this position for 7 years since 2004.

administrative village of Yangyi is a formal village-level division that includes 5 natural villages.

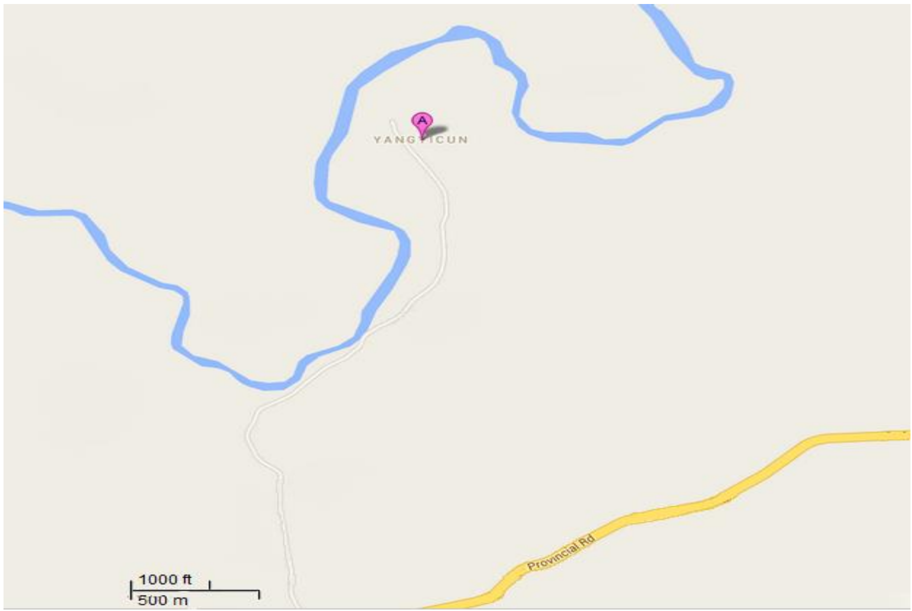



Figure 5.3 Geographical location of the village (point A)
Source: https://maps.google.com (access on April 7, 2014)


Figure 5.4 Satellite picture of the village (Central pin-point)
Source: access on https://www.google.com/earth/ on Feb. 6, 2015 (taken on Oct. 15, 2014)
Note concerning the map: The central pin-point surrounded by the river is the natural village of Yangyi where my survey was conducted. The four small villages (villages 1, 2, 3 and 4) scattered around the central pin-point are excluded from my survey, although they are affiliated to the administrative village of Yangyi.

Most of the population is concentrated in the natural village of Yangyi, which consists of ten village groups. The other four natural villages are very small and consist of only one village group each. Geographically, as shown in Figure 5.4 the natural village of Yangyi is located on the flood plain and forms a clearly demarcated agricultural community that is surrounded by the river on three sides. Four other villager groups are scattered outside the flood plain. These four small natural villages, each consisting of a single villager group, are excluded from this survey. However, I also found some villagers from the other four natural villages who lived in the natural village of Yangyi in recent years, although their farm land and other economic relations were still in other natural villages. I do not include these households. Two exceptions are households from other natural villages which are included in the survey because they rent a lot of farmland in the village of Yangyi. The natural village of Yangyi, rather than the administrative village of Yangyi, is more suitable to be considered as an independent rural community. In the following analysis, Yangyi is used to represent the natural village of Yangyi rather than the administrative village of Yangyi, unless specified otherwise.

The administrative village of Yangyi with a population of about 1,350 can be regarded as a modest size village of the mountainous areas in Jiangxi Province. According to Jiangxi Statistical Yearbook the average population of an administrative village was 1,412 persons in Jiangxi province in 2011.[34] This village size fulfills the first selection criterion of the sample village, i.e. modest in size.

The economy in Yangyi is predominantly rice production based. In fact, the village is located in one of the most important rice production regions in China. All the farmers in the village only plant rice, except for one farmer who plants 50 mu (3.3 hectare) of tobacco. In addition, there is no factory in the village. This criterion is relevant given that the objective of this case study is to determine how a village with traditional agriculture adapts to rural-urban migration, rather than examine the industrialization of the village as other studies have done (Liang et al. 2002).

Yangyi is an isolated rural community and dates back to over 1,173 years. Historical books and genealogical records indicate that ancestors emigrated from northern China in

[34] This number is available at statistic Bureau of Jiangxi on http://www.jxstj.gov.cn/Index.shtml (access on January 20, 2015)

874-879AD, in the late Tang dynasty. An obvious piece of evidence of the village's isolation lies in the fact that almost all of the household heads have the same surname, although their wives are mostly from other villages. Strictly speaking, all villagers are relatives with intricate relationships of family and clan to some degree. Some of the villagers are connected by a distant ancestor several generations before. There are more than fifty ancestral halls to remember their different direct ancestors in the village. Therefore, Yangyi is a typical Chinese closed corporate community in social and anthropological perspectives.

Although the surveyed village is difficult to access from the outside, its population has overcome transportation barriers and, to a large extent, migrated out from the 1980s onward. There are no historical data on migration in the village; however, Ningdu County where the village is located is so famous for its massive out-migration that it was twice rewarded the title of "National Model for the Transfer of Agricultural Labor Forces" by China's central government in 2007 and 2010 (Li and Zhang 2013).[35] This title is awarded by China's Ministry of Labor and Social Security to commend the county who has made distinguished progress in transferring the agricultural labor force from rural areas. In 2012, there were 345,000 workers in the county and at least 150,000 workers were migrant workers working outside their home county (Li and Zhang 2013).

### 5.3.3 Unit of land area reflects the equilibrium of agriculture

Yangyi's traditional agriculture has long been in equilibrium, as defined by Schultz (1964, 29-30). There is interesting evidence to prove it. Villagers use the output of the paddy field to measure the area of the paddy field instead of a unit of area. The official Chinese unit of area is a "mu", which equals 1/15 hectare. However, in the survey, the villagers use "one load of rice" as the unit of area. A load is a traditional Chinese food container which holds about 50 kg of rice. Every piece of land in the village is labeled by how many "loads of rice" it produces. The label for each piece of land was given by their ancestors. Nobody in the village knows exactly who generated this label for these areas from the time of their grandparents.

[35] The list of the counties rewarded this title is available on the website of Ministry of Human Resources and Social Security of the People's Republic of China on http://w1.mohrss.gov.cn/gb/zxwj/2007-04/18/content_173830.htm (access on January 20, 2015).

As we known, the same piece of paddy field can produce different amounts of rice depending on the field management, agricultural technology, soil quality, irrigation conditions and climate. Therefore, using a certain amount of production output to measure the area of a piece of land should be inaccurate. However, out of all the factors which could influence the final output of a certain area of land, only field management and technology are controllable by farmers. It is reasonable to use output to measure the area of land as long as agricultural technology does not change for a long time and all the factor inputs have reached the most efficient allocation. In this case, no matter who does the cultivation, the land cannot greatly change its output much. Using output as a unit of area can guarantee the fairness of land distribution when the quality of land varies a lot. This method of area measurement is also used in other provinces in China, e.g. Hunan (Fei and Zhang 2006, 25).

According to the village's official conversion formula, one mu equals four loads of rice. This is the conversion formula I used in the survey and the following analysis to convert the load of rice into mu. This indicates that one mu of land could produce 200 kg of rice (four loads of rice, each being 50 kg). However, this conversion formula is not accurate, because nowadays, one mu of land can produce an average of 400 kg of rice. Since no farmer in the village measured their farmland to check the accuracy of the village's conversion formula, I chose two pieces of rectangular farmland from different locations. Both pieces of farmlands are claimed by the land owners to be "ten loads of rice". The actual areas of the two pieces of farm land are 1,427 and 1,282 $m^2$, respectively. Therefore, the average area of "ten loads of rice of land" is 1,354 $m^2$ (measured by the mean of 1,427 and 1,282). As one mu of land equals 667 $m^2$, the "ten loads of rice of land" actually refers to 2.02 mu, based on my measurements. The true conversion formula should be one mu equals five loads of rice instead of four loads of rice. However, I still use the village's conversion formula in the following analysis, rather than adjust all the data based on my own measurement, because I have only measured two pieces of land and the validity of my measurement is not enough. Nevertheless, a lot of farmers felt my formula "one mu equals five loads of rice" was more plausible according to their experience of farming production.

### 5.3.4 Method of the survey

The village survey was conducted first by the author in August and September, 2011. But I revisited some of the farmers, who cultivate large amounts of land in February, 2015. The method of survey is a combination of questionnaires and in-depth interviews. In the survey, migrant households can be defined and divided into three categories as follows:

(1) Family migrant household: a household in which the entire family had migrated out to the urban areas for the whole year before my survey.

(2) Individual migrant household: one or more members of the household had migrated out to the urban areas for the whole year before my survey, but not the entire family.[36]

(3) Non-migrant household: the household with its all members has remained in the village

It is difficult to interview the households in which the entire family migrated out. The usual method is to delete those migrant samples due to the lack of direct respondents, as was done for a similar village survey in the same province by Kuiper (2005, 68). There are no data available for the migrated households in her village survey. To get the information on family migrant households, I collected indirect information from their neighbors and village officials. In general the officials in the village, especially the heads of the villager groups, are familiar with information about the household characteristics of their groups since the households of a villager group live very close to one another in a small "district" of the village. Information on the remaining households was collected through interviews with household members still at home. The respondents were heads of households or members who are responsible for their family affairs. The survey contains information about demographics and agricultural production.

## 5.4 Rural-urban migration and rice production

### 5.4.1 Demography of total population

As shown in Table 5.1, Yangyi has 232 households with 911 people in total. The male population has 49 more persons than the female population. The average household size is 4.39 people.

---

[36] A common situation is that the parents migrate into urban areas as migrant workers while their children stay at home in the village with their grandparents.

Table 5.1 Demographics of the population in the Yangyi village

| | Households | Population | Labor forces | Household size | Male | Female |
|---|---|---|---|---|---|---|
| Total | 232 | 911 | 659 | 4.39 | 480 | 431 |

Source: author's survey data, 2011

The labor force includes 659 persons according to China's official definition of labor force. The population between the ages 15-59 is classified as labor force according to the definition given by the Chinese Statistics Bureau. However, in reality most farmers still work in agriculture after age 59.

I cannot calculate the natural growth rate through this one-time survey because growth rate is a flow variable. I obtained the data of births and deaths in Yangyi for the period of January 1 to December 31, 2011 from the local government department of household registration. With 13 births and 8 deaths in Yangyi during 2011, I can calculate the natural growth rate at 5.4 per thousand in the year 2011, which was higher than the national average of 4.79 per thousand in the same year (NBSC 2012).

Figure 5.5 is a population pyramid of the total population, which shows that the young labor force aged 20-50 would be the largest section of population among all cohorts if they had not migrated out. The population pyramid is a constrictive pyramid. It shows that the birth rate is reducing and the population will decrease in the future.

