Essays on Rural-Urban Migration in China

Weijia (Vivian) Chen

Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Economics

Dennis Yang, Chair

Djavad Salehi

Amoz Kats

Jeffrey Yau

Richard Ashley

May 10, 2006

Blacksburg, Virginia

Keywords: Rural-Urban Migration, Temporary Migration, Education, Migrant Networks, China

Essays on Rural-Urban Migration in China

Weijia (Vivian) Chen

(ABSTRACT)

Since the late 1980's, China has experienced the world's largest peacetime out-migration of its rural labor force to urban areas. The temporary nature of the labor migration complicates the control on this mobile population, and its multi-faceted influence on the whole economy makes the migration policy controversial. Based on cross-sectional Chinese rural household survey data, this study analyzes the effects of migration on rural areas and explores the determinants of the participation and duration of the temporary migration.

The first chapter investigates how parental migration affects the decision of enrolling children in high school through migration's effects on household income and the opportunity cost of schooling in rural China. The opportunity cost of schooling is approximated by the marginal productivity of children imputed from family production estimation, which controls for potential endogeneity in the time allocation decisions of family members. The empirical results show that temporary migration of parents raises their children's probability of high school enrollment by 3.2%, resulting primarily from a positive income effect. These findings suggest that reductions in barriers to migration raise rural household earnings, and foster the investment in children's education.

The second chapter studies the determinants of participation and duration of temporary rural-urban migration in China highlighting the role of education and migrant networks. The Probit and Logit models are fitted to the dichotomous migration participation estimation. To correct for the sample selection bias, Heckman's two-step procedure is used to estimate the length of migratory work. Empirical results confirm the existence of a migrant network effect on both migration participation and migration length. Schooling increases migration probability non-linearly and its effect on migration length is insignificant once migration is controlled. Furthermore, the positive effect of migrant networks on migration participation is especially prominent among individuals with junior and senior high school education.

Acknowledgements

I am particularly indebted to my principal advisor, Professor Dennis Yang, for his support, help, patience and care. His influence on me, both academically and spiritually, is lifelong. The year I applied for the Ph.D. program in Economics, I got several offers from other universities, I chose to come to Virginia Tech because I was convinced by Professor Yang through emails that I could get sufficient guidance and support in this program, though at that time I had only vague notions of working on labor and development issues. It turns out that I made a right choice. I would like to thank Professor Djavad Salehi for his constant encouragement and incisive comments. Professor Salehi's insightful critiques on my presentations at the labor/development lunch group contributed to the improvement and completion of my paper. Professor Amoz Kats impressed me first by his graphic approach to the optimization problems. His humor and wisdom in teaching sparked my interest in mathematical economics. I am grateful to Professor Kats for his countless patient help both on my course work and on my theoretical model in the dissertation, for all the confidence he instilled and for every smile he inspired. Although Professor Jeffrey Yau joined Virginia Tech in my last year, his dedication as my committee member and especially the help he provided and the experience he shared when I was on my job market has led to my success in job searching. Being the director of the graduate study and my committee member, Professor Richard Ashley provided me with valuable support and advice which is central in the successful completion of my Ph. D. degree.

I would like to thank all my friends for the constant support, warm encouragement, and unlimited trust during this journey. Although they are too many to be listed, I'd like to mention at least one of them: Jungshik Jason Hur. I cordially thank Jason for all the nights he spent helping me figure out my SAS program, every complaint he listened to and all the doubts he dissipated when I felt down.

Finally, from the bottom of my heart, I would like to dedicate this dissertation to my parents: Weiqiong Cai and Qijun Chen. Their unconditional love and mental support sustained me through all the difficulties and motivated me to excel at every stage of my study. Their spiritual influence on me is profound. They taught me how to be a person with great personality and the life principles I strictly abide by, which, I believe, are the utmost important and will benefit me throughout my whole life no matter what I do and where I am.

Table of Contents

List of Tables

Chapter 1

Migration Opportunity and Children's High School Enrollment in Rural China

1.1 Introduction

The private and social benefits of schooling in developing and industrialized economies are widely recognized. Schooling has been convincingly linked to higher earnings, better health and nutrition, greater labor productivity, and faster economic growth (Psacharopoulos and Woodhall, 1985; Schultz, 1988). Improvements in the level of broad-based educational attainment are also associated with greater economic equality (Birdsall, Ross and Sabot, 1995). Thus, it is not surprising that governments have been willing to spend a substantial part of national income in promoting higher levels of educational attainment so as to keep the impetus of modern economic growth; neither is it hard to understand why parents have set aside an increasing amount of their private disposable income to school their children, foregoing the productive contribution that the children would have made to family income had they not attended school. In poor rural areas, where home production is the major family income source, the decision to enroll a child in school is significantly influenced by family budget constraints and the opportunity cost of attending school. If family members engage in new, non-family employment whose earnings are higher than conventional family production[1], the

[1] Family production includes a wide range of activities carried out at the household level, such as household agricultural production, and industrial and sideline activities including industry, construction, transportation, services, commerce, restaurants and catering, handicrafts and sideline activities.

resultant relaxed budget constraint may allow families to invest more in children's education. On the other hand, new employment activities may raise the opportunity cost of attending school because the outflow of adult labor raises the marginal productivity of children's labor in family production. Investigating the magnitude of these two competing forces for children's school enrollment is especially important for developing countries with labor market distortion and restrictions on labor mobility. Such an investigation will provide policy makers with information about the potential effect of migration on young generation's education in rural areas, which is a decisive factor in the reduction of rural-urban income disparity and a nation's long-run economic development.

As an urbanization process, the large amount of rural-urban migration in China has attracted the interest of many researchers. Compared with other developing countries, migration in China is characterized by two specific features. The first feature is the combination of gradually lifting restrictions and still--existing barriers on labor mobility. The second feature is the predominant circular (as opposed to permanent) migration pattern, which reflects the fact that most migrants in China are individuals who leave their families behind while keeping close contact with them. This homeward tendency is in line with the premise of the "new economics of labor migration" (NELM) literature, which analyzes migration as part of interrelated family decisions (Stark, 1991; Taylor et al., 1996). Most studies on migration in China focus on the determinants of migration, the effects of migration on destination areas and the consequent social impacts. Studies on the effect of migration on source communities are limited to its impact on household income, agricultural production (Alan de Brauw, J. Edward Taylor, and Scott Rozelle 2001) and capital investment. Recently, de Brauw and Giles (2005) investigated the

impact of migration on the educational attainment of youth in rural China. They used national identification cards to identify migrant opportunity and found that a decline in the cost of participating in migrant employment leads to a decrease in the probability that children will attend high school in rural China. The explanations they provided for this school enrollment decline include the low rate of return to high school education in migration work and the increasing opportunity cost when migration work becomes more accessible and local labor force decreases due to the growth of migrant workers. To my knowledge, this is the first and the only study that addresses the effect of migration on children's education. However, the authors' emphasis is on developing an instrumental variables approach to identify the impact of migration on source communities; they do not quantitatively analyze the channels through which migrant labor networks affect high school enrollment decisions.

My research further explores the effect of parental migration on children's high school enrollment in rural China. In China, the percentage of full-time students among high school age individuals is below 40%. High school tuition and other costs associated with it are among the most important reasons why parents do not enroll their children in school[2]. Furthermore, in many developing countries in which the labor market is incomplete and self-employment is the main form of labor input in home production, teenagers are important contributors to family production, wherein marginal productivity determines their time value. In this paper, I quantitatively analyze how changing family income and children's time value in family production due to parental migration affect children's high school enrollment in rural China.

[2] The proportion of high school tuition to household income ranges between 11%-15%, the total education expenditure is as high as 30%-40% of mean income in rural areas (Heckman 2003).

The literature in children's education has a rich tradition in investigating the determinants of schooling decisions. Though it is widely recognized that children's opportunity cost of time, their "shadow wage", directly affects parents' decisions on their children's labor supply and schooling[3], few studies have explicitly specified the shadow wage in empirical estimation. Among those few, only aggregate level, village or community level child wages are used. This neglect is largely due to the absence of child or teenager wages in the data sets, as most children are self-employed in household production, which does not offer an explicit wage rate. Children's shadow wages at the household level are important in the sense that they represent the time value of the children and are regarded as the opportunity cost of schooling. This paper contributes to children's schooling literature by estimating a household-level opportunity cost of schooling and examining its effect on high school enrollment decisions. In this paper, the opportunity cost of schooling is assumed to be children's contribution to family production, i.e., their marginal productivity in production, and is derived from household production estimation after correcting for the potential endogeneity in the time allocation decision of family members.

The data used in this paper comes from the Chinese Household Income Project (CHIP). In addition to detailed economic and demographic variables, this data set contains records of migration work, and sectoral time allocation and labor compensation for every working member of the households. These data allow for migration identification and activity-specific income and labor supply estimation. The empirical results indicate that on average, to be a migrant household, a household in which at least one parent is a migrant increases children's high school enrollment probability by 3.2%;

[3] See Rosenzweig and Evenson (1977).

the increased income raises the probability by 3.23%, while the increased opportunity cost decreases the probability by 0.032%.

The paper proceeds as follows. Section 2 introduces the background of the Chinese migration and education system and briefly reviews the relevant literature. Section 3 briefly outlines the theoretical framework for studying schooling enrollment. The empirical specifications and econometric issues are discussed in section 4. Section 5 describes the data and variables. The estimation results and explanation are presented in Section 6. Section 7 summarizes the findings and discusses policy implications.

1.2 Background and Literature

1.2.1 Rural-Urban Migration in China

Since the mid-1980s, rural China has experienced the largest peacetime out-migration of its labor force ever witnessed in world history (Solinger, 1999; Rozelle et al., 1999). Though the exact number of migrants is still hard to ascertain, due to both a lack of consensus in the definition of migrants and the absence of an authoritative national survey, migrants' prominent presence in Chinese urban areas and their fast-growing number are hardly disputable. A commonly-cited figure puts the number of rural migrants residing in urban areas at 50 million in the mid-1990s. Another independent survey, by the Ministry of Agriculture (MOA), puts the estimate at 75.5 million for 2000, which accounts for about 11 percent of the total labor force in China. Based on the 2000 census, Cai (2003) offered an estimate of 77 million rural-to-urban migrants for that year. Currently about 100 million, that is 15% of its rural labor force is engaged in rural-urban migration.

In the pre-reform era (before 1979), China adopted the heavy-industry-oriented

development strategy and had for decades tightly restricted rural-to-urban migration. The results were severe segmentation and large income gaps between rural and urban areas. One important method used to achieve the segregation was the denial of urban *hukou* (household registration), which designates a person's legal place of residence and work and one's access to various amenities and social services such as health care, schooling and (until recently) rationed or subsidized food products, which were provided only to urban residents.

Because of the inefficiency and the increasing rural-urban income gap associated with labor misallocation, the *hukou* system has been modified in recent years to permit more flexibility in the reallocation of labor between rural and urban markets[4], helping to spur growth in the migrant labor force (Solinger, 1999). However, despite the relaxed regulations on labor mobility, there still exist serious institutional barriers to fully free migration. First, the dual *hukou* system, which underwent some drastic changes in the past decades, has not lost its importance in Chinese cities. A nonagricultural *hukou* was made available for certain rural populations, but it was mostly only available in small towns, not cities (Wang, 1997). Migrants face a complex set of rules and regulations which require them to obtain permission from authorities in their home communities prior to their departure, in addition to getting residence and work permits in their destinations. Fees that accompany all of these procedures are sufficiently high to discourage migration (Zhao, 1995). Second, under the current Household Responsibility

[4] In 1988, the central government initiated a major policy reform that relaxed the controls over rural-urban migration. Farmers were permitted to work and to carry on business in cities provided they could secure their own staples (Forbes and Linge, 1990). This regulation gave new opportunities for rural workers to work temporarily in cities, representing improvements over the old system in which college education provided the only legitimate access to urban registration (Chan and Zhang, 1999). In the early 1990s, the end of food rationing further reduced the costs of living for temporary rural migrants in cities because they no longer had to bring food with them from the countryside. They could purchase food directly without securing rationing coupons. In 1998, the Ministry of Public Security issued another regulation loosening the control of *hukou* registration – those who moved to join their parents, spouses and children in cities could also receive urban registration (Cai, 2003).

System (HRS), farm families have land-use rights but not rights of alienation. If they permanently leave agriculture, farmers must return the land to local authorities and consequently give up a stream of potential land earnings in the future (Yang, 1997). The insecurity faced by farmers who give up their rights to land without compensation reduces their incentive to engage in family migration and splits family labor supply between farm and non-farm employment. Moreover, China's farmland arrangements under the HRS obligate the farm household to deliver a part of its grain output to the state at quantities and prices specified by the government. When rental markets are restricted, the obligation of delivering procurement quotas would reduce the flexibility of family labor allocation to alternative employment. Third, rural migrants are treated differently from their urban counterparts in terms of occupational attainment and wages[5]. Furthermore, migrants are excluded from state-subsidized benefits in urban areas, including medical insurance, housing benefits, and schooling for migrant children (Zhao, 1999; Wang & Zuo 1999; Heckman, 2003)[6].

Under the combined effect of gradually lifting restrictions and still existing barriers, Chinese rural families did not fully migrate to urban areas. This is manifested in two ways. First, the majority of workers and families do not participate in migration at all. Second, the dominant form of migration in China is circular migration: most migrants are individuals, return to their families several times per year instead of bringing their families to the cities after some initial explorations (Zhao, 2003; Hare, 1999), and seldom

[5] Meng and Zhang (2001) conducted a careful study of occupational segregation and wage differentials between urban residents and rural migrants in Shanghai based on two survey data sets containing individual information. They show that around 22 percent of urban residents who would have been better suited for blue-collar jobs were given white-collar employment, while 6 percent of rural migrants who would have been suitable for white-collar jobs were relegated to blue-collar positions. City residents also enjoyed a large wage premium.

[6] Whereas only about 10 percent of rural migrants report having any kind of medical insurance coverage, and less than 5 percent have retirement pension benefits, two-thirds of urban employees have medical insurance, and 80 percent have retirement pensions (Wang & Zuo, 1999). The *hukou* policy charges children of interregional immigrants above-normal fees that are as large as 10% of total family income just for the right to attend school (Heckman, 2003).

assimilate with the urban population. As a result of their "homeward" tendency, migrants are closely related to the welfare of their family left behind and affect other family decisions including decisions on children's education.

1.2.2 School Enrollment: General Framework and Chinese Context

The determinants of school enrollment can be classified into three categories: demand, supply and government policy. Demand indicates the school enrollment decision made at the household level comparing the costs and benefits of schooling. Costs include direct costs, such as tuition and other miscellaneous fees, and the opportunity cost of foregone home production and/or wage income. The major benefit of schooling is the higher future wage rate due to more education. When the environment changes, household decisions about children's school enrollment may also change. Rapid economic growth and new employment opportunity change both the opportunity cost and returns to schooling. The provision and the quality of education are contained in the supply side. In many developing countries, the lack of resources makes schools not available nationally and the quality of schooling is too low to make attendance feasible. The most prevalent government policy is the compulsory education law. Every developed country has its compulsory education law stipulating the starting age for and the length of, fundamental education. Other more general government policies include subsidy and education loans.

In China, the "law of Compulsory Education," established in 1986, officially institutes a system of nine-year compulsory education starting from the age of six throughout the country. The Law also prescribes: "…Local people's governments at various levels shall establish primary schools and junior middle schools at such locations

that children and adolescents can attend schools near their homes…The state shall not charge tuition for students receiving compulsory education." The implementation of a universal nine years of schooling (six years of primary school, three years of junior middle school) was carried out in different phases depending on the level of socioeconomic development (He, 1996). In the dataset, 90% of primary-school-age children and 80% of junior-middle-school-age teenagers are enrolled as full-time students. After junior middle school, families of students who enter high school are required to pay substantial tuition, as China does not heavily subsidize higher level of education, whose return, compared to that of basic primary and junior middle school education, is much lower. Furthermore, student loans are scarce, especially for before- university education, as financial markets and regulations are undeveloped in China. Many children are pulled out of high school because poor households cannot afford to finance the direct costs of education. Admittedly, there are supply considerations for high school enrollment: high schools are not located in all villages or communities so students have to travel a long way to school or even live away from home, but this paper is mainly concerned with the demand side of school enrollment.

In children's schooling literature, even though it is widely recognized that opportunity costs of time is one of the most important determinants in family decisions on children's school enrollment, studies on this issue are scant due to data availability. For labor market participants, this opportunity cost is just their wage rate, which is usually taken as exogenous to the individual in empirical studies. However, the wage rate for child labor is not recorded and reported officially as the majority of working children in rural areas of developing countries are self-employed in household production without

labor compensation. So the absence of child wage complicates the studies on children's time allocation. Studies on adult labor supply encounter a similar situation due to missing exogenous wage rates, as not all adults work in the labor market. The usual approach taken is to estimate the agricultural production function and calculate the marginal productivity of labor, which is regarded as the opportunity cost of leisure, or "shadow wage", then use the "shadow wage" in total labor supply estimation after dealing with potential endogeneity[7]. A similar approach is taken in this paper: the household production function is estimated; then, the marginal productivity of teenager labor is calculated and used as the opportunity cost of schooling, taking into consideration the endogeneity in the simultaneous time allocation decisions of family members.

Previous research on the determinants of school enrollment in China finds that mother's education, occupation and income, and the presence of siblings are all significant determinants of primary school attendance (Wang & Yuan, 1993; Connelly & Zheng 2003); girls appear to be particularly disadvantaged in primary education in some areas of Western China (Zhou, Ma, & Wu, 1995), and even in high school enrollment in a large urban area (Broaded & Liu, 1996); children in families with severe credit constraints are less likely to be enrolled and those from wealthy families have higher test scores and are less likely to repeat a grade (Brown & Park 2002); and location of residence is highly correlated with enrollment and graduation (Connelly & Zheng 2003). Some researchers use the Chinese Household Income Project (CHIP) survey, which is the

[7] Endogeneity arises because the computed marginal productivity of labor depends on the farm labor input (specifically, the more farm labor, the less the marginal product of labor and vice versa), and it is simultaneously determined with the total labor supply. To deal with this endogeneity issue, 2SLS is employed using instruments on marginal product of labor. Jacoby (1993) uses the number of adult male and female farm workers, housework production shifters (availability of electricity, plumbing, etc.) and exogenous agricultural production variables as instrumental variables. Skoufias (1994) uses household assets (value of buildings, consumer durables, land, farm implements, food stocks, financial assets owned by the household, number of bullocks and livestock owned by the household) and village-specific variables such as average rainfall and average hourly wage rates for hired labor as instrumental variables.

survey used in this paper, to investigate education attainment. Knight and Li (1996) focus specifically on the geographical differences in educational attainment, finding that the population 16 years of age and older who reside in urban areas have a substantial advantage in education attainment. Hannum (1999b; 2002a, b) shows an interaction between rural poverty and gender inequality in enrollments, and documents the significance of poverty in producing observed ethnic disparities in enrollment. Knight and Song (2000) explore an intra-household bargaining model and find that boys have a higher probability of enrollment at all levels and that children, especially boys whose mother has a higher education level than the father, are more likely to be in school.

