OSIPP Discussion Paper : DP-2018-E-010

Why People Leave Their Rural Hometown: Evidence from 8 Provinces in China

July 22, 2018

He Zhu[†]

PhD. Student, Osaka School of International Public Policy (OSIPP)

JEL Classification: J61
Keywords: Internal migration mobility, Decision making, Industry, Nested logistic model

Abstract: This paper aim to clarify what motivate people to migrate from rural to urban area in China. The focus of most previous studies of migration are restricted to the wage gap between the origin and destination. However, this study uses the RUMiC (2008) data set that has individual characteristics of migrants and stayers, combined with China Statistical Yearbook data, to explore the decision making process on China's rural to urban migration. This research provides empirical evidence that migration is a joint decision-making process characterized by the choices of migration and destination. The results also show that the living condition in hometowns pushes people to migrate. For example, the probability of moving decreases by 25% if the consumption of the rural area increases by 20% (10,000 RMB).

[†] Corresponding author. Osaka School of International Public Policy, Osaka University, Address: 1-31 Machikaneyama, Toyonaka, Osaka 560-0043, JAPAN.
Email: zhuhesi@gmail.com

1 Introduction

Since the late 1979, China experienced the world's largest migration flow from rural to urban areas according to the International Labor Organization (ILO). As a result, the urbanization of China grew from 18 percent at 1978 to 53 percent at 2015 according to the World Bank.

Although concern for investment of human capital is a fundamental motivation for the study of migration, most theoretical and empirical studies have been limitedly focused on the wage gap. The human capital model — by far the base line of theory for the paper on the decision making process of migration developed by Sjaastad (1962) — implies that the migrants' goal is to maximize utility by choosing a location that provides the highest return to human capital, or labor supply. Estimates on the wage gap that is defined as the differential between the wage from the origin and expected wage from the destination) are often used to explain why people migrate.

Other studies on migration extend the human capital model by age and, network, policy, among others (Zhao,1999 2003; Borjas, 1990; Clark et al., 2007). Considering internal migration in China, Zhao (1999) addressed that the data on wages in China's rural areas are highly ambiguous varialiable due to imprecise self-reporting and non-pecuniary income. Furthermore, there exists an extreme situation in actuality where people only have non-pecuniary wages in Chinese rural areas. This situation makes it difficult to calculate the wage in the origin area. Therefore, directly using the wage gap on rural to urban migration in China can be misleading.

In another side of view, previous migration studies, whether theoretical or empirical, fail to address the alternative destinations, and are restricted to the hometown and migrated destination. However, when people are generally considering migrating,

there can be more than one alternative location. Hence, the influence and effects from the alternative location is lacking from the migration studies. A study done by Brown and Moore (1970) discusses the location choice and suggests that the decision making process of migration is a joint decision of migration and location choice, though they did not provide any empirical results. A study on the urban relocation making process was recently conducted by Lee and Waddell (2010). Their study provides empirical evidence that the decision making process has two phases; one is to choose whether to move, and the other is to choose where to relocate.

The starting in this chapter aims to analyze the decision making process of rural to urban migration using China in two models. The first model explores migration reasons besides the wage gap as there may be other reasons to make people leave their rural hometown in China. The second model exams whether or not the decision to migrate is a joint decision making process that is combined with the decision of migration and decision of destination. In addition, the decision making process may also be characterized as a hierarchical structure where the location choice may better be described as the second-stage decision. This study uses the Chinese Longitudinal Survey Data from eight provinces, and provides empirical evidence that the industrial structure has a significant influence on people's decision on migration. These findings are based on the analysis on the migrants' individual characteristics and the aggregate data of each province from the China Statistical Yearbook provided by the National Bureau of Statistics of the People's Republic of China.

This paper is organized as follows. Section 2 provides a description of the major policies affecting migration patterns in China. Section 3 reviews literature on migration in China and countries outside China. A brief description of the data set is

provided in Section 4, and the empirical results are reported in Section 5. The last section discusses the implications of these findings.

2 Background

2.1 Hukou system

A policy known as urban Hukou (household registration system) was introduced before 1979. Based on this policy, employment, allocation of housing, food, and other necessities were contingent on the registration system, which strictly restricted rural-to-urban migration. This was the period when China recently survived from the Great Chinese Famine. The original purpose of this policy was to control the population of urban areas and to guarantee that there was an adequate population for agriculture. However, the persistence of this restrictive policy had created severe inequality regarding wage, education and, housing, among others.

