SELF-SELECTION, MIGRATION AND INEQUALTY

IN SOURCE REGIONS

Dissertation zur Erlangung des Doktorgrades

des Fachbereiches Wirtschaftswissenschaften

der Justus-Liebig-Universität Gießen

vorgelegt von Yunhe Lu

1. **Gutachter: Prof. Dr. A. Bohnet**
2. **Gutachter: Prof. Dr. J. Meckl**

Gießen, December 2015

1. Gutachter: Prof. Dr. A. Bohnet
2. Gutachter: Prof. Dr. J. Meckl

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This dissertation provides theoretical and empirical explorations of the impacts that the internal migration has on inequality in source regions. Chapter 2 identifies the lack of microeconomic contents, especially those on intra-household economic linkages as the primary reason why the Roy model may give misleading predictions on the impacts in developing countries like China. Keeping this drawback in mind, Chapters 3, 4 and 5 extend the Roy model by combining its statistical assumptions with simple yet stylized models capturing decisions of income-maximizing rural households constrained by factor endowments. Comparing to the original model, the extended models have more solid microfoundations, they relate also to a different statistical problem, i.e. censoring rather than truncation. The difference in statistical structure leads to different predictions on the impacts. Particularly, in an empirically relevant setting where different types of labor inputs are heterogeneous in productivity but perfectly substitutive in household agriculture, Chapter 4 shows that unlike the original model, the extended model admits decreasing, constant and increasing rural inequality. Chapter 4 highlights also the non-causal relationships between the pattern of selection and the direction of change of rural inequality. Chapter 6 tests these relationships with the Chinese Health and Nutrition Survey (CHNS) data. It is found that the pattern of selection of rural-to-urban migration in China could undergo two transitions between 1991 and 2009. Nevertheless, the directions of change of rural inequality predicted using patterns of migration selection could disagree with the data in certain sub-period. Further analysis attributes much of these disagreements to a well-known puzzle that the actual size of migration is often smaller than its prediction based on observed intersectoral wage gap. Chapter 6 offers thus several preliminary explanations to help resolve this puzzle.

ACKNOWLEGMENTS

In preparing this dissertation, I benefited greatly from many people. Firstly, I would like to express my immense gratitude to my supervisor, Prof. Armin Bohnet, for the valuable suggestions and guidance that he gave, for his insistence on empirical studies and for his patience and tolerance when progress of this dissertation was extremely slow and I was very frustrated by slow progress. Without his guidance and tolerance, the completion of this dissertation would be unthinkable.

I also appreciated Prof. Jürgen Meckl for his willingness to co-supervise this dissertation, for his excellent teaching, and for his great efforts to comment on drafts at various stages, which set a high standard to both thinking and writing of this dissertation. I hope this thoroughly revised dissertation could meet his standard.

I benefited greatly from the exchanges with Prof. Jingbei Hu. His passion on the economic research in particular has strong influence on me. I want to thank Dr. Caixing Zheng for her hospitality during my stay in the Chinese-German School of Graduate Studies (CDHK) at Tongji University, Shanghai. I am also grateful for professors, fellow Ph. D. students, participants of seminars and conference in Giessen and Shanghai, and friends whose feedbacks improved this dissertation substantially.

I received financial supports from the DAAD, which released me from financial burdens at the earlier stage of my Ph. D. study.

Last but by no means the least, I am particularly indebted to my parents who support me with their love. This dissertation is dedicated to them.

CHAPTER 1

INTRODUCTION

In the past three decades, China has experienced one of the largest internal migrations in the human history. In the period between 1978 and 2013, yearly around 5.3 million of rural residents left for the cities and around 7.4 million of agricultural workers got employed in non-farm industries.[1] Consequently, the percentage of rural employment fell in this period from 76.3 to 61.0, whereas the percentage of agricultural employment fell even more sharply from 70.5 to 39.6.

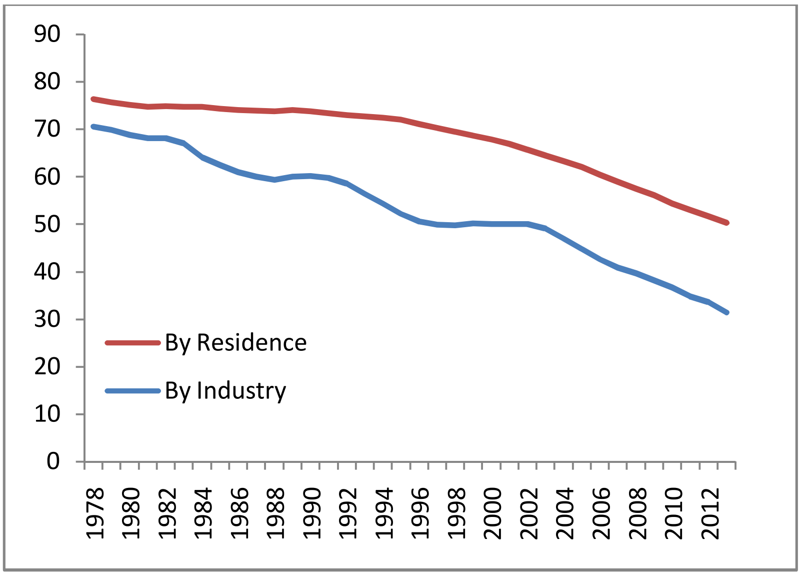


Figure 1.1: Declines in Shares of Rural and Agricultural Employments

Source: The data on the total employment, employments in the rural region and in the primary industry are available from Chinese Statistical Yearbook, 2014, Tables 4.2 and 4.3 and China Compendium of Statistics 1949-2008, Table 1-4. Calculations of the ratios are my own.

As migration occurs elsewhere, the rural-to-urban migration in post-reform China is also selective: Rural residents who choose to leave countryside and work in

---

[1] The yearly flows of migration are estimated using the formula: $M(t,t+1)=L(t)[1+g(t,t+1)]-L(t+1)$. $L(t)$ in the formula denotes the total rural/agricultural employment in the year t, $g(t,t+1)$ denotes the natural growth rate of the rural/agricultural employment between years t and $t+1$. Considering that no data on the natural growth rate in specific region or sector is available in official statistics in China, I replace them with the growth rate of total employment in the same year. For more statistical details, the reader may refer to Hu (2010).

the urban non-agricultural sector have usually different individual characteristics and household backgrounds, and hence different earning capacities in comparison with populations at the source and destination. There exist already ample empirical evidences supporting for this observation.[2]

The enormously large and selective rural-to-urban migration must have profound effects on many aspects of the Chinese economy such as growth, industrialization, international trade, income distribution. In this thesis, however, I will confine myself to just one aspect of them, namely the impacts that rural-to-urban migration has on the income distribution in rural China.

I am interested in this very specific research question mainly for two reasons: Firstly, this is because the question relates closely to the welfare of large and relatively poor population subgroup in China, namely the rural residents. Many empirical studies report a rising trend of inequality among rural residents.[3] Some of these studies argue further that rural-to-urban migration occurring at roughly the same period should be responsible for a substantial fraction of the increase in rural inequality. After reviewing these arguments, however, it is determined that most of these studies are based on flawed economic reasoning, casual observations in combination with sparse empirical findings, and thus may not be convincing. Therefore, the present thesis intends to provide a solid and more general theoretical analysis on the relationships between the rural-to-urban migration and rural inequality. Secondly, this is also because the question is interesting of its own right. In particular, it can be viewed as a counterpart to the question on the impacts of immigration on the

---

[2] Greenwood (1997) provides an excellent survey on the economic issues about internal migration in context of developed countries, which includes an extensive discussion on the individual characteristics of migrants. For the characteristics of rural-to-urban migrants in China, interested reader may refer to Zhao (2004) and Lu (2008).

[3] See for example, Benjamin et al. (2008), Li (2003), Ravallion and Chen (2007).

destination labor market and inequality in that market, which has been the focus of the scholarly debates for last few decades and has lead to noticeable progresses in understanding the functioning of labor market facing with external labor supply shocks.[4] At a higher level of abstraction, both questions are similar in all aspects except for the direction of worker flows and regions considered. The close analogy between two literatures enables us to learn from the immigration and destination literature when studying the question of this thesis. Meanwhile, considering that the question of this thesis is less frequently studied than its counterpart,[5] some lights could be shed on the immigration and destination literature at lower costs by studying the present question. I believe these reasons suffice to justify my explorations.[6]

To identify the relationship between rural-to-urban migration and rural inequality, I may take different routes. For reasons that migration is often regarded as a selective process, and that the widely used Roy model[7] offers a direct linkage between self-selective behaviors at large, including migration of individual agents and the resulted distributional impacts at some aggregate levels, this thesis adopts the Roy model as

4 See for example, Card (1990), Borjas (2003), Ottaviano and Peri (2008, 2011), Lewis (2003, 2012). For excellent reviews covering this issue, the reader may refer to Borjas (1999), Blau and Kahn (2015).

5 Mishra (2007) is a noticeable exception.

6 Although in this thesis, I focus mostly on relationships between rural-to-urban migration and its distributional impacts on the source regions with explicit reference to the post-reform Chinese experiences, my research could offer insights to investigations in other contexts, both in developing and developed countries. This is firstly because the massive migration leaving agricultural for non-agricultural sectors is by no means unique in China, it is in fact a worldwide phenomenon and is of great importance for the developing countries today. According to a report of the Food and Agriculture Organization (FAO, 2011, Table A4), the agricultural share of labor force fell from 65.3 percent to 48.2 percent in countries in developing regions during 1980 and 2010. Migration could have large impacts on income distributions in source regions. This is also because my research could be relevant to the developed countries in history and now. Historically, Kuznets (1955) uses the intersectoral migration as one of the main mechanisms to explain the change of income inequality in the process of economic growth. The large-scaled unskilled immigration, as argued in the main text, which is a counterpart for out-migration, has often caused serious social-economic problems in developed countries today and thus remains a focus of academic discussions for decades.

7 The model is named after A. D. Roy who published his seminal paper in 1951. It is also called as the Roy-Borjas model, for Borjas (1987) is the first paper that presents a simple and parametric two-sector Roy model.

the departure for all theoretical discussions.

Therefore, in the first half of Chapter 2, I review briefly the so-called standard Roy model with emphases on its basic assumptions and its key predictions on the patterns of migration selection and particularly on the impacts that the intersectoral migration has on source inequality. The prediction on the changes of source inequality is then contrasted with a well-known conjecture in literature and the observation of rising rural inequality to reveal some potential inconsistencies. The second half of Chapter 2 begins a series of theoretical efforts aiming at understanding and reconciling these inconsistencies. I find the lack of microeconomic contents, especially those relating to the intra-household economic linkages in the standard Roy model is most likely to explain these inconsistencies.