1.2.3 Migration and Children's School Enrollment

Due to the nature of migration in China, migrants affect family welfare and time allocation decisions partly by changing household income. Under the relaxation of labor mobility, some rural households in China decide to send family members to urban areas to earn the higher incomes that are available in those areas. If there is not redundant labor in family production, the outflow of adult labor to migration work may reduce the output of family production. However, as labor productivity is usually higher in migration work than in traditional family production, the income from the urban labor market may exceed the loss in family production, resulting in an increase in total family income. Empirical study shows that migration increases household per-capita income between 14 and 30 percent, taking into consideration the multiple effects of migration on farm income, self-employed income, wage and other income, and remittances (de Brauw, Taylor, Rozelle, 2001). This increase in family wealth has meant that more resources are

available for education and that old teenagers have a higher probability of being enrolled in high school, holding everything else constant.

Table 1.1 shows the income composition between two types of households with high school age children[8]. Income from family production is lower in migrant households than non-migrant households. The average migrant household has around 1,543 yuan (17%) less income from household production than the average non-migrant household. However, migrant households lead non-migrant household in parental non-family wage income by an average of 4,404 yuan (72%). So the overall household income which consists of the income from family production and parental wage income is higher in migrant households than non-migrant households[9]. The income difference pattern between two types of households with high school age children is representative of that of all households[10] in the survey.

On the other hand, the increased opportunity of migration could also change the opportunity cost of young people's time. The depletion of adult labor due to migration may increase the marginal productivity of the remaining labor in household production. In China, where the labor market is still imperfect, the predominant form of labor used in family production is self-employed labor. Given the available number of household laborers, there is now a real benefit for the family to use their children's labor as a substitute for the left adult labor. Moreover, the fact that farm households in China are obligated to deliver a specified part of their grain output to the state may even aggravate

[8] A migrant household is the one in which at least one parent is a migrant. See more explanation later.
[9] Only parents' non-family wage income is included in the total household income, as parental income has a direct effect on their children's education. However, the income from family production is generated by the work of all family members engaged in production because home production is an integrated activity, which cannot be separated on an individual basis.
[10] In appendix table 1, income and its composition are compared between all migrant and non-migrant households in the data set. The differences are similar to those exhibited in the households with high-school-year-old children.

the usage of teenager labor when adult leaves farm production. The increasing opportunity cost of schooling competes with the relaxed budget for children's high school enrollment. The data shows that the migration income effect outweighs the opportunity cost effect: 10% more high-school-year-old children are enrolled as full-time students and 10% less work on family farm or business in migrant households than non-migrant households (Table 1.2).

1.3 Theoretical Framework

Below, I present a simple model to examine the effects of migration opportunity on the decision to enroll a child in high school. The model illustrates two channels through which migration affects school enrollment: the change in household budget constraint and in the opportunity cost of schooling, which is assumed to be the value of children's time in family production. I focus the discussion of the model on the high school enrollment decision.[11]

Consider a one- period model in which a unitary rural household wishes to maximize its utility by choosing consumption and making optimal time allocation for household members. Household members are classified into two types: adults and high-school-year-old children. Adult working time is divided between family production, T_F^{ADT} , and non-family wage employment including migrant work, T_w^{ADT} .[12] Teenager's non-leisure time is spent either in school, *S*, or in household production, T_F^T . If the teenager is enrolled in school, a cost of P_S for tuition and other

[11] This model can also be used for children's school enrollment at different ages.

[12] I specify a composite time, T_w^{ADT} , for adult non-family wage employment which consists of two elements: time spent on local non-family work and time spent on migratory work. This specification is data driven as the original data set does not distinguish between these two types of non-family wage employment.

education related expenses needs to be paid by household.

There are two sources of household income: family production and non-family wage employment in the labor market. Home production may utilize labor of both adults and teenagers, $V = V(T_F^{ADT}, T_F^T; A, K)$, where A and K are household current land and stock of capital, taken as exogenous. Family production is assumed to use family members only; it does not hire outside labor and incur wage payment. The theoretical justification for this assumption includes imperfections in the labor market (Janvry, Fafchamps and Sadoulet, 1991), the high cost of monitoring agricultural production and the existence of returns to specific experience (Rosenzweig and Wolpin, 1985). The labor market wage rate, *w*, is taken as exogenous.[13] I further assume that only adults may be engaged in the labor market, and treat the wage that can be earned in the labor market as net returns to the household from non-family employment. So, the income from the labor market will be wT_w^{ADT}, and the household budget constraint is:

$$V(T_F^{ADT}, T_F^T; A, K) + wT_w^{ADT} \geq c + SP_S \qquad (1)$$

The price of consumption is normalized to one for simplicity. Furthermore, I exclude the possibility of borrowing, as the capital market and financial institutions in most developing countries are incomplete.

I abstract from demand for leisure and assume that household utility is an additively separable concave function of consumption *c*, and teenagers' schooling, *S*. φ is a vector of taste shifters and other household characteristics such as household

[13] Corresponding to the composite non-family working time, T_w^{ADT}, the labor market wage rate, *w*, is also a composite wage rate for both migratory work and local non-family work. In reality, wage rates for migratory work and local non-family work may not be the same and may depend on economic variables such as the working time, location, unoberservable personal and village characteristics.

component, which are not current choice. The household's objective function is to maximize:

$$U(c,S;\varphi),\frac{\partial U}{\partial \Omega}>0,\frac{\partial^2 U}{\partial \Omega^2}<0,\Omega=c,S \tag{2}$$

subject to the budget constraint (1) , two time constraints for adults and teenagers, in which the total non-leisure time for both is normalized to 1:

$$1=T_F^{ADT}+T_w^{ADT},1=T_F^T+S \tag{3}$$

and non-negativity constraints for adult labor days in non-family work, $T_w^{ADT}\geq 0$, and teenagers' schooling, $S \geq 0$. Substituting the time constraints into the family labor in the home production function, the Lagrangean equation is:

$$L=U(c,S)+\lambda[V(1-T_w^{ADT},1-S;A,K)+wT_w^{ADT}-c-SP_S] \tag{4}$$

The first-order conditions are:

$$\frac{\partial L}{\partial c}=U_c-\lambda=0 \tag{5}$$

$$\frac{\partial L}{\partial T_w^{ADT}}=w-V_{T_F^{ADT}}\leq 0 \tag{6}$$

$$T_w^{ADT}(w-V_{T_F^{ADT}})=0$$

$$\frac{\partial L}{\partial S}=U_S-\lambda(V_{T_F^T}+P_S)\leq 0 \tag{7}$$

$$S[U_S-\lambda(V_{T_F^T}+P_S)]=0$$

$$\frac{\partial L}{\partial \lambda}=V+wT_w^{ADT}-c-SP_S=0 \tag{8}$$

where λ is the shadow wage of income. Solving the system of equations yields a school enrollment demand function of the form:

$$S^* = S^*(\lambda, V_{T_F^T}, V_{T_F^{ADT}}, w, P_S; \varphi) \quad (9)$$

Specifically, equation (9) tells us that school enrollment depends on household budget constraints, which are represented by λ, $V_{T_F^{ADT}}$ and *w*; opportunity cost of schooling, which is the marginal productivity of teenager labor in family production $V_{T_F^T}$; direct cost of schooling, P_S and other exogenous family characters or unobserved preferences captured in φ.

Using equations (5)-(8), I can trace out the potential effect of migratory work on school enrollment decisions. First, since income earned in the non-family labor market may increase due to the new migratory work opportunity, the wealth effect eases credit constraints associated with paying high school tuition, and may facilitate high school enrollment. Second, as returns to non-family activities such as migratory work typically exceed those from family production, households may allocate more adult labor to non-family wage employment, and the outflow of adult labor in family production will change the value of teenagers' time in family production. The relationship between adult labor and teenage labor in home production, i.e., whether they are substitute or complement, will decide whether or not the marginal productivity of teenager labor rises with fewer adult laborers in household production. The net effect of migration opportunity on high school enrollment is a combination of these effects and cannot be signed a priori.

To empirically investigate the channels through which migration opportunity affects school enrollment, I simplify the enrollment demand function that I will estimate by using one variable, *Y* which contains $V_{T_F^A}$ and *w*, for total household income, because

how migration affects the different components of family income is not the main concern of this paper. Thus, the simplified enrollment demand function is:

$$S^* = S^*(Y, V_{T_F^T}, P_S; \varphi) \tag{10}$$

The way through which migration affects school enrollment is that it changes the income and opportunity cost variables, Y and $V_{T_F^T}$, used in school enrollment function.

1.4 Empirical Model and Econometric Issues

1.4.1 School Enrollment Estimation

From arguments of the enrollment demand function in equation (10), a school enrollment function of the discrete decision of a household to enroll child i in high school can be written:

$$S = \alpha_0 + \alpha_1 Y + \alpha_2 MPL^T + \alpha_3 \mathbf{Z_S} + \varepsilon_S \tag{11}$$

Schooling variable, S, is modeled as a continuous variable in the above theoretical framework. Since the CHIP survey does not report the time spent in school, and the available information about schooling is whether the individual is a full-time student, a binary schooling variable, taking the value of 1 if the child is enrolled as a full-time student in high school and 0 if otherwise, is used in the empirical estimation[14]. Y is the total household income, including the income from household production and parental non-family wage income. MPL^T is the marginal productivity of teenager labor in household production and proxies the opportunity cost of schooling. It is generated from family production below and is a household level variable. $\mathbf{Z}$ is a vector of individual,

[14] Treating schooling as a binary variable should not change the model implication.

household, and county characteristics. Specifically, it contains teenager's gender, mother's education, household composition such as household size and number of siblings, availability of school in the village, and provincial dummies. The error term, ε_s, includes unobserved factors that may affect the children's school enrollment decision, such as a family's preference for education and children's innate ability.

Family characteristics plausibly enter the model inasmuch as they affect the prices and wages that the family faces and the extent that they are systematically related to preferences. For example, a mother's educational attainment may increase her market wage, boost her productivity in home production, influence her taste for schooling her children and reduce the information costs—and therefore the price—of schooling. Family structure may affect a teenager's opportunity cost of home production.. For example, the presence of younger siblings will increase the time value of older siblings who care for them. The large size of household may reduce a teenager's marginal productivity as there are more potential adult laborers. The price of schooling will be lower in villages that have readily accessible schools.

Since S is a binary variable, a logistic model is used in the estimation. The marginal effect of household income and opportunity cost on the probability of high school enrollment will be calculated from the estimated coefficients of the logistic model at the sample mean. By combining the differences in household income and opportunity cost, which is the marginal productivity of teenager labor in family production, between migrant and non-migrant households, I can trace out the income effect and opportunity cost effect brought about by migration on the probability of a teenager's high school enrollment.

The sample used for this school enrollment estimation includes households in which head and spouse have 15-18[15] -years- old teenagers and there is valid marginal productivity of teenage labor in family production.

1.4.2 Family Production and Marginal Productivity of Teenager Labor

The purpose of estimating a family production function is to get the estimated coefficients, which reflect the input-output technology of production employed by all households in the sample, to calculate marginal productivity of teenager labor and to investigate the relationship between adult labor and teenager labor in household production. Labor inputs are classified into two types according to the purpose of this study: adult labor and teenage labor, with the age ranges between 15-18. A Translog production function is selected instead of a commonly used log transformation Cobb-Douglas form, because the interaction term in Translog production function reveals the relationship between adult and teenager labor while the log transformation C-D function implicitly assumes that the marginal productivity of one input factor does not change with the other inputs used in the production, which suppresses an important aspect in my framework. Generally, if the coefficient of the adult and teenager labor interaction term is negative, then adult labor and teenager labor are substitute in the sense that the increased use of one labor input will decrease the marginal productivity of the other and vice versa. This empirical relationship is important for us to examine how the outflow of adult labor in household production due to migration affects the marginal

[15] High school teenagers range between 15-18 years old if there is no repetition or delay for elementary and junior middle school.

productivity of teenager labor. If adult labor and teenager labor are substitutes[16], then fewer adult labor days in family production will increase the marginal productivity of the teenager, which is assumed to be the opportunity cost of schooling. The family production function takes the following form:

$$\ln V = \beta_0 + \sum \beta_i \ln X_i + \sum \beta_{ii} (\ln X_i)^2 + \sum \beta_{ij} (\ln X_i)(\ln X_j) + \beta_Z \mathbf{Z_V} + \varepsilon_V \qquad (12)$$
$$i, j = A, K, ADT, TEEN$$

where V is the aggregate value added in family production. Input factor, X, includes land (A), capital (K), adult labor (ADT) and teenager labor days ($TEEN$), in reference to the Cobb-Douglas form. $\mathbf{Z_V}$ is a vector of human capital variables including the total years of schooling completed by family workers and their total working experience, which is approximated by age-7-education years, and its square[17]. The error ε_v, which is assumed to be mean zero, encompasses not only the unobserved household characteristics that influence output but also uncontrolled disturbances, such as weather variations and unexpected shocks. The sub-sample used for family production estimation is different from the one used for teenager school enrollment. It includes all households that have family members between 15-18 years old, who are, however, not necessarily the children of the household head.

Among the households used for production estimation, only 40% have positive teenager labor input. As logarithmic operation does not allow any of the inputs to take on

[16] Literature does not have a consensus on the relation between child labor and adult labor. It is an empirical issue, which depends on the gender of parents and children, and varies among countries after controlling other variables.

[17] In the literature on education and production efficiency, researchers have used the years of schooling of the most educated family members, the education of household head, or average education measures as explanatory variables. The schooling of the most educated household member is used as a measure of managerial skill in production and the education of household head is used because the household head usually has more decision power in family production. As the main purpose of estimating a production function in this study is to impute the marginal product of teenager labor, it is proper to use total years of schooling completed by family workers to control the difference in family human capital.

a value of zero, a constant 1[18] is added to teenager labor days to keep all the observations in the estimation. The value added (V) equals the total household earnings from household production minus the costs of production including hired labor plus the potential value of the food consumed by the household. The treatment of land as an exogenous variable is a specific case in China because land is more or less fixed in size for each household and is nontransferable. Under the Chinese Household Responsibility System, land allocation is based on household population and size of labor force. The duration of land contracts is often more than 15 years. While rural households have the right to use the land they are allocated, they do not have the right to sell it. Although the central government has legalized subleasing of land in recent years, for various reasons, sub-leasing has been extremely rare. Among the sample households used for this paper, only 1.4% of households leased out land for cultivation to others, while only 2.2% of households rented land from other families in 1995. This information provides justification for treating land as an exogenous variable. Labor inputs are potential endogenous variables, which may be correlated with the disturbance term if the unobserved managerial ability that affects the total production output may also affect the family labor allocation when new non-family employment opportunities occur, or if the household adjusts the family labor according to their expectation of shocks. Provincial dummies are included to allow for the geographic condition influence on family production.

The estimated coefficients of factor inputs in the above production function are elasticities. Equation (13) is used to calculate the marginal productivity of teenager labor

[18] The choice of the additive constant is arbitrary, except for the presumption that it be 'near' zero, or at least small relative to the average input value.

in family production, which is assumed to be the opportunity cost in school enrollment estimation:

$$MPL^T = (\hat{\beta}_{TEEN} + 2 \times \hat{\beta}_{TEEN,TEEN} \ln TEEN + \hat{\beta}_{A,TEEN} \ln A + \hat{\beta}_{K,TEEN} \ln K + \hat{\beta}_{ADT,TEEN} \ln ADT) \times \hat{V} / TEEN \quad (13)$$

where $\hat{\beta}$'s are estimated coefficients from the above household production function (12) and $\hat{V}$ is the predicted value of the value added.

This generated marginal productivity of teenager labor is an endogenous variable when used in the school enrollment estimation because the disturbance term in school enrollment function, ε_s, which contains unobserved factors that affect the decision of school enrollment, may simultaneously affect adult and teenager time allocation in family production.

From (13), the relationship between adult labor and marginal productivity of teenager labor in family production can be expressed as:

$$\frac{\partial MPL^T}{\partial \ln ADT} = \hat{\beta}_{ADT,TEEN} \frac{\hat{V}}{TEEN} \quad (14)$$

As adult working time is assumed to be divided between family work and non-family wage work, equation (14) can also be considered as the effect of migration on the marginal productivity of teenager labor.

1.4.3 Identification Strategy

Estimating equation (11) and (12) using OLS would introduce endogeneity bias due to the correlation between labor inputs and the disturbance terms. One way of dealing with

endogeneity is to find instrumental variables that are correlated with the endogenous variables but unrelated to unobserved factors included in the disturbance terms. In this framework, the endogeneity is induced by the presence of new non-family employment opportunities such as migration. Specifically, the managerial ability contained in the disturbance term of family production function will adjust the labor inputs in family production, and the preference for education included in the error term of school enrollment function will reallocate the adult labor between family and non-family work and teenager time between schooling and home production as a response to this new opportunity.