In late 1979, the government began to loosen the hukou system and introduced the household responsibility system (HRS) — a practice in China first adopted in agricultural economy in 1979 and later extended to the other sectors of the economy — by which managers of local enterprises were held responsible for the profits and losses they make. This system partially supplemented the egalitarian distribution method, whereby the state assumed all profits and losses. This change made it possible for people who were born in rural areas to purchase food or other necessities for living without an urban residence. At the same time, China was on track toward economic growth and welcomed large foreign investments in the manufacturing industry in the Eastern urban areas. With the fruit of economic results in urban areas,

increasing surplus of agriculture labor, and demand for cheap labor in the manufacturing sector led to a rush of rural-to-urban migration.

Although the food shortage had been avoided, the government continues to maintain the hukou system today. The restriction on migration from rural to urban areas still has impacts on various aspect of people's livelihood including employment, social security and property rights. Since the power of the hukou system still exists in China, this makes rural to urban migration in China different compared to that of other countries.

2.2 Government's effort on protecting the migrant worker

There are several documents issued by the Chinese government to support the living of migrant workers outside their hometown provinces.

This included two policy documents issued in 2002 and 2003, named Document Number 2 of 2002 and Document Number 1 of 2003, for the purpose of initializing the elimination of labor market discrimination against migrant workers and the legitimization of their working environment. In June 2006, the State Council passed a series of measures to protect migrant workers' labor rights with Circular No. 36. The measures included these restrictions on minimum wages, solutions to wage defaults, enforcement of labor contracts, and expansion of migrant workers' social security coverage.

However, these policies and the supporting system still were implemented differently across provinces, and thus, an investigation of China's internal migration given these policies will make a valuable contribution to this line of research.

3 Literature review

3.1 International migration

Migration is a topic that has been studied around the world for the past 50 years, which is led by Australia, Canada and the United States. There are a variety of theoretical models that have been developed to explain why people migrate. From the macro side of view, migration has been explained as a result or as a process of economic development (Lewis, 1954; Ranis and Fei, 1961; Harris and Todaro, 1970). At the macro level, people move from low-to-high wage areas to seek the premium from the wage differential. Compared with the macro-level models, the micro-level models focus on individual choices (Sjaastad, 1962; Todaro, 1969, 1976, 1989; Borjas, 1990; Stark 1984). The micro-level models argue that individuals are rational, and act based on a cost-benefit calculation, which motivates them to migrate. The literature based on these micro-level models support a view that individual characteristics influence people's choice of migration.

In recent years, there are some important findings in the field of migration. Studies using the gravity modeled found the moving distance increase cost for migration decision(Karemera et al., 2000; DA Plane, 1984; Flowerdew and Aitkin, 1982; Cindy Fan, 2005). Studies based on "Push and pull" theory found that migration decision making process is characterized by destination effect and condition effect (Dorigo and Tobler, 1983; D Hare, 1999; DS Kline, 2003; Fiona et al., 2016).

3.2 Migration in China

The subject of rural to urban migration in China has been extensively investigated, and for the past years, many studies investigated why rural people migrate to urban areas.

Kevin Honglin Zhang (2003) used panel data to proved the existence of a causal relationship between the growth rate of GDP and internal migration. A study conducted by Fan (2005) using cross-country data discovered a positive relationship between the population size and people's willingness to migrate. In another study, Fan (2005) modeled China's internal migration by the gravity model with population and distance. In her latter study, she used GDP, population, and transportation distance as the explanatory variables. Zhang and Song (2003) used the provincial data from the China Statistical Yearbook to explain that economic development is caused by rural to urban migration. Therefore, the above studies indicate that the GDP growth is typically being used as factor of economic development to explain why people migrate. However, considering that the value of agricultural output is a sub-component of GDP, the results may be unreliable. To deal with this problem, this paper uses the output ratio of different industries as the factor to explain why people migrate.

Two relevant studies were conducted by Zhao (1999) regarding the motivation for migration. One discussed why rural people migrate to urban areas by decision using data from the Sichuan province. This is considered to be the foundation of migration study in China, and her findings showed that young males tend to migrate. The other study argued that people migrate form rural to urban areas as a household-level decision making process.

Some other studies emerged that discussed why people migrate at the individual level, most of which divide the destination into inter-province, intra-province and overseas. Most importantly, as far as the author has studied, none of them have ever addressed the alternatives of destination choice. To fill this gap, the second part of this study focuses on the joint decision of migration with respect to migrating and

destination. Since this study focuses on the joint decision making process, the explanatory variables do not only includ the decision of whether to migrate or not, but will also include the decisions from several alternative destinations. Due to the limitation of the data set, eight provinces were chosen for analysis in the migration decision making process.

4 Data description

This study uses two data sets, namely the Longitudinal Survey on Rural Urban Migration (RUMiC)in China from the Institute for the Study of Labor, IZA and the China Statistical Yearbook from the National Bureau of statistics of the People's Republic of China.