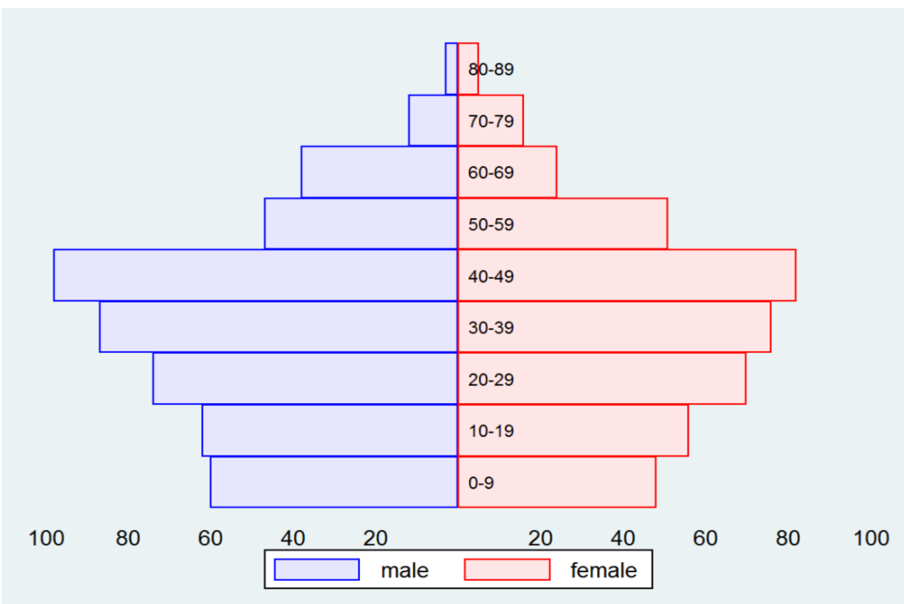



Figure 5.5 Population pyramid of total population in the village (persons)

### 5.4.2 Demography of stayers and migrants

As shown in Table 5.2. The most common household type in the village is the individual migrant household which alone accounts for half of all households (50%) in the village. The second most common type is the family migrant household with 35%. Only 15% of households are non-migrant households. The average family migrant household size is 3.9 persons, which is not so different from in non-migrant households with 3.5 persons, but almost one less person than in individual migrant households with 4.8 persons.

Table 5.2 Characteristics of migration in Yangyi village

| Category of migrant household | households | population | percent | house size |
|---|---|---|---|---|
| Family migrant household | 82 | 298 | 35% | 3.9 |
| Individual migrant household | 115 | 510 | 50% | 4.8 |
| Non-migrant household | 35 | 103 | 15% | 3.5 |
| Total | 232 | 911 | 100% | 4.39 |

Source: author's survey data, 2011

It is obvious that those who stay in the village are mostly the elderly and children. Figure 5.6 illustrates the population pyramid of stayers. Most remaining people are between the ages 0-19 and 40-69. The age group 20-29 is very small in the village and the age group 30-39 is also poorly represented. The ageing problem in the village is very serious. Figure 5.6 also shows that the agricultural labor force primarily consists of people over 40 years old. In fact, during the survey period it was difficult to find a laborer between the ages of 20-29 in the village. The trend in Yangyi can be found in most rural villages in China. According to the national survey by the development research center under the Chinese State Council in 2006, which includes observations from 1,300 villages in 13 provinces (Zhang and Li 2007), sixty percent of the young labor force under 30 have left the villages as migrant workers. Seventy-four percent of the village heads said that almost all members of the young labor force have migrated out of the villages. Many young children, up to age 10, are left in the village to live with their grandparents. They are at a disadvantage to get parental care.

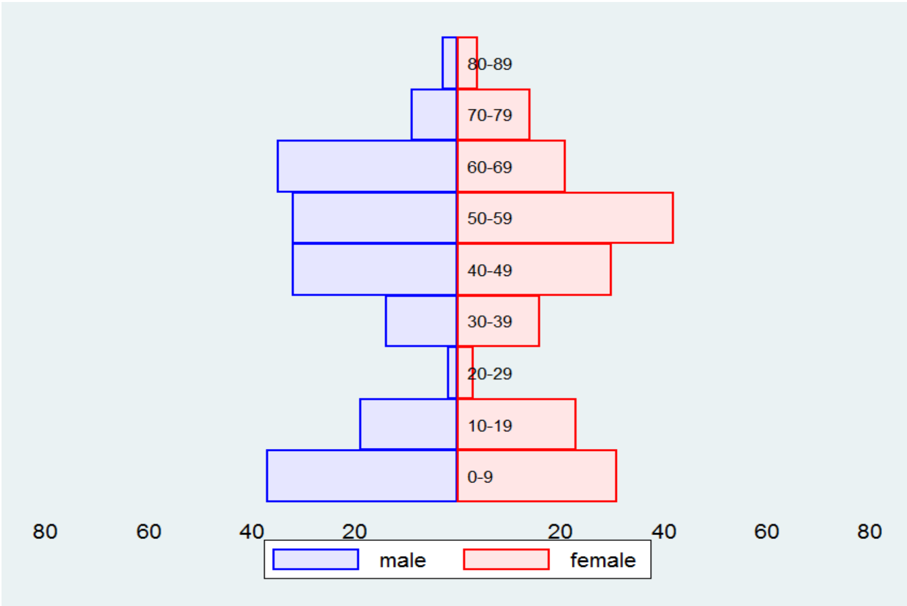



Figure 5.6 Population pyramid of stayers in the village (persons)

Figure 5.7 describes the population pyramid of migrants. As expected, the migrants are mostly between the ages 20-50. Almost all the young people in the labor age left the village. It is considered incomprehensible by the villagers that young adults around age 20 stay in the village as a farmer. In their opinion, being a migrant worker is a natural selection for the young adults. Respondents in the village often reiterate that there are no young people ready to be farmers left behind in the village.

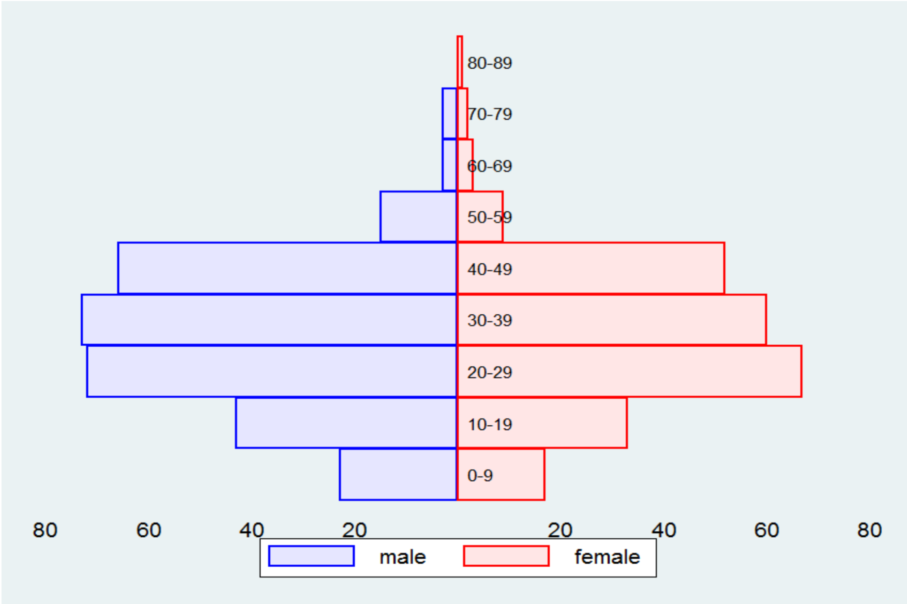



Figure 5.7 Population pyramid of migrants in the village (persons)

In terms of educational attainment of the labor force, migrants have a much higher level of education compared to the stayers. I consider only those labor forces between the ages 15-60 because compulsory education (9 years) is completed at 15 years old. Table 5.3 illustrates clearly that non-migrants received much less education than migrants. In summary, there are 83 of non-migrants with a primary school education compared to 105 of migrants with a primary education, all illiterate population (9 persons) are all non-migrants. In contrast, 77 non-migrants have middle school education while 315 migrants do. Actually, there are only two villagers with a college degree living in the village. One is the village doctor who has graduated from medical school. The other one is the owner of the largest farm in the village with 188 mu land who got a college degree through the vocational training during his farm administration.

Table 5.3 Education level of migrants and non-migrants (persons)

| Education | Non-migrant | Migrant | Total |
|---|---|---|---|
| College | 2 | 28 | 30 |
| High | 10 | 31 | 41 |
| Middle | 77 | 315 | 392 |
| Primary | 83 | 105 | 188 |
| Illiterate | 9 | 0 | 8 |
| Total | 180 | 479 | 659 |

Source: author's survey data, 2011

The migration destinations of the villagers show strong path dependence relying on social networks in Table 5.4. This means that migrants from the same village are more likely to migrate to the cities where people from their village have migrated previously. As hinted above in this chapter, most of the migrants went to three coastal provinces (75.51%) because the village is not far from the three most developed coastal provinces, Fujian province, Guangdong province and Zhejiang province. My survey results show that the villagers have a special migrant network in three cities from these three provinces: Quanzhou, a coastal city in Fujian province, Jieyang and Shantou, two cities in Guangdong province. The migrants to these three cities account for 50.8 % of the total migrants from this village. Take Quanzhou as an example, a lot of migrant villagers are working there in the catering industry, some of which run their own restaurants in Quanzhou. This social network attracts more and more villagers to find a job in the catering industry in Quanzhou.

Table 5.4 Destination of migration

| Destination of migration | Persons | Percentage |
|---|---|---|
| Local township | 36 | 6.63% |
| Local county | 37 | 6.81% |
| Local city | 13 | 2.39% |
| Other cities in local province | 17 | 3.13% |
| Zhejiang province | 17 | 3.13% |
| Guangdong province | 69 | 12.71% |
| Fujian province | 324 | 59.67% |
| Other provinces | 11 | 2.03% |
| Vietnam | 3 | 0.55% |
| Unknown | 17 | 2.95% |
| Total | 543 | 100 % |
| Top three destination cities | | |
| Quanzhou,Fujian | 195 | 35.91% |
| Shantou, Guangdong | 53 | 9.76% |
| Jieyang, Guangdong | 28 | 5.16% |

### 5.4.3 Rice production

Yangyi is located in a main rice producing area in China. Based on the climate and especially precipitation, it is technically possible to plant and harvest rice two times a year known as periods “early season rice” and “late season rice”. When the labor force is not enough or the paddy field does not have good irrigation conditions, farmers cultivate single-season rice that has a much more flexible time frame in which to arrange farm work, as shown in Table 5.5. For single-season rice, sowing and transplanting can be arranged in May, and, consequently, harvest time could be in September. Table 5.5 illustrates the schedule of the three kinds of rice production. Early season rice and late season rice grows for 90-120 days while single-season rice for 150-170 days. Partly because of the length of growth, late season rice is qualitatively better than early season and late season rice. Farmers could plant early season rice and late season rice on the same plots of land during the same year. But they could also produce only single-season rice due to loss of labor force and land condition. Although farmers only produce rice in the village, they still have to consider constraint of factor inputs to arrange the optimum combination.