Three variables are chosen as instruments for labor inputs: household size, net household financial asset, and percentage of non-family workers among all workers in the county and its square. Interpreting the number of family members as a quasi-fixed asset of the household, and net financial asset as family wealth, indicates that they are obviously highly correlated with the labor supply in family production, and at the same time, uncorrelated with the disturbance in the production function. Under the assumption that adult working time is divided between family and non-family work including migration, variables that affect adult migration, which is part of non-family work, while do not have a direct impact on children's high school enrollment, can serve as the instruments for adult working days in family production[19]. The justification for using the

[19] One problem in the literature on the effects of migration on source communities is that migrant opportunity is difficult to identify in a clean and convincing way because as part of a set of interrelated family decisions, migration opportunity is endogenous. de Brauw and Giles (2005) use the timing of national identity card distribution as an instrument for migration and employ instrumental variables generalized method of moments estimator to investigate how reductions of barriers to migration affect the decision of middle school graduates to attend high school in rural China. They show that the instrument is both related to the size of the migrant network from a village and plausibly exogenous to the school enrollment decision. However, in the data set used in this paper, there is no relevant policy information that can be used as identifying instruments for migration. By instrumenting on adult family labor days, I indirectly dealt with the migration endogeneity problem as adults simultaneously allocate time between family production and migration.

percentage of non-family workers among all workers in the county and its square as instruments for adult labor days in family production is that they affect the time allocation of adults through its network effects in non-family work, especially migration work. It is widely documented in the migration literature that referral through one's social network is a common method of migration job search[20]. Having more migrants from one's own county increases the likelihood of employment, and decreases the informational and psychic cost. Consequently, the higher percentage of county workers in non-family employment, the more likely that household workers spend more time in non-family work and vice versa. As there is no additional county level information about the percentage of non-family workers, the instrumental variable is constructed from the households included in the survey by dividing the total non-family workers by total workers in the county.

I perform OLS regression of adult labor days and Tobit regression of teenager labor days (due to the large number of zero teenager labor inputs in the observation) on the above three instrumental variables and all other exogenous variable in the production function. Then I use the predicted value of adult and teenager labor days for fitting production and school enrollment equations. The marginal productivity of teenager labor, MPL^T, is calculated at the county mean. This variable represents the time value of teenagers and acts as a proxy of the opportunity cost of schooling. It gives the idea of the potential marginal productivity of teenager labor in a specific household should this

[20] Migrant networks play an important role in labor migration in both the developed and developing world. Carrington, Detragiache, and Vishnawath (1996) explicitly show that in a model of migration, moving costs can decline with the number of migrants over time, even if wage differentials narrow between source communities and destinations. Survey-based evidence suggests that roughly 50 percent of new jobs in the US are found through referrals facilitated by social networks (Montgomery, 1991). In a study of Mexican migrants in the US, Munshi (2003) shows that having more migrants from one's own village living in the same city increases the likelihood of employment. Using household survey data from rural China, Zhao (2003) finds that experienced migrants have a positive and significant effect on subsequent migration.

family use the county average teenager labor days in its production under its current technology and other existing fixed inputs, such as land and capital.

1.5 Data and Variables

1.5.1 Description of the Dataset

The data used in this paper comes from Chinese Household Income Project (CHIP). This national household income survey has two waves: 1988 and 1995, the period during which the economic reform in China that started in 1979 was deepening and the whole economy was growing rapidly. The data provide the best information currently available for addressing the income distribution and component picture in China as a whole.

The survey samples were drawn from the national urban and rural samples used by the State Statistics Bureau for its own annual household surveys, and conducted by researchers from Chinese Academy of Social Sciences and western scholars. The CHIP questionnaires contained much more detailed information than the SSB survey. The surveys of urban and rural China were kept separate; each survey consisted of two data files, one in which households were the basic sampling units and the other with all individuals in each sampled household to be interviewed. Most questions in the questionnaires of the first survey reappeared in the second, though some new questions were added. However, these two surveys did not necessarily cover the same households, so they are not panel data.

The sample used in this paper is the second wave, the CHIP-95 survey, of rural households and individuals, as it was more recent and provided information on migration, which was not included in the previous wave. In total, there are 7,998 households and

36,025 individuals from 19 provinces in this rural sample. In addition to detailed economic and demographic variables, this data set contained records of sectoral time allocation for every working member of the households and incomes from different sources. These data provide crucial information for household production estimation and comparison between migrant and non-migrant households.

1.5.2 Summary Statistics

Since the focus of this paper is on the school enrollment of high-school-year-old children in the household, 2,386 rural households are chosen based on whether the household head has children between the ages of 15-18 from the original data set[21]. For estimation purposes, the sample is further reduced to 2,122 households that are engaged in household production as the opportunity cost of schooling is imputed from family production[22]. 467 households have more than one 15-18 years-old child, so there are 2,589 high-school-year-old teenagers in the sample.

Migrants are identified from the household survey as those who spent a significant amount of time working away from the household in the survey year--specifically, a minimum of 6 months. To explore the parental migration effects on children's high school enrollment, a household is defined as a migrant household if either or both of the parents are migrants. There are 76 (3.58%) migrant households, only 6 households in which both parents are migrants, and 95% households in which the migrant is father.

Table 1.3 reports the key variables in the empirical specification and separates the

[21] Households in which the head has two spouses are dropped, and households with sick or disabled 15-18 years old children are deleted, as their behavior will be different from others.

[22] To be able to estimate the household production function, a household should have positive net income from household production and report positive land, fixed productive assets and adult labor use in the survey year.

sample into migrant and non-migrant households. Total household income, Y, includes the income from household production and parents' wage income (if any). Family agricultural production is the major activity in household production; its income takes a 90% share of the total production income on average, though in migrant households this percentage is about 80%. Agricultural production income contains two parts: the revenue from agricultural product sales, which is equal to the gross income of grain, economic crops, forestry, animal husbandry, fishing and other agricultural activities minus the costs of the production, and the potential value of the agricultural products consumed by the household itself, which accounts for 37% of the total agricultural income. Other household production income includes the revenue from household industry, construction, transportation, services, commerce, restaurants and catering, handicrafts and sideline activities. Detailed information on variable costs is not available from the survey, except the cost of hired labor. The expense of hired labor only accounts for 1% of the total production cost, indicating that self-employment in family production is the dominating form of labor input in household production in China. On average, migrant households have 2,943 yuan more total income than non-migrant households (21.13%), though the income from total family production and agricultural production is obviously less, 16.06% and 23.97% respectively.

The input of land is measured as the amount of self-cultivated land whether irrigated or dry. The size of the cultivated land of an average farm is only 7.54 mu (1.26 acres). Migrant households have 27% less land than non-migrant households. The input of capital is calculated as 10 percent[23] of the current value of household fixed productive

[23] This is a common annual depreciation rate used in China.

assets at the end of 1995[24]. The average capital input is 308.06 yuan, and on average migrant households have 53 yuan less than non-migrant households. The average adult and teenager labor days in household production are 413 and 56 respectively. Migrant households have significantly fewer labor days of both types in household production than non-migrant households: 26.91% fewer adult labor days and 36.62% fewer teenager labor days, and 10% fewer migrant households use teenager labor than non-migrant households do. One weakness of the labor days variable is that it does not reflect labor intensity, as there is no information on daily working hours. On average, each household has 4.8 members; a migrant household has a slightly smaller household size than a non-migrant household. Net financial asset is the difference between the total value of household financial asset and debts, and is used as a proxy for household wealth. On average, non-migrant households are wealthier than migrant households by 40.26%. This finding is consistent with the assumption that migration is a way to relax family budget constraints, especially for poorer families who cannot get credit from the market.

The educational attainment in the survey is reported by seven categories based on completion levels. To convert the categorical schooling variable into a continuous variable, each completion category should be matched with a particular number of education years. Based on other studies and the education system in China, the number of years is assigned as follows: college and above (16 years), professional school (14 years), upper middle school (12 years), middle level professional, technical or vocational school (11 years), lower middle school (9 years), 4 or more years of elementary school (6 years), 1-3 years of elementary school (3 years), illiterate or semi-illiterate (0 year). The average

[24] Household fixed productive assets includes: livestock used for labor and food, large and medium sized farm tools, machinery and equipment for use in agriculture, forestry, animal husbandry and fishing, industrial machinery and equipment, transportation machinery and equipment, construction machinery and equipment, structures used for production, other productive fixed assets.

total years of schooling of family members working in family production is 19.34. Non-migrant households have two years more total schooling than migrant households. Following the convention, experience is approximated by (Age-years of schooling-7). The average total years of experience of family members engaged in family production is 65.50 years. Non-migrant households have 10 years more total years of experience than migrant households.

1.6 Estimation Results and Interpretations

1.6.1 Household Production Function

The sample used for household production estimation includes all households that have 15-18 years-old household members and are engaged in household production in the survey year, i.e., households that reported positive value added and labor use in production and valid land and capital input. 2,204[25] households are selected from the original survey. Before estimating household production function (2), I first run regressions of adult labor days and teenager labor days in family production on instrumental variables and all other explanatory variables used in the production function. As discussed above, three instruments—household size, net family financial assets and percentage of non-family workers among total workers in the county-- are used in regression. Since a significant fraction of the variable of teenager labor days takes the value of zero, a Tobit regression that accounts for the qualitative difference between zero observations and nonzero continuous observations is used for teenager labor days estimation.

[25] This number is different from the number in school enrollment estimation:-2,122 households, because for production estimation, I use households that have 15-18 years- old household members who don't have to be the children of the household head, as they do in school enrollment estimation.

Table 1.4 reports the results of first stage regression for adult and teenager labor days using instrumental variables. The dependent variable of adult labor days is in logarithmic form, while that of teenager labor days is the raw number as logarithmic operation does not allow any variables to take on a value of zero. All instrumental variables have high statistical significance in both regressions. Specifically, household size is positively associated with adult labor days while negatively related to teenager labor days. Net family financial asset has a negative effect on both adult and teenager labor days in family production. The negative coefficient of percentage of non-family workers among total workers in county (pvilnfm) and the positive quadratic coefficient indicate that the non-family job network decreases the labor supply in family production at an increasing rate. This finding is consistent with the migration network hypothesis.

Table 1.5 reports the results of instrumental variable (IV) estimation of family production after correcting for the potential endogeneity of adult and teenager labor supply (lnADT, lnTEEN). The predicted values of lnADT and lnTEEN obtained from the first stage regressions are used for the regressions. Column (1) of table 1.5 is the basic instrumental variable estimation of the log transformation Cobb-Douglas production function. An F-test is used to examine the null hypothesis that all $\beta_{ii} = 0$ and all $\beta_{ij} = 0$, the computed $F_{(10,\ 2168)} = 7.35$, strongly rejects the null at 1% significant level. So Translog is a more appropriate production functional form compared to log transformation C-D type. The Translog function (results reported in the second column of table 1.5) significantly increases the input contribution in production more than log transformation C-D production, and raises the adjusted R^2 by about 2%. The quadratic terms of land, lnA^2, and capital, lnK^2, are positive with lnA^2 achieve a significance level

at 1%, indicating the scarcity and importance of land in family production in China. The self interaction terms of adult and teenager labors, $\ln ADT^2$ and $\ln TEEN^2$, are both negative, showing that the contribution of labor in production increases at a decreasing rate. Because all of those quadratic terms are small in size, and 3 out of 4 are insignificant, a reduced Translog function that does not have self interaction terms is estimated and the coefficients are used to calculate marginal productivity of teenager labor.

In comparison with OLS results reported in appendix table 1.2, the IV approach supports the same basic conclusions. After correction, however, the coefficients of lnADT have generally increased in size and the adult and teenager labor interaction term, lnADT*lnTEEN, has become statistically significant. Furthermore, the IV approach provides a better fit for the regression, raising the adjusted R^2 by about 1%. The marginal productivity of teenager labor, MPL^T, calculated by using the estimates from IV approach and the predicted value of adult labor days, has become negatively significant in the school enrollment estimation. Because of the consideration of endogeneity, only the IV results are interpreted.

The results of estimating a reduced Translog production function are presented in column (3) of table 1.5. Land (A), capital (K) and adult labor (ADT) are main contributors in household production, with land and adult labor achieving 1% significance level and capital, 10%. The elasticities of these three inputs are 0.6, 0.28 and 0.85, respectively. The coefficient of teenager labor (TEEN) is insignificant and small in size, reflecting the fact that most households in the sample do not use teenager labor in family production, making it a relatively minor input factor. The coefficient of the adult and teenager labor interaction term (lnADT*lnTEEN) is negative at a 5% significance

level, giving evidence of the substitutability between adult and teenager labor in households that use teenager labor in family production. At the sample mean, 1% decrease in adult labor days increases marginal productivity of teenager labor (MPL^T) by 7.04 yuan, or a one- unit decrease in adult labor days increases MPL^T by 0.017 yuan. From the summary statistics table (1.3), on average, migrant households have 112.22 fewer adult labor days than non-migrant households do; consequently, marginal productivity of teenager labor in migrant households is 1.91 yuan higher than non-migrant households[26].

One additional year of total years of schooling of the family members working in family production (TFEDU) significantly increases the total output by 0.016%, while the experience and its square terms are insignificant. The 18 provincial dummies are jointly significantly different from zero, demonstrating regional differences in household production.

1.6.2 Sensitivity Analysis

A problem inherent in surveys of individual time allocation in rural areas of developing countries is the sometimes hazy distinction between production work activities and housework activities. Not only might these two sectors be difficult to distinguish in the minds of survey respondents, but, in reality, housework time may jointly produce production output and 'home-produced commodities'. This work combination is especially prevalent for young and unskilled labors. To check whether the marginal productivity of teenagers might be underestimated and the relationship between adult and teenager labor might be changed by excluding housework, days spent doing housework

[26] See the derivation of these figures in appendix A.

are added to the labor input for both categories. Re-estimating the production function yields increased coefficients of all factor inputs, especially the elasticity of teenager labor by 63.56%, though it is still insignificant. It also raises the substitutability between adult labor and teenager labor by 32.66% (Column 2 of table 1.6 reports these results, column 1 reproduces the results of a reduced Translog production estimation). Thus, excluding housework from production time underestimates the contribution of teenager labor in family production; however, this effect is not statistically significant and does not change the relation between adult and teenager labor.

As agricultural production takes the major share in total household production, it is worthwhile to estimate an agricultural production function and compare the results. Column 3 of table 1.6 reports the results of agricultural production estimation. Land, capital and adult labor are still the major contributor as they are in family production function. Adult labor and teenager labor still show some extent of substitutability though it is not significant.

1.6.3 Children's High School Enrollment Estimation

The sample used in school enrollment estimation includes all teenagers between 15-18 years-old whose parents are household head and head's spouse and who are from households with valid marginal productivity of teenager labor estimated from the household production function before. Table 1.7 presents a logistic model of children's high school enrollment. The MPL^T variable is calculated by using the coefficients from IV estimation of a reduced Translog family production function, predicted value of adult labor days and county average teenager labor days. The table also presents marginal

effects of explanatory variables on the probability of school enrollment. The marginal effect of a continuous independent variable is evaluated at the sample mean and that for the categorical variable is evaluated against the reference category. There are 18 provincial dummies in the dataset with the reference group as Beijing; I randomly choose Jiangsu province when calculating the marginal effect for the sake of simplicity[27].

School enrollment, *S*, is a binary variable, taking the value of 1 when the child is a full-time student, 0 otherwise. Suppose **X** is a vector of explanatory variables and p=Pr(S=1|**X**) is the response probability to be modeled. The linear logistic model has the form:

$$logit(p) = log(\frac{p}{1-p}) = \alpha + \beta \boldsymbol{X} \tag{15}$$

where α is the intercept parameter and β is the vector of slope parameter. The estimated coefficients reported in the second column of Table 1.7 are parameters in the above logistic model. At the sample mean, probability p can be calculated by:

$$\hat{p} = \frac{e^{\alpha+\beta \boldsymbol{X}}}{(1+e^{\alpha+\beta \boldsymbol{X}})} \tag{16}$$

Then, the marginal effect of explanatory variables on the probability can be obtained by taking derivatives of $\hat{p}$ with respect to the specific independent variable:

$$\frac{\partial \hat{p}}{\partial X} = \frac{e^{\alpha+\beta \boldsymbol{X}}}{(1+e^{\alpha+\beta \boldsymbol{X}})^2}\beta \tag{17}$$

These results are listed in the third column of Table 1.7.

[27] There are 19 provinces in the data set. The data set, however, only provides provincial codes to distinguish these provinces. The additional information that matches these codes with specific province only identifies 11 provinces, they are: Beijing, Shanxi, Liaoning, Jiangsu, Anhui, Henan, Hubei, Guangdong, Sichuan, Yunnan and Gansu.

As can be seen from Table 1.7, there is significant gender discrimination in China: boys are 7.36% more likely to be enrolled in high school than girls are. Being a Han, the largest race group in China, increases the school enrollment probability by 8.85%. Mother's education has a significant positive effect—1.64% on child's schooling as is found by many other studies. The positive significant coefficient of school availability variable reflects the supply side effect on child education, indicating the importance for the government to provide facilities to increase the school enrollment rate. The number of siblings of the same age range between 15-18 and younger school- age siblings between 6 and 14 are included in the regression to see whether there is resource competition in the household. The results show that having similar age siblings or younger school age siblings has a positive effect on school enrollment, a finding opposite to that of some previous studies[28]. Household size decreases the school enrollment probability by 3.96%, probably because of the tighter resource constraint in a bigger family[29].

Total household income has a significant positive effect on the probability of school enrollment, suggesting the income effect in children's school enrollment. The marginal productivity of teenager labor in family production negatively affects the school enrollment probability, giving evidence of the opportunity cost effect on enrollment. The marginal effects of these two key variables on the probability of school enrollment are $1.0965*10^{-5}$ and 0.000168, respectively. From the summary statistics table 1.3, on

[28] Using Peru 1991 Living Standards Survey data, Patrinos and Psadcharopoulos find that having a greater number of younger siblings implies less schooling, more age-grade distortion in the classroom and more child labor. Jacoby (1994) finds that in Peru, children who have more closely-spaced younger siblings, adjacent older male siblings, and greater child care responsibilities begin withdrawing from school earlier.

[29] Lloyd (1994) reviews the evidence on the relation between household size and children's schooling investment from developing countries and finds that larger household size reduces children's schooling participation, and progress in school and reduces parents' investment in schooling.

average, a migrant household, a household in which at least one parent is a migrant, has 2,942.94 yuan more total income than a non-migrant household has; and from the previous family production estimation, the marginal productivity of teenager labor in migrant households is 1.91 yuan higher than non-migrant households. Thus, for a migrant household, increased income raises the children's high school enrollment probability by 3.227%, while the higher marginal productivity of teenager labor, which is assumed to be the opportunity cost of schooling, decreases the enrollment probability by 0.032%. Combining these two competing forces indicates that being a migrant household increases children's high school enrollment probability by 3.195%.