The previous studies which used the IZA's data set have focused on three different aspects of migration. The first is the relationship between remittance and migration. The second, wage gaps and the equilibrium of the labor market (Klaus et al., 2016; Rachel et al., 2015; Björn et al., 2014; Hartmut et al., 2013; Klaus et al., 2014; Zhao and Qu 2014). The third, the well-being of children (Huang 2015; Meng and Yamauchi 2015) and the left-behind elders (Sylvie and Wang 2016; Zhao et al., 2016; Rachel et al., 2015; Klaus et al., 2015). However, none of them analyzed the decision making process for people to migrate from rural to urban areas.

The RUMiC consists of three parts: The Urban Household Survey, the Rural Household Survey, and the Migrant Household Survey. It was initiated by a group of researchers at the Australian National University, the University of Queensland, and the Beijing Normal University and was supported by the Institute for the Study of Labor (IZA), which provides the Scientific Use Files. The financial support for RUMiC was

obtained from the Australian Research Council, the Australian Agency for International Development (AusAID), the Ford Foundation, IZA and the Chinese Foundation of Social Sciences. The IZA data is a face to face interview data temporal covered the year of 2008. This study used the Urban Migrant Survey (UMS) and Rural Household Survey (RHS) to cover the provinces of Anhui, Hubei, Sichuan, Chongqing, Shanghai, Jiangsu, Zhejiang, Henan, and Guangdong.The sample size combined 31,791 individuals from RHS and 8,446 individuals from UMS, which totals 40,237 individuals. The coverage areas contain the main migrant supplying provinces (Sichuan, Chongqing, Anhui, Hubei, Henan), and the main recieving provinces (Guangdong, Zhejiang, Jiangsu). The macro-level data set of each province was provided by the National Bureau of statistics of the People's Republic of China.

The main purpose of this study is to analyze the decision making process why people leave their rural homes. The reason for restricting individuals between 10 to 75 years of age rather than using the whole data set is that children and the elderly are unlikely to be involved in the decisions of their family.

Table 2 presents the summary statistics for the variables used in the analysis. This study covers eight provinces. The people in the eight provinces are considered to have possibility to move to the other seven provinces. Table 3 gives the statistics of the destination provinces (including non-migrants).

Table 1. Studied Samples

	Freq.	Percent	Cum.
Stay	33,388	93.95	93.95
Move	2,150	6.05	100
Total	35,538	100	

Wave 2008

Source: 2008 survey data of RUMiC

Table 2. Destinations included in the analysis

Province	Freq.	Percent	Cum.
Jiangsu&Shanghai	6,487	18.26	24.69
Zhejiang	3,842	10.81	29.06
Anhui	4,291	12.07	41.14
Henan	4,542	12.78	53.92
Hubei	4,315	12.14	66.06
Guangdong	5,434	15.29	81.35
Chongqing	2,222	6.25	87.6
Sichuan	4,405	12.4	100
Total	35,538	100	

Wave 2008

Source: 2008 survey data of RUMiC

5 Methodology

5.1 Stay or move choice models

Migration can be triggered by income gaps between the rural and urban areas (Sjaastad 1962). The majority of the previous studies argue that the motivation of migration is because of the wage gap, however, it is almost unrealistic to gain the real income of

rural individuals due to non-pecuniary income exist. In addition, the Urban Migration Survey (UMS) 2008 reported that over 50 percent of the people chose to migrate because they are “too poor at hometown” or they have “no future in their rural hometown”. Thus, this study uses people’s living condition to explain why people migrate from rural to urban areas. Thus, this study takes into account living condition can be one motivation that caused people migrate from rural to urban areas.

This study defines migration as moving to another province, like the previous studies. This study is consistent with recent studies in the first that it uses the economic factors of migrants’ rural hometown and individual factors such as age, gender, marriage status, and educational level as the determinate factor of migration. Table 3 explains the statistical description of the variable, and Table 5 shows the results of the logit model. Table 6 provides the marginal effects of the model (3) regarding the possibility of migration.

This study shows similar results with tape in the previous studies in a point that young males tend to migrate than old females. However, this study also shows that poverty is another factor that pushes people to migrate. Considering Table 6, the marginal effects from the results imply that the percentage of moving decreases almost more than 50% if the consumption of rural areas increases by 20%. Here, a 20% increase in the consumption means that 10,000 RMB will be increased, when people in rural areas consume at an average of 50,000 RMB per year. Therefore, people are likely to choose migrate rather than staying.