Table 5.5 Schedule of agricultural production in the village

| | Early season rice | Late season rice | Single-season rice |
|---|---|---|---|
| March | Sow | | |
| April | Plough and transplant | | |
| May | | | Sow |
| June | Field management | Sow | Plough and transplant |
| July | Harvest | Plough and Transplant | Field management |
| August | | Field management | |
| September | | | Harvest |
| October | | Harvest | |
| November | | | |

Source: author's survey data, 2011

## 5.5 Transformation of traditional agriculture as a result of labor loss

Traditional agriculture has certain strong resistors to any changes in the existing state of the art. Without a tremendous external force to break the momentum of traditional agricultural production, the transformation won't happen. This tremendous external force must have destructive and unrecoverable power to distract from the traditional trajectory. In the case of Yangyi, the changes in relative scarcity of the labor force allocated to rice production due to rural-urban migration can be considered such an external force to transform traditional agriculture. Lewis's assumption of unlimited supplies of labor (Lewis 1954) does not hold for China after experiencing a long-term rural-urban migration, on the contrary, in China, labor has become the most expensive input in agricultural production.

### 5.5.1 Supplies and demands of labor force

I show the total amount of labor available to agricultural production in Yangyi and its age structure in Table 5.6. Agricultural labor or a farmer is defined as the individual who engaged in agriculture in the year my survey was conducted. The youngest farmer was 20 years old and the oldest one was 67. The total number of farmers was 137. There are many farmers who lose their work ability partially due to diseases or poor physical conditions in the village. For instance, in my survey of 137 farmers, 111 of them claimed full work ability, and 26 of them lost partial work ability. Table 5.6 shows the age structure of farmers is aging; there are 103 farmers aged 40-69 years out of 137 farmers.

Table 5.6 Age structure of farmers in the village

| Age | Male | Female | Total |
|---|---|---|---|
| 20-29 | 0 | 2 | 2 |
| 30-39 | 9 | 8 | 17 |
| 40-49 | 21 | 23 | 44 |
| 50-59 | 20 | 22 | 42 |
| 60-69 | 20 | 7 | 27 |
| 70-80 | 1 | 1 | 2 |
| 80-90 | 1 | 2 | 3 |
| Total | 72 | 65 | 137 |

Source: author’s survey data, 2011

Statistics regarding the demands of the labor force are difficult to measure throughout the survey. However, the farmers who have to employ other laborers to cultivate their fields carried out detailed bookkeeping on the cost and labor of every farm activity. Fei (Fei and Zhang 2006, 29-40) supplies a method of estimation in his village study in China. He decomposes rice production into different farming activities and calculates the accurate demand of labor forces for each activity. He can then estimate an aggregate demand of labor force in the village by knowing the total area of farmland and farming activities needed for each piece of land. Fei’s estimation is more plausible than that achieved by asking each farmer to calculate how much labor they need for their farm production. I employed Fei’s method to estimate the demands of the labor force in my village study.

The labor needed for each farm activity on one mu of land is listed in Table 5.7. The data come from the accounting books of the farmers who record the labor employed in their farm production. I discussed the data with other experienced farmers to check the reliability of data.

Table 5.7 The working time of farm activities for one mu of farm land

| Farm activity | Farm season (days) | Production tools | Labor needed person/day/mu | Labor demand for the village (person) | labor shortage |
|---|---|---|---|---|---|
| Seedling | 30 | Manual | 2 | 84 | +52 |
| Plough | 30 | Farm cattle | 2 | 84 | +52 |
| | | Tiller | 1 | 42 | +95 |
| Transplant | 15 | Manual | 2 | 168 | -31 |
| | | Transplant machine* | 0.5 | 42 | +95 |
| Harvest | 15 | Manual | 2 | 168 | -31 |
| | | Combine | 0.2 | 16 | +121 |

Source: Author's survey, 2011

Note: * indicates that transplanting machines need to prepare special seedlings with a new sowing technology. A half-day here just refers to the operational time for the transplanting machine, and does not include the time of preparing special seedlings.

Demand for agricultural labor can be roughly estimated with cultivated areas of land in Table 5.7. There are 1,262 mu of arable land in the village. But demand for labor is concentrated in the period of transplanting and harvesting if they do not use machinery, i.e. half the month each, that is, the so-called rush period. For instance, in the harvesting months of July and October, demand for agricultural labor reaches 2,524 workers/day, which is calculated by 1,262 mu of arable land multiplied by 2 people/day needed per mu of land. On the other hand, assuming all agricultural laborers work full time during this period, they can demand 168 workers (2,524 workers/day divided by 15 days). The shortage of 31 persons in the labor supply (measured by 168 minus 137 of total labor supply) indicates that 31 laborers are needed during the rush period in the village.

In Yangyi, reliance on externally hired labor in agriculture during the rush period has increased as a result of increased out migration of labor into nonfarm employment opportunities. But the shortage in agricultural labor has been occurring not only in Yangyi but also in nearby villages. The wage of agricultural laborers has been increasing since the beginning of rural-urban migration. In the village, the wage of a laborer for transplanting has increased from 70 yuan per day to 138 yuan per day from 2010 to 2014, Table 5.8 reports the changes of wage in the village from 2010 to 2014. Wages suddenly rose to 138 yuan/day in 2014, because a highway started to be built near the village.

Highway building increases the demand for labor and pushes wages up to a very high position, which hints that there was a shortfall of labor in the village.

Table 5.8 Wage rising in the village

| year | 2010 | 2011 | 2012 | 2013 | 2014 |
|---|---|---|---|---|---|
| wage (yuan/day) | 70 | 76 | 97 | 117 | 138 |

Source: Author's survey, 2011 and 2015.

Note: Deflated into the price in 2010 with rural CPI.

Therefore, farmers in Yangyi need to develop a labor-saving farming strategy if they are to maintain or even extend their agricultural production, which inevitably leads to more and more capital-intensive agriculture by means of substitution of labor for machinery. For instance, if farmer use transplant machines to transplant seedlings or combine machines to harvest, the labor force is surplus for both farm activities in the village.

**5.5.2 Intensification**

Different from the traditional rice producers, modern rice producers use more fertilizer, more pesticides, more land, more machinery, less labor in rice production. To illustrate two different kinds of production, I conducted over 10 in-depth interviews throughout the survey period with experienced farmers who have produced rice for at least 10 years. Data on costs and revenues of rice production collected from these interviews are compiled in Table 5.9. The table includes all costs and revenues of planting either the early season rice or late season rice for one mu (0.07 hectare) land in 2011.

As shown in Table 5.9, even in the absence of a natural disaster, farmers' profit change substantially when yield estimates fluctuate about 10 percent. Take the last column of value in Table 5.9 as an instance. The profit for a household of the traditional rice production (the last column) is 119 yuan when the yield is 900 jin (450 kg) and 224 yuan when the yield is 1000 jin (500 kg). Similarly, the profit can fluctuate from 107 yuan to 232 yuan with modern rice production. The other important risk of farming is from price fluctuations of factor inputs, i.e. seeds, fertilizers and pesticides. However, the biggest increase came from the cost of labor. At least 10 yuan (USD 1.5) of wage has increased each year as shown in Table 5.8.

A modern farmer has a similar profit per mu of land as the traditional farmer, but the source of their profit is different from the traditional producer. The costs for land and labor have been considered in a modern farmer's cost analysis, while traditional farmers do not normally consider these costs. The profit for the modern farmers actually reflected the return on farmers' management, because each input factor has been deducted except for the farmers' management. By contrast, traditional farmers still count the return from their labor and land as profit.

It shows that traditional farmers don't consider such farm work like spraying pesticide, transporting rice, and drying rice to be labor costs in production. In their accounting, the cost of labor is used mainly for peak-season activities such as rice transplanting and harvesting. Both activities demand large amounts of labor over short periods. The family labor forces are not enough to finish these farm jobs in time. Besides, the output of the farm jobs, like transplanted areas and harvested quantities, are visible and easy to measure. Researches on farmers' behavior in other Asian countries, like the Philippines (Hayami and Kikuchi 2000, 167-168), have also found that family labor is used mainly for the tasks that require care and judgment without immediate visible outcomes. The task of family labor includes water and pest control, fertilizer application, seed-bed preparation and land preparation.

On the other hand, in recent years, Modern rice producer would count all of these as labor cost and they has increasingly been replaced by farm machinery. Modern factors are introduced into the production: modern farmers particularly use more fertilizers and pesticides to improve their output, while traditional farmers rely on intensive cultivation.

This cost analysis is important in order to understand the future direction of modern agriculture. The traditional farmers who increase their income by working more instead of employing external labor can only keep their life at a subsistence level, because, no matter how hardworking they are, the profit from a small area of land is limited. Modern farmers can only increase their income through enlarging the land area under cultivation.

Table 5.9 Cost-revenue accounting of planting early season rice and later season rice in 2011

| | Unit | Modern | | | Traditional | | |
|---|---|---|---|---|---|---|---|
| | | quantity | price | value | quantity | price | value |
| Rice seed | jin | 3 | 30 | 90 | 3 | 30 | 90 |
| Basal fertilizer | jin | 100 | 1.11 | 111 | 75 | 1.11 | 83 |
| Urea fertilizer | jin | 50 | 0.9 | 45 | 20 | 0.9 | 18 |
| Pesticide | time | 6 | 25 | 150 | 4 | 25 | 100 |
| Land rent | mu | 1 | 110 | 110 | | | |
| Total material cost | yuan | | | 506 | | | 291 |
| Plough(manual) | person/day | with tiller | | 160 | with ox | | 190 |
| Transplanting | person/day | 2 | 80 | 160 | 2 | 80 | 160 |
| Harvest | sack | with combine | | 84 | with labor | | 240 |
| Spraying pesticide | person/day | 1 | 80 | 80 | | | |
| Transporting rice | sack | 12 | 3 | 36 | | | |
| Drying rice | sack | 12 | 1 | 12 | | | |
| Total labor/machinery cost | yuan | | | 512 | | | 590 |
| Total cost | yuan | | | 1018 | | | 881 |
| Revenue ( if low yield) | jin | 900 | 1.25 | 1125 | 800 | | 1000 |
| Revenue ( if high yield) | jin | 1000 | 1.25 | 1250 | 900 | | 1125 |
| Profit( if low yield) | jin | | | 107 | | | 119 |
| Profit (if high yield) | jin | | | 232 | | | 244 |

Source: Author’s survey data, 2011, 2014

Note: (1) Mu is Chinese measurement unit for area, 1 mu=0.07 hectare. (2) Jin is Chinese measurement unit for weight, 1 jin=0.5 kilogram. (3) Person/day means a labor force working a day. (4) Yuan is the unit of Chinese currency. 1 yuan=USD 0.15 in 2011.

