Migration Decision and Rural Income Inequality in Northwestern China

Hua, Yue

Hunan Univeristy

10 July 2014

Migration Decision and Rural Income Inequality in Northwestern China

Abstract: This paper examines the decision between internal migration and home production for rural households and its impact on rural income distribution. By constructing counterfactual scenarios under which households are allowed to switch freely between internal migration and home production, this study finds that the migrant households in the studied region could have earned more had they choose not to migrate and work in local sectors, given the results that show remittances earned by the migrant households are less than their simulated home production earnings. The findings also illustrate that there would also be less income inequality in this area if migrants choose to work locally. These results are compatible with the fact that the internal migration in the study area is very likely to be involuntary, due to the lack of arable land and insufficient local nonfarm job opportunities, usually provided by township and village enterprises.

Key Words: Internal Migration; Inequality; Job Decision

JEL Classification: O15; O18

The author acknowledge this paper to be published on *China Economic Journal*

1.1 Introduction

Internal rural-urban migration in China has become a highly noticeable social and demographical phenomenon since the beginning of the economic reform in 1978. Urban areas, especially large cities, have been developing disproportionately, which is the consequence of the unbalanced domestic development strategy focusing on urban industrialization since 1949.[1] This rural-urban divide and growth disparity motivates huge amounts of rural surplus laborers to seek jobs in urban areas for both pecuniary and nonpenuciary payments. A migrant worker in China can be defined as a rural laborer who works in urban area but does not have a formal urban resident identity, or urban *hukou*.[2] Thus, migrant workers are treated very differently from urban workers in terms of remuneration, both by the government and by the urban employers. They receive lower wage than their equally-productive urban colleagues, and are often excluded from the basic social security programs.

As the world's largest floating population, the number of the Chinese migrant worker was 180 million at the end of the 1990s. In 2012, this number has exceeded 250 million,[3] and it has been predicted that the number of migrant workers will continue to increase swiftly, reaching approximately 300 million by 2015 (Yang, 2009). Most of the migrant workers are low-skilled young workers with limited educational attainment. These

[1] The adoption of this development strategy reflects the doctrine of learning from the Soviet Union and Stalinist Communism. Mao believed that the heavy machinery sector should always be the first priority in economic development to enhance the national defense.
[2] China's modern *hukou* System (Household Registration System) started in 1950. It divides residents formally into rural residents and urban residents based on their residence and biological kinship.
[3] Reported in *China Statistic Yearbook* by the National Bureau of Statistics of China.

workers are employed mainly in mining, manufacturing, construction and low-skilled service sectors which often require intensive physical work.

The characteristics of the Chinese migrant workers can be attributed to both rural labor supply and demand. On the supply side, China has 674 million rural residents, with more than 150 million registered rural households. Traditional rural Chinese household makes fertility decision that aims at maximizing the household's productivity, thus girls are often discriminated against and boys with more physical working abilities are preferred. On the other hand, the "one-child policy" is difficult to implement in many rural areas due to the lack of monitoring and judicial resources, so rural households often choose to have multiple children. Children within one household would naturally compete for limited parental cares and educational resources with their siblings. Poorly educated parents and undeveloped rural school system can also adversely affect the children's intellectual development and school performance, especially those children who are left behind. The result is high dropout rates and low average educational attainment.

On the demand side, shortage of arable land has always been a binding constraint on agricultural production in China (Zhu and Luo, 2009). As reported annually by the Ministry of Land and Resources, China's arable land per capita has been declining for more than 10 years. In 1997, arable land per capita was 1.58 mu, and this number declined to 1.38 mu in 2011.[4] This is not only because of the increasing rural population, but also because of the shrinking amount of total arable land due to urban sprawl and

[4] 1 mu = 667 m^2

environmental deterioration. Adoption of advanced agricultural technologies and machineries in some area leads to a high marginal product of labor, which also decrease the demand for agricultural laborers. When both rural farm and nonfarm sectors are unable to absorb all these redundant laborers due to limited capacity, migrating to urban areas becomes the only option to seek better pecuniary payments for these rural surplus laborers. Urban firms have strong incentive to hire these migrant workers who would accept a much lower wage relative to their equally-productive urban counterparts.

Internal migration is not barrier free and costless, especially for the relatively poor households. Successful migration often requires contacts, know-how and some capital or at least the ability to borrow (Davin, 1998). In China, transportation costs are high for the undeveloped areas that are remote and inaccessible. Living costs (rents and utility fees) in urban areas are also increasing due to the long-lasting domestic inflation.

One of the most important social welfare issues in cotemporary China is the income inequality in the rural area. The Gini coefficient of the rural area is 0.39 in 2011, very close to 0.4, which is the recognized security line of uneven distribution.[5] Large income inequality in China often induces severe hierarchical conflicts between the rich and the poor, resulting in discrimination against the poor, crimes against the rich and distrust of the government. While most of the minority population resides in rural areas, income inequality across different minorities can even create secessionism which may threaten national security. Thus, rural income inequality can significantly affect the social stability in China via different channels and should receive some serious inspections. Remittances,

[5] Reported by the National Bureau of Statistics of China.

defined as money transmitted to rural households by migrant workers (Adams, 1989), are very likely to change the income distribution in the source communities. Hence, examining the possible effects of remittances on rural income inequality is critical to explore the welfare implication of internal migration, and can also provide useful guidance for policy makers who aim at constructing a more stable and harmonious society.