As responses to the drawback of the standard Roy model, Chapters 3, 4 and 5 build several simple yet highly stylized models to capture income-maximizing rural households' interrelated production, labor supply and migration decisions in three settings of increasing complexity where labor inputs are assumed to be homogeneous, heterogeneous and perfectly substitutive, heterogeneous and imperfectly substitutive. These microeconomic models combine then with the largely statistical assumptions of the standard Roy model. Because these extended Roy models take the intra-household economic linkages that are totally missing in the standard Roy model into account, they serve as proper analytical frameworks for studying the relationships between out-migration and change of rural inequality in Chapters 3 and 4, but to a lesser extent in Chapter 5. Particularly, in the setting discussed in Chapter 4 that turns out to have the highest empirical relevance, I find the approximate non-causal relationships between the pattern of migration selection and direction of change of rural inequality as follows. Out-migrations exhibiting positive sorting tend to coexist with increase in rural inequality; out-migrations exhibiting negative sorting tend to coexist with decrease in rural inequality; migrations exhibiting non-hierarchical sorting are compatible to both increase and decrease in rural inequality.

Chapter 6 tests these theoretical relationships using datasets collected in the Chinese Health and Nutrition Surveys (CHNS), 1991-2009. The preliminary results suggest some disagreements between theory and data. I give thus several possible explanations for the disagreement. Empirical explorations in this chapter yield also insights for broader researches. Particularly, evidences show that different types of labor inputs in agriculture in China during the period 1991-2009 can be best described as heterogeneous and perfect substitutive labor inputs. Moreover, Chapter 6 reports for the first time in literature that the pattern of migration selection could undergo two transitions in the period. Chapter 7 summarizes this thesis and offers preliminary discussions on policy implications of my theoretical and empirical explorations.

CHAPTER 2

THE ROY MODEL AND ITS MICROECOMOMIC CONTENTS

In this Chapter, I provide firstly a brief review on the theoretical aspect of the Roy model.[8] This review emphasizes **basic assumptions** of the Roy model and its **key predictions** on the pattern of selection and on the impacts that intersectoral migration has on the inequality at the source. Notice that the review is more than a collection of standard results available from existing studies.[9] By reviewing the literature, I demonstrate that the prediction of the Roy model on the impacts of migration on inequality at the source contradicts with a well-known theoretical conjecture and the prediction is potentially inconsistent with the observation of rising inequality in rural China. These inconsistencies motivate the following theoretical explorations starting from the second half of this chapter and continuing through Chapters 3, 4 and 5.

2.1 A Review on the Roy Model

The basic assumptions shared by most of studies that use the Roy model as main analytical framework can be summarized as follows:

(1) The Roy model assumes that the underlying distribution of the log wages that a heterogeneous population of workers face in the source sector 0 and the host sector 1 is jointly bivariate normal such that

$$(\log w_0, \log w_1) \overset{i.i.d.}{\sim} N(\mu_0, \mu_1, \sigma_0, \sigma_1, \rho_{01}), \tag{2.1}$$

[8] This review does not cover the empirical aspect of the Roy model, since they are often technical and thus beyond the scope of this review. For a comprehensive survey on the empirical issues of the Roy model and its various extensions, readers may see French and Taber (2011).

[9] Most of these materials in this review are standard, thus results in the text are often given without complete proof. Readers may see Appendix A for mathematical details. Readers may also refer to original studies such as Borjas (1987), Heckman and Sedlacek (1985) and Gould (2002).

where μ_0, μ_1 denote the mean log wages in sectors 0 and 1, σ_0, σ_1 denote the standard deviations of log wages in sectors 0 and 1, while ρ_{01} denotes the correlation coefficient between log wages in both sectors.

(2) The Roy model assumes further that all workers in the population of interest are wage-income maximizers in the sense that they choose to work in the sector where they receive the highest wages. To be concrete, if a worker chooses to be employed in sector 1, he must earn higher wage in sector 1 than what he could earn in sector 0. Consequently, the observed log wages resulted from the self-selection takes the form of

$$\log w = \max\{\log w_0, \log w_1\}. \tag{2.2}$$

(3) Last but not the least, the Roy model assumes that the post-selection log sectoral wage distribution can be obtained by truncating underlying log sectoral wage distribution. Correspondingly, the level of post-selection inequality of sectoral wage distribution can be measured by a truncated variance.[10] For instance, the observed post-selection wage inequality in the source sector 0 is given by $Var(\log w_0 \mid \log w_0 \geq \log w_1)$.

To differentiate from various extensions of the Roy model developed in subsequent chapters, I call henceforth the model satisfying all three basic assumptions listed above as the standard Roy model henceforth.

Following Borjas (1987), I categorize all intersectoral migrations into three patterns of selection, namely the positive sorting, the negative sorting and the non-

[10] See Greene (2007), Chapter 24 for discussions on the truncated distribution and its moments.

hierarchical sorting.[11,12] Roughly speaking, the positive sorting refers to cases in which out-migrants from sector 0 outperform on average other workers from the same population when they were employed in both sectors. The negative sorting refers to cases in which out-migrants are outperformed on average by other workers in both sectors. The non-hierarchical sorting refers to the rest of cases in which out-migrants are outperformed on average by other workers on average in the source sector 0, but outperform on average other workers in the host sector 1. The sufficient and necessary conditions associating with each of the three patterns of selection in terms of distributional parameters are as follows:[13]

$$\text{Positive Sorting:} \quad \rho_{01} > \sigma_0 / \sigma_1, \text{ and } \sigma_0 < \sigma_1; \tag{2.3}$$

$$\text{Negative Sorting:} \quad \rho_{01} > \sigma_1 / \sigma_0, \text{ and } \sigma_0 > \sigma_1; \tag{2.4}$$

$$\text{Non-hierarchical Sorting:} \quad \rho_{01} < \min\{\sigma_0 / \sigma_1,\ \sigma_1 / \sigma_0\}. \tag{2.5}$$

Considering that the standard Roy model admits three patterns of selection, one may conjecture that intersectoral migrations of different patterns of selection could result in different impacts on the inequality in the source sector 0. However, it turns out that the standard Roy model rejects such conjecture. Under its basic assumptions, formal derivation relying heavily on the properties of truncated normal distribution shows clearly that the level of post-selection inequality must be lower than its

---

[11] In Borjas (1987), the third pattern of migration selection is called as the "refugee sorting". However, considering that this pattern of selection is not restricted to the migration of refugees, the non-hierarchical sorting would be a more appropriate name for most cases. Note that unlike other patterns of selection, if the migration exhibits non-hierarchical sorting, then the "skill" of workers must be multi-dimensional. Otherwise, as long as the wages in both sectors are positively monotonic transformations of the one-dimensional skill, a group of workers who outperform others in one sector must outperform other workers in another sector, too.

[12] There is no fourth pattern of migration selection for the reason that its existence would imply the correlation coefficient between wages earned in both sectors exceeds unity, which is clearly untrue.

[13] See Appendix A for details.

pre-selection level, that is,[14]

$$Var(\log w_0 \mid \log w_0 \geq \log w_1) \leq Var(\log w_0). \tag{2.6}$$

In other words, the standard Roy model predicts unambiguously that intersectoral migration, regardless of its pattern of selection, always reduces the wage inequality in the source sector.

The result given by equation (2.6) has been derived in a quite general setting. Thus, it could have large potential to be applied to the concrete setting focused by this thesis. For the present setting, this result should be read as the selective rural-to-urban migration always reduces the wage inequality in rural regions of China, and this result is unaffected by the pattern of migration selection.

This prediction is clearly inconsistent with a well-known theoretical conjecture shared by Lipton (1980) and Li (2003), among others, namely the pattern of migration selection and change of rural inequality caused by migration are often closely related. Out-migrations of different patterns are likely to have qualitatively different impacts on the rural inequality. In particular, some authors argue further that the positive sorting out-migration tends to increase the rural inequality, while other patterns of out-migrations tend to decrease the rural inequality.

Furthermore, if this prediction could be justified, then unlike the belief holding by many scholars, rural-to-urban migration cannot explain the rising inequality observed in rural China. To the contrary, the selective rural-to-urban migration should be considered as a counterforce that offsets in part the increase in rural inequality

[14] See Appendix A for details.

resulted from other factors, notably the skill biased technical change.

Both inconsistencies need to be better understood and reconciled. In fact, all of the theoretical explorations below can be viewed as efforts made in this direction.

2.2 The Microeconomic Contents of the Roy model

The standard Roy model has been employed to study a wide range of self-selective behaviors, notably individual agents' occupational, regional, educational choices and their aggregate market implications. The popularity of this model may reflect two of its features. On the one hand, it demonstrates that the model is indeed a unified and analytically convenient framework;[15] on the other hand, it may also suggest that the model provides little beyond that framework – otherwise, we should not expect self-selective behaviors with such varying natures could all fit into the same framework. As the review of the standard Roy model shows, two of three of its basic assumptions, i.e. assumptions (1) and (3) are mainly statistical. They describe the distribution of log sectoral wages and how it changes during the migration. Only one of its basic assumptions, i.e. assumption (2) concerns the microeconomic contents of the model, but it is still based on a choice problem of the simplest kind.[16] Therefore, it is legitimate to criticize the standard Roy model for its lack of microeconomic contents – at least, most of its microeconomic contents hide behind the statistical descriptions. The lack of or say, the ambiguity in microeconomic contents of the standard Roy model could be partially responsible for the two inconsistencies mentioned above.

---

[15] Empirical inclined economists find the Roy model attractive, mainly because this simple model can be used to motivate procedures to correct the selection biases.

[16] The term "microeconomic contents" is borrowed from the survey of Neal and Rosen (2000). The following theoretical exploration that emphasizes the microeconomic contents of the Roy model is partly inspired by their evaluation.

To clarify and enrich the microeconomic contents of the standard Roy model, I can rely nothing but its basic assumptions, especially its two mainly statistical assumptions (1) and (3).