Table 3. The explanation of variables for stay or move model

Variables	Mean	Std. Dev.	Min	Max
Male	0.53	0.50	0.00	1.00
Age	37.04	15.88	10.00	74.00
Married	0.64	0.48	0.00	1.00
No school	0.056	0.23	0.00	1.00
Average rural living standard	0.48	0.21	0.32	1.13
Erural	0.16	0.18	0.02	0.73
Unemployment	3.69	0.61	2.60	4.60
Wage-gap wave 2008	14795.81	4034.67	11097.66	23622.38

Source: Survey Data and China Statistical Yearbook in 2008

Table 4. Data description

Variables	Description
Average rural living standard (million RMB)	Consumption of rural areas divided by Population of rural areas
Erural	Electricity consumption of rural areas divided by Population of rural areas
Unemployment	unemployment rate of rural provinces
No school	Never attended school
Male	Gender for individuals (1 for male, 0 for female)
Age	Age for individuals
Married	Married status for individuals (1 for married and didn't get divorced, 0 for others)
Wage-gap wave 2008	Income of urban minus consumption of rural for individuals (million)

Table 5. Results of logit model for the stay-or-move model

VARIABLES	(1)	(2)	(3)
Male	0.317***		0.314***
	(0.0457)		(0.0845)
Age	-0.0555***		-0.0644***
	(0.00239)		(0.00463)
Married	1.177***		0.746***
	(0.0663)		(0.110)
No school			0.497*
			(0.2561)
Average rural living standard		-1.037**	-18.33***
		(0.519)	(0.806)
Erural Per person		-1.653**	-2.088**
		(0.654)	(1.016)
Unemployment		0.241***	2.676***
		(0.0546)	(0.109)
Wage-gap			0.00122***
			(2.61e-05)
Constant	-1.783***	-2.991***	-21.12***
	(0.0642)	(0.196)	(0.589)
Observations	35,538	35,538	30,021

Note: Figures in parentheses are t statistics.
*** Significant at the 1% level
** Significant at the 5% level
* Significant at the 10% level

Source: author's calculation

Table 6. Marginal effect for the Logit model

	Delta-method					
	dy/dx	Std. Err.	z	P>z	[95% Conf.	Interval]
Male	0.005	0.001	3.91	0.000	0.002	0.007
Age	-0.001	0.000	-15.84	0.000	-0.001	-0.001
Married	0.011	0.002	7.02	0.000	0.008	0.014
No school	0.007	0.004	1.94	0.053	-0.000	0.015
Average rural living standard	-0.271	0.018	-15.32	0.000	-0.305	-0.236
EruralPerperson	-0.030	0.015	-2.00	0.045	-0.059	-0.001
Unemployment	0.040	0.002	16.74	0.000	0.035	0.044
Wag-gap	0.000	0.000	20.21	0.000	0.000	0.000

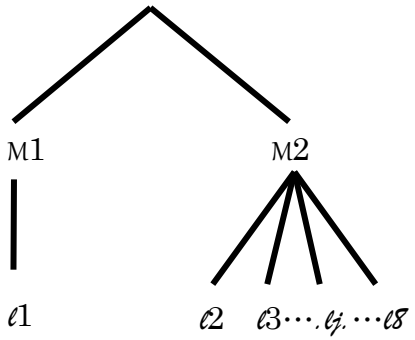
Source: Survey Data and China Statistical Yearbook in 2008

5.2 Joint decision model with destination choice

In China, there are few studies regarding destination choice. Even though Shuming (2007) and Fan (2005) showed that moving distance(distance from the origin to the destination) has influenced migration decision, although their studies did not specifically consider the presence of alternative destinations. Brown and Moore (1970) discussed location choices among alternative destinations. Their study decomposed location choice into two phases: the first phase considers both external and internal factors for the stay-or-move decision; and the second considers a choice of destinations. This two-stage decision structure motivate the use of the hierarchical model to model the migration decisions.

Thus, in this study, the migration choice is considered as a joint decision making process by individuals. The nesting structure is assumed to include a binary mobility choice of 'stay' or 'migrate' at the top level, and a multiple choice of location at the bottom level. The 'stay' nest stands for hometown (hukou province), and the 'migrate' nest includes the other seven provinces as alternative destinations.

Figure 1. The Structure of nested logit model



Mobility choice (m):

M1= stay

M2= migrate

Location choice (l):

$l_1 = stay$

$l_j = destination\ \ provinces$ (j= 2,…8)

the probability of an individual choosing location l in nested logit model(NLM) is defined as,

$$P(l) = P(l|m) * P(m) \quad (1),$$

where, $P(\mathrm{l}|\mathrm{m})$ is the conditional probability of choosing location from the mth nest. $\mathrm{P(m)}$ is the marginal probability of choosing mth nest. The bottom level conditional probability is equivalent to the standard MNL equation and is written as,

$$P(l|m) = \frac{e^{V_l}\ \mu_l}{\sum_{l' \in L_m} e^{V'_l \mu_l}} \quad (2),$$

where, V_l represents the observable components of the utility function for each elemental eta alternative and $\mu_l$is the corresponding scale parameter. The marginal choice probability of choosing mth nest is,

$$P(m) = \frac{e^{V'_m \mu_m}}{\sum_{m' \epsilon M} e^{V'_{m'} \mu_m}} \quad (3),$$

where, V'_m is the logsum ,or the inclusive value, associated with the mth nest, and $\mu_m$is the top level scale parameter. The logsum represents the expected value of the maximum within the random utilities of all the alternatives in the mth nest. For random sampling, the expanded logsum can be written as,

$$V'_m = \left(1/\mu_l\right) \ln \left\{ \sum_{l' \in L_m} \left[\left(1/R'_l\right) e^{V'_\mu \mu_l} \right] \right\} \quad (4),$$