Using data from a poverty monitoring survey of rural households in Gansu and Inner Mongolia province in northwestern China, this study contributes to the existing literature by performing quantitative analyses on remittances and income inequality in this area for the first time. This region is worth studying because it is geographically and economically different from other macroregions in China (Naughton, 2007), which can lead to significant differences in the household decision between migration and home production. Findings show that while initially less-endowed migrant households are better off by sending out migrant workers and receiving remittances in absolute term, they can actually be even more economically improved if they choose to work in local sectors, assuming such job opportunities are always available. Also, to reduce the overall income inequality in this region, local government should focus either on increasing migrant workers' net earnings in the urban area, or on creating more local job opportunities.

The rest of the paper is organized as follows. Section 2 illustrates the importance of remittances for rural households and provides a brief review of some representative empirical evidence. Section 3 presents the methodology used for estimation and

simulation. Section 4 describes the data. Section 5 reports and interprets the estimation and simulation results. Section 6 focuses on inequality analysis and policy implications. Section 7 concludes.

2. Remittances in Rural China and Literature Review

In China, it is conventional for migrant workers to send or bring part of their annual earnings back to their rural households as remittances. Prior studies showed that the remittances share of total household income is generally low in Asian countries. One study on 12 villages in northern India finds that the average remittance share in total household income is 6.9% (Tumbe, 2011). Other studies on Thailand, Malaysia and Indonesia find even lower remittance shares (Zhou, 2006). China, however, has a much higher remittance share than other Asian countries due to both economic and cultural reasons. Migrant workers, under the influence of collectivism culture, possess a strong sense of responsibility to earn both pecuniary and non-pecuniary resources for their family members. The remittances also serve as an income smoothing tool by diversifying the source of income for rural households, especially for those households whose income depends solely on farming activities, which is very sensitive to natural conditions. As reported by the Consultative Group to Assist the Poor (CGAP), the overall remittances share in total migrant household income in China is 30%-50%. Remittances serve as a critical source of rural household expenditures on education, job training, medical care and daily life.

Remittances, as a critical part of the total household income, can also affect the rural income distribution conditional on the position of the migrant households in the income spectrum. If the first households to make a migration decision are from the upper end of the income distribution, they are more likely to receive high returns from this high-risk "investment" (Taylor, 1986). However, if the households who choose to migrate are at the lower end of income distribution and are forced to do so because they have surplus laborers that cannot be absorbed by the local labor market, the remittances earned by these migrant workers may help the migrant households to catch up with their wealthier nonmigrant neighbors and the local income inequality could be reduced..

Stark (1986) takes remittances as an "extra" net income of the migrant households that has no substitutes if migrant households do not participate in migration. Adams (1989) issued a study on remittances and inequality in rural Egypt, in which he criticized and improved the Stark approach. The Adams approach no longer conceives remittances as purely exogenous, but treats remittances as a substitute for home production earnings, which highlights the existence of the opportunity cost of migration. Enlightened by the Adams approach, Bradford and Boucher (1995) further contribute to the research of this topic by the improving the methodology. They formally introduce the migration participation equation, which enables them to control for the sample selection problem (incidental truncation) when estimating the impact of remittances on rural household income. The theoretical framework on this topic came to its maturity before studies on rural China started in the late 1990s. Zhao (1999) first conducted a case study on migration decision and earning differences in Sichuan province in southwestern China.

He finds that earnings received by migrant workers are much higher than what they would receive as workers in local farm or nonfarm sectors. But internal migration in the studied area has high fixed costs. Du, Park and Wang (2005) restudied this issue using data from four western provinces. They observe that the rural policy reform initiated in the late 1990s benefits rural households disproportionately, in the sense that the poorest rural households still find migration extremely costly, while middle-class rural households benefit a lot from lower agricultural taxes, more transportation subsidies, better working conditions and higher labor insurance coverage. Thus, migration can enlarge income inequality under this scenario. Zhu and Luo (2009) study the impact of migration on income inequality in Hubei province, situated in central China. In their sample, nonmigrant households have richer land resources and higher initial household wealth, but remittances received by migrant households are more than home production earnings of nonmigrant households, thus participating in rural-urban migration is both individually and socially optimal.

3. Data Description

The data used in this study comes from a China regional poverty monitoring survey, lasted consecutively from 1999 to 2004. The survey was supported by the World Bank and conducted jointly by the National Bureau of Statistics of China and the China Poverty Relief Office in Gansu and Inner Mongolia Province, which are two large and less-developed provinces in northwestern China (See Appendix 1 for a poverty map). The

dataset actually used is a two-period pooled cross sectional dataset composed of the 2001 and 2004 survey (Wave 3 and Wave 6).[6]

The 2001 and 2004 survey covered 700 and 1500 rural households, respectively. The surveyed households were located in 150 administrative villages in 15 poverty-stricken counties, subject to independent random sampling with no time or spatial correlation between two years. This sample is representative of the low-income rural population in this region, which is the population of interest in this study. Among the 700 households surveyed in 2001, 295 households participated in rural-urban migration, while the other 405 households did not. Among the 1500 households surveyed in 2004, 633 households sent out at least one migrant worker, while the other 867 households opted to stay in rural area. The share of the migrant households in 2001 and 2004 was 42.1% and 42.2%, respectively. All the 928 migrant households in the sample received a strict positive amount of remittances, while the 1272 nonmigrant households received no remittances.