Let’s consider an average farming household’s annual income. The average land area per household is 5.4 mu (0.36 hectare). Since the expected income for traditional producer is 181 yuan per mu (mean of 119 and 244) and the expected income for modern producer is 170 yuan per mu (mean of 107 and 232), an average traditional household can generate 1,958 yuan per year while an average modern producer can have 1,831 yuan per year if farmers plant double season rice. For most households, the income from rice production is their sole income source.[37] It is impossible to satisfy the basic living condition for the four-person household with this income according to China's poverty line of 9,200 yuan.[38]

As discussed in the costs and revenues presented above, if the land transfer is impossible, the only possible way for traditional farmers to make more money is to reduce the cost of labor in production because the other factor costs are exogenous variables. Farmers can do nothing but accept them, e.g. they can't change the price of factor inputs and rice. However, they can save costs from labor. The general approach to reducing labor costs is working more in the peak season instead of elumploying external labor, which is known as "self-exploration" (Ellis 1993, 105-20). If they don't employ external labor, they can retain 1180 yuan of labor costs per year (590 yuan for one season, double season cost 1180 yuan). However, even when a household does all the farm work themselves, the annual income is only 3138 yuan per year (1958 plus 1180) which is still under the poverty line of China, 9,200 yuan. This is a typical situation of a traditional farming household. They work on the farm land all the day, but they get only very low returns from agriculture and live in subsistence.

The example above of farm production in the village shows that the farmers who practice traditional agriculture live in subsistence. Their income is so low that they struggle in the subsistence level. They can't afford any losses to their production. This property of strong risk averseness is a typical characteristic of traditional farmers which keeps traditional agriculture in a very stable state in many rural areas. It is often observed that the farmers refuse to change their productive behavior from low-productivity traditional production, like the village of Panajachel, Guatemala (Schultz 1964, 90-94) which had access to new profitable productive factors, but yet they denied using them.

### 5.5.3 Mechanization

As labor migrates out of the rural areas and wages for agricultural hired labor increase, substituting labor for machinery is the most straightforward approach to increase income

---

[37] Animal production is not common in this village. I did not collect data on animal production because their domestic animals are normally several chickens for their home consumption.

[38] In 2011, China's poverty line was 2,300 yuan (USD 345) per capita per year, which indicates 9,200 yuan (USD 1,380) per year for a four-person household. This poverty line has remained the same to date.

given that it is difficult for farmers to substitute other agricultural inputs, e.g., seeds, fertilizers, pesticides. Huang et al (2013) found that rising labor costs were the most important factor to drive mechanization based on the national survey.

Take maintaining domestic buffalo as an example. In Chinese traditional rice cultivation, a domestic buffalo is indispensable for plowing paddies. Only a decade ago, most households in Yangyi owned a buffalo for themselves. But in the year of my survey, out of 72 farm households, only 10 households still raised their own farm cattle and young calves to plough the paddies. Other households depend on machines for ploughing. The main reason of giving up farm cattle, according to the villagers, is that they cost too much time and labor to take care of them. Labor cost is so high that farmers raise young calves to sell as an income stream in order to compensate their labor cost on maintaining the farm cattle. It is also less efficient to plough a large acre of land with farm cattle compared with machines. However, the owners of these farm cattle are very old and do not know how to use machines like tiller or tractors. The youngest owner is 57 years old in 2011. Figure 5.8 shows pictures of farm cattle and agricultural machines in the village.

As the farmers in the survey emphasized that the cost for raising a buffalo is too expensive, I can calculate the cost of a buffalo and a mini-tiller to prove this intuition. A mini-tiller is one of the agricultural machines that can replace the function of farm cattle in ploughing paddies. Other common options with the function of ploughing are walking tractors and tiller. The reason why I chose a mini-tiller in the following analysis is that it is the cheapest machine. It is reasonable to assume that farmers choose an agricultural machine based on the principle of minimizing costs. If a farmer can choose between several different machines to do the same job, i.e. ploughing land, I assume that they choose the cheapest option.

Figure 5.8 Ploughing machine-Mini Tiller (left), buffalo (upper right), and combined harvester (low right)
Source: Photographs taken by author, Yangyi, 2011.

In 2010, the general price of a buffalo in the village was 4,000 yuan compared to the average price of 2,054 yuan for a mini-tiller. Table 5.10 compares the cost of a buffalo and a ploughing machine. A buffalo can serve for 15 years normally while a mini-tiller can only work for 5 years. The cost of a barn for the farm cattle is 1,000 yuan in total. If I use the depreciation-straight line method and don't consider the inflation rate, the annual depreciation cost of a buffalo is 333 yuan, which is calculated from the total fixed cost (5,000) divided by 15 years. The annual depreciation cost for a machine is 410 yuan (2,054/5).

Table 5.10 The annual cost of farm cattle and mini-tiller for an average household

| | Purchasing cattle | Purchasing mini-tiller | Renting buffalo | Renting machine |
|---|---|---|---|---|
| Annual fixed costs | | | | |
| annual depreciation value | 333 | 410 | | |
| annual maintenance | | 100 | | |
| annual raising labor cost | 5,840 | | | |
| Variable costs | | | | |
| diesel | | 81 | | |
| labor cost | 800 | 160 | | |
| Renting cost | | | 600 | 432 |
| Total | 6973 | 751 | 600 | 432 |

Source: author's survey data, 2011
Note: the household owns 4.3 mu (0.3 hectare) of average area of land in the village. Generally, working life of a buffalo is 15 years and that of a ploughing machine is 5 years.

The more important cost is the variable cost of raising a buffalo. The fodder crops are normally grass in the spring and summer, the stems of rice in the autumn and winter. Normally, these can be gotten free from rice harvest in village. So I don't calculate the cost of fodder crops. But it takes at least two hours every day to feed and pature cattle, which is equal to 73 working days per year if a work day is 10 hours. As I stated in Table 5.8, the wage of farm labor is 80 yuan per day on average in 2011, hence 73 days of work would cost 5,840 yuan. Let's consider a normal household with an average area of land, 5.4 mu. It takes 10 days for a buffalo to plough this 5.4 mu land. It would cost 800 yuan for labor and the total cost for the buffalo is 6,973 yuan per year.

On the other hand, a mini-tiller is also more efficient than a buffalo. Only two days is needed for a ploughing machine to plough 5.4 mu land. So the working speed of a machine is five times faster than that of a buffalo. The costs of the ploughing machine include the cost of diesel (15 yuan/mu) 81 yuan for 5.4 mu land. Maintenance costs are about 100 yuan per year and the labor cost for the driver is 80 yuan per day. So the total cost for the ploughing machine is 751 yuan per year. The above cost calculation shows that the annual cost of the buffalo is almost ten times higher than using a mini-tiler for an average household.

The other options for the farmer are to rent a machine or a buffalo to plough rather than to purchase it. The cost for renting a machine includes 432 yuan of the renting cost for 5.4 mu (80 yuan/mu). , totally cost 301 yuan. On the other hand, the cost for renting a buffalo is 600 yuan (60 yuan/ day). In total it takes 600 yuan for ploughing 5.4 mu of land by renting a buffalo. But the last case is very few. Either when the farmers are very old and do not know how to use mini-tillers, or when the paddy field is too deep to use machines, they rent a buffalo to plough.

After comparing the annual cost of these four options, owning a buffalo is a much more expensive way to plough. This is the reason why most farmers no longer want to raise a buffalo by themselves when they calculate their labor costs. Compared to renting a machine, renting a buffalo is obviously an unworthy option and inefficient.[39] Not only does it cost more but it takes five times longer to finish the farm work.

For purchasing or renting a machine, it would depend on a household's land size to decide to purchase a machine or to rent one. An ordinary household with 5.4 mu of land would prefer renting a machine because it costs less than what it would to purchase a machine. However, the difference is very small. If the household cultivates more land, they may opt to purchase a machine by themselves at some point of land size because the average cost of purchasing machine falls with the land size.
The cost functions of renting a machine (5.1) and purchasing machine (5.2).

$$Y = 80X, \tag{5.1}$$

$$Y = 510 + 45X. \tag{5.2}$$

As shown in Equations (5.1) and (5.2), purchasing a machine has a fixed cost of 510 yuan while renting a machine does not have any fixed cost. The variable cost for purchasing a machine is 45 yuan per mu while that for renting is 80 yuan per mu. The solution of these two cost functions is at 14.6 mu land. The cost of renting a machine overtakes that of purchasing machine from the land scale of 14.6 mu. Although the cost is not the only consideration for the farmer in decision making on renting or purchasing machines, it should be a main cause for the behavior of a rational individual. Therefore,

[39] There are still a few farmers raising or renting water buffalo because some paddy fields are contoured to the hilly terrain and not suitable to plough with machinery.

this cost analysis indicates that 14.6 mu land is the transition point of land size for farmers from renting machinery to purchasing ones.

In the survey, I found that many households owned different machinery to fit different farm operations. The machines listed in Table 5.11 are all in use in the corresponding work seasons. Table 5.11 shows the type, number, price, actual payment and subsidy of all the machinery in the village. Here the actual payment is the actual money farmers paid and excludes subsidies from the government. For instance, a rice mill is 857 yuan, which does not include the subsidy of 250 yuan because the subsidy is given to the dealers directly from the government department, namely the bureau of county agricultural machinery. Especially the threshing machine, the tiller and the mini-tiller used for threshing and ploughing are very popular with the villagers. As I discussed before, when the cost of labor is high, farming is efficient with a combination of several machines according to the land size.

Table 5.11 Type, number, mean of price and subsidy of machinery in 2011

| Type | Number | Price | Actual payment | Subsidy | Subsidy ratio |
|---|---|---|---|---|---|
| Rice Mill | 6 | 1,107 | 857 | 250 | 0.23 |
| Feed mill | 1 | 1,180 | 680 | 500 | 0.42 |
| Irrigator | 1 | 2,100 | 1,600 | 500 | 0.24 |
| Ridger | 2 | 2,400 | 1,700 | 700 | 0.29 |
| Sprayer | 6 | 405 | 255 | 150 | 0.37 |
| Threshing machine | 21 | 1,070 | 665 | 405 | 0.38 |
| Centrifugal pump | 1 | 2,000 | 1,400 | 600 | 0.30 |
| Walking tractor | 4 | 5,800 | 4,125 | 1,675 | 0.29 |
| Mini-tiller | 14 | 2,769 | 2,054 | 714 | 0.26 |
| Tiller | 33 | 4,663 | 3,278 | 1,385 | 0.30 |
| Transplanting machine | 4 | 17,300 | 10,050 | 7,250 | 0.42 |
| Combined harvester* | 4 | 43.667 | 34,666 | 9,000 | 0.20 |
| Total cost | 107 | 467,667 | 360,011 | 141,306 | 0.30 |

Source: author’s survey data, 2011

Note: one of the combined harvesters is a second hand machine with 10,000 yuan and is not included into the calculation.

As shown in Table 5.11. The total cost for machinery is 360,011 yuan. The government’s subsidies take up 30% (141,306 yuan). Subsidies are directly distributed to the dealers. The subsidy for machinery which farmers receive is significantly positively related to the total cost of their machinery. It may seem self-explanatory, but other agricultural subsidies in the village do not have a positive relationship with the cost of farm production. The machinery subsidy is the only one with this positive relationship.