In this study, the stay nest is the *hukou* province, l_1, thus the u_m is degenerated and restricted as one location. By implementing the NLM as the methodology to analyze the destination province characteristics, the attractiveness of destination l for an individual n is expressed by the utility function,

$$V_l = \alpha X_l + \beta(D_n) + \gamma I_n X_l,$$

where X_l is the choice-specific explanatory variable containing an array of characteristics describing alternative l, D_n is the moving distance for each individual as the interaction term, and I_n is an array of attributes for individual n. This utility function reveals that when an individual makes a rational decision of location, he or she takes into consideration both characteristics of the potential alternative destinations and their own characteristics. $y_{mn}$is 1 if individual n chooses nest m, and 0 otherwise. And y_{lmn} is 1 if individual n chooses alternative l in nest m, and 0 otherwise.

In this study, the NLM is estimated using the Survey of 2008. In the model, Shanghai is included in the Jangsu province because the area of Shanghai is too small when compared with other provinces. Another reason is that Shanghai is a city belonging to Jiangsu before 1927. The explanatory variables for the NLM are as follows in Table 8 for each province. Because this study covers 8 out of the 34 provinces in China, it is

would be fair to admit that the result may be partially because of the movement both within and outside of the 8 providences. Nevertheless, both the main receiving provinces (Jiangsu, Zhejiang, Guangzhou) and the main sending provinces (Sichuan, Anhui) are included in the data set, resulting in significant coefficient. The explanatory variables are listed in Table 8 for both individual and province levels. Note that each province means the *hukou* provinces of individuals before migration occurs. The distance in this study stands for the geographic distance between individuals' *hukou* provinces and their destination provinces. The distance variable has been calculated by the author according to Vincenty's (1975) equations (by assuming a sphere with a radius of 6,371km to approximate the shape of the earth and computing great-circle distances). The reason for using the industrial structure (the ratio of the secondary industry output value and the ratio of the tertiary industry output value) as the explanatory variables in the analysis is twofold. One reason is that the previous studies suggest economic factors are the main factor for migration, and the second is that few migrants work in the primary industry. Thus, I set up the following hypotheses: first, the ratio of the secondary industry to GDP for each alternative location is the attracting factor for individuals; second, the ratio of the tertiary industry to GDP for each alternative location is the attractiveness factor for individuals; third, the moving distance for each alternative location is the cost factor for individuals.

Table 7. Mobility matrix

	Hukou Province							
	Anhui	Chongqing	Guangdong	Henan	Hubei	Jiangsu	Sichuan	Zhejiang
Destination								
Anhui	83.46%	0.08%	0.05%	0.72%	0.17%	0.40%	0.48%	0.24%
Chongqing	0.00%	96.13%	0.00%	0.00%	0.06%	0.00%	2.93%	0.00%
Guangdong	0.34%	0.88%	99.82%	2.32%	2.78%	0.09%	2.18%	0.13%
Henan	0.95%	0.04%	0.00%	92.86%	0.14%	0.00%	0.04%	0.40%
Hubei	0.11%	0.42%	0.05%	0.82%	95.03%	0.05%	0.21%	0.11%
Jiangsu	9.65%	0.84%	0.02%	1.58%	0.80%	98.30%	1.06%	1.65%
Sichuan	0.05%	0.88%	0.04%	0.06%	0.10%	0.05%	92.40%	0.00%
Zhejiang	5.44%	0.72%	0.02%	1.65%	0.91%	1.11%	0.71%	97.47%
Total	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Source: 2008 Survey Data

The mobility matrix from the 2008 survey is shown in Table 7. The numbers represent the percentage of individuals who chose a particular province among the eight alternative destinations. The mobility matrix shows that Anhui, Henan, Hubei, and Sichuan are the net sending provinces, whereas the net receiving provinces are Guangdong, Jiangsu, and Zhejiang. Also, Chongqing is described as a province which is neutral in terms of sending and receiving province.

Table 8 presents the definition of the variables that are used in the Nested Logit Model. This study uses the log moving distance of each individual to seven other alternative destinations (location of potential destination) and industrial structure measures in the destination (secondary and tertiary) as the choice specific explantory variables. The results of the NLM are shown in Table 9. Table 10 summarizes the results of Tables 8 and 9.