The survey provides quite comprehensive household level information, but does not collect detailed individual information for every member in the household. The numbers of household laborers in three age brackets (age 16-30, age 31-50 and older than 50) are used to control for household's age structure. We use the number of household members with an educational attainment above 9th grade to control for household's education level and unobserved preference for education. In rural China, children receive governmental subsidies that cover part of their expenditures on education until 9th grade. After 9th grade,

[6] The data of this program is non-public and can only be used by authorized personnel. The 2001 and 2004 data were released for pedagogical purposes, other data remains undisclosed and is unavailable upon request. An introduction of the program in Chinese can be found at http://www.wl.cn/895306.

they are no longer eligible for the benefit and the households will face a different budget constraint if they make further investment in child education. So household's education preference can be revealed by the number of relatively well-educated members in the household. Table 1 reports the mean of the key variables for the full sample and for the two subsamples characterized by migration decisions.

Table 1 Selected Descriptive Statistics

	All Households	Nonmigrant Households	Migrant Households
Total Household Income (CNY) ***	9648.75	10646.29	8284.62
Laborers of Age 16 to 30***	0.61	0.47	0.79
Laborers of Age 31 to 50*	0.76	0.74	0.79
Laborers of Age above 50***	0.55	0.49	0.62
Members with Schooling above 9***	0.16	0.18	0.14
Members with Job Training**	0.14	0.16	0.13
Arable Land Area (m^2) ***	1933.98	2458.99	1209.79
Number of Children under 6**	0.34	0.44	0.19
Remittances (CNY) ***	1157.84	0	2744.87
Home Production Earnings (CNY) ***	7622.53	8638.29	4592.16
Number of Observations	2200	1272	928

Note: (1) 1 CNY = 0.16 USD
(2) Home production earnings are defined as earnings from local units, including both the income from farming activities and the income from being employed by local enterprises.
(3) ***Difference significant at 1% , **significant at 5%, *significant at 10%

It is shown that migrant households in this sample are poorer than nonmigrant households with 28.5% less total household income on average. Migrant households have 68.1% more young laborers with age between 16 and 30, but own 103.3% less arable lands than nonmigrant households on average. Migrant households also receive less

education and vocational training. These are important indicators that migrant households in this sample are initially less endowed and might be forced to participate in migration by sending out surplus laborers. The fertility decision of nonmigrant households to bear more children further reflects that they are located at the upper half of the income distribution.

4 Estimation Strategies

The substituting relationship between home production earnings and remittances is plausible for the rural households in China, which verifies the applicability of the Adams approach. Participating in migration always incurs a high fixed cost and is typically considered as a critical alternative for working at home, including both farm and nonfarm productive activities (Zhao, 1999). Because most of the migrant workers are young males with relatively high physical working ability, the household would have to face a large decline in the household productivity if it chooses to send out a migrant worker (Li, 2011). Thus, there are obvious opportunity costs of getting remittances.

In this study we admit the fact that households must sacrifice some positive amount of home production earnings to participate in migration and receive remittances, and vice versa. Two economic counterfactual groups are created. In the first group, the 928 migrant households are assumed to be nonmigrant households and receive home production earnings instead of remittances. In the second group, the 1272 nonmigrant households are assumed to be migrant households and receive remittances at the expense

of some home production earnings. Thus remittances and home production earnings are treated as substitutes, which is the core of the Adams approach. These constructions are based on the assumption that rural households would always make working decisions between migration and home production, so there are no idle laborers, which is not implausible given the working motivation and ability of the rural young laborers in China.

In the first counterfactual group, we know the income of the actual 1272 nonmigrant households, but not the potential home production earnings for the 928 migrant households. Therefore this part of (counterfactual) income for the migrant households needs to be simulated. By the same argument, we need to simulate the potential remittances for the 1272 nonmigrant households in the second counterfactual group. By comparing the income level and income distribution of the actual group with those of the two counterfactual groups, we can determine whether participating in migration is the first-best choice for migrant households, and symmetrically, whether staying in rural areas is optimal for nonmigrant households. [7]

To simulate the home production earnings under the first counterfactual scenario, the unbiased estimators of the coefficients of the nonmigrant household income model are required. The nonmigrant household income model takes a linear form and is written as

$$\log(Y_i^{nr}) = \rho D_i + \beta X_i + \delta D_i G_i + e_i \qquad (1)$$

[7] An alternative approach is to use propensity score matching, which requires many details to be filled in such as how to model the score and how to do inference. However, the implementation of matching is not yet standardized and different researchers might reach very different conclusions, even when using the same data and covariates (Angrist and Pischke, 2009). We thus believe regression should still be the starting point for this work.

where Y_i^{nr} is the total household income for household i, the superscript nr indicates that all households are assumed to be nonmigrant households and receive no remittances. D_i represents the year dummy set to one if the household was surveyed in 2004, to allow for different intercepts for two periods. X_i is a vector of control variables, including household's information on initial endowments and its demographics. G_i is a subset of X_i and contains variables that may affect household income differently across the two periods, such as educational level and job training. e_i is the disturbance term. Because we are now in the first counterfactual scenario, the actual Y_i^{nr} is unknown for the 928 migrant households. We first estimate the actual nonmigrant subsample with 1272 observations to get the unbiased coefficients for the income equation, and then apply these coefficients to predict the home production earnings for the 928 migrant households, giving their characteristics. Since the nonmigrant households might be self-selected into home production activities, this subsample estimation may incur the sample selection problem which may prevent us from obtaining unbiased estimates. Following the Heckman's two-step method for selection bias correction, a participation equation that incorporates the exclusion restrictions is estimated, to test for the existence and magnitude of potential sample selection bias. The participation equation is written as:

$$R_i = \rho D_i + \alpha Z_i + \delta D_i K_i + v_i \qquad (2)$$

R_i is the home production dummy which equals unity if the household engages in home production and receives no remittances, and equals 0 otherwise. Z_i is a vector of

exogenous variables that contains all the control variables in X_i and some variables that affect household's home production decision but not their income directly. K_i contains all the variables in G_i plus the exclusive restrictions. The Inverse Mills Ratio (IMR) λ_i is then calculated from the participation equation and added into Equation (1) as a regressor, which yields:

$$\log(Y_i^{nr}) = \rho^c D_i + \beta^c X_i + \delta^c D_i G_i + \lambda_i + u_i \quad (3)$$