By looking at its basic assumption (1), it becomes clear that this assumption states only that the log underlying wages are jointly normally distributed, but it remains silence about the source of heterogeneity in earning capacity. Different from the conventional practice that attributes all earning heterogeneity indifferently into the heterogeneity in vaguely defined skills, inspired by the econometric literature, I propose in this thesis to decompose the log underlying wage that an agent *i* could earn in sector *j*, $\log w_{ji}$ into several components, each of which has relatively clear economic interpretation such as

$$\log w_{ji} = \mu_j + \varepsilon_{ji}, \tag{2.7}$$

where μ_j denotes the mean of log wages in sector *j*; ε_{ji} denotes the demeaned error term that measures the deviation of the wage that an individual agent *i* could earn in sector *j* from μ_j. The demeaned error term can be further decomposed as:

$$\varepsilon_{ji} = \varphi_i + \xi_{ji} + \eta_{ji}. \tag{2.8}$$

The first term φ_i in equation (2.8) refers to the agent-specific fixed effect. The skill components associating with φ_i are valued the same in both sectors and thus can be perfectly transferred across sectors. Typical examples for φ_i include the earning heterogeneity relating to skill components acquired from elementary and secondary educations as well as to the innate ability. The second term ξ_{ji} refers to the earning heterogeneity relating not only to individual agents' characteristics, but also to the sector where they work. The skill components associating with ξ_{ji} are often valued differently across sectors and thus cannot be perfectly transferred. Typical examples for ξ_{ji} are often related to the sector-specific physical or human capital. The third term η_{ji} in equation (2.8) is the residual of the conceptual decomposition and it refers to pure noises or unexplained components in the wage data.[17] Thus, η_{ji} has often no clear economic interpretation. To simplify the discussions below, I ignore the residual term η_{ji}, while focus on φ_i and ξ_{ji}. Hence, the demeaned error term ε_{ji} can be decomposed approximately as:

$$\varepsilon_{ji} \approx \varphi_i + \xi_{ji}. \tag{2.9}$$

Hereafter, equations (2.7) and (2.9) are called as the wage decomposition formulas. These formulas provide useful guidance for the theoretical explorations in Chapters 3, 4 and 5. In Chapter 3, I begin by considering the question on the impact that out-migration has on rural inequality in a simple homogeneous labor setting in which the error term φ_i and components of ξ_{ji} relating to individual agents' skills are assumed the same for all rural workers. In Chapters 4 and 5, the assumption of labor homogeneity will be gradually relaxed. Both φ_i and ξ_{ji} are allowed to be univariate random variables with non-degenerate distributions. I will revisit the same question in heterogeneous labor settings with increasing complexity.

Furthermore, according to assumption (3) of the standard Roy model, the post-migratory wage inequality in rural regions should be measured by the truncated

---

[17] In Chapters 3, 4 and 5, η_{ji} includes also errors introduced by first-order Taylor's approximation.

variance $Var(\log w_0 \mid \log w_0 \geq \log w_1)$, that is, the wage inequality among those who choose to stay in rural regions. Given that in the standard Roy model the economic linkages among individual agents are almost entirely absent, the usage of truncated variance seems justified. This is because in absence of inter-personal linkages, out-migration of some agents leaves no impact on wages of staying agents, and wages earned by out-migrants do not count in measuring the rural inequality. Nevertheless, for numerous other situations, economic linkages among individual agents at the source, especially those among migratory and staying members from the same households could be so strong that the neglect of these linkages would lead to misleading predictions.

The economic linkages between migratory and staying members of rural households can be sorted roughly into two categories. For the first category, the withdrawal of labor services provided by migratory members will affect staying members' agricultural productivities and hence wage incomes by changing the amounts and composition of inputs of farms. For the second category, there exists usually income pooling to different degrees between migratory and staying members. In fact, the remittance emphasized in literature can be viewed as a special case of the income pooling mechanisms. A simple way to incorporate both categories of intra-household economic linkages into analysis would be to assume that rural households behave as if they were unities attempting to maximize given objectives such as total incomes adjusting for the perfectness of intra-household income pooling.[18]

As for the context focused by this thesis, namely rural-to-urban migration and rural regions in China, a plenty of evidences suggest that these intra-household

[18] The perfectness of intra-household income pooling can be interpreted alternatively as the degree of altruism of migratory members towards staying members. Usually, the degree of altruism has its value between 0 and 1.

economic linkages could be of particular importance. Firstly, like in many developing countries, household farms undertake the majority of agricultural production in post-reform China. Since factor markets in rural regions are often incomplete, households have limited access to these markets and thus have to rely heavily on the largely predetermined cultivated land and labor services provided by their members in agricultural production. Consequently, out-migrations of some household members caused by the emergence of urban employment are likely to have big effects on the amounts and composition of agricultural inputs of household farms and hence on staying members' wage incomes. Secondly, a large fraction of Chinese rural migrants are temporal migrants in the sense that they will eventually return to rural regions. Besides, household members such as elders and children are often left behind during their urban employments. Therefore, migratory members may have strong incentives to share incomes with staying members. Hence, the income pooling among household members could be quantitatively important. To avoid introducing the arbitrariness by allowing the perfectness of income pooling to vary, unless stated otherwise, this thesis assumes perfect income pooling for all rural households.

For the rest of theoretical explorations in Chapters 3, 4 and 5, I make efforts to build a series of simple yet highly stylized models to capture income-maximizing rural households' agricultural production, labor supply and migration decisions made subjective to various factor market environments and in different settings regarding the heterogeneity and substitution among labor inputs. Then I combine these micro-economic models with the largely statistical framework of the standard Roy model, especially with its powerful assumption of joint normality of log sectoral wages in the hope of achieving a thorough understanding on the microeconomic contents of the Roy model. More importantly, by doing so, intra-household economic linkages that are missing in the standard Roy model can be incorporated into the extended Roy models. In comparison with the standard Roy model, these extended Roy model have more solid microfoundations and thus they can offer appropriate frameworks for formal analysis of the research question.

To close the model, I introduce additional assumptions on the urban sector. Note that to keep the discussions in subsequent chapters manageable, simple assumptions are preferred. In particular, this thesis always adopts a succinct assumption on the urban sector, namely equally skilled rural workers face the exogenously given wage rate in that sector. This thesis abstracts the migration costs entirely from theoretical explorations, partly because until recently, the economic profession knows very little about the size, functional form and distribution of them.[19] This is also because this thesis intends to examine how far one can go without resorting to the puzzling migration costs.[20] Moreover, for the reason that the Hukou system[21] can be largely understood as an institution that imposes restrictions on rural-to-urban migration by raising migration costs faced by rural migrants, and that migration costs are ignored in theoretical explorations, the Hukou system will not be discussed explicitly below.

---

[19] The intense exchange between Borjas (1987) and Chiswick (1999) reminds us the importance of the functional form of migration costs in determining the pattern of migration selection and other related issues. It also shows the ignorance of the economic profession about migration costs. Borjas (2014) offers a sharp yet highly interesting criticism on the reverse engineering practices in choosing functional form of migration costs.

[20] Katz and Murphy (1992) use a similar strategy to study the change of wage structure in the United States.

[21] The Hukou system can be translated literally as the "household registration system". For details on this system, interested readers may see Cheng and Selden (1994) or a recent survey by Chan (2009).

CHAPTER 3

THE IMPACTS OF OUT-MIGRATION: THE MODEL WITH HOMOGENEOUS LABOR

In this chapter, I begin the theoretical exploration of the impacts of out-migration on income distribution in rural regions using an extended Roy model as analytical framework. To keep the exploration tractable, throughout this chapter I assume all rural workers are intrinsically homogeneous in productivity. Further discussions allowing for various kinds of labor heterogeneity, which is an issue that complicates the exploration will be left as the main task for next two chapters.

3.1 The Unconstrained Maximization Problem

As suggested in Chapter 2, I consider firstly a simple maximization problem faced by each rural household. In this model, households try to maximize total incomes by allocating labor endowments L^0 between the agricultural production on their own farms and emerging urban non-agricultural employments. Thus, the maximization problem of any rural household is given by

$$\max_{L} f(A^0, L) + w_1(L^0 - L), \qquad \text{(problem 3.1)}$$

where the production function $f(.)$ represents the technology adopted by rural household that combines two types of inputs, namely land (A) and homogeneous labor (L) into the final agricultural output whose price is normalized at unity. $f(.)$ is assumed to be twice continuously differentiable, increasing, concave and linearly homogeneous. To reflect the observation that a large fraction of rural households do not abandon agricultural production entirely after the emergence of opportunities for rural residents to be employed in urban sector, or in short, to reflect the feature of partial out-migration, $f(.)$ is further assumed to satisfy one of the Inada conditions of the form $\lim_{L \to 0+} \partial f(A^0, L) / \partial L \to +\infty$.[22]

Moreover, the land cultivated by any rural household remains unchanged during out-migration, that is, A=A^0.[23] At any time point, total labor services offered by household members are largely predetermined. The optimal agricultural labor input L may or may not be smaller than L^0, depending on the concrete specification of rural labor markets.[24] This section focuses on problem 3.1 in which rural household's income-maximization problem is not constrained by its labor endowment.

The model further that the wage rates faced by all rural workers when they were employed in the urban non-agricultural sector are given exogenously by a constant w_1. It follows from the succinct assumption introduced in Chapter 2 that equally skilled rural workers face the same wage rate in urban sector, along with the presumption that throughout this chapter all rural workers are homogeneous and hence equally skilled.

---

[22] Indirect evidences on the prevalence of the partial migration in the Chinese context are provided in Appendix B.

[23] This assumption is relevant for rural China nowadays. Several empirical studies point out that the land market in rural China is still underdeveloped. According to Jin and Deininger (2004), in the period of 1995-2000, only 5.4 percent of rural households in their sample obtained additional land through official land redistribution. Around 10 percent of rural households participated in land rental market. Ye et al. (2005) shows that in their sample collected in 17 provinces in 2005, 81.8 percent of households did not obtain any land from other households.

[24] To my knowledge, there is no survey on the development of the Chinese rural labor market that is comparable to that given by Jin and Deininger (2004). Instead, in Chinese literature, the development of rural labor market is often evaluated indirectly by percentages of out-migration and non-farm activities among rural residents, rate of return to schooling, etc. Therefore, it would be impossible to decide precisely the relevance of the unconstrained and constrained maximization problems. As a response, I discuss in this chapter both problems, though admittedly, emphasis is put on the constrained maximization problem that could be of greater relevance when rural labor market is less developed in the sense that rural households have to rely heavily on labor services of their own members. In subsequent chapters, however, I will consider the constrained maximization problems only.

Solving problem 3.1 gives the first-order condition as follows

$$\partial f(A^0, L)/\partial L = w_1, \tag{3.1}$$

which suggests that in the unconstrained optimum, the marginal product of labor in agriculture evaluating at (A^0, L^*) equals to the given urban wage rate w_1.

Thanks to the linearly homogeneity of agricultural technology, the first-order condition given by equation (3.1) can be written in a more compact form such that

$$MPL(a) \equiv \partial f(a,1)/\partial L = w_1, \ with\ a \equiv A^0 / L. \tag{3.2}$$

The existence and uniqueness of a finite solution a^* to equation (3.2) are guaranteed by properties of the marginal-product-of-labor function $MPL(a)$, namely $MPL(a)$ is increasing in a and $\lim_{a \to +\infty} MPL(a) > w_1$, which is a consequence of the Inada conditions. The unconstrained optimal land-labor ratio is given by $a^* \equiv A^0 / L^*$.