I do not have data on the capital source. However, According to survey responses, it is common for the household to pay dealers a down payment at the first time which normally makes up one third of the total price. These funds come from farmer savings. The rest expense of the machinery comes from their relatives or debt to dealers. They may pay off the debts to dealers at the end of winter because their cash flow closely related with crop season. In winter, they have more money after selling their harvest. No household can borrow credit from the bank except the farmer who owns 188 mu of land. He borrowed 20,000 from the bank to purchase a combined machine.

I revisited some of farmers in the village in February, 2015. According to this survey, twenty new farming machines had been purchased during 2011–2014; the machines purchased most were tillers (11), while the number of farm cattle had decreased from 10 head in 2011 to 9 head in 2015. Four new combine harvesters and a transplanting machine were purchased by the farmers who cultivated a large amount of land. This indicates that the farmers working on a large area are capable of increasing their investment in machinery. The trend of modern agriculture is that the number of farmers will decrease and the farmers who farm on a large-scale will invest more in agricultural machinery.

### 5.5.4 Consolidation

Newly introduced agricultural machinery will not be used efficiently if the average area of arable land remains very small. With massive labor outmigration a number of paddies have been transferred to other remaining farmers in Yangyi since its users have moved far away from the home village in search of employment.

In China, farm households only have the use-right of the land distributed to them for cultivation. The land owner is the “collective” in the form of villager group. However, individual farmers can rent out the use-right of their land to others as they wish. The land transfer enables the remaining farmers to expand their land. The price of land transfer in the village is 50 kg of rice per person or 120 yuan per mu in nine villager groups out of ten. These two prices are equivalent, because the value of 50 kg of rice is equal to 120 yuan, and a person can be distributed about one mu land from the collective of village. The price of land transfer is different in one village group whose paddy field is too deep to operate with a machine. Therefore, their price is 25 kg of rice per person, which is equal to only half of the normal price.

In Yangyi, transfer of land initially proceeded completely spontaneously. Migrant families were ready to let their idle land be cultivated by their close relatives in the village first as nobody seemed to be interested in renting them. However, as machinery was introduced by a few farmers, the demand for land increased and rent paid. The rent for farm land was first in form of a certain quantity of grain, and later of cash.

Land demand and supply refer to the land farmers want to rent in and rent out respectively. Table 5.12 shows the land transfer market in the year of my survey. At the time, the village had 232 households and the mean area of land was 5.4 mu. The land-demanding households got 19.4 mu on average through land transfer, which is almost 4 times more than the original size of a household's land. The largest rent-in happened to a family who had rented 180 mu of land from other households. He was the biggest rice farmer in Yangyi who owned a combined harvester and four tractors. He was the only farmer in the village who got loans from the state banks to purchase the heavy machineries due to his large scale of cultivated land. However, the total area of his land is 188 mu, but 80 mu land is outside the area of Yangyi village defined in this study. I have separated those 80 mu land from Yangyi to keep the figures for the village accurate for the study. The area of farmland rented by this farmer in Yangyi village is 100 mu; That is the number I have used as the basis to analyze the land circulation in the village.

Note that the land demand is 19 mu larger than the land supply. This number can be counted as the error of measurement during the survey. In most cases, farmers do not have a clear concept of the area of their farmland, as their land is not measured by units of mu. They only know how many people in their households get land from the collective. They are not familiar with using mu as the unit of measurement. When they convert their unit of area, namely loads of rice, to mu, their conversation is not very accurate.

Table 5.12 Land demand and supply in the village (mu)

| Variable | Households | Mean area | Min | Max | Total area |
|---|---|---|---|---|---|
| total land | 232 | 5.4 | 1.4 | 17.3 | 1,262 |
| land demand | 44 | 19.4 | 2 | 100 | 851.8 |
| land supply | 160 | 5.2 | 1.4 | 12.2 | 832.2 |

Source: author's survey data, 2011
Note: the unit of land is mu, 1 mu=0.07 hectare.

The farming households can be defined into three kinds in the village. The first is the big farming household who cultivates not only their own land but also others' lands. The second is called as small farming household who cultivate only their own land. The third is known as non-farming household who rent their land out and don't cultivate land. Table 5.13 shows that there were 44 big farming households who rented land from others through land transfer. These households had 25.4 mu land on average and cultivated 1,117.2 mu of land in total. There are 28 small farming households planting rice on their own land in the village. Average area of own arable land per household was 5.9 mu. They cultivate 164.9 mu of land in total. A hundred and sixty farming households don't cultivate land and become non-farming households. They supply 832.2 mu of land to the big farming households. Because of migration these non-farming households transfer their land to the big farming households and increase their average cultivated land fivefold. Table 5.13 indicates that the cultivated farm land for the big farming households in the village increased due to migration. This makes the machinery operation become possible in the traditional small farm.

Table 5.13 Category of farming households (mu)

| Category of farming household | Household | Mean of area | Min | Max | Total areas |
|---|---|---|---|---|---|
| Big farming household | 44 | 25.4 | 5 | 108.6 | 1,117.2 |
| Small farming household | 28 | 5.9 | 2.8 | 11.7 | 164.9 |
| Non-Farming household | 160 | 5.2 | 1.4 | 12.2 | 832.2 |

Source: author's survey data, 2011

The survey found that the farmers with big farms are increasing their farmland in 2015. For instance, the farmer with largest farm land is now renting 300 mu of land compared with 188 mu in 2011. Land consolidation is an inevitable trend of modern agriculture in this village.

### 5.5.5 Inefficient agricultural subsidies

The Chinese government abolished agricultural taxes and fees in 2003 and, by contrast, farmers can now get subsidies from the government. There are four types of subsidies, namely grain subsidy (in Chinese *liangshi butie*), input subsidy (*nonzi zonghe butie*), seed subsidy (*liangzhong butie*) and agricultural machinery subsidy (*nongjiju butie*). The main subsidies, grain subsidy and input subsidy, accounted for 82 % of the total farming subsidies in 2008 (Huang et al. 2011). Based on my survey, the first three subsidies – for grain input and seed – are mostly not granted to the farmers who are tilling, because these three subsidies are appropriated based on the registered area of contracted land, while most farmers with contracted land do not cultivate the land. My findings on the subsidization in China's agriculture are consistent with the conclusions of Huang et al.(2011, 2013) obtained from a nationwide set of household data. The tiller does not receive more subsidies than those contractors who are not tillers. The subsidies are non-distorting. With the exception of the machinery subsidy which is based on tillers, other grain, seed and input subsidies are distributed equally based on the population of the rural households.

The non-distorting subsidies cannot help to promote agricultural production. As the subsidies are based on the contracted areas which equal the registered population in every household, the subsidies actually work as income transfer programs for the rural population. No matter whether this population is involved in agricultural production or not, they can get these agricultural subsidies. An important implication from my survey for the policymakers is that an alternative subsidy policy would be to subsidize farmers according to the actual grain they have produced or give a subsidized purchase price. Both of these subsidy options benefit the grain producer directly based on their grain output, which indicates that the farmer who produces more grain can benefit more from the subsidies.

### 5.5.6 Institutional innovation on land tenure

Institutional innovation is to facilitate technological innovation by organize collective action (Hayami and Godo 2005, 21). My case study supplied a good example on how a farmer organized and coordinated the others to consolidate their land together.

Given that land in Yangyi was originally distributed to individual households equally, land fragmentation was an acute problem for land consolidation. According to the distribution procedure, all available arable land owned by a villager group is first classified into several grades of quality. Then the land of the same quality is distributed to individuals. Therefore, everyone has different pieces of land of every grade of quality in different locations. Through this procedure, a household is distributed several piece paddies, which are often not located in the adjacent field. Every household has 2-4 pieces of land from different locations. It is not enough for agricultural mechanization operation to get use-right of arable land from migrant households. If all these lands from different households are located in the very scattered places, it is still very difficult for efficient mechanized operations. In 2010, an institutional innovation came up spontaneously during the process of land consolidation that consolidated many scattered land into several adjacent large lands. As stated before, land fragmentation is an acute problem in this village. Every household has 3-4 pieces of land. Most of this land is in different locations. This could result in the high cost of mechanized operation.

The above-mentioned biggest farmer in this village came up with an unprecedented solution to land fragmentation. He consolidated these fragmented lands into several large adjacent pieces of land so that the cost for mechanization operations decreased tremendously. According to land regulations in the village, the land is reallocated to the villagers equally every 5 years. Taking advantage of this opportunity of land reallocation in 2010, he rented 100 mu of land from around 20 households in the village. He communicated with all of these households before the land reallocation and persuaded them to let him act as a single agent on behalf of all households to take part in this land reallocation. He ended up getting these 100 mu lands in three adjacent locations. One location has about 20 mu of land, and another two locations have about 40 mu of land respectively. He signed a contract with every household and paid them rent of 50 kg unhusked rice per mu land per year, worth about 120 yuan per year during 2010-2015. Using rice as the rent is better than cash because it can protect these small-scale land holders from the fluctuations in prices of unhusked rice every year. His innovation is very effective; I revisited him three years later in February, 2015. There is not one case of breach of contract during his five-year contract. The advantage of this method is that it saves a lot of transaction costs at the time point of land distribution. If the land has been distributed, the cost of consolidating these lands from every household is very expensive.

It is notable that this land locational consolidation happens voluntarily through communication among villagers, rather than from administrative guidance. This village doesn't even have a formal land market. However, through private negotiation they can decrease the transaction costs to a very low level.

**5.5.7 Complementary infrastructure investment**

Mechanized operations require some complementary infrastructure. These infrastructural investments come mostly from public investment. In Yangyi, these public investments are from the central government's direct financial allocation. In 2010, the county government spent about 2 million yuan in the project of farm land leveling in Yangyi. This project was to smooth and shape the field surface to level the field. Farm land leveling is considered to be a project to change the farm land into its best condition with minimal earth movement and variation in irrigation (Jat et al. 2009). The project in the village also included building roads to enable all agricultural machinery to arrive to most farm fields and irrigation ditches to improve the irrigation system. Figure 5.9 shows the pictures of the infrastructure construction of roads and irrigation ditches (left), land without leveling (upper right), and land after leveling (low right).

These infrastructural projects financed by the government provide important complementary platform for mechanized operations. It is necessary for agricultural modernization. But its function should not be overstated. We can imagine that the effect of these infrastructure projects actually relies on the degree of machinery usage and land consolidation. Without massive labor migration and land usage transfer, mechanization and land consolidation can't be achieved solely by public investment in infrastructure.

Figure 5.9 Roads and irrigation ditch (left), land without leveling (upper right) and land after leveling (low right)
Source: Photographs taken by author, 2011.