The selection-bias-corrected coefficients ρ^c,β^c and δ^c can now be used to simulate the deterministic part of potential home production earnings for migrant households. For the unobserved residuals, a random value is generated based on the observed error term of nonmigrant households:

$$\mu_i = \varphi_i \Phi^{-1}(r) \quad (4)$$

where φ_i is the estimated disturbance of nonmigrant households, r stands for a random number between 0 and 1, and Φ^{-1} is the inverse of the cumulative probability function of the standard normal distribution (Zhu and Luo, 2009).

The estimating process is much the same for the second counterfactual group in which all households are assumed to be migrant households. We acquire unbiased coefficients of the wage equation from the actual migrant subsample and apply those coefficients to simulate the unknown household income for nonmigrant households, assuming they received some remittances. The migrant household income model takes the form:

$$\log(Y_i^r) = \rho D_i + \beta X_i + \delta D_i G_i + e_i \quad (5)$$

Y_i^r now stands for the total household income with superscript r indicating all the households are now assumed to be migrant households. Using the actual migrant household subsample with 928 observations, a first stage migration participation equation is estimated, the inverse mills ratio is then calculated and added into Equation (5), which yields:

$$\log(Y_i^r) = \rho^w D_i + \beta^w X_i + \delta^w D_i G_i + \lambda_i + u_i \quad (6)$$

Subsequently, the selection-bias-corrected coefficients ρ^w, β^w and δ^w are used to simulate the potential remittances for nonmigrant households. Unobserved residuals are constructed using the observed residuals of migrant households and following the same procedures in the first counterfactual.

5. Estimation Results

We first examine the counterfactual scenario in which all households are assumed to be nonmigrant households, and receive some positive amount of home production earnings. This counterfactual scenario is designed to explore whether migrant households would earn more had they chose to and work locally and not to migrate. Table 2 presents the selected estimation results for Equation (2). Recall that the econometric purpose of estimating this equation is to obtain the inverse mills ratio for the subsequent selection-bias correction. Furthermore, the estimators can be used to verify the demographic

patterns of migrant households (and symmetrically nonmigrant households), complementary to the sample statistics in Table 1.

The number of adult laborers is shown to be negatively related to the probability of engaging in home production, regardless of the laborers' age. This relationship can be well-explained by the large amount of surplus laborers in rural China area at all age levels. Since rural industrial sectors are undeveloped and have very limited capacity to absorb surplus laborers, rural-urban migration is the only option for surplus laborers to get a job with better payment. Hence, for an average rural household, more laborers often imply a higher probability of sending out migrant workers.

The positive relationship between the number of relatively well-educated members and the probability of home production is consistent with the fact that migrant households are less educated on average in the studied area, which can make them disadvantageous in competing with nonmigrant households on local resources. Rural job training programs serve as a less costly substitute for formal education for these migrant households. These vocational training programs mainly focus on low-skill service jobs, such as truck driving, mail sorting and haircutting (Li, 2011). The trainings are fairly job-oriented and most of the trainees can be employed in urban servicing sectors after they graduate, resulting in a larger proneness of participating in rural-urban migration.

Table 2 Results of the Home Production Participation Equation

	Coefficient	Std. Error
Laborers of Age 16 to 30	-0.46***	0.05
Laborers of Age 31 to 50	-0.82***	0.07
Laborers of Age above 50	-0.32***	0.07
Members with Schooling above 9	0.64***	0.11
Members with Job Training	-0.99***	0.29
Marriage (=1 if married)	1.12***	0.25
Number of Children under 6	- 0.08	0.06
Arable Land Area (m^2)	0.0005***	0.0001
Distance to the nearest bus station	0.26***	0.06
Distance to the nearest town center	0.12***	0.05
Year (=1 if surveyed in 2004)	-3.06***	1.31
Constant	4.06	0.52
Observation	2200	
Log-likelihood	-662.61	
Pseudo R^2	0.25	

Note: (1) Other control variables include interactions of year dummy with education, training, arable land, and two distance variables.
(2) Standard errors in parentheses; * significant at 10%, ** significant at 5%, *** significant at 1%.
(3) The two exclusion restrictions passed the F-test for weak instruments with a F statistic of 13.26.

Households who possess more arable land tend to input more human capital resources in home production, especially in farming activities. In China, lands are nationalized and non-tradable. The only way for rural households to obtain arable land is to sign the land leasing contract with local government through the "Housing Responsibility System". Thus, households cannot purchase or sell arable lands to change their initial "land endowment" in the short run. Lastly, marriage tend to keep the workers

stay in rural area, while number of young dependents does not seem to affect one's working decision significantly.