Table 8. Variables and descriptions in The destination choice model

VARIABLES	Description
Log-distance	The logit of the geographic distance for each individual to each alternatives
second	The ratio of secondary industry output value for each alternative
tertiary	The ratio of tertiary industry output value for each alternative
choice	The choice of individuals (1 for specific location, 0 for others)
age	Age of individuals
male	Gender of individuals (1 for male 0 for female)
married	Marrital status of individuals (1 for married, 0 for others)

Table 9. The results of Nested Logit Model for each province

	Chongqing	Jiangsu	Sichuan	Zhejiang	Anhui	Guangdong	Henan	Hubei
Choice specific								
Logdistance	-3.070*	-2.852***	-2.252***	-2.525***	-0.408***	-2.853**	-0.000529***	-0.00109***
	(1.587)	(0.428)	(0.514)	(0.26)	(0.0764)	(1.311)	(0.00013)	(0.000174)
Tertiary	44.81	42.85***	41.13***	5.888		3.54		
	(30.28)	(5.548)	(10.96)	(4.675)		(13.61)		
Second	21.63**	17.44*	12.38*	31.93***	3.715***			
	(9.513)	(9.13)	(6.908)	(7.709)	(0.786)			
Migrant								
Age	-0.0317***	-0.0791***	-0.0434***	-0.0413***	-0.0480***	-0.0738**	-0.0470***	-0.0466***
	(0.00865)	(0.0105)	(0.00477)	(0.00968)	(0.00409)	(0.0359)	(0.00536)	(0.00642)
Male	-0.0271	0.344*	0.352***	0.425**	-0.309***	-0.0636	0.426***	0.340**
	(0.214)	(0.197)	(0.108)	(0.215)	(0.0831)	(0.636)	(0.11)	(0.136)
Married	0.696**	2.362***	1.109***	1.292***	1.036***	1.392	1.289***	0.779***
	(0.307)	(0.329)	(0.159)	(0.348)	(0.129)	(1.035)	(0.166)	(0.197)
Miggrat_tau	2.198	1.545***	1.689***	1.348***	0.192***	1.254	-0.894***	-0.765***
Observations	19,000	51,720	41,544	30,096	35,296	44,776	43,112	38,784

Note: Figures in parentheses are t statistics.
*** Significant at the 1% level
** Significant at the 5% level
* Significant at the 10% level
Source: Author's estimation.

Table 10. The signs of the coefficient estimates of the Nested Logit Model

VARIABLES	Chongqing	Jiangsu	Sichuan	Zhejiang	Anhui	Guangdong	Henan	Hubei
Destination-specific variables								
Log-distance	-	-	-	-	-	-	-	-
Tertiary		+	+					
Second	+	+	+	+	+			
Migrant-specific variables								
Age	-	-	-	-	-	-	-	-
Male		+	+	+	-		+	+
Married	+	+	+	+	+		+	+

Source: Survey Data and China Statistical Yearbook in 2008

From Tables 9 and 10, the general results are that the individual variables have similar results with both prior studies and the prior results: young males tend to migrate. The new findings from the destination choice model show that the choice specific variables (moving distance, industry mechanism) have varied outcomes where distance is a cost, and the second and tertiary industry ratios are attractiveness factors for migrants.

In order to confirm the two stage structure of the nested logit model, the author conducted a comparison test among the multinomial logit model, conditional logit model, and nested logit model. Table 11 shows the results of the multinomial logit

model and nested logit model and presents the results of independence of irrelevant alternatives (IIA) test between the conditional logit model and nested logit model. Although the Table 11 do not show supportive results for all of the provinces, especially for the receiving provinces, the main sending provinces have the supportive results of this study.

Table 11. Test for specification of the choice specific variables and LR test for Independence of Irrelevant Alternatives (IIA)

	Chongqing	Jiangsu	Sichuan	Zhejiang	Anhui	Guangdong	Henan	Hubei
significance of migrants' tau	0.768**	1.545***	1.689***	1.348***	0.192***	1.254	0.725***	-0.765***
Hypothesis :nested logit model better than multi-nominal logit model	yes	yes	yes	yes	yes		yes	yes
(tau = 1), Prob > chi2	0.24	3.05*	4.87*	2.32	146.41***	0.03	-718.89***	551.15***
Hypothesis :nested logit model better than conditional logit model	yes		yes	yes			yes	yes

Note: Figures in parentheses are t statistics.
*** Significant at the 1% level
** Significant at the 5% level
* Significant at the 10% level

6 Conclusion and Discussion

In recent years, as a response to the migration and the urban development boom, the Chinese government made a considerable effort towards labor policy making. With the government's continuous efforts, the situation of migrant workers has been improved, though it still varies across provinces.