At rural community level, suppose further that all households have the identical access to the production technology, the $\log MPL(a)$-loci and thus their intersections with $\log w_1$-line must be the same. Moreover, it is common that initial land-labor ratios vary among rural households and thus they can be represented by a random variable with non-degenerate distribution. The first-order condition given by equation (3.2) describes the optimal plans of out-migration for rural households: all rural households, regardless of their initial land-labor ratios will adjust agricultural labor inputs until the marginal products of labor in agriculture equal to the given urban wage. However, in achieving the optimum, households whose initial land-labor ratios a^0 are smaller than a^* rent part of their members' labor services to urban employers or to other rural households at the given urban wage rate. By contrast, households whose initial land-labor ratios a^0 are greater than a^* hire labor services from other rural households at the given urban wage rate.

Consequently, as Figure 3.1 shows, all rural households will end up with the same land-labor ratio a^* and log marginal products of labor $\log MPL(a^*) = \log w_1$.

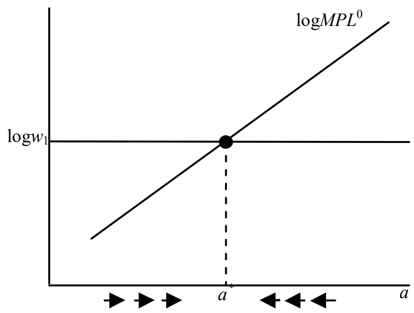


Figure 3.1: The Unconstrained Maximum

In the present setting, the impacts that out-migration has on rural inequality can be easily determined: As long as the initial distribution of land-labor ratios among rural households is non-degenerate, the pre-migratory rural inequality exceeds the post-migratory one. In other words, if problem 3.1 is the right problem to be maximized, then out-migration tends to reduce the rural inequality to zero. Note that so far the analysis does not resort to the joint normality assumption maintained by the standard Roy model.

3.2 The Constrained Maximization Problem

In developing countries like China, rural labor markets are often less complete than that presumed in Section 3.1. In particular, rural households have often limited accesses to these markets so that agricultural production undertaken by them has to rely on services provided by their members. For a polar case considered hereafter, rural households use exclusively services of their members. Hence, the predetermined L^0 serve as upper bounds for households' agricultural labor inputs and households' pursuits of maximum income are constrained by labor endowments. Notice that such assumption is not unique in my thesis. It is shared in fact by many researches in a large literature originating from Chayanov (1966) and known now as the Agricultural Household Models.

Therefore, rural households' income-maximization problem subjective to the constraint of labor endowment $L \le L^0$ can be formulized as follows

$$\begin{aligned} &\max_L f(A^0, L) + w_1(L^0 - L), \\ &s.t.\ L \le L^0. \end{aligned} \qquad \text{(Problem 3.2)}$$

Henceforth, problem 3.2 is called as the household's constrained maximization problem. Note also that for this problem, the optimal agricultural labor input should be strictly positive, which is guaranteed by one of the Inada conditions. In this regard, all out-migrations considered in the present setting are partial out-migrations.

Solving problem 3.2 gives a set of the first-order conditions such that

$$\begin{aligned} &\text{(i)}\quad \partial f(A^0, L)/\partial L - w_1 \ge 0; \\ &\text{(ii)}\ \ [\partial f(A^0, L)/\partial L - w_1](L^0 - L) = 0; \\ &\text{(iii)}\ L \le L^0. \end{aligned} \qquad (3.3)$$

Solution to the first-order conditions is denoted by $L^{\#}$. Note that throughout this thesis, the superscript $^{\#}$ is reserved for the constrained maxima, while superscripts ^{0} and * are reserved for the initial values and unconstrained maxima.

After rearrangements, the first-order conditions can be rewritten as

$$\log MPL^{\#} = \begin{cases} \log MPL^{0}, & if \ \log MPL^{0} > \log w_1; \\ \log w_1, & if \ \log MPL^{0} < \log w_1, \end{cases} \tag{3.4}$$

or equivalently,

$$\log MPL^{\#} = \max\{\log MPL^{0}, \log w_1\}\,. \tag{3.5}$$

Interestingly, equation (3.5) obtaining by solving rural households' constrained maximization problem takes a similar form to equation (2.2), which is an a priori assumption of the standard Roy model that has a strong individualistic flavor. The close resemblance of both equations suggests the possibility of combining the micro-economic constrained maximization model with the statistical assumptions of the standard Roy model. It suggests further possibility of employing the extended Roy model as framework to quantify the impact of out-migration on rural inequality.

In literature, the impact is usually captured by the difference of the pre- and post-migratory rural inequalities. More specifically, the pre-migratory rural inequality can be measured by $Var(\log MPL^{0})$, while the post-migratory rural inequality can be measured by $Var(\log MPL^{\#})$ instead of the truncated variance used in the standard Roy model, i.e. $Var(\log MPL^{0} \mid \log MPL^{0} > \log w_1)$. $Var(\log MPL^{\#})$ is likely to be the right choice because under the Inada conditions all rural households are still engaged at least partially in agricultural production after migration. As a result, the agricultural wages, i.e. shorthand for the marginal products of labor in agriculture earned by partially migratory households continue to contribute to the distribution of agricultural wages among rural households and hence to the post-migratory overall inequality in rural communities. As suggested by the law of total variance of the form

$Var(\log MPL^{\#}) = E_I[Var(\log MPL^{\#} \mid I)] + Var_I[E(\log MPL^{\#} \mid I)]$, where the event of out-migration is captured by $I \equiv 1\{\log MPL^0 < \log w_1\}$, $Var(\log MPL^{\#})$ takes fully account of the influences of agricultural wages earned by migratory households on the total inequality and thus should be preferred.

Before quantifying the impact of out-migration on rural inequality by introducing the joint normality assumption of the standard Roy model, I would like to highlight the differences between the model developed in this section and the standard Roy model. At a higher level of abstraction, both models differ in one fundamental way, namely they associate with two interrelated but different statistical problems. Using the terminologies of statistics and econometrics, out-migrations in the first model result in **censoring**, while out-migrations in the second model result in **truncation** on the underlying distribution of log agricultural wages among households.[25] To be concrete, in the first model, owing to the partial feature of out-migration, the post-migratory distribution of agricultural wages can be obtained by combining (i) a continuous distribution above the level of $\log w_1$ with unchanged probability density function and (ii) a discrete distribution at the level of $\log w_1$ with a usually positive probability $\Pr\{\log MPL^0 < \log w_1\}$. By contrast, in the standard Roy model, since wages earned by migratory agents do not affect the wage distribution in the source sector, the post-migratory wage distribution can be obtained by adjusting underlying wage distribution relating to staying agents upwards. The censored and truncated variances $Var(\log MPL^{\#})$, $Var(\log MPL^0 \mid \log MPL^0 > \log w_1)$ measure the levels of inequality of two distributions mentioned above.

[25] For details on concepts of censoring and truncation, moments of the censored and truncated normal distribution, reader may refer to Greene (2007), Chapter 24.

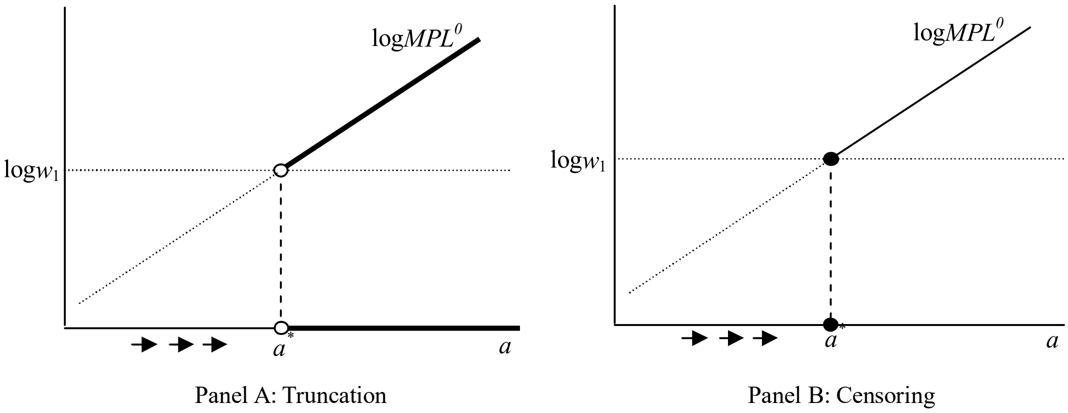


Figure 3.2: Truncation versus Censoring

Note that the thick lines in Panel A suggest that the probability density functions associating with staying households have been scaled up so that they integrate to one. The solid dots in Panel B suggest that the probabilities associating with them are often strictly positive and equal to the population share of migratory households in the rural community of interest.

By asserting that $\log MPL^0$ among different rural households is independently, identically and normally distributed such as

$$\log MPL^0 \overset{i.i.d.}{\sim} N(\mu_0, \sigma_0^2), \tag{3.6}$$

the microeconomic model developed in this section can be combined with the analytically convenient distributional assumption maintained by the standard Roy model. The rest of this section uses the resulted extended Roy model to address questions concerning the impacts of out-migration on income distribution in rural regions both qualitatively and quantitatively.

Together with other assumptions including the first-order condition given by equation (3.5) and $\log w_1 = const$, the changes of mean log agricultural wages ($\Delta E^{\#}$) and of rural inequality ($\Delta V^{\#}$) can be formulized as follows:[26]

[26] Equations (3.7) and (3.8) follow directly from Theorem 24.3 in Greene (2007).

$$\begin{aligned}\Delta E^{\#} &\equiv E(\log MPL^{\#}) - E(\log MPL^{0}) \\ &= \sigma_0[c\Phi(c) + \phi(c)] \\ &\equiv \sigma_0 J(c);\end{aligned} \tag{3.7}$$

$$\begin{aligned}\Delta V^{\#} &\equiv Var(\log MPL^{\#}) - Var(\log MPL^{0}) \\ &= \sigma_0^2\{(1-\Phi(c))[(1-\delta(c)) + \Phi(c)(c-\lambda(c))^2] - 1\} \\ &\equiv \sigma_0^2 K(c),\end{aligned} \tag{3.8}$$

where c denotes the standardized wage gap defined as $c \equiv (\log w_1 - \mu_0)/\sigma_0$; $\Phi(.)$ and $\phi(.)$ denote the cumulative and probability density functions of a standard normal distribution. $\lambda(.)$ is the Inverse Mills ratio defined as $\lambda(.) \equiv \phi(.)/(1-\Phi(.))$ and $\delta(.) \equiv \lambda'(.)$.

The signs of $J(c)$ in equation (3.7) and $K(c)$ in equation (3.8) are difficult to be determined analytically. Instead, they are determined by using numerical evidences summarized in graphical form. Panel A of Figure 3.3 shows that for a wide range of c, $J(c)$ and $\Delta E^{\#} = \sigma_0 J(c) \geq 0$, which means out-migration tends to increase the mean of log agricultural wages earned by households in rural regions. More importantly, Panel B shows that for a wide range of c, $K(c)$ and $\Delta V^{\#} = \sigma_0^2 K(c) \leq 0$, which means out-migration tends to decrease rural inequality. Therefore, qualitatively speaking, the extended Roy model gives the same prediction on the impacts that out-migration has on rural inequality as that given by the standard Roy model.