The exclusion restrictions used in the participation equations (the first stage of the Heckit Method) are the distance to the nearest bus station and the distance to the nearest town center. In the two surveyed provinces, infrastructure, including the road system, is extremely undeveloped, due to historical and geographical reasons. Thus the transportation cost is fairly high for migrant workers, considering their circulatory pattern. This indicates that the distance to the nearest bus station and town center can substantially affect household's migration decision, which is proven by the positive and statistically significant results in Table 2. It should be admitted that, in theory, these two distance variables could be expected to affect rural household income. But the surveyed area is remote and home production is largely characterized by traditional and autarky agriculture, so it is economically valid to ignore the feeble influences that the distance variables might have on household incomes through channels other than internal migration. Test of overidentifying restrictions is shown in Appendix 2.

Table 3 reports the selected estimation results of Equation (1) and Equation (3) for the subsample composed of 1272 nonmigrant households, with and without the IMR. The statistical significance of the inverse mills ratio indicates the existence of the sample selection bias, albeit the fact that the selection-bias corrected coefficients are not substantially different from the basic OLS coefficients. The positive IMR coefficient indicates that the nonmigrant households are above average in their unobservable characteristics with regard to the migrant households in the sample. Most of the

statistically significant results are consistent with the economic reality in rural China. More laborers lead to more household income, while the middle-aged laborers (often the household heads) tend to contribute the most to the household income. The number of young dependents has negatively affects the household income. More arable lands are indicators of higher yields from home production, which naturally results in higher household income. Job training and marital status show no significant impacts on household income in this sample.

Table 3 Selected Results of the Nonmigrant Household Income Model

	OLS without IMR Equation (1)		OLS with IMR Equation (3)	
Laborers of Age 16 to 30	0.02	(0.02)	0.07***	(0.02)
Laborers of Age 31 to 50	0.17***	(0.02)	0.15***	(0.03)
Laborers of Age above 50	0.14***	(0.03)	0.12***	(0.03)
Members with Schooling above 9	0.11***	(0.03)	0.11***	(0.03)
Members with Job Training	-0.05	(0.31)	-0.09	(0.31)
Marriage (=1 if married)	-0.21***	(0.06)	-0.14	(0.09)
Number of Children under 6	-0.13***	(0.03)	-0.13***	(0.03)
Arable Land Area (m^2)	0.0003***	(0.00004)	0.0003***	(0.0004)
Year (=1 if surveyed in 2004)	-0.17**	(0.08)	-0.26**	(0.12)
Inverse Mills Ratio	N/A	N/A	0.25***	(0.07)
Constant	8.95***	(0.87)	9.05**	(0.15)
Observation	1272		1272	
Adjusted R^2	0.34		0.34	

Note: (1) Other control variables include interactions of year dummy with education and training,
(2) Standard errors in parentheses; * significant at 10%, ** significant at 5%, *** significant at 1%.

The selection-bias-corrected coefficients from Equation (3) can now be applied to predict the potential total household income for the 928 migrant households had they received some home production earnings instead of remittances.

For the second counterfactual group in which all nonmigrant households are assumed to be migrant households, the estimation results for the first-stage migration participation equation are exactly the same as the results from Equation (2) in magnitudes, but with opposite signs, because the only change is the dependent dummy variable, which now equals 1 if household participated in internal migration. Analogously, the inverse mills ratio is obtained and added into the second-stage household income regression as an independent variable.

Table 4 presents the selected estimation results of Equation (5) and Equation (6) for the 928 observed migrant households, with and without the IMR.

Table 4 Selected Results of the Migrant Household Income Model

	OLS without IMR Equation (5)		OLS with IMR Equation (6)	
Laborers of Age 16 to 30	0.09***	(0.02)	0.11***	(0.02)
Laborers of Age 31 to 50	0.11***	(0.03)	0.15***	(0.03)
Laborers of Age above 50	0.11***	(0.03)	0.14***	(0.03)
Members with Schooling above 9	0.19***	(0.04)	0.28***	(0.03)
Members with Job Training	0.03	(0.05)	-0.09	(0.31)
Marriage (=1 if married)	-0.09	(0.19)	0.05	(0.09)
Number of Children under 6	0.003	(0.14)	-0.13***	(0.03)
Arable Land Area (m^2)	0.0002***	(0.00004)	0.0003***	(0.00004)
Year (=1 if surveyed in 2004)	-0.24**	(0.14)	-0.27*	(0.19)
Inverse Mills Ratio	N/A	N/A	-0.42***	(0.09)
Constant	10.45***	(0.31)	11.68***	(0.46)
Observation	928		928	
Adjusted R^2	0.32		0.32	

Note: (1) Other control variables include interactions of year dummy with education and training,
(2) Standard errors in parentheses; * significant at 10%, ** significant at 5%, *** significant at 1%.

Sample selection bias is similarly obvious for migrant households given the statistical significance of the inverse mills ratio. By comparing the magnitude of the coefficients in Table 3 and Table 4, it is clear that the effect of the number of laborers on household income is largely the same for both migrant and nonmigrant households. Same argument applies for the effect of arable land areas on household income. This confirms that migrant households are not less productive than the nonmigrant households on the margin or facing a structurally different production function. For them, the most plausible reason of a lower household income is they could not make the surplus laborers more productive by allocating them to local jobs.