1.7 Summary and Policy Implications

Due to the loosening restrictions on labor mobility, China is experiencing an unprecedented labor migration from rural to urban areas. As part of a set of interwoven economic choices, migration affects a household's decision on enrolling a child in high school. In this paper, I decomposed the migration effect into income and opportunity cost effect and suggested a strategy to estimate the opportunity cost of schooling, taking into consideration the endogeneity due to simultaneous decisions on adult labor allocation between family and migration work and child time allocation between family production and schooling. Based on cross-sectional Chinese rural household data, the empirical results indicate that the increased household income in a migrant household raises the probability of high school enrollment by 3.23%, while the increased opportunity cost--the marginal productivity of teenager labor in family production as a result of lost adult labor in family production-- decreases the enrollment probability by 0.032%. The distinct

income effect confirms the fact that budget constraint is one of the most important reasons that parents do not enroll their children in high school. Overall, children in migrant households have 3.20% higher probability to be enrolled in high school than those from non-migrant households. These results are all statistically significant.

The recent increase in migration has left policy makers concerned about the potential consequences of growing inequality between urban and rural areas. More educated and more able persons benefit more from the new economy, at least in the short run. Currently, there is a low level of public support for education in most provinces of China. Since schooling is mostly funded at the local level and schooling finances are tied to the level of wealth in a region, rich provinces tend to produce more human capital than do poor provinces. School fees are a substantial fraction of household income in rural areas. Access to education is not uniform across rural and urban areas. This is a powerful source of inequality in Chinese society across people. The results presented in this paper suggest that freeing up labor market for rural migrants is a way to foster human capital that entails less direct cost to governments. In the long run, inequality across regions and over generations will be reduced as the population becomes more skilled and as opportunities for education investment are spread more widely throughout Chinese society.

At present, tuition policies for secondary school students discriminate against the children of the poor and the children of migrants. Ending segregation against migrants may create frictions between urban residents and rural migrants in the short run, but may promote education for rural youth and have the salutary effect of achieving intergenerational social equality in the future. Human capital is the asset that ultimately

determines the wealth of China. The results presented in this paper call for reassessment of migration policies and their long- term effect on human capital accumulation.

Appendix A: Derivation of the Difference in Marginal Productivity of Teenager Labor Between Migrant and Non-migrant Households

Formula (3) and (4) in the paper give the calculation of marginal productivity of teenager labor and the unit change in MPL^T in response to percentage change in adult labor days:

$$MPL^T = (\hat{\beta}_{TEEN} + \hat{\beta}_{A,TEEN} \ln A + \hat{\beta}_{K,TEEN} \ln K + \hat{\beta}_{ADT,TEEN} \ln ADT) \times \hat{V} / TEEN$$

$$\frac{\partial MPL^T}{\partial \ln ADT} = \hat{\beta}_{ADT,TEEN} \frac{\hat{V}}{TEEN}$$

At the sample mean, V = 9447.85, TEEN = 55.72, ADT = 412.96. So 1% decrease in adult labor days increases MPL^T by 7.035 yuan:

$$\frac{\partial MPL^T}{\partial \ln ADT} = \hat{\beta}_{ADT,TEEN} \frac{\hat{V}}{TEEN} = -0.04149 \times 9447.85 / 55.72 = -7.035$$

Furthermore, a unit decrease in adult labor days increases MPL^T by 0.017 yuan:

$$\frac{\partial MPL^T}{\partial ADT} = \frac{\frac{\partial MPL^T}{\partial \ln ADT}}{ADT} = -7.035 / 412.96 = -0.017$$

On average, non-migrant households have 112.22 more adult labor days than migrant households have, so MPL^T in migrant households is 1.91 yuan higher than non-migrant households.

Table 1.1: Income and Its Composition in Migrant and Non-Migrant households With Children In 15-18 Age Group

Variable	Migrant Households (n=85)	Non-migrant Households (n=2274)
Family Production Income	7761.70 (4065.81)	9304.96 (5576.28)
Total Household Income	13873.98 (10022.76)	11070.79 (8664.90)
Parents Non-family Wage Income (All HH)	6112.28 (8968.06)	1708.88 (6746.66)
Households with Non-family Wage Income Parents (Percentage)	88.23%	37.7%
Parents Non-family Wage Income (HHs with Non-family Wage Income Parents)	6927.25 (9250.55)	4390.84 (10259.02)

Notes: Standard Deviation is in parentheses.
Differences between Migrant and Non-migrant households are all statistically significant at the 1% level

Table 1.2: Status of Children Between 15-18 in Migrant and Non-Migrant Households (%)

Status	Migrant Households (n=85)	Non-migrant Households (n=2274)
Full-time student	60.0%	50.04%
Work on Family Farm or Business	21.18%	31.05%
Employed Outside the Household	11.61%	12.94%

Table 1.3: Summary Statistics:
Household Mean and Standard Deviations for Migrant and Non-migrant Households

Variables	Symbol	Total HH	Migrant HH	Non-Migrant HH
Total household income (yuan)	Y	11096.53	13930.89	10987.95
		(7547.17)	(10228.35)	(7407.17)
Household production income	V	9447.85	7976.42	9502.51
		(5602.81)	(4096.06)	(5644.41)
Household agricultural income		8562.50	6566.93	8636.62
		(5118.94)	(3096.25)	(5164.52)
Household industrial and sidelines income		885.35	1409.49	865.88
		(2646.34)	(2933.91)	(2633.84)
Cultivated land (mu)[30]	A	7.54	5.51	7.61
		(6.18)	(3.90)	(6.23)
Capital input (yuan)	K	308.06	256.45	309.98
		(411.03)	(339.02)	(413.41)
Human capital and household demographics				
Adult labor days in household production per year	ADT	412.96	304.76	416.98
		(222.43)	(173.77)	(223.05)
Teenager labor days in household production per year	TEEN	55.72	35.99	56.46
		(93.46)	(73.72)	(94.05)
% of households that use teenager labor		40.48%	30.26%	40.86%
Total years of schooling of family members working in family production	TFEDU	19.34	17.18	19.42
		(9.82)	(9.96)	(9.80)
Total years of experience of family members working in family production	TFEXP	65.50	54.59	65.90
		(20.61)	(20.92)	(20.49)
Household size	HHSIZE	4.80	4.70	4.81
		(1.20)	(1.14)	(1.21)
Net financial asset	FIN	4120.62	2497.54	4180.91
		(8135.31)	(11282.61)	(7992.04)
Number of households		2,122	76	2,046

[30] 1 mu=0.0667 hectares=1/6 acre.

Table 1.4: First Stage Regressions of Adult and Teenager Labor Days in Family Production on Instrumental Variables

Independent Variables	lnADT (OLS)	TEEN (TOBIT)
LnA	0.02863**	-7.597
	(0.01559)	(5.844)
LnK	0.04607***	8.427**
	(0.00957)	(3.769)
TFEDU	0.01248***	8.492***
	(0.00122)	(0.4921)
TFEXP	0.01472***	1.187*
	(0.00181)	(0.723)
$TFEXP^2$	-0.00005386***	-0.0019
	(0.00001118)	(0.0043)
#HHSIZE	0.04509***	-17.87***
	(0.01049)	(4.052)
#FIN	-0.00000353***	-0.0018***
	(0.00000145)	(0.0007)
#‡pvilnfm	-0.9158***	-350.34***
	(0.2212)	(82.52)
#‡$pvilnfm^2$	0.6144***	316.56***
	(0.2089)	(78.24)
Constant	3.275***	-1044.47
	(0.2055)	(306706.1)
Provincial Dummy	Yes†	Yes†
Adjusted R^2	0.3157	-
Sample size	2,204	2,204
Predicted Value of Dependent Variable, Mean	5.867	55.88

NOTE: The figures in parentheses are standard errors of the estimated coefficients

† The F-test strongly rejects the null hypothesis that the provincial dummies are jointly 0 at 1% significant level

Instrumental variables of labor days in production function

‡ Instrumental variables of labor days in school enrollment function

*Significant at the 10% level

**Significant at the 5% level

***Significant at the 1% level

Table 1.5: Instrumental Variable Estimation of the Production Function, Predicted Value of lnADT and lnTEEN

Independent Variables	Dependent Variable = lnV		
	Log transformation C-D (1)	Full Translog (2)	Reduced Translog (3)
lnA	0.06612***	0.8142***	0.6009***
	(0.01454)	(0.2597)	(0.2383)
lnK	0.09015***	0.2811*	0.2836*
	(0.00937)	(0.1716)	(0.1653)
lnADT	0.4227***	0.8984	0.8482***
	(0.110)	(1.3872)	(0.2051)
lnTEEN	-0.2177***	0.027	0.06409
	(0.02903)	(0.2915)	(0.1035)
ln^2A		0.03322***	
		(0.00685)	
ln^2K		0.00367	
		(0.00527)	
ln^2ADT		-0.0007899	
		(0.1353)	
ln^2TEEN		-0.0002973	
		(0.0046)	
lnA*lnK		-0.0256**	-0.01959*
		(0.01109)	(0.01101)
lnA*lnADT		-0.1233***	-0.07785*
		(0.05052)	(0.04612)
lnA*lnTEEN		0.01341	0.00534
		(0.01010)	(0.00963)
lnK*lnADT		-0.02355	-0.01779
		(0.03318)	(0.03033)
lnK*lnTEEN		-0.01599***	-0.01661***
		(0.00506)	(0.00483)
lnADT*lnTEEN		-0.033	-0.04149**
		(0.05592)	(0.01881)
TFEDU	0.0138***	0.01469***	0.01631***
	(0.00256)	(0.00276)	(0.00275)
TFEXP	0.0006365	-0.00136	-0.0009125
	(0.00226)	(0.00243)	(0.00228)
$TFEXP^2$	-0.00001077	5.387162E-7	-0.00000143
	(0.00001106)	(0.00001237)	(0.00001128)
Constant	2.13***	0.172	0.1089
	(0.7268)	(3.6724)	(1.2164)
Provincial Dummy	Yes†	Yes†	Yes†
Adjusted R^2	0.2783	0.2987	0.2922
Sample size	2,204	2,204	2,204

NOTE: The figures in parentheses are standard errors of the estimated coefficients

† The F-test strongly rejects the null hypothesis that the provincial dummies are jointly 0 at a 1% significant level

*Significant at the 10% level

**Significant at the 5% level

***Significant at the 1% level

Table 1.6: Sensitivity Analysis of Instrumental Variable Estimation of the Translog Production Function, Predicted Value of lnADT and lnTEEN

Independent Variables	(1)	(2)	(3)
LnA	0.6009***	0.95506***	0.3539*
	(0.2383)	(0.3062)	(0.1962)
LnK	0.2836*	0.4726**	0.3508**
	(0.1653)	(0.2155)	(0.1527)
LnADT	0.8482***	1.1101***	0.6225***
	(0.2051)	(0.2397)	(0.1974)
LnTEEN	0.06409	0.1759	-0.1385
	(0.1035)	(0.1138)	(0.09211)
lnA*lnK	-0.01959*	-0.01301	-0.00804
	(0.01101)	(0.01121)	(0.01124)
lnA*lnADT	-0.07785*	-0.1387***	-0.01416
	(0.04612)	(0.05492)	(0.03913)
lnA*lnTEEN	0.00534	0.0118	-0.03357***
	(0.00963)	(0.0097)	(0.00907)
lnK*lnADT	-0.01779	-0.04926	-0.04791*
	(0.03033)	(0.03743)	(0.02936)
lnK*lnTEEN	-0.01661***	-0.01316***	0.00185
	(0.00483)	(0.00508)	(0.00474)
lnADT*	-0.04149**	-0.06161***	-0.00104
lnTEEN	(0.01881)	(0.01933)	(0.01867)
TFEDU	0.01631***	0.01823***	0.01404***
	(0.00275)	(0.00272)	(0.00251)
TFEXP	-0.0009125	-0.00038798	0.00277
	(0.00228)	(0.00206)	(0.00213)
$TFEXP^2$	-0.00000143	0.00000369	-0.0000172
	(0.00001128)	(0.00000923)	(0.00001104)
Constant	0.1089	-1.279	1.287
	(1.2164)	(1.436)	(1.1015)
Provincial Dummy	Yes†	Yes†	Yes†
Adjusted R^2	0.2922	0.2958	0.2823
Sample size	2,204	2,204	2,200

NOTE: The figures in parentheses are standard errors of the estimated coefficients

† The F-test strongly rejects the null hypothesis that the provincial dummies are jointly 0 at a 1% significant level

*Significant at the 10% level

**Significant at the 5% level

***Significant at the 1% level

Table 1.7: Logistic Model of Children's High School Enrollment:
Dependent variable: enrolled as a full-time student=1; otherwise=0;

Independent Variable	Coefficient Estimate	$\partial(prob)/\partial x$	Definition of Variable
Gender	0.3154*** (0.0892)	0.0736	Child gender dummy: reference=girl
Minority	0.3794** (0.2044)	0.0885	Child Nationality dummy: reference=Non-han
Total Household Income	0.000047*** (0.000013)	1.0965E-5	Income from household production and parental wage income
$MPL^{T\#}$	-0.00072*** (0.000289)	-0.000168	Marginal productivity of teenager labor estimated from household production function
Mother Education	0.0702*** (0.0134)	0.0164	Years of mother's education
School	0.3025* (0.1901)	0.0706	Dummy variable for school in the village: reference=no school in village
Siblings between 15-18 years old	0.2574*** (0.0882)	0.06	Number of siblings between 15-18 years old
Siblings between 6-14 years old	0.2484*** (0.0635)	0.0580	Number of siblings between 6-14 years old
HHSIZE	-0.1696*** (0.0436)	-0.0396	Number of household members
Provincial Dummies	Yes†		Dummy variables for provinces: Reference=Beijing
Intercept	-0.6243 (0.9179)	-0.146	Intercept
Sample Size	2,325		

Note: *Significant at the 10% level
**Significant at the 5% level
***Significant at the 1% level
This variable is calculated by using coefficients from IV estimation of production function, predicted value of adult labor days and county average teenager labor days to calculate
† The F-test strongly rejects the null hypothesis that the provincial dummies are jointly 0 at the 1% significant level

Appendix Table 1.1: Income and Its Composition in Migrant and Non-Migrant households

Variable	Migrant Households (n=323)	Non-migrant Households (n=7648)
Family Production Income	5550.44 (2863.42)	7543.42 (5214.17)
Total Income	11311.95 (8298.34)	9267.95 (7634.78)
Parental Off-Farm Income (All HH)	5702.69 (7703.24)	1642.10 (5875.99)
Households with Off-Farm Income Parents (Percentage)	85.76%	36.83%
Parental Off-Farm Income (HHs with Off-Farm Income)	6382.08 (7879.30)	4223.29 (8826.91)

Notes: Standard Deviation is in parentheses.
Differences between Migrant and Non-migrant households are all statistically significant at the 1% level

Appendix Table 1.2: OLS Estimation of the Translog Production Function

Independent Variables	Dependent Variable = lnV
LnA	0.4371***
	(0.1054)
LnK	0.3129***
	(0.07261)
LnADT	0.3114***
	(0.0668)
LnTEEN	0.06935
	(0.04521)
lnA*lnK	-0.02828***
	(0.01047)
lnA*lnADT	-0.03153**
	(0.01606)
lnA*lnTEEN	0.00015603
	(0.00452)
lnK*lnADT	-0.02634**
	(0.01274)
lnK*lnTEEN	-0.01015***
	(0.00357)
lnADT*lnTEEN	-0.00464
	(0.00724)
TFEDU	0.00529***
	(0.00115)
TFEXP	0.0006986
	(0.00159)
$TFEXP^2$	-0.00000527
	(0.00000964)
Constant	6.095***
	(0.3962)
Provincial Dummy	Yes†
Adjusted R^2	0.2835
Sample size	2,204

Chapter 2

Participation and Duration of Temporary Rural-Urban Migration in China: The Role of Migrant Networks and Education

2.1 Introduction

Under the market liberalization and the lifting of restrictions on labor mobility, the massive migration of rural labor to urban areas has been one of the most important and controversial issues in China. From an economic development perspective, the movement of labor out of agriculture is a manifestation of the efficiency of a free market. The labor mobility from the low-return traditional agricultural sector to the modern sector with higher wages balances the marginal productivity of labor between the two sectors, integrates the rural and urban labor markets, and consequently raises the economic growth of the whole society. The labor flow between rural and urban areas also brings funding, physical and human capital to the less-developed rural areas (i.e., the previous chapter investigates the monetary inflow into rural households that facilitates the investment in children's education, contributing to the accumulation of human capital in rural areas).

However, despite all these positive aspects, there exist both economic and political concerns on migration. In the destination urban areas, there is fear that the employment opportunities in the urban areas are not sufficient to absorb the magnitude of rural labor swarming into cities. The consequent high unemployment rate would become a destabilizing effect on the cities and towns to which workers migrate. The temporary

nature of much of the migration makes the government control difficult, leading to the concern of public security. Furthermore, the transportation bottleneck, which is especially severe during the Spring Festival holidays when seasonal migrants return home, has always been a tough issue for the Chinese government. Policy makers are also worried about the potential shortage in food supply due to the outflow of farmers. There is also political fear that the increasing urban-rural income gap in the short run could even spill over into political unrest.