Nevertheless, the difference in statistical problems of two models does lead to quantitatively different predictions on the impacts of out-migration on rural income distribution. To see this, I compare the means and variances of the truncated and censored normal distributions. Holding other things constant, the differences between the truncated and censored moments are $DE \equiv E(\log MPL^{\#}) - E(\log MPL^{+})$ and

$DV \equiv Var(\log MPL^{\#}) - Var(\log MPL^{+})$, where $\log MPL^{+}$ denotes the truncated log agricultural wage such that $\log MPL^{+} = \log MPL^{0}$, *if* $\log MPL^{0} \geq \log w_1$. Henceforth, $E(\log MPL^{+})$ and $Var(\log MPL^{+})$ denote truncated mean and variance.

After some rearrangements, we have

$$DE = \Phi(\log w_1) - \Phi(E(\log MPL^{+})) . \tag{3.9}$$

Since $\Phi(.)$ is increasing and $\log w_1 \leq E(\log MPL^{+})$, otherwise members of non-migratory households would migrate, we know $DE \leq 0$. This result suggests that in comparison with the standard Roy model, the extended Roy model predicts a smaller increase in mean log agricultural wages resulting from out-migration.

Likewise, we have also[27]

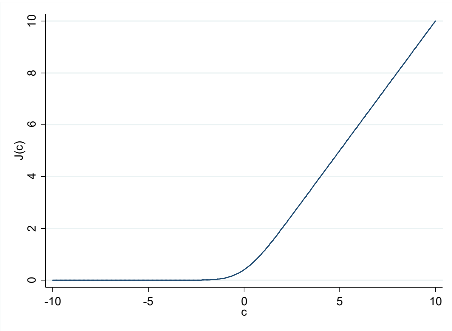
$$\begin{aligned} DV &= \sigma_0^2 \Phi(c)\{[1-\Phi(c)][c-\lambda(c)]^2 - [1-\delta(c)]\} \\ &\equiv \sigma_0^2 \Phi(c) L(c). \end{aligned} \tag{3.10}$$

Since both σ_0^2 and $\Phi(c)$ are positive, the sign of DV depends on the sign of $L(c)$. According to Panel C of Figure 3.3, the sign of $L(c)$ could be either positive or negative. Particularly, $L(c)$ oscillates violently around $L(c) = 0$ for large and positive c. Nevertheless, for the majority cases of interest, say $c < 4$,[28] it is safe to conclude that $L(c) > 0$ and hence $DV > 0$. This result suggests that holding other things equal, the extended Roy model predicts a smaller decrease in rural inequality than the standard Roy model does. Panel D of Figure 3.3 provides further supportive

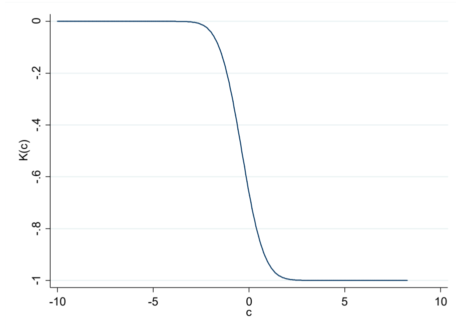
[27] Equation (3.10) follows directly from Theorems 24.2 and 24.3 in Greene (2007).

[28] Note that for the standard normally distributed random variable $z \sim N(0,1)$, we have Pr(z>4)=0.0000, which means {z>4} is an event that almost never happens.

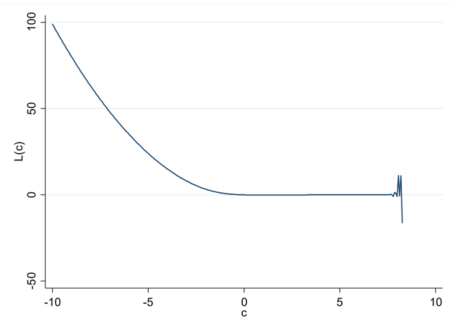
numerical evidences on the magnitudes of truncated and censored variances relative to the variance of underlying distribution.



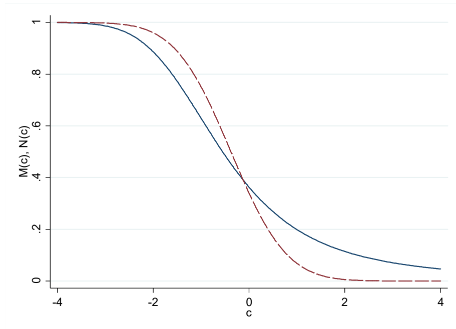
Panel A: $J(c)$



Panel B: $K(c)$



Panel C: $L(c)$



Panel D: $M(c)$ and $N(c)$

Figure 3.3: Numerical Evidences on the Truncated and Censored Moments

Note that in Panel D $M(c)$ is defined as $M(c) \equiv Var(\log MPL^{+}) / \sigma_0^2$ and is illustrated by dashed curve; $N(c)$ is defined as $N(c) \equiv Var(\log MPL^{\#}) / \sigma_0^2$ and is illustrated by solid curve. Definitions for $J(c)$, $K(c)$ and $L(c)$ can be found in equations (3.7), (3.8) and (3.10). Software: Stata 12.0.

Until now, no attempt has been made to identify the source of rural inequality. In what follows, I return to this issue briefly. Following the suggestion in Section 2.2, I expand $\log MPL^0$ around some $\overline{a}$ as the following first-order Taylor's series

$$\log MPL(a^0) \approx \log MPL(\overline{a}) + \frac{d \log MPL(\overline{a})}{d \log a^0} \log(\frac{a^0}{\overline{a}}). \qquad (3.11)$$

If $\bar{a}$ is chosen so that $\mu_0 = \log MPL(\bar{a})$, then equation (3.11) can be rewritten in a form closely related to the wage decomposition formulas in Chapter 2 such that

$$\log MPL_h^0 = \mu_0 + \varepsilon_{0h}, \tag{3.12}$$

where the demeaned error term ε_{0h} can be approximated by

$$\varepsilon_{0h} \approx \xi_{0h} = \frac{d \log MPL(\bar{a})}{d \log a_h^0} \log(\frac{a_h^0}{\bar{a}}) \equiv k \log(\frac{a_h^0}{\bar{a}}). \tag{3.13}$$

This decomposition clarifies to a certain extent the source of pre-migratory rural inequality. According to equation (3.13), the rural wage inequality can be attributed to the inequality of land distribution among rural households. Such understanding is still incomplete, but it would be superior to the conventional practice to attribute all wage inequality indifferently to the heterogeneity of vaguely defined skills. Nevertheless, it must be admitted that the source of rural inequality plays only a minor if not no role in determining the impacts of out-migration on rural income distribution, as long as the distribution of underlying sectoral wages are assumed to be jointly normal.

Careful readers may have noticed that up to this point, the analysis concerns primarily with how out-migration affects wage inequality among different households. However, the existing literature emphasizes more on the wage inequality among different individual workers. Thus, I will adapt previous discussions for the latter question. As will be clear soon, such extension is possible mainly because household's income-maximization has already implied wage-maximization of its members, as long as the externalities of some household members' out-migration on others members' agricultural wages are taken into account.

One way to justify such extension can be illustrated as follows: We may think the whole process of labor reallocation made by a rural household whose $a^0 < a^*$ as a series of small adjustments.[29] At the very beginning of this process, the revenue from reallocating the marginal unit of labor input away from agriculture is given by the urban wage w_1, while the cost equals to the loss caused by withdrawal of this unit of labor input from agricultural production, i.e. $MPL(a^0)$. Since $w_1 > MPL(a^0)$, it is in the best interest of both income-maximizing household and wage-maximizing worker as labor supplier to reallocate this unit of labor input away from agriculture. In fact, once the land-labor ratio faced by this household and all its members is given, external observers cannot separate two maximization problems pursued by different agents from one another. Similarly, for the rest of household agricultural labor inputs, income-maximizing household will decide in sequence whether they should be reallocated by comparing the revenues with costs evaluating at the constantly updating ratio between land and labor services of staying members. The reallocation will stop only when the cost of small adjustment catches up with its revenue.

Note particularly that for already reallocated labor inputs, they receive incomes exclusively in the form of urban non-agricultural wages. Hence, it makes no sense to consider the agricultural wages earned by them, though both wages equalize at the margin. To avoid such ambiguity, a new notation $\tilde{w}$ is introduced to denote the post-migratory wage rates received by all labor inputs offered by members of rural households. According to the discussions above, for each unit of the labor inputs (ℓ), the wage rate it receives takes the form of $\log \tilde{w}_\ell = \max\{\log MPL_\ell^0, \log w_1\}$, $\forall \ell$.

[29] The dynamic adjustments are not compatible to the static framework in this chapter in strict sense. Nevertheless, these adjustments provide justification for the solution to the static optimization problem.

If an individual worker is viewed as supplier of a bundle of units of labor inputs, then the equation above can be also applied approximately to capture the log post-migratory wages earned by individual workers such that

$$\log \widetilde{w}_i = \max\{\log MPL_i^0, \log w_1\}, \ \forall i. \tag{3.14}$$

This equation takes a similar form as that of equation (3.5). Moreover, as explained above, both equations are also related through their economic meanings. Therefore, equation (3.14) can be seen as a straightforward extension of equation (3.5) at the level of individual workers.

Suppose that migratory and staying workers belonging to the same households relate closely through income pooling, we need to consider not only the agricultural wages earned by staying workers, but also the urban wages earned by migratory workers in determining the level and change of wage inequality among individual workers. Consequently, setting aside the complex issues such as redistribution within household, the post-migratory distribution of wages of individual workers can be obtained by censoring the underlying distribution at the urban wage rate.

Suppose further that the distribution of $\log MPL^0$ among individual workers is also normal, together with the assumption $\log w_1 = const$ and equation (3.14), the post-migratory individual wage inequality must be smaller than its original level, that is, $Var(\log \widetilde{w}) \leq Var(\log MPL^0)$. In other words, out-migration tends to decrease the wage inequality among these individual workers as well.