The selection-bias-corrected coefficients for the migrant household income equation can now be applied to predict the potential household income for nonmigrant households had they received remittances at the expense of less home production earnings. Table 5 reports the household income for the actual sample and the simulated household income for the two counterfactual samples.[8]

Table 5 Average Household Income in Three Scenarios

	Total	Nonmigrant	Migrant
Actual Sample	9648.75	10646.29	8284.62
Counterfactual 1	10228.31	10646.29	9655.37
Counterfactual 2	8931.42	9403.31	8284.62

[8] The simulations are based on the assumption that migrant households would have the same marginal productivity in home production as nonmigrant households, and vice versa. This assumption is arguable and the simulated income of nonmigrant households could be underestimated, because in the sample, they are better endowed and more likely to be well-informed and risk-resilient, which could make them earn more from migration than the actual migrant households. Symmetrically, the simulated income of migrant households could be overestimated.

The reported numbers suggest that migrant households could earn 16.5% more on average by switching to local jobs. In contrast, the nonmigrant household income would drop by 11.7% if they chose to participate in internal migration. Overall, the entire region is better off with higher total household income if counterfactual 1 is realized. It should be noted that the reported results ignore the human capital accumulation of the migrant workers and the long-run effect that migration might have on the development of the source areas. It has been documented that migrants may bring back to the home country increased skills and knowledge that could only be picked up in cities but are transferable to the home environment (Stark *et al*. 1997). But it is more likely that it will be high-skill individuals working in creative and dynamic sectors of the economy that will contribute, upon return, to the development of the home areas (Rosenzweig, 2005). Hence, the results should still be reliable given the fact that most of the migrant workers in northwestern China are low-skilled individuals.

We next present evidence that shows the switching from internal migration to home production is not only individually optimal, but also socially desirable in terms of income distribution.

1.6 Inequality Analysis and Policy Implications

Income inequality is studied by decomposing the Gini coefficient into three easily-interpretable terms, following the seminal works by Lerman(1985) and Stark(1986). The differences between these inequality indicators of the observed group and those of the

two counterfactual groups can be used to illustrate the impact of different job choices on rural income distribution, while also provide evidences for its underlying reasons.

We write the Gini coefficient for rural household income as a function of the covariance between income and its cumulative distribution, which is:

$$G = \frac{2\mathrm{Cov}[y,\mathrm{F}(y)]}{\mu} \tag{7}$$

where G is the Gini coefficient of total household incomes for all the 2200 surveyed households. $\mathrm{F}(y)$ is the cumulative distribution of total incomes y, and μ denotes the average income for 2200 households. Utilizing the properties of the covariance, Equation (7) can be written as

$$G = \frac{2\sum_{k=1}^{K} \mathrm{Cov}[\mathrm{y_k},\mathrm{F}(y)]}{\mu} = \sum_{k=1}^{K} \mathrm{S}_k\mathrm{G}_k\mathrm{R}_k \tag{8}$$

Where S_k is the share of income component k in total household income, i.e. $S_k = y_k/y$; G_k is the Gini index corresponding to income component k; and R_k is the Gini correlation of component k with total income. Equation (8) enables us to decompose the role of remittances and home production earnings in inequality into three terms:

(a) S_k : the magnitude of remittances/home production earnings relative to total income;

(b) G_k : the inequality of remittances/home production earnings, and

(c) R_k : the correlation of remittances/home production earnings with total income.

Table 6 shows the results of Gini decomposition for the three groups. The first panel of Table 6 shows the result of the traditional Stark (1985) approach where we can

obtain a measure of the overall impact of remittances upon village income inequality. Total Gini coefficient is at the level of 0.33. The remittances are distributed quite unevenly across migrant households, with a Gini coefficient of 0.74. Referring to the third panel, the simulated remittances are even more unequally distributed, with a Gini coefficient of 0.9, and the total income distribution with Gini coefficient of 0.4 is significantly less desirable than the actual income distribution with a Gini coefficient of 0.33.

Table 6 Decomposition of the Rural Household Income Inequality

	Share in total income (S)	Gini for income component (G)	Gini correlation with total income (R)	Contribution to Gini coefficient (SGR)
Observed Group				
Remittances	0.12	0.74	0.02	0.04
Home Production Earnings	0.79	0.41	0.90	0.28
Other Income	0.09	0.41	0.42	0.01
Total Income	1.00	0.33	1.00	0.33
Counterfactual Group 1				
Remittances	N/A	N/A	N/A	N/A
Home Production Earnings	0.91	0.34	0.90	0.28
Other Income	0.09	0.41	0.49	0.03
Total Income	1.00	0.31	1.00	0.31
Counterfactual Group 2				
Remittances	0.72	0.90	0.32	0.21
Home Production Earnings	0.19	0.51	0.88	0.17
Other Income	0.09	0.41	0.45	0.02
Total Income	1.00	0.40	1.00	0.40

Note: Other income is calculated by subtracting remittances and home production earnings from total household income.

With a dominating share of 72% in the total income, remittances now contribute 0.21, which is over 50%, to the total income inequality. The second panel shows that without the income component of remittances, the Gini coefficient would drop to a more desirable level of 0.31. A 0.02 decrease in the Gini might not be of strong economical importance, but an increase of the Gini from 0.33 to 0.4 can surely be devastating. The results confirm that by switching from migration to home production, the entire region would be better off (or at least remain stable) in terms of income distribution, and more importantly, we should expect to see a sizable increase in the income inequality if more households choose to participate in rural-urban migration.

Home production earnings are highly correlated with the total household income in all the three scenarios, with the Gini correlation around 0.9, indicating that these earning opportunities are not well-diffused across the entire region and are only taken advantage of by those at the upper end of income distribution. Internal migration, instead, is widely accessible in the studied region with a Gini correlation of 0.02 in the observed sample. This further suggests that households are taking migration as an alternative when local jobs are unavailable.