The first part of this study used a logit model to analyze the decision whether to stay in people's hometown or migrate. The results indicate that the percentage for moving decreases almost 25% when the consumption of the rural area increases 20%. Other results show that people's choice of migration is forced rather than a personal choice. A

significant reason that is pushing people to move is an economic factor: the rural living condition, which has long been neglected. This result suggests that the willingness to move is low in relatively developed rural areas.

The second part of this study provides empirical evidence that the decision on migration and destination is a joint decision making process. The effect pf determination characteristics vary across provinces. On the one hand, the industry mechanism (the ratios of secondary and tertiary industry to GDP) motivates migrants to move from rural to urban areas. On the other hand, the geographic distance has a negative effect on the decision to move, i.e. the longer the distance is from the hometown to the destination, the lower the willingness to move will be.

While this paper suggests a new structure for study on China's migration, there are still opportunities for further studies. Since the nested logit model has been used for each province separately, we suggest that future studies explore individual provinces to address their own special characteristics that influence migrants' decision of destination choice.

Reference

Brown, L.A., Moore, E.G.: The intra-urban migration process: a perspective. *Geogr. Ann*. B 52, 1–13 (1970)

Connelly, R., & Maurer-Fazio, M. (2015). Left Behind, At Risk, and Vulnerable Elders in Rural China: What the RUMIC Data Reveal about the Extent, Causes, and Consequences of Being Left Behind.

Dorigo, G., & Tobler, W. (1983). Push-Pull Migration Laws. *Annals of the Association of American Geographers,73*(1), 1-17.

Fan, C. C. (2005). Modeling interprovincial migration in China, 1985-2000. *Eurasian Geography and Economics*, 46(3), 165-184.

Fan, C. C. (2005). Interprovincial migration, population redistribution, and regional development in China: 1990 and 2000 census comparisons. *The Professional Geographer*, 57(2), 295-311.

Frank Qu, Z., & Zhao, Z. (2014). Evolution of the Chinese rural-urban migrant labor market from 2002 to 2007. *China Agricultural Economic Review*, *6*(2), 316-334.

Institute for the Study of Labor (IZA), Australian National University, the University of Queensland and the Beijing Normal University (2014), Longitudinal Survey on Rural Urban Migration in China (RUMiC) (2008 years of). IDSC of IZA. http://dx.doi.org/ 10.15185/izadp.7680.1

Lee, B. H., & Waddell, P. (2010). Residential mobility and location choice: a nested logit model with sampling of alternatives. *Transportation*, 37(4), 587-601.

Massey, D., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., & Taylor, J. (1993). Theories of International Migration: A Review and Appraisal. *Population and Development Review*, 19(3), 431-466.

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

Shuming Bao, Örn B. Bodvarsson, Jack W. Hou, Yaohui Zhao(2007). Interprovincial Migration in China: The Effects of Investment and Migrant Networks. *Chinese Economy*, *42* (4), 7-29.

Taylor, J. E., Rozelle, S., & De Brauw, A. (2003). Migration and Incomes in Source Communities: A New Economics of Migration Perspective from China. *Economic Development & Cultural Change*, 52(1), 75-101.

Vincenty, T. (1975) Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations, Survey Review 22(176): 88-93.

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

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

Zhao, Y.. (1999). Labor Migration and Earnings Differences: The Case of Rural China. *Economic Development and Cultural Change*, *47*(4), 767–782.

Zhao, Yaohui (2003). The Role of Migrant Networks in Labor Migration: The Case of China. *Contemporary Economic Policy* **21** (4): 500–511

Appendix

The education levels

The explanatory variables in the ‘migrate’ or ‘stay’ choice model numbers (from 1 to 8 stand for the levels of education).

Table A1: Levels of Education

Education	Freq.	Percent	Cum.
Never attended school	2,143	6.72	6.72
elementary school	7,719	24.2	30.92
Junior middle school	15,826	49.63	80.55
Senior middle school	3,877	12.16	92.71
Vocational school	1,483	4.65	97.36
Polytechnic college	672	2.11	99.46
Undergraduate	161	0.5	99.97
Postgraduate	10	0.03	100
Total	31,891	100	

Source: RUMiC Survey Data in 2008

The macroeconomic data of provinces

Table A2: The macroeconomic data for the destination choice model

Province	Secondary industry	Tertiary Industry	GDP	Second ratios	Tertiary ratios
Jiangsu	23079.18	19760.76	45051.84	0.512281	0.438623
Zhejiang	11567.42	8799.31	21462.69	0.538955	0.409982
Anhui	4198.93	3234.64	8851.66	0.474366	0.365428
Henan	10259.99	5099.76	18018.53	0.569413	0.283029
Hubei	5082.07	4466.85	11328.92	0.448593	0.394287
Guangdong	18502.2	16321.46	36796.71	0.502822	0.443558
Chongqing	3057.8	2160.48	5793.66	0.527784	0.372904
Sichuan	5823.39	4561.69	12601.23	0.462129	0.362004