This result turns out to be robust. In fact, it holds even when the distribution of $\log MPL^0$ is not normal, while other specifications, especially $\log w_1 = const$ keep unchanged.[30,31] The intuition behind is straightforward: According to the law of total variance, the overall wage inequality can be decomposed into weighted within-group wage inequalities and between-group wage inequality. We can determine easily the directions of impacts of out-migration on each of these components of decomposition. Previous discussions suggest that (1) out-migration tends to decrease the wage inequality within the group of migratory agents, since all migratory households and workers end up with earning the same wage rate w_1; (2) out-migration has no impact on the wage inequality within the group of staying households. Moreover, if agricultural wages earned by workers are evaluating using the constantly updating land-labor ratios, then out-migration has no impact on the wage inequality within the group of staying workers, either; (3) meanwhile, out-migration tends to narrow the average wage gap between migratory and staying households as well as that between both groups of workers. Therefore, the overall impact of out-migration is to decrease the rural inequality. The reasoning above requires no normality of $\log MPL^0$.

[30] This is generally not the case for truncation. Heckman and Honoré (1990) show that to ensure that truncation results in lower variance, the original distribution has to be log-concave. Normal distribution is one of the log-concave distributions.

[31] Similar reasoning can be used to deal with the impacts of out-migration on the rural inequality defined over alternative incomes such as the average products of labor in agriculture or per capita household income. It turns out that for rural inequalities defined using all income concepts above, out-migration tends to reduce the rural inequality. This alternative approach has the virtual that it does not rely on strong distributional assumption. Nevertheless, it offers only qualitative predictions on the change of rural inequality in the course of out-migration. Moreover, this approach is not readily to be extended to consider problems with heterogeneous labor.

3.3 Necessity of Allowing for the Labor Heterogeneity

Although substantial efforts have been made in this chapter to extend the standard Roy model and then to use the extended Roy model to study the impacts that out-migration has on inequality, two inconsistencies pointed out in Chapter 2 cannot be satisfactorily reconciled. According to equation (5.38), just like the standard Roy model, the extended Roy model also predicts that out-migration decreases rural inequality and it leaves still no room for changes of rural inequality in other directions. The only notable difference between two models lies in that the extended model predicts a smaller decrease in rural inequality resulted from out-migration than the standard model does (see equation 3.10 and Figure 3.3, Panel C).

Discussions in this chapter have shown also that the prediction that out-migration always decrease the rural inequality is very robust. It applies to rural inequalities defined among rural households and among workers. Moreover, it applies to cases where $\log MPL^0$ is normally distributed and where it is not. The robustness seems so impressive that some doubts need to be cast on it. One may suspect whether the robustness is derived mostly from some overly strong assumptions that has been imposed, for instance, the succinct assumption $\log w_1 = const$. As suggested by the analysis without assuming the normality, by adopting that succinct assumption the wage distribution of migratory agents is squeezed into a constant, which could explain to a certain extent to the decrease of overall rural inequality. As a response, that assumption will be relaxed in subsequent discussions and the distribution of $\log w_1$ will be non-degenerate. Given that throughout the theoretical exploration, I stick to the assumption introduced in Chapter 2, namely equally skilled rural workers face with the same urban wage rate, relaxation of that assumption would require some kind of heterogeneity of skills among individual workers and their households. This serves as a major motivation for introducing the labor heterogeneity into theoretical analyses in next two chapters.

There is also a minor motivation for introducing of the labor heterogeneity. As readers may have seen, in previous discussions on the level and change of rural inequality among individual workers, these workers' agricultural wages relate not only to the agricultural technology, urban wage and their households' initial land-labor ratios a^0, but also to their sequence of migration within households. However, since in the present setting, all rural workers are by assumption homogeneous and little is known about whether all household members have the same access to the land, there exists no sound explanation for such sequence. Allowing the rural workers to be heterogeneous and hence to have different relative advantages in both sectors could help understand that sequence.

CHAPTER 4

THE IMPACTS OF OUT-MIGRATION: THE MODEL WITH HETEROGENEOUS AND PERFECTLY SUBSTITUTIVE LABOR

In the next two chapters, I continue the theoretical exploration of the impacts that out-migration has on income distribution in rural regions. A significant departure from Chapter 3 is that rural workers are now assumed to be heterogeneous in earning capacity. As will be shown later on, introducing of the labor heterogeneity complicates the exploration greatly so that it is not always possible to obtain unambiguous predictions. Nevertheless, the theoretical exploration below deepens our understanding on the impacts of out-migration on rural inequality. Furthermore, it points out the limitation of the Roy models in general, especially when they are applied to the settings with imperfect labor substitution.

4.1 Model Setup

To begin with, I list and discuss briefly the assumptions essential for the following two chapters. These assumptions can be sorted into three groups.

4.1.1 The Agricultural Technology

The first challenge is to find a proper way of modeling the labor heterogeneity among different types of rural workers and labor services. This decision could affect the subsequent exploration.

This thesis employs the nested CES framework to model the labor heterogeneity. Within that framework, different types of labor inputs, together with non-labor inputs are organized by a production function in the nested CES form according to the ease of substitution. This choice relates to two literatures. The first literature adopts the multi-level nested CES as the empirical specification to study the impacts of immigration on the destination (See Borjas, 2003, Ottaviano and Peri, 2008, 2012)[32]. The second literature emphasizes the role of imperfect substitution in determining the changes of wage structure and inequality (See Katz and Murphy, 1992; Katz and Autor, 1999; Acemoglu, 2002).

Following both literatures, the model assumes that all households in the rural community of interest have accesses to the same two-level nested CES production technology given below: At the higher level of the nested CES function, land (A) and effective labor (QL) inputs are combined by a Cobb-Douglas function to produce the final agricultural output (y) such that

$$y = TA^{\alpha}QL^{1-\alpha}, \tag{4.1}$$

where T denotes the hicksian technical factor. At the lower level of this nested CES function, the effective labor input is in turn the outcome of a CES function combining different types of agricultural labor inputs.[33] Mainly for the analytical convenience it brings with, this model assumes further continuum types of labor inputs,[34] each of

[32] The methodological innovation is primarily made in Card and Lemieux (2001).

[33] The Cobb-Douglas functional form is often criticized for it imposes overly strong restrictions on the technology. In particular, under the Cobb-Douglas, the elasticity of substitution among inputs must be unity. Nevertheless, in the literature that uses the nested CES specification as its empirical framework, the Cobb-Douglas form is the leading specification for the highest level of the nesting structure. Moreover, throughout Chapters 4 and 5, I assume that at the lower level of the nesting structure, all types of labor inputs enter the CES function symmetrically – There is no one labor type that can be separated from other labor types. The symmetric CES specification is also restrictive, since it requires the elasticities of substitution are the same for any pair of labor types. This shortcoming does not affect much the popularity of such a specification in theoretical and empirical studies, partly because it offers a simple and manageable way to incorporate both the labor heterogeneity and imperfect substitution among labor inputs into the analysis.

[34] It may seem at odd to assume continuum of labor types in the context of household farm, since (1) the average size of workforce in household farms is often small and (2) skills of different household members are often highly correlated. A rationale might be used in justifying this assumption is that skill level of workers relate not only to the largely fixed characteristics, such as education, age, gender etc.,

which is indexed with a one-dimensional skill level in the range of $s \in [0,1]$. Hence, the effective labor input takes the form of

$$QL = [\int_0^1 q(s) L(s)^\rho ds]^{1/\rho}, \tag{4.2}$$

where $L(s)$ denotes the amount of agricultural labor input with skill s, $q(s)$ can be loosely interpreted as the quality of $L(s)$.[35] The parameter ρ relates closely to the elasticity of substitution among all types of labor inputs denoting by $\sigma_E \equiv 1/(1-\rho)$.

To ensure rural households' optimization problem is well defined, it is often assumed that $\rho \leq 1$.[36] For the purpose of this thesis, however, we need to assume further that $\rho \geq 0$. In literature, $\rho \geq 0$ corresponds to the probably more relevant cases in which the skill biased technical change raises the skill premium.

Note that among all admissible values of ρ, a special case $\rho = 1$ deserves a separate treatment first, because it relates to an analytically simple and empirically relevant setting in which in spite of the differences in earning capacity, all types of labor inputs offered by rural workers are perfect substitutes in household agricultural production.[37] In the present chapter, I focus on this specific setting and explore the

---

but they may be also contingent on working conditions, such as season, weather and daily working hours. If these working conditions vary over a wide range continuously, the actual skill level may vary accordingly.

[35] For a detailed discussion on the economic interpretation of parameters $\{q(s)\}$, reader may see Appendix C.

[36] More explicitly, to ensure the income-maximizing problem is well defined, the production technology must be locally concave in the neighborhood of the optimal choice. The condition $\rho \leq 1$ guarantees both local and global concavity of the CES technology. However, it becomes less clear whether the same condition applies to the more complicated nested CES technology.

[37] Admittedly, it seems difficult to find concrete examples in which different inputs (or consumption goods) are perfect substitutes in producing the final output (or utility). However, several empirical evidences do suggest the relevance of this setting. For example, in an empirical study using the firm-level microdata, Hellerstein et al. (1999) finds no evidence suggesting that workers with different characteristics are imperfect substitutes. In Chapter 6, I cannot reject the original hypothesis that rural

impacts of out-migration on rural inequality. Some preliminary discussions on more general setting with the parameter $0 \le \rho < 1$, in which agricultural labor inputs are heterogeneous and imperfectly substitutive in household agricultural production will be delayed until the next chapter.

4.1.2 The Urban Sector

The model adopts again the succinct specification for the urban non-agricultural sector, namely rural workers are price-takers in that sector: All rural workers endowed with skill level s face the exogenously given urban wage rate $w(s)$.

4.1.3 The Factor Markets in Rural Regions

Similar to Chapter 3, this chapter assumes the land cultivated by each rural household keeps unchanged during migration, i.e. $A = A^0$. At any point of time, household's labor endowments $\{L^0(s)\}$ are largely predetermined: They depend on both the quantity and quality of labor services offered by their members. Suppose that the rural labor market works perfectly, individual households' decisions are unaffected by their labor endowments. By contrast, suppose the rural labor market works less perfectly, which is often the case for developing countries like China, households' decisions are likely to be constrained by their labor endowments. For all discussions below, I will focus on a setting with imperfect rural labor markets and thus households' constrained maximization problem. Such choice can be partly justified by its higher relevance to the developing countries. Equally important, as demonstrated in Chapter 3, this choice can be further justified by the larger potential that the constrained

---

workers belonging to different education-age cells are perfect substitutes in household agricultural production at 5% significance level. These empirical findings might suggest that the imperfect substitution may be only of importance outside individual economic agents.

maximization problem has to combine with the joint normality assumption of the standard Roy model.