Given the above findings, two types of policies should be designed to increase the total average income and to decrease the income inequality in the studied area. The first is to aim at achieving the first-best outcome by encouraging households to participate in home production. However, this outcome can only be obtained when rural laborers are allowed to switch freely between home production and migration. Since the agricultural land endowments of rural households are fixed in the short run due to institutional

reasons, as discussed above, many rural laborers cannot participate in agricultural production due to the lack of land. Moreover, the rural low-skill service industries have very limited capacity to absorb surplus laborers. Thus, surplus laborers of migrant households can hardly switch to home production due to the binding constraint of limited agricultural resources and local job opportunities. This constraint also tends to result in a low marginal return to labor input on home production for migrant households, by the law of diminishing marginal returns. Policies of this type should focus on creating equal opportunities for rural households to participate in home production, mainly by developing township and village enterprises (TVEs). TVEs have been proved to be efficient in absorbing rural surplus laborers since most of them are very labor intensive (Naughton, 2007). TVEs often cluster to focus on the entire processing industry for agricultural products in the target area, which can add sizable extra values to agricultural products. Rural household can increase their incomes by selling fresh agricultural products directly to the TVEs at the contracted prices, which also reduce the market risk they may encounter.[9]

The second set of policies, which is more enforceable in the short run, should focus on increasing remittances for the migrant households, mainly by increasing migrant workers' nominal earnings and reducing their migration costs. There are two general ways to increase migrant worker's earnings in China: increasing their wage rate and preventing them from being cheated by the employer. The latter case is by no means a

[9] The Chinese central government issued the "Decision on accelerating the development of the TVEs in middle and western region" in 1993 to officially confirm the benefits and encourage the development of the TVEs. See http://www.mofcom.gov.cn/article/b/bf/200207/20020700031377.shtml for the complete policy.

trivial one in reality. Every year, large number of migrant workers suffers from delayed payments and even no payments due to the dishonesty of their employers and little legal protections they can rely on. Most of the migrant workers have no knowledge on law, so it is the government's responsibility to guarantee a legal employment contract to be signed before any de facto employment takes place. Employers should be legally convicted and punished for their dishonest and unlawful behaviors.

The government is also responsible for reducing the costs of migration, chiefly by reducing the living cost and the transportation cost. In China, the largest part of the living cost for migrant workers comes from the housing rents in urban areas. In the metropolitan areas such as Beijing, Shanghai and Guangzhou, the monthly rent of a common apartment is very likely to exceed a migrant worker's monthly earnings, resulting in the formation of urban slums with poor and unhygienic living conditions for the occupying migrant workers. Thus, government should invest more in constructing the "economically affordable house" which is specifically designed for poor urban households and migrant workers. Migrant workers should also receive rent subsidies and transportation subsidies as received by less-endowed urban residents.[10]

It is important to note that these two types of policies are not mutually exclusive. They can function together to provide both the "equality of opportunity" and the "equality of outcome" for current migrant households in the surveyed area.

[10] The suggested policies are concerned by the "three-nong" problem since 2001 and are included explicitly in the 2004-2009 "No.1 Central Document" issued annually by the Chinese central government.

1.7 Conclusion

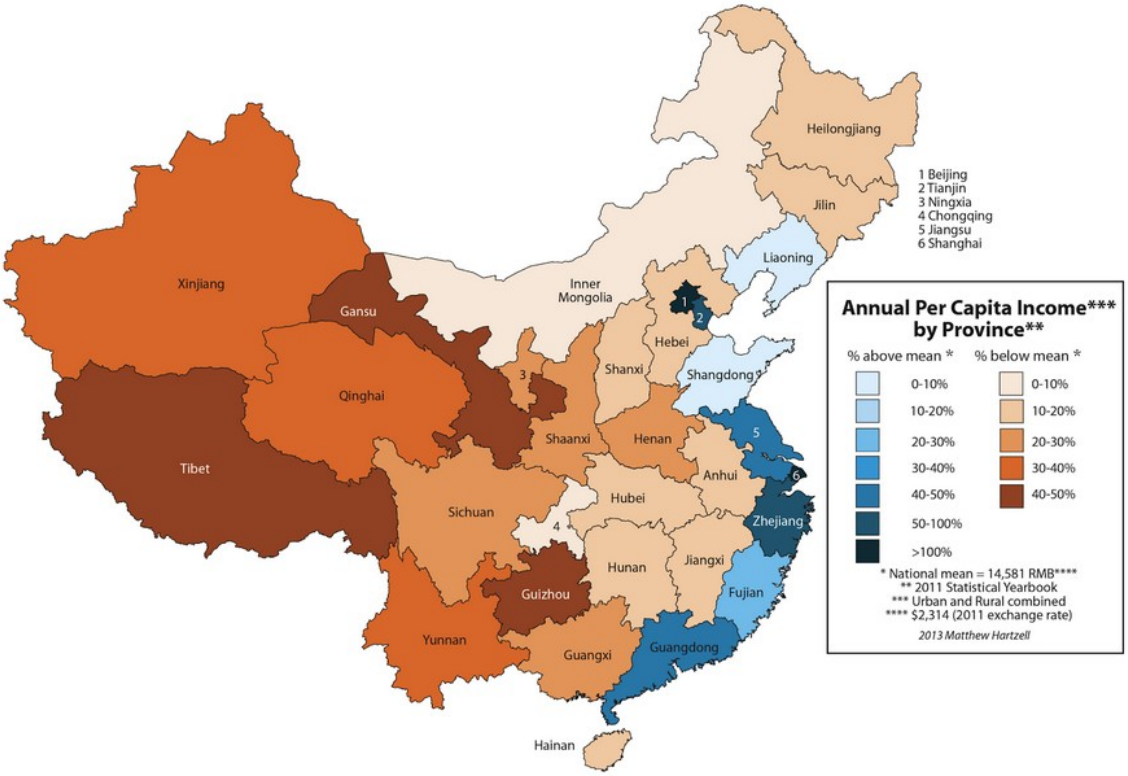
Internal migration in China after the economic reform has played a critical role in increasing household income level and changing income distribution in rural areas. Remittances are the most direct pecuniary rewards from migration, and serve as an important source of income for migrant households. However, there is no consensus on the impact of internal migration, since nonmigrant households and migrant households have very different initial endowments and face different job choice constraints in different areas, while the remittances and home production earnings they receive varies significantly across different regions. This is why most of the studies on this topic are place-based studies.