The sheer volume of the labor flow in China, its multi-faceted influence on society, and its economic meaning in development context motivate my study in rural-urban migration. Emphasizing the household's role in migration decision (Stark, 1991) and the interrelationships between migrants and source communities, I study migration effects on children's education in rural areas in the first chapter. Specifically, I investigate two competing effects on children's school enrollment due to parental migration: income effect and opportunity cost effect. Due to the fact that most households in rural China undertake family production, the opportunity cost of schooling is approximated by the marginal productivity of child labor in family production as it represents children's time value and reflects children's forgoing productive contribution to household income had they not attended school. In the theoretical framework, migration is modeled as a household decision and as one form of non-family wage employment. As returns to non-family activities typically exceed those from family production, participation in migration may increase total household income, allowing more investment in children's education. On the other hand, the outflow of adult labor may increase the demand for child labor and children's time value in family production, leading to a negative

opportunity cost effect. Based on the theoretical framework, I specify a school enrollment function with total household income and opportunity cost of schooling as two major explanatory variables to quantitatively analyze these two competing forces. The opportunity cost of schooling, i.e. the marginal productivity of child labor in family production, is estimated from family production function using the adult and child labor days spent in family production together with the estimated coefficients and predicted output level. Given the structure of the model, the opportunity cost of schooling is an endogenous variable, e.g., unobserved household characteristics (such as preference for education) would affect both the opportunity cost of schooling and children's school enrollment, violating the orthogonality condition of the error term. Specifically, if a household puts a high value on children's education, it would enroll the child in school, reduce children's working time and meanwhile allocate more adult labor on family production than on non-family work to maintain the necessary output level.[31] Therefore, it is necessary to estimate school enrollment function using instrumental variables. In the migration literature, migration participation is closely related to the number of the migrants in the same village because of the migrant network effect. Assuming the similar network effect in non-family activities, I use percentage of non-family workers among all workers in the village and its square as instrument variable as they may affect family members' labor allocation, but not the unobserved household preference. Estimating the school enrollment function using the predicted opportunity cost of schooling from the first-stage regressions, I find a dominant positive income effect on children's school enrollment due to the temporary migration of the parents. The results add weight on the

[31] China's farmland arrangements under the Household Responsibility System obligate the farm household to deliver a certain quantity of its grain output to the state at prices specified by the government to maintain the ownership of the land assigned by local government.

positive side of rural-urban migration and lend support to a favorable policy toward labor mobility.

In the first chapter, migration is modeled as a household decision and as a form of non-family wage employment that may increase household income and compete with family production for adult working time. Focus is on migration effects on source communities. Migrants themselves or individual migration behavior is not described. Furthermore, the percentage of non-family workers among all workers in the village is used as an instrumental variable for the opportunity cost of schooling in school enrollment estimation based on the assumed network effect proposed in the migration literature without further elaboration. The detailed analysis of the determinants of migration and migrant network effect is the task of this chapter.

Migration behavior is studied at two levels: first, rural workers decide whether to participate in migration, i.e., whether to stay and work in rural areas or to leave for cities; second, after the migration participation decision has been made, migrant workers decide the length of migratory work. In addition to demographic variables, this chapter focuses on the effects of two important variables on migration, namely migrant networks and education. Migrant networks play an important role in labor migration in both the developed and developing world. In China, the dominant form of rural-urban labor migration is temporary migration, featured by the close connection between migrants and the source areas they are from.[32] This temporary nature of migration is an important channel through which migrant networks facilitate out-migration in China. "Migrant networks are sets of interpersonal ties that connect migrants, former migrants, and

[32] Hare (1999) reports that an average migrant returns home two to three times a year. According to a sample survey in Shanghai, Wang and Zuo (1999) suggest that migrants plan to spend a specific amount of time away form home before returning to the village.

non-migrants in origin and destination areas through ties of kinship, friendship, and shared community origin" (Massey et al., 1993, p. 448). Migrant networks affect migration by reducing the information and psychic cost of labor migration, especially in the early stages of the migration cycle (Schwartz, 1973). Migrant networks can reduce the information cost by providing the potential migrants in rural areas with job market information and job search assistance in urban areas, and can reduce the psychic cost by providing mutual support among migrants from the same village in the destination areas.

Education is postulated to enhance farmers' ability to perceive and interpret market information so as to better respond to economic disequilibra (e.g. Schultz, 1975). The economic reform in China changed the choices faced by rural workers by boosting local non-farm jobs and loosening restrictions on labor mobility.[33] Better educated farmers may respond to these new job opportunities by allocating more labor input to non-farm work in which the return to labor and education is higher than the conventional agricultural production. Although the positive role of education in shifting labor input from farm production to non-farm work is well recognized,[34] its effect on the choice between local non-farm work and migratory work is ambiguous.

Depending on their education levels, rural workers may also interpret and respond to migrant networks differently. One view is that as the education level increases, potential migrants may be better at utilizing the migrant networks to facilitate migration. An opposing view is that it may also be possible that better educated rural workers rely more on their own abilities to analyze the new opportunity in a changing environment thus

[33] Starting in 1983, the government announced a series of policies that encouraged the development of private enterprises in rural China and loosened restrictions on labor mobility out of agriculture. (Yang and Chen, 2005)

[34] Based on cross-sectional Chinese farm household data, Yang (1997) finds that the better educated farmers work in non-farm activities. Using household survey data between 1986-1995 from Sichuan province, Yang and Chen (2005) find that education plays an important role in allocating productive inputs and expanding non-farm productions under the factor market liberalization.

depending less on migrant networks. Because of the ambiguous effect of education on migrant networks, it is interesting to further explore the interactive effect of education and migrant networks empirically.

Although a lot of work has already been done to investigate the determinants of the dichotomous migration participation decision, work on the determinants of the length of temporary migration is scant, with an exception in Hare (1999).[35] Hare uses duration analysis to examine migrants' return migration decision in rural China by specifying the number of days elapsed before the migrant returns home as a dependent variable. However, the return migration estimation uses a migrant sub-sample and the analysis is silent on the sample selection bias. To my knowledge, no work has been done to explore the determinants of the length of temporary migration, taking into account the sample selection bias. Understanding the length of migrants' employment spells is important because much of the policy concern regarding migrant workers in China is related not only to the size of the migration per se, which depends on individuals' migration participation decision, but also to their frequent movement and their short terms of residence in their job locations. It is a typical view frequently expressed by officials in destination communities that the transient nature of the temporary migrants is at the root of many of the problems attributed to their presence (Field interview by Hare, July 1996).

Thus, one contribution of this chapter is to investigate the determinants of the length of migratory work spells given the choice of migration participation. The potential sample selection bias in the migration duration estimation is dealt with by Heckman's two step procedure, and the results are contrasted with those of a Tobit model, which utilizes

[35] Hare (1999) investigates both out migration and return migration decisions. The analysis of return migration decision bears some similarity to the analysis of the decision on the length of migratory work as the empirical specification for both decisions deals with the censored data issue.

information on both migration choice and migration length in estimation, but does not separate the marginal effects between these two.

The data set used in this paper is the rural household and individual survey in 1995 from the Chinese Household Income Project (CHIP). The question on the length of migratory work allows migration identification and the estimation of migration duration. Other rich demographic information enables the control for individual and household characteristics. The empirical findings of this paper confirm the existence of a migrant network effect in the migration decision. Schooling has a significant non-linear effect on migration participation choice but is insignificant on the migration length. Furthermore, when choosing migration participation, individuals with junior and senior high school education rely more on migrant networks than people with elementary, college and above education. After migration decision is made, there is no significant difference in migrant network effect on migration duration among different education levels.

The rest of the paper proceeds as the follows. The next section briefly reviews the relevant literature which studies the determinants of migration participation choice with an emphasis on education and migrant networks. Section III discusses and compares different statistical models to be used in the empirical analysis and the reasons of using each model are justified. Data description and empirical results are presented in Section IV. Section V concludes the paper.

2.2 Brief Literature Review

Education is widely included as an important determinant in the migration choice estimation. One channel through which education affects migration is the wage difference

in rural and urban areas. The Harris-Todaro two sector model of migration (1970) dominated economists' thinking about migration in developing countries in the 1970s. The model suggests that the expected income difference between rural and urban areas motivates the internal migration in the developing countries. The relation between wage income and education is captured in the typical Mincerian earnings function. Because the returns to education in modern sectors in urban areas is higher than that in the conventional agricultural sector where the low level routine work does not require much education, education is expected to have a positive effect on migration. Other sources of the education effect come from the information advantage of education in the job search (Shwartz, 1973) and the role of schooling in reducing psychic costs of migration (Sjastaad). However, empirical studies do not unanimously support the positive relation between education and migration participation. Using the household survey in Sichuan province, Zhao (1998, 1999) finds the insignificant and even negative effect of education on migration. This negative effect of schooling comes mainly from the unwillingness of high school graduates to migrate as they had an advantage in selecting local non-farm work than migratory work. Further evidence is provided by de Brauw and Giles (2005), who find a non-linear return to schooling for migrants—primary, middle school and post-secondary years of schooling have a higher return than high school does in the migratory work in cities.

The effect of migrant networks is closely related to the temporary nature of migration in China—instead of assimilating into the place they are working, most migrants keep close contacts with their source communities by returning home several times a year. This connection between migrants and the villages they are from facilitates

the migrant network effect through the information flow between rural and urban areas. Aside from the extensively documented evidence that a large percentage of migrants found their first job through the help of relatives, friends and co-villagers who are migrants, the causality between migrant networks and migration is explored by Rozelle et al. (1997) and Zhao (2003). Using village level data, Rozelle et al. approximate migrant networks by lagged village migration and obtain a positive and significant coefficient. Approximating migrant networks by the number of early migrants from the village, Zhao shows that experienced migrants have a positive and significant effect on subsequent migration through practical assistance in the process of migration.

2.3 Econometric Models

The variable representing the migratory work length was recorded as the months an individual worked away from home in 1995. For reasons to be explained below, Tobit, Heckman and Poisson models are used to estimate the determinants of the length of migratory work instead of the commonly used OLS. In addition to the migration duration, the migratory work length variable also allows us to define a migrant in the absence of a formal definition for migrants in the literature. By defining migrants as those who engaged in migratory work for any positive months, a binary variable representing migration status is constructed. In econometric applications, the Probit and Logit models have been used almost exclusively to explain a binary dependent variable. In this section, I will provide theoretical backgrounds of these models and explain the relevance of using these models in analyzing the participation and duration of migration.

2.3.1 Tobit model

Tobit model is developed to deal with the censored dependent variable. The dependent variable representing migratory work length is a left censored variable in the sense that the months of the migratory work for those who did not participate in migration were all reported as zero. In the data set used in this study, about 90 percent of the observations are zero. Ordinary Lease Squire fails to account for the qualitative difference between limit (zero) observations and nonlimit (continuous) observations. Tobit regression model utilizes the information on both the migration participation and the length of migratory work to estimate the marginal effects of education, migrant networks and other variables on the length of migratory work.

The general formulation of a Tobit regression model (left censored) is usually given in terms of an index function,

$$\begin{aligned} y_i^* &= \boldsymbol{\beta}' \mathbf{x}_i + \varepsilon_i \\ y_i &= 0 \quad \text{if } y_i^* \le 0, \\ y_i &= y_i^* \quad \text{if } y_i^* > 0 \end{aligned} \tag{1}$$

where ε_i is a normal error term with zero mean and standard deviation σ. The index variable y_i^* is sometimes called the latent variable.

There are differences in the marginal effects in the model. For the index variable, i.e. the length of migratory work, a change in $\boldsymbol{x}_i$ has two effects: it affects the probability that the person will engage in migration work and the length of the migratory work given the migration participation. The combined marginal effect is:

$$\frac{\partial E[y_i^* \mid \mathbf{x}_i]}{\partial \mathbf{x}_i} = \boldsymbol{\beta}. \tag{2}$$

But for y, the length of migratory work given the migration participation, the marginal effect is only:

$$\frac{\partial E[y_i \mid \boldsymbol{x}_i]}{\partial \boldsymbol{x}_i} = \boldsymbol{\beta}\Phi(\frac{\boldsymbol{\beta}' \mathbf{x}_i}{\sigma}) \tag{3}$$

where Φ represents the normal cumulative distribution function.

Due to these different marginal effects, interpretation of the estimated coefficients should be cautious. In the Tobit procedure used below, the reported estimated coefficients are the marginal effects on the index variable, i.e., the total effects on the probability of choosing migration and the length of migratory work given migration choice.

2.3.2 Heckman's two-step sample selection procedure

Different from the above Tobit model, which uses information on both the migration participation decision and the length of the migratory work to estimate the combined marginal effects, the focus of Heckman's sample selection procedure is on the pure marginal effects of various variables on the length of migratory work given participation. So the essence of the model is to estimate a migratory work length function using an inverse Mills ratio to correct the sample selection bias.

The length of the migratory work is observed only if the individual chooses to participate in the migration. Least squares regression using the observed data—for instance, OLS regression using only data for those who are migrants—produces inconsistent estimates because the data is truncated and consequently the mean of the disturbance term is no longer zero—a violation of the OLS assumption.

To put the problem in a general framework, let the equation that determines the sample selection, i.e., migration participation be:

$$z_i^* = \gamma' w_i + u_i, \tag{4}$$

where z_i^* is the selection variable. Theoretically, z_i^* can be regarded as the difference between a person's expected income associated with participating in migration[36] and the expected income from local work. The individual will chose to participate in migratory work only when the difference is greater than zero.[37] However, in the survey used in this study, only the sign of z_i^* is observed, i.e., whether an individual migrated or not. Let z_i = 1 if the individual is a migrant and z_i = 0 otherwise:

$$\begin{aligned} &z_i = 1 \quad if\ z_i^* > 0, \\ &z_i = 0 \quad if\ z_i^* \le 0, \qquad [38] \\ &prob(z_i = 1) = \Phi(\boldsymbol{\gamma}' \mathbf{w_i}) \\ &prob(z_i = 0) = 1 - \Phi(\boldsymbol{\gamma}' \mathbf{w_i}) \end{aligned} \tag{5}$$

where Φ is the normal cumulative density function. Let the equation of primary interest, i.e., the determinants of the lengths of migratory work, be

$$y_i = \beta' x_i + \varepsilon_i, \qquad \text{observed only if } z_i = 1 \tag{6}$$

Suppose that ε_i and u_i have a bivariate normal distribution with zero means and correlation ρ, the disturbance variance in the primary equation is σ_ε. Suppose, as well, z_i and $\mathbf{w_i}$ are observed for a random sample of individuals, but y_i is observed only when z_i = 1, then

$$E[y_i \mid z_i = 1] = \boldsymbol{\beta}' \mathbf{x} + \rho\sigma_\varepsilon \frac{\phi(\boldsymbol{\gamma}' \mathbf{w}_i)}{\Phi(\boldsymbol{\gamma}' \mathbf{w}_i)} \tag{7}$$

---

[36] The expected income from migratory work is adjusted for the cost of migration and the different living expense in the urban and rural areas.

[37] Though there are many other reasons why people migrates, the income difference is one of the most important reason among all and is used here as a sample selection criterion for simplicity. Some other reasons for migration proposed by New Economics of Labor Migration include the risk diversification, response to relative deprivation etc.

[38] Since there is no information on the scale of z^*, the disturbance variance in the selection equation cannot be estimated.

The parameters of the above sample selection model (7) can be estimated by maximum likelihood. However, this is quite cumbersome, so in practice, Heckman's two-step sample selection procedure is usually used instead. The procedure proceeds as follows:

Step 1: Estimate the Probit equation by maximum likelihood to obtain estimates of γ. For each observation in the selected sample, compute the sample selection term, also called inverse Mills ratio (IMR)

$$\hat{\lambda}_i = \frac{\phi(\hat{\gamma}' \mathbf{w}_i)}{\Phi(\hat{\gamma}' \mathbf{w}_i)}. \tag{8}$$

Step 2: Include the inverse Mills ratio (8) in the length of migratory work equation as a right-hand variable and run OLS regression. The estimated coefficient β is unbiased, reflecting only the marginal effect on the length of the migratory work given the migration choice.

2.3.3 Count regression (Poisson)

The variable representing the length of migratory work is a count variable with the following characteristics: first, the length of the migratory work is non-negative; furthermore, a large percentage of individual (about 90%) did not participate in migration, leading to a preponderance of zero observations; finally, the length of migratory work ranges between 1 month to 12 months, taking the values of 12 integer numbers. The Poisson regression model has been widely used to account for these factors. It has several advantages over other potential estimators: it takes into account the difference between people participated in migration and those did not; it does not lead to negative prediction, as a linear specification would; it also accounts for the discrete nature of the dependent

variable. Thus a Poisson model is also used to estimate the migration length function. The Poisson regression model takes the following form:

$$y_i = e^{\boldsymbol{\beta}' \mathbf{x}_i} + \varepsilon_i \tag{9}$$

2.3.4 Logistic model—a comparison with Probit model

The empirical analysis on the migration participation choice involves the explanation of a binary dependent variable. Aside from Probit model, which is the first step in the above Heckman's two-step estimation procedure, another widely used model is Logit Model. Statistically, Probit model arises from a normal distribution while Logit model comes from a logistic distribution. These two models usually give very similar results and it is a theoretically unsolved question on the choice between these two models. However, Logit model is used more often in the migration literature to identify the determinants of migration participation because of its convenience in interpreting the marginal effects compared with the Probit model, which involves the interpretation of the inverse of the standard normal cumulative distribution function on dependent variable. In the next section, both Probit and Logit models will be fit into the data, the estimated coefficients of both models will be reported, and the corresponding marginal effects will be calculated and compared.

2.4 Data and Empirical Results

2.4.1 Data and Summary Statistics

The data used in this study come from the rural survey of the Chinese Household Income Project 95 (see more detailed description about the Chinese Household Income Project in

the previous chapter). The survey is a sub-sample from approximately 65,000 rural households used by the State Statistical Bureau. A total of 7,998 rural households representing 34,739 individual household members in 19 provinces were interviewed.

Household interviews collected data at both individual and household levels. Individual respondents reported on their personal information (age, sex, education, minority, marital status), employment status, time allocation and detailed sources of income. For people who left the household for at least a month in order to work or to look for work in 1995, total months of migration were recorded. Though the survey also made an attempt to elicit the information on the migration destination and the remittance sent back by migrants, the majority of the respondents did not answer these questions. At the household level, information was collected on the character of the household and residence, costs and income of household operation, different categories of household expenditure, assets and debts. As there is no separate village survey, variables representing community characteristics used in the study are constructed from the aggregation of the individuals and households in the survey.[39]

As there is no consensus in the definition of migrants nor does the survey distinguish migrant workers from local non-farm workers, to simplify the sample selection step in the Heckman's two-step procedure, I define migrants as those who reported a positive number of months in migratory work.[40] In the sample survey, 9.37% of all laborers (24,335 between the ages of 16 to 65 years) engaged in some migratory work in 1995

[39] As the CHIP survey is a random sub-sample of the survey maintained by State Statistical Bureau, the households and individuals in the survey can be regarded as representatives in the counties where they are from.