4.2 The Constrained Maximization Problem

Under all assumptions listed above, the constrained maximization problem faced by any rural household takes the form of

$$\begin{aligned}&\max_{\{L(s)\}} T(A^0)^{\alpha}[\int_0^1 q(s)L(s)ds]^{1-\alpha} + \int_0^1 w_1(s)[L^0(s) - L(s)]ds \\ &s.t. \quad L(s) \le L^0(s), \forall s \in [0,1] \\ &\text{where } A^0, \{L^0(s)\} \text{ are predetermined;} \\ &\text{and } \{w_1(s)\} \text{ are exogeneously given.}\end{aligned} \qquad \text{(Problem 4.1)}$$

Note that the nested CES specification used in problem 4.1 requires the effective agricultural labor input in the constrained optimum, i.e. $QL^{\#} \equiv \int_0^1 q(s)L^{\#}(s)ds$ to be strictly positive. However, since all types of labor inputs are by assumption perfect substitutes, some of them may take zero values in the constrained optimum, which means that labor services at these skill levels can be completely reallocated away from household agriculture.[38]

Solving the constrained maximization problem gives a set of the complimentary-slackness conditions as follows:

$$\begin{cases} \log MPL(s,...) < \log w_1(s) \Leftrightarrow L(s) = 0; \\ \log MPL(s,...) = \log w_1(s) \Leftrightarrow L(s) < L^0(s); \\ \log MPL(s,...) \ge \log w_1(s) \Leftrightarrow L(s) = L^0(s). \end{cases} \qquad (4.3)$$

By taking partial derivatives and natural logarithms, we obtain the formula for

[38] Another possibility is that $L^0(s) = 0$ for some s. For the discussion in this chapter, the existence of some zero skill-specific labor inputs cause no particular problem.

the log agricultural wage rate earned by a marginal worker at skill level s in the rural household such that

$$\log MPL(s,...) = \log T(s) + \alpha \log a, \tag{4.4}$$

where the technical term $T(s) \equiv (1-\alpha)q(s)T$ and household's effective land-labor ratio a is defined as $a \equiv A^0 / \int_0^1 q(s)L(s)ds$.

Hence, the pre- and post-migratory log agricultural wages can be obtained by evaluating equation (4.4) using the pre-migratory and constrained optimal land-labor ratios, namely $a \equiv a^0$ and $a = a^\#$.

$$\log MPL^0(s,...) = \log T(s) + \alpha \log a^0; \tag{4.5}$$

$$\log MPL^\#(s,...) = \log T(s) + \alpha \log a^\#. \tag{4.6}$$

The difference between them is captured by ς defined as

$$\begin{aligned} \varsigma &\equiv \log MPL^\#(s,...) - \log MPL^0(s,...) \\ &= \alpha(\log a^\# - \log a^0). \end{aligned} \tag{4.7}$$

The additional term ς measures the effects that out-migration has on household farm's effective land-labor ratio and thus on the log agricultural wages earned by staying members. Moreover, ς summarizes all corrections to be made upon the pre-migratory log agricultural wage $\log MPL^0(s,...)$ to get the true reservation wage of out-migration defining as

$$\begin{aligned} \log w^r(s,...) &= \log MPL^0(s,...) + \varsigma \\ &= \log MPL^\#(s,...). \end{aligned} \tag{4.8}$$

Considering that in the constrained optimum, certain types of rural workers and their services can be completely reallocated from agriculture to urban sector, these workers earn no longer agricultural wages. Thus, analogous to Chapter 3, a new notation $\log \widetilde{w}(s,...)$ is introduced to denote the post-migratory log wage rate earned by marginal workers at skill level s in the household of interest, regardless of the sector where they are employed.

With the help of notations $\log \widetilde{w}(s,...)$ and $\log w^r(s,...)$, the complimentary-slackness conditions given by equation (4.3) can be rewritten elegantly as

$$\begin{aligned} &\log \widetilde{w}(s,...) = \max\{\log w^r(s,...), \log w_1(s)\}, \ \forall s, \\ &where \ \log w^r(s,...) = \log MPL^0(s,...) + \varsigma. \end{aligned} \tag{4.9}$$

As readers may have noticed, equation (4.9) resembles to equation (3.5), except that in the presence of labor heterogeneity the log reservation wage $\log w^r(s,...)$ deviates from pre-migratory log agricultural wage $\log MPL^0(s,...)$ by the term ς.

Strictly speaking, all previous discussions deal with the determination of wages earned by the last units of labor services or marginal workers at different skill levels in household. These discussions may not be easily adapted for the wage determination for all units of labor services and rural workers, since even for those with the same skill, their agricultural wages differ. In fact, the agricultural wages earned by a marginal worker would be the lowest among all workers at his skill level, since out-migration of marginal workers tends to raise agricultural wages of non-marginal workers. Nevertheless, since the amounts of labor services and workers at each skill level in rural households are usually small, the agricultural wages earned by marginal workers offer good approximations to agricultural wages earned by non-marginal workers at corresponding skill levels. In this regard, we have the approximate relation for all labor services and rural workers as follows:

$$\begin{aligned}&\log \widetilde{w_i} = \max\{\log w_i^r, \log w_{1i}\},\ \forall i,\\ &\textit{where}\ \log w_i^r = \log MPL_i^0 + \varsigma_i.\end{aligned} \tag{4.10}$$

As will be shown below, ς is of importance for the subsequent exploration, yet its functional form remains unknown. Thus, before proceeding I take a moment to address this issue briefly.

According to its definition given by equation (4.7), the functional form of ς relies on the unknown constrained optimal household's effective land-labor ratio $a^{\#}$, which relies in turn on the values of constrained optimal labor inputs $\{L^{\#}(s)\}$.

As suggested by the complimentary-slackness conditions given by equation (4.3), the values of $\{L^{\#}(s)\}$ are determined by comparisons between $\log MPL^{\#}(s,...)$ and $\log w_1(s)$. To simplify the discussion, it is assumed that the loci of $\log MPL^{\#}(s,...)$ and $\log w_1(s)$ are increasing in *s*, and the $\log w_1(s)$-locus intersects only once at $s = \tilde{s}$ with $\log MPL^{\#}(s,...)$-locus from below.[39]

Under these two assumptions, the continuum of skill $[0,1]$ is divided into three intervals: On the left of $s = \tilde{s}$, since $\log MPL^{\#}(s,...) > \log w_1(s)$, the complimentary-slackness condition predicts that $L^{\#}(s) = L^0(s)$; At the point of $s = \tilde{s}$ where both loci intersect, the condition predicts that $0 < L^{\#}(s) < L^0(s)$; while on the right of $s = \tilde{s}$,

[39] Both assumptions are likely to be fact-based. The first assumption can be restated as the rate of return to skill in both sectors are positive, while the second assumption is inevitable, given the first assumption and the observation that a large fraction of out-migrations are usually partial and selective. These assumptions simplify the discussion mostly because by imposing them, the probability that a type of labor input being partially migratory can be ignored.

since $\log MPL^{\#}(s,...) < \log w_1(s)$, the condition predicts that $L^{\#}(s) = 0$. Therefore, after taking logarithms, the definition of $a^{\#}$ can be rearranged as follows:

$$\begin{aligned} \log a^{\#} &= \log A^0 - \log \int_0^1 q(s) L^{\#}(s) ds \\ &= \log A^0 - \log \int_0^{\tilde{s}} q(s) L^0(s) ds. \end{aligned} \tag{4.11}$$

Besides, the fact that $\log MPL^{\#}(s,...)$ and $\log w_1(s)$ intersect at $s = \tilde{s}$ implies

$$\log T(\tilde{s}) + \alpha \log a^{\#} = \log w_1(\tilde{s}). \tag{4.12}$$

The solution to the system of equations (4.11) and (4.12) is $(\log a^{\#}, \tilde{s})$.

Unfortunately, even in this simplified setting, we can only obtain an implicit solution to $\tilde{s}$ such that

$$G(\tilde{s}) \equiv \alpha \log \int_0^{\tilde{s}} q(s) L^0(s) ds + \log w_1(\tilde{s}) - \log T(\tilde{s}) - \alpha \log A^0 = 0, \tag{4.13}$$

and thus an implicit solution to ς of the form

$$\varsigma = \alpha[\log a^{\#}(\tilde{s}) - \log a^0]. \tag{4.14}$$

4.3 Exploring the Impacts Using the Extended Roy Model

Analogous to Chapter 3, this chapter builds an extended Roy model by combining household's constrained income-maximization problem with two further assumptions, namely (1) all households in the community have same technical accesses, and (2) the log wages faced by rural workers in both sectors, i.e. $(\log MPL^0, \log w_1)$ are jointly normally distributed. Nevertheless, because of the existence of ς, we cannot determine the precise impacts of out-migration on rural inequality using the extended Roy model.

There are two difficulties associating with the unknown term ς: Firstly, since for all rural households, the values of ς are always non-negative, the random variable ς at community level cannot be normally distributed in any strict sense. Therefore, if the distribution of log sectoral wages $(\log MPL^0, \log w_1)$ is jointly normal, the distribution of $(\log w^r, \log w_1)$ is at best approximately jointly normal. Secondly, even if the distribution of $(\log w^r, \log w_1)$ is a reasonable approximation to a jointly normal distribution, it is still impossible to determine even qualitatively the impacts that out-migration has on rural inequality without knowing about the distributional properties of ς, especially the covariance matrix of $(\log MPL^0, \varsigma)$.

To illustrate the second difficulty and facilitate further exploration, I collect the key assumptions maintained by the extended Roy model as follows:

(1) The distribution of log underlying sectoral wages $(\log MPL^0, \log w_1)$ faced by rural workers is jointly normal; The distribution of $(\log w^r, \log w_1)$ can be approximated by another jointly normal distribution such that

$$(\log w^r, \log w_1) \overset{i.i.d.}{\sim} N(\mu_0^{'}, \mu_1, \sigma_0^{'}, \sigma_1, \rho_{01}^{'}). \tag{4.15}$$

(2) The log observed wage of any rural worker is given by equation (4.10)

$$\begin{aligned} &\log \widetilde{w} = \max\{\log w^r, \log w_1\}, \\ &where\ \log w^r = \log MPL^0 + \varsigma. \end{aligned}$$

(3) This thesis concerns with the inequality of wages earned by workers who are

members of rural households. Thus, the pre-migratory inequality should be measured by $Var(\log MPL^0)$, while the post-migratory inequality should be measured by $Var(\log \widetilde{w})$, which is the variance of a distribution obtaining by censoring the distribution of $\log w^r$ at corresponding $\log w_1$.

Having had all three assumptions, as will be shown in the remainder of this section, we may be able to determine the relative magnitudes of the pre-censored and censored variances, i.e. $Var(\log w^r)$ and $Var(\log \widetilde{w})$. Unluckily, since the pre-censored and pre-migratory variances, i.e. $Var(\log w^r)$ and $Var(\log MPL^0)$ take usually different values in the sense that

$$\begin{aligned} Var(\log w^r) &= Var(\log MPL^0 + \varsigma) \\ &= Var(\log MPL^0) + Var(\varsigma) + 2Cov(\log MPL^0, \varsigma), \end{aligned} \tag{4.16}$$

the lack of knowledge on ς prevents us from determining the relative magnitudes of pre- and post-migratory variances, i.e. $Var(\log MPL^0)$ and $Var(\log \widetilde{w})$ and thus the direction of change of rural inequality resulting from out-migration.