**5.5.8 Housing investment in the village**

I have discussed why farmers chose to purchase agricultural machinery as substitutes for labor due to out migration. However, as I stated in Chapter 4, investments in agriculture are not a priority for farm households. This conclusion can be proved in this village survey, housing investments instead of agricultural investments are what villagers generally choose to invest in with the money from migrant worker income. The survey did not collect data on villager's housing investments. However, as the pictures shown in Figure 5.10, under unified planning, since 2009 a new district has been built with more than sixty four-storey houses standing in the center of the village. Many new houses were been building during my survey in 2011. According to my interview, each of these houses cost between 200,000-300,000 yuan. The total housing investment mainly came from migrant wages and is over 12-18 million yuan.

Throughout the survey I found that most of these houses had been empty and owners only return to live a few days at Chinese Spring Festival. Compared to the total agricultural machinery investment in Table 5.11 with about 0.36 million yuan, while the investment in housing is around 12-18 million. Obviously, investments in housing are a priority for farm households. More and more migrant workers consider returning to their village as a farmer when they cannot enjoy the same social welfare or find a suitable job in the cities (Zhang 2006). This is the reason why the income from migrant workers is firstly invested in housing for the rural households.

Figure 5.10 A newly built district of private houses, Yangyi, 2011.
Source: Photographs taken by author, 2011.

**5.6 Conclusions**

This study emphasizes the important function of labor outmigration in the process of transforming into a modern agricultural system, even though the government still plays a significant role, such as through public investment in land leveling in my case study. In the context of Chinese villages, the function of government intervention is often overemphasized. Accordingly, this study aims to illustrate a possible approach towards modern agriculture triggered by labor outmigration.

The village study in this chapter illustrates the response of a typical agricultural village to the loss of labor force due to massive rural-urban migration. A decline of the agricultural labor force would be a challenge for agriculture in China in the long term.

This case study sheds light on a possible direction of Chinese agriculture. Rural-urban migration changes the relative price of factor inputs in agriculture. The allocative function of factor inputs has adjusted to the changing prices of the input factors, which result to the substitution of manual labor for machinery and land consolidation.

The process of transforming to modern agriculture is mainly the result of marketization of factors that function as price signals to induce technological advances (induced innovation). Marketization, such as integrating rural markets, disseminating economic information about products and factors, reducing imperfections in the capital market would be a major engine for the process of transforming traditional agriculture. In this village study, even without formal land and labor markets, there are the unified prices for farm labor and arable land. By the guidance of price signals, farmers are looking for an efficient way to allocate their resources.

With more and more farmers leaving agriculture through rural-urban migration, traditional agriculture has been experiencing the process of transformation through intensification, mechanization and consolidation. However, farmers make long-term investments in farmland only if land tenure is unambiguous and stable. The current ambiguous land tenure is a key issue to restrict a farmer’s agriculture investment. The agriculture subsidies, especially the grain subsidy, seed subsidy and input subsidy, subsidize the land contractors instead of the tillers, which also hampers the land consolidation. In addition to reducing the number of farmers through rural-urban migration, the reform of land tenure is an important priority to promote the agricultural transformation.

## References

Acosta, P. 2006. “Labor supply, school attendance, and remittances from international migration: The Case of El Salvador.” *World Bank Policy Research Working Paper* (3903).

Akay, A., O. Bargain, and K.F. Zimmermann. 2012. “Relative concerns of rural-to-urban migrants in China.” *Journal of Economic Behavior & Organization* 81(2):421–441.

Angrist, J.D. 2001. “Estimation of limited dependent variable models with dummy endogenous regressors: Simple Strategies for Empirical Practice.” *Journal of Business & Economic Statistics* 19(1):2–28.

Antman, F.M. 2011. “International migration and gender discrimination among children left behind.” *American Economic Review* 101(3):645–649.

Ashton, B., K. Hill, A. Piazza, and R. Zeitz. 1984. “Famine in China 1958-61.” *Population and Development Review* 10(4):613–645.

Au, C., and J.V. Henderson. 2006. “How migration restrictions limit agglomeration and productivity in China.” *Journal of Development Economics* 80(2):350–388.

Banister, J., and J.R. Taylor. 1989. “China: surplus labour and migration.” *Asia-Pacific Population Journal* 4(4):3–20.

Bao, S., Ö.B. Bodvarsson, J.W. Hou, and Y. Zhao. 2011. “The regulation of migration in a transition economy: China’s hukou system.” *Contemporary Economic Policy* 29(4):564–579.

Batista, C., A. Lacuesta, and P.C. Vicente. 2007. “Brain drain or brain gain? Micro evidence from an African success story.” *IZA Discussion Paper* No. 3035.

Batista, C., A. Lacuesta, and P.C. Vicente. 2012. “Testing the ‘brain gain’ hypothesis: Micro evidence from Cape Verde.” *Journal of Development Economics* 97(1):32–45.

Beine, M., F. Docquier, and C. Oden-Defoort. 2011. “A panel data analysis of the brain gain.” *World Development* 39(4):523–532.

Beine, M., F. Docquier, and H. Rapoport. 2001. “Brain drain and economic growth: theory and evidence.” *Journal of Development Economics* 64(1):275–289.

Beine, M., F. Docquier, and H. Rapoport. 2008. “Brain drain and human capital formation in developing countries: Winners and Losers.” *The Economic Journal* 118(528):631–652.

Beine, M., F. Docquier, and H. Rapoport. 2010. "On the robustness of brain gain estimates." *Annals of Economics and Statistics / Annales d'Économie et de Statistique* (97/98):143–165.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. "How much should we trust differences-in-differences estimates?" *The Quarterly Journal of Economics* 119(1):249–275.

Bhattacharyya, A., and E. Parker. 1999. "Labor productivity and migration in Chinese agriculture A stochastic frontier approach." *China Economic Review* 10(1):59–74.

Biao, X. 2007. "How far are the left-behind left behind? A preliminary study in rural China." *Population, Space and Place* 13(3):179–191.

Borjas, G.J. 2003. "The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market." *The Quarterly Journal of Economics* 118(4):1335–1374.

Brown, P.H., and A. Park. 2002. "Education and poverty in rural China." *Economics of Education Review* 21(6):523–541.

Bouman, B. 2014. "Modernizing Asian rice production." *Rice Today* 13(1):32–33.

Cai, F. 2010. "Demographic transition, demographic dividend, and Lewis turning point in China [Renkou Zhuanbian, Renkou Hongli Yu Liu Yisi Zhuanzhedian]." *Economic Research Journal.* 45(4):4–13 (in Chinese).

Cai, F., J. Giles, P. O'Keefe, and D. Wang. 2012. *The Elderly and Old Age Support in Rural China*. Washington, DC, The World Bank Press.

Cai, F., A. Park, and Y. Zhao. 2008. "The Chinese labor market in the reform Era." In L. Brandt, T. Rawsi (Eds.), *China's Great Economic Transformation.* New York and Cambridge, Cambridge University Press: 167–214.

Cai, Y. 2008. "An assessment of China's fertility level using the variable-r method." *Demography* 45(2):271–281.

Cai, Y. 2013. "China's new demographic reality: Learning from the 2010 census." *Population and Development Review* 39(3):371–396.

Cao, K.H., and J.A. Birchenall. 2013. "Agricultural productivity, structural change, and economic growth in post-reform China." *Journal of Development Economics* 104:165–180.

Card, D., and A. Krueger. 1990. "Does school quality matter? Returns to education and the characteristics of public schools in the United States." *National Bureau of Economic Research Working Paper Series* No. 3358.

Carrillo, B. 2004. “Rural-Urban migration in China: Temporary migrants in search of permanent settlement.” *PORTAL Journal of Multidisciplinary International Studies* 1(2).

Carter, C.A., J. Chen, and B. Chu. 2003. “Agricultural productivity growth in China: farm level versus aggregate measurement.” *China Economic Review* 14(1):53–71.

Chan, K. W. 2010. “A China paradox: Migrant labor shortage amidst rural labor supply abundance.” *Eurasian Geography and Economics* 51(4):513–530.

Chan, K. W. 2012. “Crossing the 50 percent population Rubicon: Can China urbanize to prosperity?” *Eurasian Geography and Economics* 53(1):63–86.

Chan, K.W., and L. Zhang. 1996. “The hukou system and rural-urban migration in China: processes and changes.” *Chinese Sociology & Anthropology* 29(1):15–26.

Chen, X., and X. Zhang. 2009. “The distribution of income and well-being in rural China: a survey of panel data sets, studies and new directions.” *Munich Personal RePEc Archive Papers* No. 20587.

Chiquiar, D., and G.H. Hanson. 2002. “International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States.” *National Bureau of Economic Research Working Paper Series* No. 9242.

Chi, W., R.B. Freeman, and H. Li. 2012. “Adjusting to really big changes: The labor market in China, 1989-2009.” *National Bureau of Economic Research Working Paper Series* No. 17721.

Coleman, D. 2008. “The demographic effects of international migration in Europe.” *Oxford Review of Economic Policy* 24(3):452–476.

Connelly, R., and Z. Zheng. 2003. “Determinants of school enrollment and completion of 10 to 18 year olds in China.” *Economics of Education Review* 22(4):379–388.

Croll, E.J., and H. Ping. 1997. “Migration for and against agriculture in eight Chinese villages.” *The China Quarterly* 149:128–146.

Crook, F.W. 1985. “The Baogan Daohu incentive system: translation and analysis of a model contract.” *The China Quarterly* 102:291–303.

De Brauw, A., and J. Giles. 2008. “Migrant opportunity and the educational attainment of youth in rural China.” *World Bank Policy Research Working Paper* (4526).

De Brauw, A., J. Huang, S. Rozelle, L. Zhang, and Y. Zhang. 2002. “The evolution of China’s rural labor markets during the reforms.” *Journal of Comparative Economics* 30(2):329–353.

De Brauw, A., and S. Rozelle. 2008. “Migration and household investment in rural China.” *China Economic Review* 19(2):320–335.

Deininger, K., S. Jin, and F. Xia. 2012. “Moving off the farm: Land institutions to facilitate structural transformation and agricultural productivity growth in China.” *World Bank Policy Research Working Paper No.* 5949.

Djajić, S. 2003. “Assimilation of immigrants: Implications for human capital accumulation of the second generation.” *Journal of Population Economics* 16(4):831–845.

Drèze, J., and A.K. Sen. 1989. *Hunger and Public Action*. Oxford, Clarendon Press.

Dustmann, C., and A. Glitz. 2011. “Migration and education.” *Handbook of the Economics of Education* 4:327–439.

Edwards, A.C., and M. Ureta. 2003. “International migration, remittances, and schooling: evidence from El Salvador.” *Journal of Development Economics* 72(2):429–461.

Ellis, F. 1993. *Peasant Economics: Farm Households in Agrarian Development. 2nd ed*. Cambridge, UK: Cambridge University Press.

Fan, C.S., and O. Stark. 2007. “The brain drain,‘educated unemployment’, human capital formation, and economic betterment1.” *Economics of Transition* 15(4):629–660.

Fan, S., and C. Chan-Kang. 2005. “Is small beautiful? Farm size, productivity, and poverty in Asian agriculture.” *Agricultural Economics* 32(S1):135–146.