[40] Different studies have different definitions for migrants to serve the purpose of study. For example, Zhao (1999) defined migrants as those who took part in migratory work, i.e. work out of the home county for at least 3 months. This time limit is consistent with the SSB's definition of non-farm worker; Zhao (2003) defined a worker as a migrant if he or she spent more time in migratory work than either faming or local non-farm work or if he or she derived more income from migratory work than farming or local non-farm work.

(defined as worked away from the household). This percentage is a little higher than the national trend—a commonly cited figure puts the number of rural migrants at 50 million in the mid-1990s, that is, about 7-8% of the total rural labor force. According to the sample survey, the distribution of the length of migratory work is even among migrants (see Table 2.1), indicating that the time limit chosen to define a migrant would not significantly change the estimation results.[41]

Table 2.2 summarizes individual, household and village characteristics for all workers and separates the whole sample into migrants and non-migrants. Compared with non-migrants, migrants are mostly male (71.5% vs. 48.5%), less likely being household head (30.6% vs. 32.0%) and married (49.2% vs. 74.2%). On average migrants are 8 years younger than non-migrants. The education attainment is reported by seven categories based on completion levels in the survey. Table 2.2 compares both the categorical and continuous education levels (see p.28 for the conversion of education categories to a continuous variable) between migrants and non-migrants. On average, migrants have 1.3 years more schooling than non-migrants. A higher percentage of migrants completed the junior and senior high school than non-migrants (54.6% vs. 39.6%, 13.5% vs. 11.3% respectively). Looking at household characteristics, migrant families have a slightly larger household size (4.86 vs. 4.64). The two types of workers do not differ much in terms of the family composition, i.e. the number of kids and old people. The difference in land size is noticeable—migrant households have 23% less land than non-migrant households do (1.24mu vs. 1.52mu). The net financial assets, defined as the difference between total value of all financial assets and total debts at the end of 1995, are higher in

[41] Other criteria, together with the length of the migratory work, have been tried to identify migrants, but no further information is gained for migrant identification based on a second criterion. See Appendix for the percentage of workers engaged in farm work, local non-farm work and sent remittance back home for different migratory work months.

non-migrant households (4405.57 yuan vs. 3547.50 yuan). To summarize, migrant workers in the sample are mostly young, single males, better educated and are from larger families with less land and fewer financial asset. These results are consistent with many other studies of Chinese labor migration (Li and Han, 1994; Zhao, 1998, 1999, 2003). Regarding the county characteristics, migrants are from the counties with lower per capita income (1449.37 yuan vs. 1573.53 yuan). The ratio of local non-farm labors in total labor is used to approximate the local non-farm activities, there is almost no difference in this variable between migrants and non-migrants. Migrants are from the counties in which there are 37.26 migrants on average, while non-migrants live in the counties in which there are only 21.69 migrants. This difference roughly shows the effect of migrant networks in migration.

Table 2.3 provides more descriptive information regarding the pool of migrants. Our migrants sub-sample sent a significant amount of remittance home (859.89 yuan on average),[42] and worked 4.8 days/week and 7 hours/day on average, indicating they are almost full-time workers. Because of the short length of the migratory work chosen to define a migrant, a significant fraction of the migrants still engaged in some household agricultural production. Specifically, in the survey 12.80% of the migrants took a second job aside from their non-farm work, and migrants spent about two months on household farm work. This two-month farm work usually happens during the harvest season when the family is short of labor on farm and hiring outside labor for agricultural production is not a usual practice in rural China. The part-time family farm work among migrants shows the connection between migrants and their families in rural areas.

[42] Roughly only half of the migrants remitted money back, so the average amount of remittance among those who sent money home is much higher. For those who did not remit home, it may be because the higher living expense in urban areas offsets the higher migratory work income.

Aside from agriculture, industry and construction work employs by far the greatest number of migrants—25.16% and 19.6% respectively. A nationwide sample in the same year (Han, 1995) displays a similar distribution of migrant workers among employment sectors: 33 percent of the migrants were employed in construction, 22 percent in manufacturing. Hare (1999) using a county survey in Henan province finds a higher concentration of migrants in construction (56.5%) and a similar percentage in manufacturing (23.4%). With respect to the employment ownership, private and collective enterprises take a relative larger share. Most migrant workers who are not identified as farm labors are ordinary workers and temporary or short-term contract workers—17.45% and 28.63% respectively. These two largest occupational categories are related to the employment sectors where most migrant are employed. As most construction work is contracted by project teams, the contractor hires migrants as short-term contract workers when there is need for labor to work on a project. By employing temporary workers, the contractor saves the fixed labor cost after the project is over. In industry, migrants are mainly engaged in manual work, which is classified as ordinary workers, in contrast to the skilled, professional or technical workers.

The predominance of contract and ordinary workers employed in the industry and construction sectors reflects a number of facts associated with the rural-urban migration in China: first, most migrants take up low level and physically demanding jobs that are characterized by long hours, poor working conditions, low and unstable pay—jobs which urban residents are unwilling to take;[43] second, as most of the low level jobs taken by

[43] According to a survey in Shanghai, a third of all local employees work in white-collar occupations, as professionals, technical staff, leaders of government organizations and enterprises, and office clerical staff. Only about 3 percent of rural migrants work in such occupations (Wang and Zuo, 1999). Nationally, only 3 percent of all long-term temporary migrant employees are in professional/cadre/clerical positions, compared with 24 percent for urban permanent residents (Quanhe Yang and Fei Guo, 1996).

migrants are heavy manual work, it does not require a high education level though the migrant sub-sample displays higher educational attainments, on average, than those who do not migrant. Nevertheless, the overall low levels of educational achievement in the county may influence the employment sector where migrants receive the wage offers. In turn, the low level education required in migratory work discourages the human capital accumulation in rural areas;[44] third, due to the short-term and temporary work character, migrants seldom assimilate with the destination areas, instead they keep a close contact with the families in rural areas;[45] fourth, as a consequence of both low level jobs and the connection between migrants and rural areas, attainment of migrant jobs relies more heavily on the introduction by relatives, friends and co-villagers.

The migrant network effect is further confirmed by the job search method among migrants: most migrants (35.64%) landed the job through the introduction of relatives or friends, though 31.74% claimed to get the job by own search. The role of migrant networks in internal migration in China has been extensively documented in other studies. Drawing a sample survey of 2,838 rural households in seven provinces in 1993, Mallee (2000) records: "Of the jobs held by migrants, 45.1 percent had been found through the migrants' own effort, 26.2 percent with the help of relatives and friends". Using the random sample collected in Xiayi county of Henan province, Hare (1999) finds 83% migrants found jobs through personal introductions and the majority of those providing the introductions were other workers in the same firm and were either an extended relative or friend of the individual. In a sample of 1,304 migrants working in Jinan

[44] Using the 2003 round of the Research Center for Rural Economy survey, de Brauw and Giles estimate the net returns to education for migrants and find a higher return to primary and middle school of schooling than to high school, a finding consistent with an observed decline in high school enrollment with expanding migrant opportunity.
[45] According to the Ministry of Labor survey, only 10% of migrant workers expressed a willingness to remain in their destination areas permanently.

Municipality, 70.5 percent of the migrants reported that they had arranged their job before migrating and 81 percent reported that relatives, friends, and co-villagers served as migration information sources (Liu, 1995). Similar results are found in a sample of 706 migrant workers in Shanghai and southern Jiangsu Province collected in 1995, where more than 75.6% of the migrants got help from relatives and friends during their first trip out of the village (Zhao, 2000); And in a sample survey of 15,000 migrants in Shangdong Province in 1995, more than 70% of surveyed migrants had prearranged jobs before migration (Meng, 2000).

2.4.2 Estimation Results—Migration Choice

The results of estimating the Probit and the Logit models for migration participation choice are presented in Table 2.4. In both models, the dependent variable is a binary variable, taking the value of one for those individuals who are migrants, namely, worked at least one month away from home in 1995, and zero otherwise. The same set of independent variables including individual, household and village characteristics is included in two models. Both the estimated coefficients of the models and the marginal effects of the independent variables on the probability of migration are reported. For continuous variables, marginal effect is a probability change in response to an infinitesimal increase in the value of the independent variable evaluated at mean values. For dummy variables, the marginal effect is computed as the difference in probabilities of migration between the group with designated value 1 and the reference group with the value 0.

The first two models in Table 2.4 use discrete schooling variables in estimation.

Comparison on the marginal effects between the Probit and Logit models shows that the results are very similar in both the magnitudes and the significance levels, an outcome consistent with the theoretical prediction. Therefore, for simplicity without losing generality, the following discussion uses the Probit model only.

The results confirm the pattern shown in the descriptive statistics before that migrant workers tend to be male, younger and unmarried. A male worker is 5.2 percent more likely to participate in migration than a female worker. Marriage reduces the probability of migration by 5.3 percentage points. The probability of migration increases with age but at a declining rate, consistent with the fact that most migrants are between the ages of 18-30. All these three variables are statistically significant at 1 percent level. Individual demographic variables reflect the individual's position and responsibilities within the household, and therefore influence the reservation wage of the individual. Furthermore, the age and the sex of the individual may also affect wage offers as well. Formal education, measured by completion levels, controls for the individual's human capital and is expected to affect wage offers. With illiterate and semi-illiterate being the reference group, all other categories of education are statistically significant in influencing migration decision. Specifically, elementary, junior and senior high school degrees increase the probabilities of migration by 2.7%, 3.2% and 2.4% respectively. Considering the magnitude, junior high school education is most effective in promoting labor migration and its effect far exceeds the effect of senior high school education. As the education level reaches college and above the coefficient becomes insignificant and even negative. The result on how various schooling levels affect migration participation differently lends support to the fact that most migratory work are low-level physical work,

which is does not require a high education and is undesirable to urban residents.

With respect to the variables of household characteristics, household size has a significant positive effect on migration—one additional family member increases the probability of migration by 0.4 percent. Among variables of household composition, only the number of children between the ages of 7-15 has statistically significant effect on migration. One more kid in this age range decreases the migration probability by 0.6 percent. The number of pre-school children does not deter migration as is usually expected. This phenomenon may reflect the role of the extended family plays in raising young children. China has a tradition of grandparents helping to raise children, thus freeing parents to make migration decision independently. Though the coefficient of elder people in the household is insignificant, the negative sign may suggest some potential care needed from the family members given that the old-age-security system is not well established in China. It may also imply the effect of Household Responsibility System (HRS) and the absence of agricultural labor market in rural China. China's farmland arrangements under the HRS obligate the farm household to deliver a part of its grain output to the state at quantities and prices specified by the government. The reduced adult labor on family agricultural production due to the ageing of family members and the scarcity of hiring outside labor may oblige adult workers to give up the migration opportunity to guarantee the agriculture output so as to maintain the ownership of the land. The effect of land availability on migration decision is statistically significant. Individuals from land-scarce households tended to have higher probabilities of migration. Reducing the per capita land in the household by one mu increases the probability of a family worker's migration by 0.4 percent. The per capita allocation of land may affect the

migration decision in the following way. Land is a significant determinant of rural agricultural income because more land is expected to increase the individual's labor productivity in agricultural production, reduced land size tended to reduce rural income, which led to increased motivation to migrate.

The role of migrant networks in facilitating subsequent labor migration is represented by variables representing the number of migrants in the village. The quadratic and cubic terms are used to capture the non-linear effect and to increase the fitness of the model. All these three terms are highly significant at 1 percent level. As the results show, the more migrants a village has, the more likely it is for an individual in this village to choose migratory work. Increasing the number of migrants in the village by one increases the probability of migration by 0.7 percentage point. Considering the large number of rural labor force and the dynamic process, the effect of the migrant networks is noticeable. The declining rate of migrant network effect indicates the importance of the initial presence of the migration in the village. The pioneers open the door to migration by disseminating information about destination labor market and introducing co-villagers to the jobs there, it is those pioneers who play a crucial role in promoting the sequent migration of the rural workers.

The estimated coefficient of the village per capita income is insignificant and minimal in magnitude. [46] The proportion of labor force in local non-farm sector for the whole village is used to approximate the availability of non-farm local employment. Zhao (1999) finds that the non-farm local work is preferred to migratory work. The negative sign of the estimated coefficient, though statistically insignificant, tends to lend support

[46] Income is generally found to have an inverted U shape in the migration literature because incentives for migration are positively related to expected income gains from immigration, but the poorest people may not be able to afford the cost of migration.

to this finding.

In the third and fourth models in Table 2.4, the discrete schooling variables are replaced by a continuous schooling variable and its quadratic term. The significant coefficients exhibit a concave schedule. Combined with the results from the first and second models, the schooling effect is nonlinear, and its marginal effect on migration is positive and is increasing at low education levels. After junior high school, the effect abates and even declines to negative at college and above education, though the coefficient is not significant. The coefficients of the other independent variables are robust to the change of the schooling variables for both models. Comparison between Model 3 and Model 4 again confirms the highly similar results of the Probit and Logit models.

2.4.3 Estimation Results—Length of Migratory Work

Table 2.5 reports the results of estimating the length of migratory work using Tobit model, Poisson regression and Heckman's two-step sample selection procedure. The dependent variable in these three models is the months of migratory work in 1995 reported by individuals. It takes an integer between 0 to 12. As a significant fraction of the respondents did not participate in migration, the dependent variable is left censored with 90% observations being zero. The availability of this variable allows us to further investigate the relative magnitudes of various factors in influencing the migration duration in addition to the stay or leave decision.

The first and second column in Table 2.5 presents the results by fitting a Tobit model and a Poisson regression model respectively to the whole data set. Because the Poisson

regression is a non-linear model, both the estimated coefficients and the marginal effects are reported in the table. All independent variables achieve the same significance levels and exhibit the same signs in two models, though the magnitudes differ with Tobit having a larger marginal effect than Poisson model does. Furthermore, the directions that various factors affect the length of migratory work in these two models are the same as the directions they affect the migration participation choice in the Probit and Tobit models in Table 2.4.

Taking the results of the Tobit model for example, being a male significantly increases the length of migratory work by almost 5 months. Being one year older increases the length by half month, and this increase slows down as people get older. Marriage reduces the length by 4.7 months. Education initially encourages longer migratory work: elementary, junior and senior high school education extend the migratory length by 2.6, 3.4 and 2.6 months respectively. College education and above decreases migratory work by 2.1 months though this effect is not statistically significant. With respect to household characteristics, household size, number of children between 7 to 15 years and per capita land have significant effect of migration length. Having one more family member increases migratory work by 0.56 months, one more kid between the ages of 7-15 and one more mu per capita land in the household decrease the migration work by 0.7 and 0.47 months respectively. The migrant network effect remains highly significant--one more migrant from the village increases the migration work by 0.77 months.

The results of the Heckman's two-step sample selection procedures are totally opposite to those from the Tobit and Poisson models. First sight against intuition, further

consideration may justify some of these unexpected results. The reported marginal effects in the Heckman's procedure are computed after migration participation is controlled. That is, they represent the effects of explanatory variables only on the length of migratory work given migration participation. Though the estimation method employed in the Tobit model utilizes information on both the migration participation and the length of migration, the estimated coefficients do not distinguish the effect of various factors on migration length from their effects on the migration probability. So the reported estimates can be regarded as a combined effect of the independent variables on both migration length and migration probability.

Looking at the estimates of the Heckman's procedure, given migration participation, being a household head reduces migration length by 2.4 months. The reduction may be explained by the important role taken by the household head in the family. Contrary to the migration participation, being a male decreases the length of migratory work after the migration choice is made. The reason why males worked less than females as migrants may be that males are the major labor input in agricultural production, thus they may have to go back home several time a year to help the production during the busy seasons. Younger migrants worked longer than old ones—one year older decreases the migration length by 0.5 months. As most migratory work is manual physical work, young migrants are physically more capable than old ones. Married people spent more time on migration work than the unmarried. The explanation may be the more responsibility and heavier family burden taken by those married migrants. Once the migration decision was made, married migrants worked longer than their unmarried counterpart so as to get more income to support their families. The schooling variables all become insignificant in

affecting the work length after migration was chosen. This is consistent with the fact that most migrants are working in construction and industry where they are not differentiated by their education levels. Variables representing household characteristics are all insignificant in the Heckman's procedure. With respect to the migrant networks, it significantly reduces the working length given migration participation. This result is counter intuitive and the potential reason might be that the migratory work obtained through migrant networks are mostly short-term and temporary jobs. "In the construction sector, experienced migrants often contract in construction projects and then recruit workers from home villages to carry them out. The service industry is another important area of employment for migrants. When help is wanted, employers usually ask their trusted employees to refer workers from their hometowns." (Zhao 2003) The significance of the inverse Mills ratio in the second step of the Heckman's procedure gives evidence to the existence of sample selection problem and validates the choice of Heckman's two-step sample selection procedure to explore the pure effect on the length of migratory work given migration participation.

2.4.4 Estimation Results—Interaction of Education and Migrant Networks

It would be interesting to further examine whether the migrant network effect changes with the education levels. If the migrant networks are more prevalent among low level jobs then it is expected that an increase in the education level of a migrant would reduce the network effect. On the other hand, if better educated migrants are better at interpreting the information in the market then higher education can both enhance and weaken the migrant network effect. Better educated migrants may rely more on their own abilities to

interpret the market information and find a job through their own efforts instead of following their precedents and working under undesirable condition. It may also be possible that the better educated migrants are better at utilizing the migrant networks to explore more and better job opportunities in the city, as an old Chinese saying goes: “Having one more friend means one more opportunity”. Overall, there is no a priori expectation on how the migrant network effect changes with education levels.

To investigate whether migrant network effect differs under various education levels, the interaction terms of education and migrant networks are included in the migration estimations. The discrete variables of education are used with illiterate and semi-illiterate being the reference group to reveal how different schooling level affects migrant network effect. Because the Probit and Logit models give similar results, and because Probit model is the first step in Heckman’s sample selection procedure, only the Probit model is fitted in the migration participation estimation. Heckman’s two-step sample selection procedure is employed to investigate the pure effect of various variables on the length of migratory work given migration participation choice. Tobit model is also used to compare with Heckman’s procedure.