Using data from two large and undeveloped provinces in northwestern China, this empirical study finds that migrant households in this area are initially less endowed, and the remittances they received by participating in migration are lower than potential home production income they could have earned if they switched to home production. Thus, rural-urban migration is not the optimal decision for migrant households if they can switch costlessly to local jobs. The decomposition of the Gini coefficient also shows that encouraging internal migration is also not the first-best choice for a social planner if she aims to achieve a higher average income and lower income inequality in the studied region. Policies should be implemented to increase migrant worker's remittances by increasing their nominal wages and reducing their migration costs, and protect their rights to be fully and punctually paid. More importantly, government should spend more

budgets on protecting existing arable lands and encouraging the development of the TVEs to absorb the surplus laborers more adequately in the studied area.

Appendix 1: Poverty Map of China

This poverty map is extracted from the World Bank GeoiQ System. It is shown that Gansu province and Inner Mongolia province located at the northwest part of China have annual per capita income below national average, especially Gansu Province.



Appendix 2: Test of Overidentifying Restrictions

Analogous to the test of the validity of overidentifying instruments in an overidentified model using the Hansen-Sargan test, we should perform an overidentifying restrictions (OIR) test in the Heckit model to account for potential biases that" can do more harm than good" (Bound, Jaeger and Baker, 1995). The test is conducted by explicitly including the endogenous decision variable R_i into the structural income model and perform a Hansen-Sargan test, with the two exclusion restrictions (distance to the nearest bus station and distance to the nearest town center) treated as two instruments for R_i. The model takes the form:

$$\log(Y_i) = \rho D_i + \gamma R_i + \beta X_i + \delta D_i G_i + e_i \qquad \text{(A.1)}$$

Y_i now stands for the total household income for all the 2200 surveyed households. In STATA we use the postestimation "estat overid" command following the "ivregress gmm" command to produce the Hansen-J statistic. The test statistic is χ^2 (1) distributed because the number of overidentifying restrictions is 1. The yielded Hansen-J statistic is 1.604 and its p-value is 0.342. Thus we do not reject the null hypothesis and conclude that the overidentifying restriction is valid.

References

Adams, R. 1989. “Worker Remittances and Inequality in Rural Egypt.” *Economic Development and Cultural Change* 38(1): 45-71.

Adams, R. 1994. “Non-farm Income and Inequality in Rural Pakistan: A Decomposition Analysis.” *Journal of Development Studies* 31(1): 110-133.

Barham,B., and S. Boucher.1998. “Migration, Remittances, and Inequality: Estimating the Effects of Migration on Income Distribution.” *Journal of Development Economics* 55(2): 307-331.

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

Du,Y., P.Albert, and S.Wang. 2005. “Migration and Rural Poverty in China.” *Journal of Comparative Economics* 33(4): 688-709.

Davin, D. 1999. “Internal Migration in Contemporary China.” ST. Martin’s Press, New York.

Edward,T. 1999. “The New Economics of Labor Migration and the Role of Remittances.” *International Migration* 37(1): 63-88.

Lorene, Y. 1976. “Internal Migration and Economic Development in Brazil.” *Quarterly Journal of Economics* 76 (2): 218-245.

Li, Q. 2011. “Urban Migrant Workers and Social Stratification in China.” Beijing: Social Sciences Academic Press.

Heckman,J. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica* 47(1): 153-161.

Rosenzweig.M., 2005. " Consequences of Migration for Developing Countries." Paper of the United Nations Expert Group Meeting on International Migration and Development. UN/POP/MIG/2005/08

Stark,O., T.Edward., and Y.Shlomo. 1986. “Remittances and Inequality.” *The Economic Journal* 96(383): 722-740.

Stark,O., T.Edward., and Y.Shlomo. 1988. “Migration Remittances and Inequality: A Sensitive Analysis using the Extended Gini Index.” *Journal of Development Economics* 28(3): 309-322.

Stark, O., C. Helmenstein and A. Prskawetz (1997). A Brain Gain with a Brain Drain." *Economics Letters* 55: 227-34.

Tumbe, C. 2011. “Remittances in India: Facts and Issues.” Indian Institute of Management Bangalore Working Paper, No. 331

Zhao,Y. 1999. “Labor Migration and Earning Differences: The Case of Rural China.” Working Paper, University of Chicago.

Zhu,N., and Xubai.L. 2009. “The Impact of Migration on Rural Poverty and Inequality: A Case Study in China.” *Agriculture Economics* 41:191-204.

作者信息

华岳

助理教授

湖南大学经济与贸易学院

电话：(0731)88684750

电子邮件： huayue_econ@126.com