Fan, S., and P.G. Pardey. 1997. “Research, productivity, and output growth in Chinese agriculture.” *Journal of Development Economics* 53(1):115–137.

Fei, H. 2001. *Peasant Life in China [Jiangcun Jingji-Zhongguo Nongmin De Shenghuo]*. Beijing: The Commercial Press (in Chinese).

Fei, H. and Z., Zhang. 2006. *Earthbound in China [Yunnan Sancun]*. Beijing: Social Sciences Academic Press (in Chinese).

Feng, W. 2013. “The silver and white economy: the Chinese demographic challenge.” Preliminary version of chapter 6, in forthcoming report "Fostering resilient economies: demographic transition in local labour markets." *OECD Local Economic and Employment Development|(LEED) Working Paper*.

Frisbie, P. 1975. “Illegal migration from Mexico to the United States: A longitudinal analysis.” *International Migration Review* 9(1):3–13.

Fu, Y., and S.A. Gabriel. 2012. “Labor migration, human capital agglomeration and regional development in China.” *Regional Science and Urban Economics* 42(3):473–484.

Giles, J., D. Wang, and C. Zhao. 2010."Can China's rural elderly count on support from adult children? Implications of rural-to-urban migration." *Journal of Population Ageing* 3(3-4):183–204.

Giulietti, C., J., Wahba, and K. F. Zimmermann. "Entrepreneurship of the left-behind." *Research in Labor Economics* 37 (2013): 65-92.

Glewwe, P., and H.G. Jacoby. 2004. "Economic growth and the demand for education: is there a wealth effect?" *Journal of Development Economics* 74(1):33–51.

Greene, W.H. 2012. *Econometric analysis. 7d ed.* Boston, MA: Prentice hall.

Hanson, G.H., and C. Woodruff. 2003. "Emigration and educational attainment in Mexico." *Documento de Trabajo del IR/PS. Disponible Discussion Papers*:1-38

Harris, J.R., and M.P. Todaro. 1970. "Migration, unemployment and development: A two-sector analysis." *The American Economic Review* 60(1):126–142.

Hayami, Yūjirō. 1976. *Anatomy of peasant economy*. Manila, Philippines: International Rice Research Institution (IRRI) Press.

Hayami, Y., and Y. Godo. 2005. *Development economics: from the poverty to the wealth of nations*. Oxford; New York: Oxford university press.

Hayami, Y., and M. Kikuchi. 2000. *A rice village saga: three decades of Green Revolution in the Philippines*. Basingstoke, England: Macmillan Press in Association with International Rice Research Institute.

Hayami, Y., and V.W. Ruttan. 1971. "Induced innovation and agricultural development." *Center for Economics Research University of Minnesota Discussion Paper*: No.3

Hess, P.N. 2013. *Economic Growth and Sustainable Development*. New York: Routledge.

Hicks, J.R. 1932. *The Theory of Wages*. London: Macmillan.

Hildebrandt, N., D.J. McKenzie, G. Esquivel, and E. Schargrodsky. 2005. "The Effects of Migration on Child Health in Mexico [with Comments]." *Econom ú* 6(1):257–289.

Holz, C.A. 2008. "China's economic growth 1978–2025: What we know today about china's economic growth tomorrow." *World Development* 36(10):1665–1691.

Huang, J., X. Wang, and S. Rozelle. 2013. "The subsidization of farming households in China's agriculture." *Food Policy* 41:124–132.

Huang, J., X. Wang, H. Zhi, Z. Huang, and S. Rozelle. 2011. "Subsidies and distortions in China's agriculture: evidence from producer-level data." *Australian Journal of Agricultural and Resource Economics* 55(1):53–71.

Huffman, W.E. 1980. “Farm and Off-Farm Work Decisions: The Role of Human Capital.” *The Review of Economics and Statistics* 62(1):14–23.

Huffman, W.E. 2001. “Chapter 7 Human capital: Education and agriculture.” In Bruce L.Gardner and Gordon C. Rausser, ed. *Handbook of Agricultural Economics*. Elsevier, 333–381.

Hussain, A. 2002. “Demographic Transition in China and its Implications.” *World Development* 30(10):1823–1834.

Ito, J. 2010. “Inter-regional difference of agricultural productivity in China: Distinction between biochemical and machinery technology.” *China Economic Review* 21(3):394–410.

Jat, M.L., M.K. Gathala, J.K. Ladha, Y.S. Saharawat, A.S. Jat, V. Kumar, S.K. Sharma, V. Kumar, and R. Gupta. 2009. “Evaluation of precision land leveling and double zero-till systems in the rice–wheat rotation: Water use, productivity, profitability and soil physical properties.” *Soil and Tillage Research* 105(1):112–121.

Jin, S., J. Huang, and S. Rozelle. 2010. “Agricultural productivity in China.” In Alston,J, Babcock,B, and Pardey,P (Eds.),*The Shifting Patterns of Agricultural Production and Productivity Worldwide.* Lowa State University, Ames, Lowa: Midwest Agribusiness Trade and Research Information Center (MATRIC), 229–277.

Joseph, A.E., and D.R. Phillips. 1999. “Ageing in rural China: Impacts of increasing diversity in family and community resources.” *Journal of Cross-Cultural Gerontology* 14(2):153–168.

Kamal-Chaoui, L., E. Leman, and Z. Rufei. 2009. "Urban Trends and Policy in China." *OECD Regional Development Working Papers* 2009/1.

Khandker, S.R., G.B. Koolwal, and H.A. Samad. 2010. *Handbook on impact evaluation: quantitative methods and practices*. Washington, DC: World Bank Publications.

Knight, J., and R. Gunatilaka. 2010. “Great expectations? The subjective well-being of rural–urban migrants in China.” *World Development* 38(1):113–124.

Kuiper, M.H. 2005. *Village modeling: a Chinese recipe for blending general equilibrium and household modeling*. Ph.D Thesis. Wageningen, Holland: Wageningen University. Available on: http://edepot.wur.nl/121621[Accessed August 21, 2014].

Kusek, J.Z., and R.C. Rist. 2004. *Ten Steps to a Results-based Monitoring and Evaluation System: A Handbook for Development Practitioners*. Washington, DC: World Bank Publications.

Le Bras, H. 1991. “Demographic impact of post-war migration in selected OECD countries.” *Migration: The Demographic Aspects. Organization for Economic Cooperation and Development*:15–28.

Leete, R., and I. Alam. 1993. *The Revolution in Asian fertility: dimensions, causes, and implications*. Oxford: Clarendon Press.

Lewis, W.A. 1954. “Economic development with unlimited supplies of labour.” *The manchester school* 22(2):139–191.

Li, J. 2013. “Private capital encouraged to invest in railways.” *Xinhuanet*. Available on: http://news.xinhuanet.com/english/china/2013-08/20/c_132645764.htm[Accessed August 21, 2014].

Li, G., F. Chao, and L. Fan. 2009. “Is the small-sized rural household more efficient? The empirical evidence from Hubei province. [Xiaonong Huzhende Gengjiajuyou Xiaolü Ma? Laizi hubeisheng de jingyan zhengju]." *China Economic Quarterly* 9(1): 95-124. (in Chinese).

Liang, Z. 2001. “The age of migration in China.” *Population and Development Review* 27(3):499–524.

Liang, Z., Y.P. Chen, and Y. Gu. 2002. “Rural industrialization and internal migration in China.” *Urban Studies* 39(12):2175–2187.

Liang, Z., and Z. Ma. 2004. “China’s floating population: new evidence from the 2000 census.” *Population and development review* 30(3):467–488.

Lin, J.Y. 1989. “Rural reforms and agricultural productivity growth in China.” *UCLA Working Paper* (576).

Lin, Y. 2004. “Why is it difficult to recruit workers in the province of labor surplus? [Laoli Fusheng Yanhe Chuxian Zhaogongnan?].” *Xinhuanet*. Available on: http://news.xinhuanet.com/focus/2004-06/18/content_1527269.htm[Accessed August 21, 2014]. (in Chinese).

Li, S. 2010. “The economic situation of rural migrant workers in China.” *China Perspectives* 4:4–15.

Li, W., and L. Zhang. 2013. “Based on human resources to improve the quality of employment- Survey and research on migrant workers in Ningdu county [Lizu Ziyuanyoushi Tisheng Jiuyeziliang-Dui Ningduxiang Nongming Waichuwugongxiangzhuande Diaochayuyuanjiu].” Jiangxi Agriculture [*JIANGXINONGYE*] (1):35–36. (in Chinese).

Loh, S., and M.V. George. 2007. “Projected population size and age structure for Canada and provinces: with and without international migration.” *Canadian Studies in Population* 34(2):103–127.

López-Córdova, J.E. 2006. “Globalization, Migration, and Development: The Role of Mexican Migrant Remittances.” *INTAL-ITD Working Paper* (20).

Lu, Y. 2012. "Education of Children Left Behind in Rural China." *Journal of Marriage and Family* 74(2):328–341.

Lu, Z., and S. Song. 2006. "Rural–urban migration and wage determination: The case of Tianjin, China." *China Economic Review* 17(3):337–345.

Marchiori, L., I.-L. Shen, and F. Docquier. 2013. "Brain Drain in Globalization: A General Equilibrium Analysis from the Sending Countries' Perspective." *Economic Inquiry* 51(2):1582–1602.

Mason, A., and W. Feng. 2005. "Demographic dividends and China's post reform economy." In *XXV IUSSP General Population Conference. Tours, France*. 18–23.

McKenzie, D., and H. Rapoport. 2011. "Can migration reduce educational attainment? Evidence from Mexico." *Journal of Population Economics* 24(4):1331–1358.

Mckenzie, D., and H. Rapoport. 2007. "Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico." *Journal of Development Economics* 84(1):1–24.

McKenzie, D., and M.J. Sasin. 2007. "Migration, remittances, poverty, and human capital: conceptual and empirical challenges." *Social Science Research Network*, No. 999482.

Mendola, M. 2008. "Migration and technological change in rural households: Complements or substitutes?" *Journal of Development Economics* 85(1–2):150–175.

Mi, H. 2008. "A statistical method on rural surplus labor force [Duinongcun Shengyu Laodongli Tongji Fangfa, Zhijie Jisuanfa De Xiuzheng]."*Population Research*. 32(6):76–83. (in Chinese)

Mishra, P. 2007. "Emigration and wages in source countries: Evidence from Mexico." *Journal of Development Economics* 82(1):180–199.

Mountford, A. 1997. "Can a brain drain be good for growth in the source economy?" *Journal of Development Economics* 53(2):287–303.

Mullan, K., P. Grosjean, and A. Kontoleon. 2011. "Land tenure arrangements and rural–urban migration in China." *World Development* 39(1):123–133.

Munshi, K. 2003. "Networks in the modern economy: Mexican migrants in the U. S. labor market." *The Quarterly Journal of Economics* 118(2):549–599.

Murphy, R. 2002. *How migrant labor is changing rural China.* Cambridge, UK: Cambridge University Press.

NBSC. 2011. *China Statistics Yearbook.* Beijing: China Statistics Press.