Comparison on the results between the Probit model in Table 2.6 and that in Table 2.4 shows that except for the schooling variables, other independent variables are robust to the inclusion of the interaction terms. With respect to the schooling variables, the inclusion of the interaction terms decreases the effect of junior high school on migration choice though it is still significant. The senior high school education becomes insignificant now and the sign of college education and above changes to positive. Looking at the interaction terms, except for the elementary school*migrant networks term,

all other three interaction terms are statistically significant. The migrant network effects for individuals with junior or senior high school education are equally stronger than that for individuals with no education. This shows that better educated people is better at utilizing the favorable resources such as the migrant networks around them to take advantage of the emerging opportunity in the economy. The migrant network effect becomes significantly negative if the people received college education and above. The negative coefficient reveals some substitutability between education and migrant networks. Highly educated people tend to rely more on their own knowledge than on migrant networks when making a migration choice.

The differences between the two Heckman's procedures in Table 2.4 and 2.6 lie in the education variables and migrant networks variable. Although all the schooling variables in both models are insignificant, the sign of senior high school changes to positive and the sign of college and above changes to negative with the inclusion of interaction terms. The effect of migrant networks reduces in magnitude though still significant. None of the schooling-networks interaction terms are significant. Looking at the signs of these interaction terms only, elementary, junior and senior high school education tend to decrease the migrant network effect on the length of migratory work while college and above degree inclines to increase the network effect given the migration participation choice.

The changes in the schooling variables between two Tobit models are similar to the changes in the Probit model before: the effect of junior high school decreases though still significant, the effect of senior high school becomes insignificant and the sign of college and above turns from negative to positive. Regarding the interaction terms, having a

junior high school degree significantly increases the positive migrant network effect on the length of migratory work while a degree of college and above decreases the effect. All the effects in the Tobit model are combined effects on the length of migratory work and the migration probability. Comparison on the results between the Tobit and Heckman's procedure in Table 2.6 again confirms the clear difference and calls our attention to the proper model choice depending on the research purpose and the accurate results interpretation.

2.5 Conclusion

Using 1995 Chinese rural household survey data, this paper investigates the determinants of participation and duration of temporary rural-urban migration highlighting the role of education and migrant networks. A few findings on the determinants of migration participation are consistent with those found in the existing literature--migrants are more likely to be single, young, male, from larger families with less land, and fewer kids between the ages of 7-15. These results are invariant to the choice of discrete or continuous education variables and to the inclusion of the interaction terms of education and migrant networks. Though being positive and significant on average, education has different effects on migration participation at different levels. In particular, junior high school education significantly improves the likelihood of migration over no education, and its magnitude exceeds that of both primary school and senior high school education, though both are also significant in increasing the migration probability. College education and above does not have a significant effect on migration participation and the sign of the coefficient is negative. Migrant networks are measured by the number of migrants in the

village. The result shows a significant migrant network effect on migration participation decision. Specifically, one more existing migrant leads to 0.7 percent higher probability for an individual in the same village to migrate. The temporary nature of migration is an important channel through which migrant networks facilitate out-migration in China. The close connection between the migrants and the village they are from facilitates the information dissemination and reduces migration cost for subsequent migrants.

The examination of the determinants of the length of migratory work reveals sharp difference on the results between the Tobit model and Heckman's two stage sample selection procedure. Both the Tobit and Heckman's two stage sample selection procedure are developed to deal with the censored dependent variable—data with a significant portion of observations taking one value. In this study, the length of migratory work is a censored variable because for the 90% of the respondents in the survey who did not participate in migration, the length of migratory work is reported as zero. The Tobit model utilizes the months involved in migratory work as well as the migration participation decision to estimate the effects of various factors on migratory work length. The estimated coefficients represent a mixed effect on migration participation and migratory work length. On the other hand, he estimated coefficients of Heckman's sample selection procedure show the pure marginal effect on the length of migratory work given migration participation. This is achieved by fitting a Probit model in the first stage and including an inverse Mills ratio--which is calculated using the estimated coefficients in the Probit model--in the second stage OLS regression to correct for the sample selection bias. The theoretical difference between the Tobit model and Heckman's procedure leads to different estimation results of the two models. The results of the Tobit

model resemble those of the migration participation models in terms of the significance levels and the signs of the independent variables. However, the results of the Heckman's procedure are opposite. Specifically, migrants who are younger, married, female and non-household head worked longer. Education is insignificant on the migratory work length and the migrant networks have a negative effect on working time once migration is chosen.

Looking at the interaction terms of education and migrant networks, the migrant network effect on individual's migration participation decision changes with different education levels. Junior and senior high school education promotes the migrant network effect while the college education and above weakens this effect. Once migration is chosen, the migrant network effect does not vary with education levels significantly.

Distinguishing between the determinants of migration participation choice and the length of migratory work given migration participation helps the policy makers better understand the different aspects of migration. The examination of the determinants of migration participation gives policy makers the guidelines to control the volume of the labor mobility between rural and urban areas. The investigation of the determinants of the length of migratory work given migration choice sheds light on the temporary nature of migration in China—the frequent movement of migrants between rural and urban areas and the short length of residence in their working place.

Finally, the existence of migrant networks justifies the choice of migrant networks as an instrumental variable in the previous chapter. Furthermore, it has important policy implications. First, with migrant networks, the subsequent migrants are better informed about the destination job market than their precedents. As a significant proportion of

migrants landed a job through introduction even before they migrated, the threat of the destabilizing force due to the potential unemployment faced by migrants is less stringent. Second, the reduced individual's migration cost due to the migrant networks is also a reduced cost to the government as less funding is needed to accommodate migrants who cannot find a job especially in the beginning of migration. Third, once the initial migration is triggered, migration will become a self-sustaining and self-enforcing process. Given the positive benefits of migration such as its contribution to children's education in rural areas discussed in the previous chapter, more assistance and information should be provided to those first migrants.

Table 2.1: Migration Participation and Length in Rural China, 1995
(age between 16 to 65 years, N=24,335)

Participation in Migration	Length of Migratory Work			
	1-2 months	3-5 months	6-9 months	10-12 months
9.37	19.07	22.13	24.51	34.29

Table 2.2: Comparison of Individual, Household and Village Characteristics of Migrants and Non-migrants in Rural China, 1995 (age between 16-65 years)

	All Workers N=24,335	Migrants N=2,281	Non-Migrants N=22,054
Individual Characteristics			
Household head (%)	31.89	30.60	32.03
Male (%)	50.68	71.55	48.52
Married (%)	71.93	49.23	74.28
Age	35.52	28.27	36.27
Years of education	7.1	8.25	6.98
Illiterate and Semi-illiterate (%)	17.83	5.39	19.12
Primary school (%)	28.94	26.17	29.22
Junior high school (%)	41.0	54.63	39.59
Senior high[47] (%)	11.46	13.46	11.25
Professional schools and above (%)	0.77	0.35	0.81
Household Characteristics			
Household size	4.66	4.86	4.64
# Kids under 7 years	0.29	0.31	0.29
# Kids 7-15 years	0.67	0.59	0.68
# Elderly 65 years and above	0.18	0.20	0.18
Per capita land (mu[48])	1.49	1.24	1.52
Net financial asset (yuan[49])	4325.14	3547.50	4405.57
County Characteristics			
County per capita income (yuan)	1561.89	1449.37	1573.53
Ratio of local non-farm labors in total labor (%)	0.67	0.67	0.66
# Migrants in county	23.15	37.26	21.69

[47] Including middle level professional, technical or vocational school
[48] 1 mu = 0.0667 hectares
[49] 1 yuan = $US 0.12

Table 2.3: Characteristics of Migratory Employment (N=2,281 Migrants)

Percentage of migrants worked in a second job in 1995	12.80%
Average number of work days per week	4.82
Average number of work hours per day	6.97
Average number of days spent in family farm work	60.09
Average amount of remittance sent back home (yuan)	859.89
Sector of Employment	*Percentage*
Agriculture	29.86
Industry	25.16
Construction	19.60
Transportation, communications, posts and telecommunications	2.89
Commerce and trade	3.20
Restaurants and catering	4.82
Personal services or consulting services	2.89
Others[50]	9.87
Employer Ownership Types	*Percentage*
Farming household	35.82
Private enterprise	16.22
Non-farming individual enterprise	5.57
Township or village enterprise	7.01
Other collective enterprise	10.13
State-owned enterprise of institution	6.31
Sino-foreign joint venture	6.88
Foreign owned enterprise	0.88
Others	9.64
Occupational Category	*Percentage*
Farm labor	30.21
Ordinary worker	17.45
Skilled worker	4.12
Professional or technical worker	1.01
Owner or manager of enterprise	0.44
Village cadre	0.26
Official of party or government office or institution	0.83
Ordinary cadre in an enterprise	0.61
Temporary or short-term contract worker	28.63
Non-farm individual enterprise (such as retailer, driver, etc.)	5.13
Others	9.34
Job Search Methods	*Percentage*
Own search	31.74
Introduced by relative or friend	35.64
Assigned by local labor bureau	1.49
Assigned by township/village collective	0.75
Introduced by labor service company	1.18
Other	3.59

[50] Others include: Forestry, animal husbandry, fishing or water conservancy (1.36%); mining and geological survey and prospecting (0.48%); materials supply and marketing, warehousing (0.75%); real estate (0.04%); public utilities (0.18%); public health, sports and social welfare (0.39%); education, culture, arts and broadcasting (0.53%); scientific and technical service (0.09%); finance, insurance (0.31%); party, government or social organization (0.48%); others (5.26%).

Table 2.4: Probit and Logit Model of Migration Choice
(Dependent Variable: Migrant = 1; Non-migrant = 0)

	Probit Model		Logit Model		Probit Model		Logit Model	
	Discrete Schooling (1)		Discrete Schooling (2)		Continuous Schooling (3)		Continuous Schooling (4)	
	Coefficient Estimate	Marginal Effect	Coefficient Estimate	Marginal Effect	Coefficient Estimate	Marginal Effect	Coefficient Estimate	Marginal Effect
Intercept	-3.449***	-	-7.391***	-	-3.618***	-	-7.797***	-
	(0.326)		(0.807)		(0.331)		(0.815)	
Household head	0.13***	0.014	0.356***	0.0163	0.125***	0.013	0.340***	0.015
	(0.044)		(0.085)		(0.044)		(0.086)	
Male	0.503***	0.052	0.884***	0.039	0.505***	0.052	0.886***	0.038
	(0.033)		(0.062)		(0.033)		(0.062)	
Age	0.058***	0.006	0.148***	0.006	0.059***	0.006	0.150***	0.006
	(0.01)		(0.019)		(0.010)		(0.019)	
Age^2	-0.001***	-0.0001	-0.003***	-0.0001	-0.001***	-0.0001	-0.003***	-0.0001
	(0.0001)		(0.0003)		(0.0001)		(0.0003)	
Married	-0.439***	-0.053	-0.875***	-0.046	-0.438***	-0.052	-0.872***	-0.045
	(0.047)		(0.088)		(0.047)		(0.089)	
Elementary School	0.245***	0.027	0.549***	0.026	-		-	
	(0.052)		(0.109)					
Junior high school	0.302***	0.032	0.650***	0.030	-		-	
	(0.052)		(0.109)					
Senior high school	0.204***	0.024	0.450***	0 .023	-		-	
	(0.061)		(0.124)					
College and above	-0.22	-0.019	-0.595	-0.020	-		-	
	(0.178)		(0.391)					
Years of education	-		-		0.105***	0.011	0.245***	0.010
					(0.016)		(0.035)	
Years of $education^2$	-		-		-0.006***	-0.0006	-0.015***	-0.0006
					(0.001)		(0.002)	

Household size	0.044***	0.004	0.089***	0.004	0.044***	0 .004	0.088***	0.004
	(0.011)		(0.022)		(0.011)		(0.022)	
# elderly 65 years and above	-0.006	-0.0006	-0.014	-0.0006	-0.001	-0.0001	-0.005	-0.0002
	(0.023)		(0.055)		(0.030)		(0.055)	
# Kids 7-15 years	-0.055***	-0.006	-0.112***	-0.005	-0.052***	-0.005	-0.105***	-0.004
	(0.018)		(0.034)		(0.018)		(0.034)	
# Kids under 7 years	0.001	0.0001	-0.013	-0.0006	0.003	0.0003	-0.010	-0.0004
	(0.027)		(0.052)		(0.028)		(0.053)	
Household per capita land	-0.042***	-0.004	-0.085***	-0.004	-0.041***	-0.004	-0.084***	-0.004
	(0.014)		(0.027)		(0.014)		(0.027)	
# Migrants in the village	0.072***	0.007	0.145***	0.006	0.073***	0.007	0.146***	0.006
	(0.006)		(0.011)		(0.006)		(0.011)	
# Migrants in the village^2	-0.001***	-0.0001	-0.002***	-0.00009	-0.001***	-0.0001	-0.002***	-0.00009
	(0.0002)		(0.0003)		(0.0002)		(0.0003)	
# Migrants in the village^3	0.00***	0.00	0.00001***	0.000	0.000***	0.000	0.00001***	0.000
	(0.00)		(0.000)		(0.000)		(0.000)	
Village per capita income	-0.00	-0.000	-0.00004	-0.000	-0.000	-0.00	-0.00005	-0.000
	(0.00)		(0.000)		(0.000)		(0.000)	
Ratio of local nonfarm laborers in total labor	-0.07	-0.007	-0.091	-0.004	-0.048	-0.005	-0.063	-0.003
	(0.076)		(0.153)		(0.077)		(0.154)	
Province dummies	Yes		Yes		Yes		Yes	
# of observations	24335							
# of migrants	2281							
Pseudo R^2	0.197		0.199		0.198		0.2	

Note: To calculate the marginal effect, the continuous variables are evaluated at mean values, dummy variables are evaluated against reference group
The reference group for the education variables is illiterate and semi-illiterate.
The figures in parentheses are standard errors of the estimated coefficients
*Significant at the 10% level
**Significant at the 5% level
***Significant at the 1% level

Table 2.5: Tobit Model, Poisson Regression and Heckman's Two-Step Sample Selection Procedure on the Length of Migratory Work
(Dependent Variable: Months of Migratory Work in 1995)

	Tobit	Poisson		Heckman
		Coefficient Estimate	Marginal Effect	
Intercept	-36.015***	-4.999***	-	47.730***
	(3.321)	(0.243)	-	(17.609)
Household head	1.141***	0.204***	0.055	-2.351***
	(0.451)	(0.030)		(0.778)
Male	4.996***	0.570***	0.150	-4.183**
	(0.339)	(0.020)		(1.907)
Age	0.602***	0.152***	0.040	-0.509**
	(0.098)	(0.007)		(0.257)
Age^2	-0.011***	-0.003***	-0.0007	0.009**
	(0.001)	(0.0001)		(0.005)
Married	-4.679***	-0.783***	-0.249	3.094*
	(0.483)	(0.029)		(1.739)
Elementray School	2.642***	0.575***	0.172	-1.478
	(0.547)	(0.043)		(1.176)
Junior high school	3.444***	0.753***	0.215	-1.430
	(0.549)	(0.043)		(1.365)
Senior high school	2.598***	0.660***	0.226	-0.136
	(0.634)	(0.047)		(1.107)
College and above	-2.054	-0.455***	-0.096	3.227
	(1.852)	(0.149)		(2.316)
Household size	0.556***	0.105***	0.027	-0.160
	(0.117)	(0.007)		(0 .216)
# elderly 65 years and above	-0.254	-0.074***	-0.019	-0.375
	(0.305)	(0 .018)		(0.337)
# Kids 7-15 years	-0.704***	-0.145***	-0.038	0.167
	(0.183)	(0.012)		(0.297)
# Kids under 7 years	-0.174	-0.091***	-0.024	-0.241
	(0.282)	(0.018)		(0.317)
Household per capita land	-0.468***	-0.096***	-0.025	0.175
	(0.143)	(0.010)		(0.227)
# Migrants in the village	0.765***	0.131***	0.034	-0.566**
	(0.060)	(0.004)		(0.293)
# Migrants in the $village^2$	-0.011***	-0.002***	-0.0005	0.009**
	(0.002)	(0.0001)		(0.005)
# Migrants in the $village^3$	0.0001***	0.00001***	0.000	-0.00005**
	(0.000)	(0.000)		(0.00003)
Village per capita income	-0.0002	-0.000	0.000	0 .0005*

	(0.0002)	(0.00002)		(0 .0003)
Ratio of local nonfarm laborers in total labor	-1.120	-0.241***	-0.063	-0.607
	(0.789)	(0.056)		(0.929)
Inverse Mills Ratio	-	-	-	-10.298***
				(4.757)
Province dummies	Yes		Yes	Yes
# of observations		24,335		
(Pseudo) R^2	0.11	0.246		0.171

Note: The R^2 in the Heckman procedure is for the second stage OLS estimation
The figures in parentheses are standard errors of the estimated coefficients
*Significant at the 10% level
**Significant at the 5% level
***Significant at the 1% level

Table 2.6: Probit, Tobit Models and Heckman's Two-Step Sample Selection Procedure (with the Interaction Terms of Education and Migrant Network)

	Probit		Tobit	Heckman
	Estimated Coefficient	Marginal Effect		
Intercept	-3.261***		-34.265***	48.102***
	(0.336)		(3.414)	(18.680)
Household head	0.133***	0.014	1.179***	-2.447***
	(0.044)		(0.450)	(0.847)
Male	0.501***	0.053	4.968***	-4.408**
	(0.033)		(0.339)	(2.068)
Age	0.058***	0 .006	0.595***	-0.532**
	(0.010)		(0.098)	(0.277)
Age^2	-0.001***	-0.0001	-0.011***	0.010**
	(0.0001)		(0.001)	(0.005)
Married	-0.447***	-0.055	-4.760***	3.374*
	(0.047)		(0.482)	(1.923)
Elementray School	0.233***	0.026	2.883***	-1.170
	(0.094)		(0.992)	(1.753)
Junior high school	0.161*	0.017	2.276***	-0.507
	(0.091)		(0.957)	(1.561)
Senior high school	0.064	0.007	1.434	1.190
	(0.106)		(1.103)	(1.482)
College and above	0.582**	0.094	6.355**	-4.609
	(0.288)		(3.020)	(4.760)
Household size	0.044***	0.005	0.559***	-0.184
	(0.011)		(0.116)	(0.236)
# elderly 65 years and above	-0.011	-0.001	-0.299	-0.324
	(0.030)		(0.304)	(0.362)
# Kids 7-15 years	-0.051***	-0.005	-0.656***	0.153
	(0.018)		(0.182)	(0.306)
# Kids under 7 years	0.004	0.0004	-0.147	-0.262
	(0.027)		(0.281)	(0.337)
Household per capita land	-0.044***	-0.005	-0.495***	0.216
	(0.014)		(0.142)	(0.251)
# Migrants in the village	0.044***	0.005	0.470***	-0.346*
	(0.003)		(0.034)	(0.196)
# Migrants in the $village^2$	-0.0003***	-0.00003	-0.003***	0.002*
	(0.000)		(0.0003)	(0.001)
Elementary school* # migrants in the village	0.0004	0.00005	-0.006	-0.013
	(0.002)		(0.024)	(0.030)
Junior high school * # migrants in the village	0.005**	0.0005	0.038*	-0.035
	(0.002)		(0.023)	(0.031)