Source: China Statistical Yearbook in 2008

The calculation of the distance

The distance from the home to destination provinces of migrants between "*Hukou*" (their hometown), and destination provinces are calculated using the gaining the longitude and latitude information from Google Maps. In this study, the geographic distances is caulculated as the length of the shortest curve between two points along the surface of a mathematical model of the Earth. By default, this paper used the input coordinates that are assumed to be based on the WGS 1984 datum (the same used by Google Earth/Maps and GPS devices), and calculated the ellipsoidal distances using Vincenty's (1975) equations (Given the coordinates of the two points (Φ1, L1) and (Φ2, L2), to calculate the ellipsoidal distance s. $s=bA(\sigma-\Delta\sigma)$). This assumes a sphere with a radius of 6,371km to approximate the shape of the earth and can computes great-circle distances. The distance is shown in Table 1.

Table A3: Provinces of geographic information

Province	Google Maps GPS Coordinates	
	N	E
Henan	34.7655	-113.754
Jiangsu	32.0617	-118.763
Sichuan	30.6512	-104.076
Hubei	30.5466	-114.342
Anhui	31.8611	-117.285
Shanghai	31.2304	-121.474
Zhejiang	30.2674	-120.153
Guangdong	23.1322	-113.267
Chongqing	29.563	-106.552

Source: Google Maps

In this research, the size of each province is shown as follows. For those who chose to stay, the moving distance has been calculated as the radius of their *hukou* province.

Table A4: Size of each province

Province	Size(km^2) of provinces From National Bureau of Statistics of the People's Republic of China
Henan	167,000
Jiangsu	102,600
Sichuan	485,000
Hubei	185,900
Anhui	139600
Shanghai	7,037
Zhejiang	101,800
Guangdong	177,900
Chongqing	82,300

Source: Google Maps and National Bureau of Statistics of the People's Republic of China

Table A5: The calculated moving distance for those who chose to stay

Province	Radius(km)
Henan	230.5596
Jiangsu	180.7169
Sichuan	392.9126
Hubei	243.2567
Anhui	210.7986
Zhejiang	180.011
Guangdong	237.965
Chongqing	161.8546

Source: Author's calculation using the coordinates information in Google Maps

Statistics for Nested-logit model

Table A6: The explanatory variable for nested logit model for each province

	Sichuan					Zhejiang				
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Age	41544	37.47006	18.68832	1	97	30096	39.64115	18.53832	1	96
Male	41560	0.524735	0.499394	0	1	30096	0.523392	0.499461	0	1
Married	41560	0.620019	0.485388	0	1	30096	0.653642	0.475817	0	1
Choice	41560	0.125	0.330723	0	1	30096	0.125	0.330724	0	1
Second	41560	0.504543	0.038464	0.448593	0.569413	30096	0.504543	0.038464	0.448593	0.569413
Tertiary	41560	0.383727	0.048035	0.283029	0.443558	30096	0.383727	0.048035	0.283029	0.443558
Log_distance	41560	6.462717	1.248251	3.457217	7.339688	30096	6.245932	0.971373	4.24604	7.339688

	Chongqing					Jiangsu				
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Age	19000	38.73642	18.79175	1	92	51720	37.38237	18.2958	1	97
Male	19000	0.522105	0.499524	0	1	51720	0.516628	0.499728	0	1
Married	19000	0.660211	0.47365	0	1	51720	0.636814	0.480923	0	1
Choice	19000	0.125	0.330728	0	1	51720	0.125	0.330722	0	1
Second	19000	0.504543	0.038464	0.448593	0.569413	51720	0.504543	0.038464	0.448593	0.569413
Tertiary	19000	0.383727	0.048035	0.283029	0.443558	51720	0.383727	0.048035	0.283029	0.443558
Log_distance	19000	6.428889	0.928027	4.303762	7.17964	51720	6.074296	0.993703	4.369928	7.245877

	anhui					guangdong				
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Age	35296	34.77403	18.15296	1	95	44776	33.53904	18.78814	1	102
Male	35304	0.511897	0.499866	0	1	44880	0.520856	0.49957	0	1
Married	35304	0.586902	0.492397	0	1	44880	0.495722	0.499987	0	1
Choice	35304	0.125	0.330724	0	1	44880	0.125	0.330723	0	1
Second	35304	0.504543	0.038464	0.448593	0.569413	44880	0.504543	0.038464	0.448593	0.569413
Tertiary	35304	0.383727	0.048035	0.283029	0.443558	44880	0.383727	0.048035	0.283029	0.443558
Log_distance	35304	5.843096	1.264499	3.063435	7.140438	44880	6.661319	0.843075	4.456937	7.165823

Source: 2008 Survey Data and China Statistical Yearbook