NBSC. 2012. *China Statistics Yearbook.* Beijing: China Statistics Press.

O'Neill, B.C., D. Balk, M. Brickman, and M. Ezra. 2001. "A guide to global population projections." *Demographic Research* 4:203–288.

Peng, X. 2011. "China's demographic history and future challenges." *Science* 333(6042):581–587.

Philipov, D., and J. Schuster. 2010. "Effect of migration on population size and age composition in Europe." *Vienna Institute of Demography, European Demographic Research Papers*,No. 2. Available on http://www.oeaw.ac.at/vid/download/edrp_2_10.pdf [Accessed August 21, 2014].

Qiao, X., and Y. Huang. 2013. "Floating populations across provinces in China-Analysis based on the sixth census." *Population and development* 19(1).

Ranis, G., and J.C.H. Fei. 1961. "A Theory of Economic Development." *The American Economic Review* 51(4):533–565.

Rapoport, H., and F. Docquier. 2006. "Chapter 17 The Economics of Migrants' Remittances." In Serge-Christophe Kolm and Jean Mercier Ythier, ed. *Handbook of the Economics of Giving, Altruism and Reciprocity*. Applications. Elsevier, 1135–1198.

Rong, Z., L. Yang, and Y. Yuan. 2012. "Labor migration choice and its impacts on households in rural China." In Annual Meeting of Agricultural and Applied Economics Association. Seattle, Washington.

Rozelle, S., J.E. Taylor, and A. de Brauw. 1999. "Migration, Remittances, and Agricultural Productivity in China." *The American Economic Review* 89(2):287–291.

Runsheng, D. 2010. "Narratives of Chinese Economic Reforms: How Does China Cross the River?" In *The course of China's rural reform*. Washington, DC: International Food Policy Research Institution, 15–29.

Schmook, B., and C. Radel. 2008. "International labor migration from a tropical development frontier: globalizing households and an incipient forest transition." Human Ecology 36(6):891–908.

Schultz, T.W. 1964. *Transformation tranditional Agriculture*. London: The University of Chicago Press.

Scott, J.C. 1976. *"The moral economy of the peasant: Subsistence and rebellion in Southeast Asia*." New Haven, CT: Yale University Press.

Shen, J., N. Spence, A. Bigsten, Z. Jiang, Y. Zeng, H. Pandey, R. Bora, H. Beladi, S. Marjit, K. Yadava, and others. 1996. “Modelling urban-rural population growth in China.” *Environment and Planning A* 28(8):1417–44.

Smith, S.K., J. Tayman, and D.A. Swanson. 2001. *State and local population projections: Methodology and analysis*. Springer.

Stark, O. 1991. *The Migration of Labor*. Blackwell Oxford.

Stark, O. 2005. “The new economics of the brain drain.” *World Economics* 6(6):137–140.

Stark, O., and D.E. Bloom. 1985. “The new economics of labor migration.” *The American Economic Review* 75(2):173–178.

Stark, O., and A. Dorn. 2013. “International migration, human capital formation, and saving.” *Economics Letters* 118(3):411–414.

Stark, O., C. Helmenstein, and A. Prskawetz. 1997. “A brain gain with a brain drain.” *Economics Letters* 55(2):227–234.

Stark, O., C. Helmenstein, and A. Prskawetz. 1998. “Human capital depletion, human capital formation, and migration: a blessing or a ‘curse’?” *Economics Letters* 60(3):363–367.

Stark, O., and R.E.B. Lucas. 1988. “Migration, Remittances, and the Family.” *Economic Development and Cultural Change* 36(3):465–481.

Stark, O., and Y. Wang. 2002. “Inducing human capital formation: migration as a substitute for subsidies.” *Journal of Public Economics* 86(1):29–46.

*Stock, J.H., and M. Yogo. 2005. “Testing for Weak Instruments in Linear IV Regression.” Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg:80.*

Tang, J., and Z. Ma. 2007. "Migration selectivity in China: An analysis based on the 2000 census data." *Population Research* 31(5): 42-50.(in Chinese).

Taylor, E.J. 1999. “The new economics of labor migration and the role of remittances in the migration process.” *International Migration* 37(1):63–88.

Taylor, J.E. 2001. “Microeconomics of Globalization: Evidence from Mexico, China, El Salvador, and the Galapagos Islands.” *Report to the Latin America and Caribbean Regional Office, World Bank, Washington, DC*.

Taylor, J.E., and A. López-Feldman. 2007. “Does migration make rural households more productive? Evidence from Mexico.” Agricultural and Development Economics Division of the Food and Agriculture Organization of the United Nations (FAO - ESA). No. 07-10.

Taylor, J.E., and P.L. Martin. 2001. "Human capital: migration and rural population change." *Handbook of agricultural economics* 1:457–511.

Taylor, J.E., S. Rozelle, and A. de Brauw. 2003. "Migration and incomes in source communities: A new economics of migration perspective from China." *Economic Development and Cultural Change* 52(1):75–101.

The Economist. 2014a. "Building the dream." *The Economist*. April 19th, 2014. Available on: http://www.economist.com/news/special-report/21600797-2030-chinese-cities-will-be-home-about-1-billion-people-getting-urban-china-work [Accessed August 21, 2014].

The Economist. 2014b. "China's $6.8-trillion hole?" *The Economist.* Nov. 28th, 2014. Available on: http://www.economist.com/blogs/freeexchange/2014/11/wasted-investment [Accessed January 24, 2015].

The New York Times, 2013. "China's Great Uprooting: Moving 250 Million into Cities." June 16th, 2013. Available on: http://www.nytimes.com/2013/06/16/world/asia/chinas-great-uprooting-moving-250-million-into-cities.html?pagewanted=all [Accessed August 21, 2014].

Todaro, M.P. 1969. "A Model of Labor Migration and Urban Unemployment in Less Developed Countries." *The American Economic Review* 59(1):138–148.

Tsang, M.C. 1996. "Financial reform of basic education in China." *Economics of Education Review* 15(4):423–444.

Tuan, F., A. Somwaru, and X. Diao. 2000. "rural labor migration characteristics, and employment patterns a study based on China's agricultural census.pdf." *TMD DISCUSSION PAPER* No. 63.

Tyers, R., J. Golley, and I. Bain. 2006. "Projected economic growth in China and India: the role of demographic change." *The Australian National University working papers in Economies and Econometrics* No.477.

Van den Berg, M.M., H. Hengsdijk, J. Wolf, M.K. Van Ittersum, W. Guanghuo, and R.P. Roetter. 2007. "The impact of increasing farm size and mechanization on rural income and rice production in Zhejiang province, China." *Agricultural Systems* 94(3):841–850.

Von Braun, J. 1979. "Analyse und Projektion der sektoralen und regionalen Arbeitskräfteentwicklung in der Landwirtschaft der Bundesrepublik Deutschland." Hannover: Alfred Strothe Verlag. (in German).

Wang, F., and A. Mason. 2008. "The demographic factor in China's transition." Cambridge, UK: Cambridge University Press.

Wang, F., and X. Zuo. 1999. "Inside China's Cities: Institutional Barriers and Opportunities for Urban Migrants." *The American Economic Review* 89(2):276–280.

Wolf, E.R. 1957. "Closed corporate peasant communities in Mesoamerica and Central Java." *Southwestern journal of Anthropology* 13(1):1–18.

Wooldridge, J.M. 2002. *Econometric analysis of cross section and panel data*. Cambrige, Massachusetts: MIT Press.

Wu, B., and S. Yao. 2010. "'Empty villages' in poor areas of China: A case study of rural migration in North Shannxi?." *The University of Nottingham China Policy Institute Discussion Paper* No.56.

Wu, Z., and S. Yao. 2003. "Intermigration and intramigration in China: A theoretical and empirical analysis." *China Economic Review* 14(4):371–385.

Yang, D. 2008. "International migration, remittances and household investment: Evidence from Philippine migrants' exchange rate shocks." *The Economic Journal* 118(528):591–630.

Yang, J.,Wang, H., Jin, S.,Chen, K., Riedinger, J. and Chao, P. 2014. "Migration, local off-farm employment, and agricultural production efficiency evidence from China." *International Food Policy Research Institute (IFPRI) Discussion Paper* No. 01338.

Zepeda, L. 2001. "Agricultural investment and productivity in developing countries." *FAO Economic and Social Development Paper* No. 148.

Zhang, Z. 2006. "Staying in the cities or go back to the rural area? [Jincheng Dingju Haishi Huixiang Fazhan]." *Chinese Rural Economy* 2006 (7):23–31. (in Chinese).

Zhang, G., and Z. Zhao. 2006. "Reexamining China's fertility puzzle: Data collection and quality over the last two decades." *Population and Development Review* 32(2):293–321.

Zhang, J., and Z. Zhao. 2011. "Social-family network and self-employment: evidence from temporary rural-urban migrants in China." *IZA Discussion Paper* No. 5446.

Zhang, K.H., and S. Song. 2003. "Rural–urban migration and urbanization in China: Evidence from time-series and cross-section analyses." *China Economic Review* 14(4):386–400.

Zhang, L., J. Huang, and S. Rozelle. 2002. "Employment, emerging labor markets, and the role of education in rural China." *China Economic Review* 13(2–3):313–328.

Zhang, S., and G. Luo. 2012. "China's migrant workers: How far from being citizens?" *Asian Social Science* 9(1).

Zhang, X., J. Yang, and S. Wang. 2011. "China has reached the Lewis turning point." *China Economic Review* 22(4):542–554.

Zhang, Z., and J. Li. 2007. "Movement of farmers-turned-laborers and rural middle-aged unemployment [Mingong Liudong Yu Xiangcun Zhongnian Shiye]." *Market & Demographic Analysis* 13(4):1–9. (in Chinese).

Zhao, Q., X. Yu, X. Wang, and T. Glauben. 2012. "The impact of parental migration on children's school performance in rural China." In International Association of Agricultural Economists Triennial Conference. Unpublished manuscript. Foz do Iguccu, Brazil.

Zhao, Y. 1997. "Labor Migration and Returns to Rural Education in China." *American Journal of Agricultural Economics* 79(4):1278–1287.

Zhao, Y. 1999. "Leaving the Countryside: Rural-To-Urban Migration Decisions in China." *The American Economic Review* 89(2):281–286.

Zhao, Y. 2000. "Rural-to-urban labor migration in China: the past and the present." *Rural Labor Flows in China*:15–33.

Zhong, F., Q. Li, J. Xiang, and J. Zhu. 2013. "Economic Growth, Demographic Change and Rural-Urban Migration in China." *Journal of Integrative Agriculture* 12(10):1884–1895.

Zhong, F., and J. Xiang. 2012. "Demographic transition in Rural China: a field survey of five provinces." In Huw Mckay and Ligang Song (Eds.) *Rebalancing and Sustaining Growth in China*. Canberra, Australia: ANU Press, 251–263.

Zimmermann, A., T. Heckelei, and I. Pérez. 2006. "Literature review of approaches to estimate structural change." *working paper SEAMLESS Report* (16):1–45.