Senior high school * # migrants in the village	0.004*	0.0005	0.038	-0.044
	(0.003)		(0.028)	(0.036)
College and above * # migrants in the village	-0.039***	-0.004	-0.407***	0.391
	(0.014)		(0.141)	(0.261)
Village per capita income	-0.00*	-0.000	-0.0004*	0.0007**
	(0.00)		(0.0002)	(0.0004)
Ratio of local nonfarm laborers in total labor	-0.074	-0.008	-1.144	-0.548
	(0.076)		(0.784)	(0.993)
Inverse Mills Ratio	-		-	-10.955**
				(5.196)
Province dummies	Yes		Yes	Yes
# of observations			24,335	
(Pseudo) R^2	0.197		0.110	0.165

Note: The R^2 in the Heckman procedure is for the second stage OLS estimation
The figures in parentheses are standard errors of the estimated coefficients
*Significant at the 10% level
**Significant at the 5% level
***Significant at the 1% level

Appendix Table 2.1: Percentage of Workers Engaged in Household Farm Operation, Local Non-farm Work, Sent Back Remittance for Different Migratory Work Length (Age Between 16 to 65 Years, N=24,335)

Length of Migratory Work	Engaged in Household Farm Operation (%)	Engaged in Local Non-farm Work (%)	Sent Back Remittance (%)
0 month	84.84	60.58	0.40
1 month	88.52	63.93	32.79
2 months	91.05	71.88	24.60
3 months	86.67	64.89	55.56
4 months	82.55	58.39	50.34
5 months	77.10	47.33	54.96
6 months	80.43	48.26	46.52
7 months	77.65	41.18	54.12
8 months	58.44	33.77	46.10
9 months	56.67	32.22	58.89
10 months	44.48	26.27	48.66
11 months	23.33	13.89	53.33
12 months	23.97	14.61	44.94

References

Basu, K. and Van, P.H. "The Economics of Child Labor." *The American Economic Review*, Vol. 88, No. 3 (Jun., 1998), 412-427.

Beegle, K., Dehejia, R. and Gatti, R. "Child Labor, Crop Shocks, and Credit Constraints." NBER Working paper 10088, (2003).

Benjamin, D. and Brandt, L. "Property Rights, Labor Markets, and Efficiency in a Transition Economy: The Case of Rural China." Working paper (2000), Department of Economics, University of Toronto.

Binder, M. "Family Background, Gender and Schooling in Mexico." *The Journal of Development Studies*, Vol. 35, No. 2 (Dec., 1998), 54-71.

Bowles, S. "Migration as Investment: Empirical Tests of the Human Investment Approach to Geographical Mobility." *The Review of Economics and Statistics*, Vol. 52, No. 4 (Nov., 1970), 356-362.

Birdsall, N., Ross, D. and Sabot, R. "Inequality and Growth Reconsidered: Lessons from East Asia." *World Bank Economic Review*, Vol. 9, No. 3 (1995), 477-508.

Broaded, C.M. and Liu, C. "Family Background, Gender and Educational Attainment in Urban China." *The China Quarterly*, Vol. 145 (1996), 53-86.

Brown, P. and Park, A. "Education and Poverty in Rural China." *Economics of Education Review*, Vol. 21, No. 6 (2002), 523-541.

Cole, W.E and Sanders, R.D. "Internal Migration and Urban Employment in the Third World." *The American Economic Review*, Vol. 75, No. 3 (Jun., 1985), 481-494.

Connelly, R. and Zheng, Z. "Determinants of School Enrollment and Completion of 10 to 18 Year Olds in China." *Economics of Education Review*, Vol. 22, (2003), 379-388.

De Brauw, A. and Giles, J. "Migrant Opportunity and the Educational Attainment of Youth in Rural China." Working paper (Jul., 2005).

DeGraff, D.S., Bilsborrow, R.E. and Herrin, A.N. "The Implications of High Fertility for Children's Time Use in the Philippines." In C.B. Lloyd (ed.).

Diamond, C. and Fayed, T. "Evidence on Substitutability of Adult and Child Labour." *The Journal of Development Studies*, Vol. 34, No. 3 (Feb., 1998), 62-70.

Duleep, H.O. and Regets, M.C. "Immigrants and Human-Capital Investment." *The American Economic Review*, Vol. 89, No. 2 (May, 1999), 186-191.

Du, Y., Park, A. and Wang, S. “Is Migration Helping China’s Poor?” mimeograph, Department of Economics, University of Michigan (2004).

Fafchamps, M. and Quisumbing, A.R. “Human Capital, Productivity, and Labor Allocation in Rural Pakistan.” *Journal of Human Resource*, Vol. 32, No. 2 (Spring, 1999), 369-406.

Fleisher, B. and Yang, D. “Problems of China’s Rural Labor Markets and Rural-Urban Migration.” Working paper, (2004).

Giles, J. “Is Life More Risky in the Open? Household Risk-Coping and the Opening of China’s Labor Markets.” *Journal of Development Economics*, (forthcoming).

Giles, J. and Yoo, K. “Precautionary Behavior, Migrant Networks and Household Consumption Decisions: An Empirical Analysis Using Household Panel Data from Rural China.” Department of Economics, Michigan State University, mimeo.

Glewwe, P. and Jacoby, H. “Economic Growth and the Demand for Education: Is There a Wealth Effect?” *Journal of Development Economics*, Vol. 74, (2004), 33-51.

Glomm, G. “Parental Choice of Human Capital Investment.” *Journal of Development Economics*, Vol. 53 (1997), 99-114.

Griliches, Z. “Estimates of the Aggregate Agricultural Production Function From Cross Sectional Data.” *Journal of Farm Economics*, Vol. 45 (1963a), 419-428.

Griliches, Z. “Research Expenditures, Education, and the Aggregate Agricultural Production Function.” *American Economic Review*, Vol. 54 (1964), 961-974.

Hannum,E. “Political Change and the Urban-Rural Gap in Basic Education in China.” *Comparative Education Review*, Vol. 43, No. 2 (1999a), 193-211.

Hannum, E. “Poverty and Educational Access in Rural China: Constraints in the Household and Village.” Paper presented at the Annual Meeting of the Population Association of America, New York, March, 1999b.

Hannum, E. “Ethnic Differences in Basic Education in Reform-era Rural China.” *Demography*, Vol. 39, No. 1 (2002a), 95-117.

Hannum, E. “Explaining Educational Gender Inequality in Reform-era Rural China.” Working paper (2002b), University of Pennsylvania.

Hare, D. “‘Push’ Versus ‘Pull’ Factors in Migration Outflows and Returns: Determinants of Migration Status and Spell Duration Among China’s Rural Population.” *The Journal of Development Studies*, Vol. 35, No. 3 (Feb., 1999), 45-72.

Harrise, J. R. and Todaro, M.P. "Migration, Unemployment and Development: A Two-Sector Analysis", *American Economic Review,* Vol. 60 (1970), 126-42.

He, D. *Educational in Contemporary China* (in Chinese). Beijing: Contemporary China Press.

Heckman, J. "China's Investment in Human Capital." Manuscript, Nov., 2003.

Jacoby, H.G. "Productivity of Men and Women and the Sexual Division of Labor in Peasant Agriculture of the Peruvian Sierra." *The Journal of Development Economics*, Vol. 37 (Dec., 1990), 265-287.

Jacoby, H.G. "Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra." *The Review of Economic Studies*, Vol. 60, No. 4 (Oct., 1993), 903-921.

Jacoby, H.G. "Borrowing Constraints and Progress Through School: Evidence from Peru." *The Review of Economics and Statistics*, Vol. 76, No. 1 (Feb., 1994), 151-160.

Jalan, J. and Ravallion, M. "Are the Poor Less Well Insured: Evidence on Vulnerability to Income Risk in Rural China." *Journal of Development Economics*, Vol. 58 (1999), 61-81.

Janvry, A., Fafchamps, M. and Sadoulet, E. "Peasant Household Behavior with Missing Markets: Some Paradoxes Explained." *The Economic Journal*, Vol. 101, No. 409 (Nov., 1991), 1400-1417.

Knight, J. and Li, S. "Educational Attainment and the Rural-Urban Divide in China." *Oxford Bulletin of Economics and Statistics*, Vol. 58, No. 1 (1996), 83-117.

Knight, J. and L, Song. "The Rural-Urban Divide: Economic Disparities and Interactions in China." *Oxford University Press*, Oxford (1999).

Knight, J. and Song, L. "Differences in Educational Access in Rural China." Working paper (2000), Department of Economics, University of Oxford.

Kochar, A. "Urban Influences on Rural Schooling in India." *Journal of Development Economics*, Vol. 74, (2004), 113-136.

Levy, M.B. and Wadycki, W.J. "Education and the Decision to Migrate: An Econometric Analysis of Migration in Venezuela." *Econometrica*, Vol. 42, No. 2 (Mar., 1974), 377-388.

Li, F and Han, X. 1994, "Age Groups and Educational Composition of Peasants Leaving for Nonagricultural Work-Studies of Rural Labor Migration" (in Chinese). *Zhongguo Nongcun Jinji* (China's Rural Economy), No. 8 (Aug., 20 1994), 10-14; translated in Joint Publications Research Service-CAR-95-054 (15 Nov. 1994. pp.61-9).

Li, H. “Economic Transition and Returns to Education in China.” *Economics of Education Review*, Vol. 22 (2003), 317-328.

Liu, Q. “Rural-Urban Migration Sample Survey in Jinan Municipality of China: Sample Design and Preliminary Results”, Chinese Academy of Social Sciences, unpublished manuscript (1995).

Lloyd, C. B. “Investing in the Next Generation: The Implication of High Fertility at the Level of the Family.” Research Division Working Paper No. 63. New York, The Population Council, 1994.

Mallee, H. “Agricultural Labor and Rural Population Mobility: Some Observations,” in *Rural Labor Flows in China*, edited by L.A. West and Y. Zhao. Berkeley: Institute of East Asian Studies, University of California at Berkeley, 2000.

Massey, D.S., J. Arango, G.Hugo, A. Kouaouci, A. Pellegrino, and J.E. Taylor. “Theories of International Migration: A Review and Appraisal.” *Population and Development Review*, Vol. 19, No. 3 (1993), 431-66.

Meng, X. *Labor Market Reform in China.* Cambridge: Cambridge University Press, 2000.

Patrinos, H.A. and Psacharopoulos, G. “Family Size, Schooling and Child Labor in Peru—An Empirical Analysis.” *Journal of Population Economics*, Vol. 10 (1997), 387-405.

Psacharopoulos, G. “Child Labor Versus Educational Attainment: Some Evidence From Latin America.” *Journal of Population Economics*, Vol. 10 (1997), 377-386.

Psacharopoulos, G and Woodhall, M. “Education for Development.” New York: *Oxford University Press*, 1985.

Ray, R. “Child Labor, Child Schooling, and Their Interaction with Adult Labor: Empirical Evidence for Peru and Pakistan.” *The World Bank Economic Review*, Vol. 14, No. 2, 347-67.

Riskin, C., Zhao, R and Li, S. “Chinese Household Income Project, 1995. Codebook and Data Collection Instrument.” ICPSR 3012.

Rosenzweig, M.R. and Evenson, R. “Fertility, Schooling, and the Economic Contribution of Children of Rural India: An Econometric Analysis.” *Econometrica*, Vol. 45, No. 5 (Jul., 1977), 1065-1079.

Rosenzweig, M.R. “Neoclassical Theory and the Optimizing Peasant: An Econometric Analysis of Market Family Labor Supply in a Developing Country,” *The Quarterly*

Journal of Economics, Vol. 94, No. 1 (Feb., 1980), 31-55.

Rosenzweig, M.R. "Educational, Subsidy, Agricultural Development, and Fertility Change." *The Quarterly Journal of Economics*, Vol. 97, No. 1 (Feb., 1982), 67-88.

Rosenzweig, M.R. and Wolpin, K.I. "Specific Experience, Household Structure, and Intergenerational Transfers: Farm Family Land and Labor Arrangements in Developing Countries." *The Quarterly Journal of Economics*, Vol. 100, Supplement (1985), 961 – 987.

Rozelle, S., Li, G., Shen, M., Hghare, A. and Giles, J. "Leaving China's Farms: Survey Results of New Paths and Remaining Hurdles to Rural Migration." *The China Quarterly*, Vol. 158 (1999), 367-393.

Schultz, T.P. "Education Investments and Returns." Handbook of Development Economics, Vol. 1, Ch 13.

Schwartz, A. "Interpreting the Effect of Distance on Migration." *The Journal of Political Economy*, Vol. 81, No. 5 (Sep. – Oct., 1973), 1153-1169.

Siddiqi, F. and Patrinos, H.A. "Child Labor: Issues, Causes and Interventions." Working Paper.

Skoufias, E. "Using Shadow Wages to Estimate Labor Supply of Agricultural Households." *American Journal of Agricultural Economics*, Vol. 76, No.2 (May, 1994), 215-227.

Solinger, D. "Contesting Citizenship in Urban China: Peasant Migrants, the State, and the Logic of the Market." University of California Press (Berkeley), 1999.

Stark, O. "The Asset Demand for Children During Agricultural Modernization." *Population and Development Review*, Vol. 7, No. 4 (Dec., 1981), 671-675.

Stark, O. "The Migration of Labor." Blackwell Publishers, 1991.

Stark, O., Helmenstein, C. and Prskawetz, A. "Human Capital Depletion, Human Capital Formation, and Migration: a Blessing or a "Curse"?." *Economics Letters*, Vol. 60 (1998), 363-367.

Taylor, J.E. and T.J. Wyatt. "The Shadow Value of Migrant Remittances, Income and Inequality in a Household Farm Economy." *Journal of Development Studies*, Vol. 32, No. 6 (1996), 899-912.

Taylor, E. and Yunez, N.A. "The Returns from Schooling in a Diversified Rural Economy." *American Journal of Agricultural Economics*, Vol. 82 (May, 2000), 287-297.

Taylor, J. E., Rozelle, S. and de Brauw, A. “Migration and Incomes in Source Communities: A New Economics of Migration Perspective from China.” *Economic Development and Cultural Change*, Vol. 52, No. 1 (2003), 75-101.

Todaro, M.P. “A Model of Labor Migration and Urban Unemployment in Less Developed Countries.” *The American Economic Review*, Vol. 59, No. 1 (1969), 138-148.

Udry, C. “Child Labor.” Discussion paper No. 856. Economic Growth Center, New Haven, CT.

Wang, F. and Zuo, X. “Inside China’s Cities: Institutional Barriers and Opportunities for Urban Migrants.” *The American Economic Review*, Vol. 89, No. 2, Papers and Proceedings of the One Hundred Eleventh Annual Meeting of the American Economic Association (May, 1999), 276-280.

Wang, S. and Yuan, M. “Children’s Education and Problems in China (in Chinese).” In *1992 Children Survey of China*. Beijing: China Statistics Press.

Welch, F. “Education in Production.” *The Journal of Political Economy*, Vol. 78, No. 1 (Jan.- Feb., 1970), 35-59.

Williamson, J.G. “Migration and Urbanization.” In *Handbook of Development Economics*, Vol.1, Ch 11. Edited by Hollis B. Chenery and T.N.Srinivasan. Elsevier Science B.V., 1997.

Yang, D.T. “Education and Off-Farm Work.” *Economic Development and Cultural Change*, Vol. 45 (1997a), 613-632.

Yang, D.T. “Urban-Biased Policies and Rising Income Inequality in China.” *The American Economic Review*. Vol. 89, No. 2 (May, 1999), 306-310.

Yang, D.T. “Determinants of Schooling Returns during Transition: Evidence from Chinese Cities.” Working paper.

Yang, D and Chen, V. “Policy Reform, Private Enterprises, and Rural Household Earnings,” in *Private Enterprise and Economic Development in China*, edited by Shuanglin Lin, World Scientific Press, 2005.

Yang, Q and Guo, F. “Occupational attainment of Rural to Urban Temporary Economic Migrants in China, 1985-1990.” *International Migration Review*, Vol. 30, No. 3 (Fall 1996), 771-87.

Zhao, S. “Organizational Characteristics of Rural Labor Mobility in China,” in *Rural Laobr Flows in China*, edited by L.A. West and Y. Zhao. Berkeley: Institute of East Asian Studies, University of California at Berkeley, 2000.

Zhao, Y. “Labor Migration and Returns to Rural Education in China.” *American Journal of Agricultural Economics*, Vol. 79, No. 4 (Nov., 1997), 1278-1288.

Zhao, Y. “Leaving the Countryside: Rural to Urban Migration Decisions in Mainland China.” China Center for Economic Research, Beijing University, (Dec., 1998), Working paper.

Zhao, Y. “Labor Migration and Earnings Differences: The Case of Rural China.” *Economic Development and Cultural Change*, Vol. 47, No. 4 (Jul., 1999), 767-782.

Zhao, Y. “The Role of Migrant Networks in Labor Migration: The Case of China.” *Contemporary Economic Policy*, Vol. 21, No. 4 (Oct., 2003), 500-511.

Zhou, W., Ma, Y.Q. and Wu, S.Q. “The Analysis of Girls Educational Problems in Ling Sha Hui Minority (in Chinese).” *Nationalities Education Research* (1995), 18-19.