For the reason that the difficulties caused by ς cannot be easily overcome, for the preliminary discussion below, I have no choice but to ignore ς. Note that by ignoring ς, the first category of intra-household economic linkages mentioned in Chapter 2 is excluded from subsequent discussions. The agricultural wages earned by staying workers are thus unaffected by out-migration. Nevertheless, I continue to emphasize the second category of linkages, namely the intra-household income pooling. The income pooling mechanism requires that the wages earned by migratory rural household members must be taken into account in determining the post-migratory wage distribution and inequality.

Once ς is ignored, for any rural worker, his log reservation wage of out-migration $\log w^r$ equals to his log pre-migratory agricultural wage $\log MPL^0$. For expository simplicity, I use $\log w_0$ to denote $\log MPL^0$ below. Therefore, the three assumptions of the extended Roy model can be simplified as:

(1') The distribution of log underlying sectoral wages $(\log w_0, \log w_1)$ faced by all rural workers is jointly normal such that

$$(\log w_0, \log w_1) \overset{i.i.d.}{\sim} N(\mu_0, \mu_1, \sigma_0, \sigma_1, \rho_{01}). \tag{4.17}$$

(2') The log observed wage of any rural worker is given by

$$\log \tilde{w} \approx \max\{\log w_0, \log w_1\}. \tag{4.18}$$

(3') The post-migratory wage inequality is measured by $Var(\log \tilde{w})$.

By comparing the assumptions of the standard and the simplified extended Roy models, it is found that both models are similar except for one notable distinction, namely, the post-migratory wage inequality in the extended Roy model is measured by the censored variance $Var(\log \tilde{w})$ instead of its truncated counterpart suggested by the standard Roy model, i.e. $Var(\log w_0 \mid \log w_0 > \log w_1)$.

Given the three assumptions above, I attempt to determine the impacts that out-migration has on rural inequality. According to assumption (3'), such impacts could be determined by comparing the pre- and post-migratory wage inequalities, i.e. $Var(\log w_0)$ and $Var(\log \tilde{w})$.

By assumption, the pre-migratory wage inequality equals to σ_0^2. Applying the law of total variance to the post-migratory wage inequality $Var(\log \widetilde{w})$ yields:

$$
\begin{aligned}
Var(\log \widetilde{w}) &= E_I(Var(\log \widetilde{w} \mid I)) + Var_I(E(\log \widetilde{w} \mid I)) \\
&= [1-\Phi(c)]Var(\log w_0 \mid I=0) \\
&+ \Phi(c)Var(\log w_1 \mid I=1) \\
&+ \Phi(c)[1-\Phi(c)]\{[E(\log w_0 \mid I=0)]-[E(\log w_1 \mid I=1)]\}^2,
\end{aligned}
\tag{4.19}
$$

where $I = 1\{\log w_0 < \log w_1\}$ denotes the migration status of rural workers, c denotes the standardized intersectoral wage gap with $c \equiv (\mu_1 - \mu_0)/\sqrt{\sigma_0^2 + \sigma_1^2 - 2\rho_{01}\sigma_0\sigma_1}$, and $\Phi(c)$ denotes the population share of migratory rural workers.

Substituting the expressions of truncated means and variances into equation (4.19) gives the formula below[40]

$$
\begin{aligned}
Var(\log \widetilde{w}) &= [1-\Phi(c)]\sigma_0^2[1-\beta^2\delta(c)] + \Phi(c)\sigma_1^2[1-\gamma^2\delta(-c)] \\
&+ \Phi(c)[1-\Phi(c)]\{[\mu_0 - \sigma_0\beta\lambda(c)]-[\mu_1+\sigma_1\gamma\lambda(-c)]\}^2,
\end{aligned}
\tag{4.20}
$$

where $\lambda(c)$, $\delta(c)$ are the inverse Mills ratio and its derivative with respect to c, β and γ are correlation coefficients defined as $\beta \equiv (\rho_{01}\sigma_1 - \sigma_0)/\sqrt{\sigma_0^2 + \sigma_1^2 - 2\rho_{01}\sigma_0\sigma_1}$ and $\gamma \equiv (\sigma_1 - \rho_{01}\sigma_0)/\sqrt{\sigma_0^2 + \sigma_1^2 - 2\rho_{01}\sigma_0\sigma_1}$.

As suggested by equation (4.20), the functional form of $Var(\log \widetilde{w})$ is so complex that the impacts of out-migration on rural inequality may not be determined analytically. Therefore, for the rest of this section, graphics based on numerical evidences are used intensively to assist the exploration. To facilitate graphical presentations, the exploration below measures the distributional impacts using the

[40] See Appendix A for a discussion.

ratio (instead of difference) between $Var(\log \widetilde{w})$ and $Var(\log w_0)$.

$$\begin{aligned} R_V &\equiv Var(\log \widetilde{w}) / Var(\log w_0) \\ &= [1-\Phi(c)][1-\beta^2\delta(c)] + \Phi(c)(\sigma_1/\sigma_0)^2[1-\gamma^2\delta(-c)] \\ &\quad + \Phi(c)[1-\Phi(c)]\{[(\mu_0/\sigma_0)-\beta\lambda(c)]-[(\mu_1/\sigma_0)+(\sigma_1/\sigma_0)\gamma\lambda(-c)]\}^2, \end{aligned} \tag{4.21}$$

mainly because the ratio depends only on dimensionless parameters. To reduce the dimensionality, all parameters in equation (4.21) are expressed as combinations of four independent parameters $\{\sigma_1/\sigma_0, \mu_1/\mu_0, \mu_0/\sigma_0, \rho_{01}\}$ such that[41]

$$\begin{cases} c = \dfrac{[(\mu_1/\mu_0)-1](\mu_0/\sigma_0)}{\sqrt{1+(\sigma_1/\sigma_0)^2-2\rho_{01}(\sigma_1/\sigma_0)}}; \\ \beta = \dfrac{\rho_{01}(\sigma_1/\sigma_0)-1}{\sqrt{1+(\sigma_1/\sigma_0)^2-2\rho_{01}(\sigma_1/\sigma_0)}}; \\ \gamma = \dfrac{(\sigma_1/\sigma_0)-\rho_{01}}{\sqrt{1+(\sigma_1/\sigma_0)^2-2\rho_{01}(\sigma_1/\sigma_0)}}; \\ (\mu_1/\sigma_0) = (\mu_1/\mu_0)(\mu_0/\sigma_0). \end{cases} \tag{4.22}$$

Based on equations (4.21) and (4.22), I give some discussions on the following three questions about the redefined rural inequality and its changes during rural-to-urban migration. They are (1) whether out-migration always decreases rural inequality; (2) if the answer to the first question is negative, then under which conditions the rural inequality tends to decrease and increase; and (3) whether the pattern of migration selection and the (direction of) change of rural inequality are causally interrelated. Subsequent discussions are organized around these questions.

Since R_V is a function of parameters $\{\sigma_1/\sigma_0, \mu_1/\mu_0, \mu_0/\sigma_0, \rho_{01}\}$, interactions among these parameters make it often difficult to approach these questions. Therefore,

[41] These four independent parameters form a basis, but the basis is by no means the only basis.

I isolate one parameter at a time, holding other parameters constant. In other words, I consider several simpler questions of comparative statics as intermediate steps. The discussion below pays particular attention on comparative statics of R_V with respect to σ_1 / σ_0 and μ_1/μ_0.

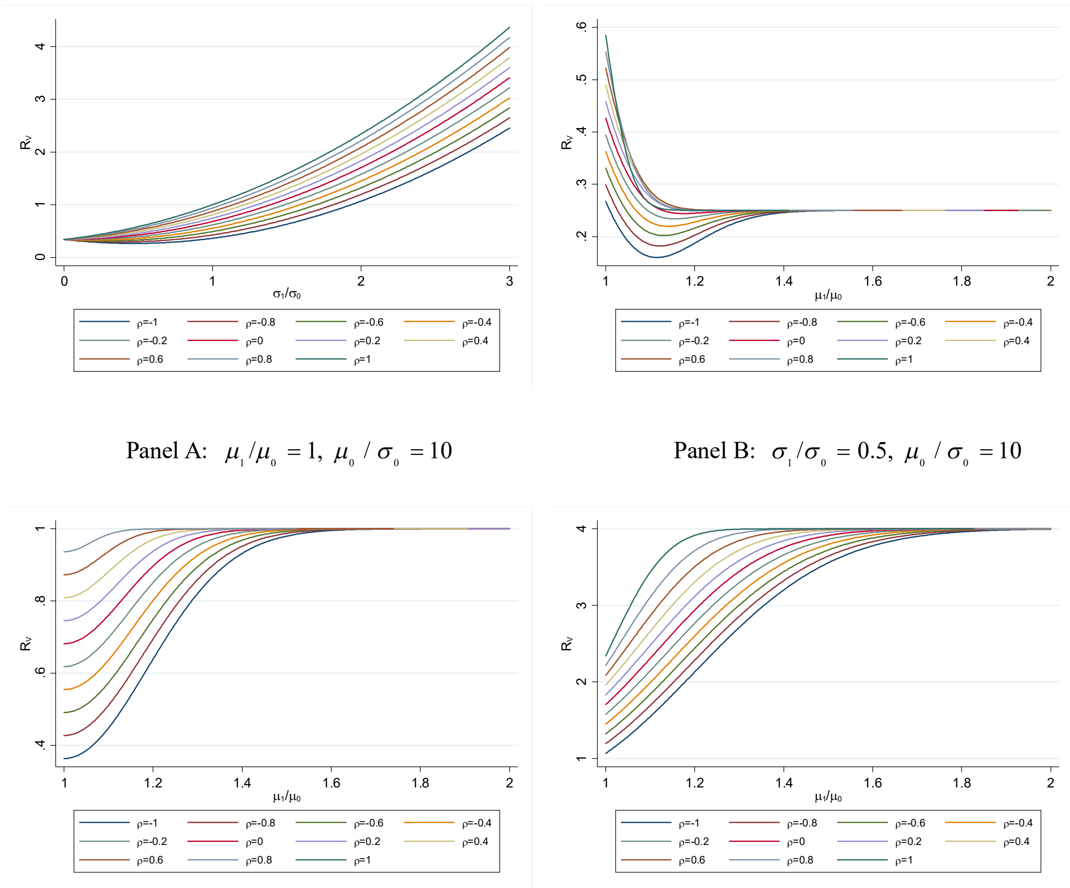
Firstly, I provide evidences about comparative statics concerning σ_1 / σ_0. To visualize the evidences, specific values have to be assigned to parameters μ_1/μ_0 and μ_0/σ_0 in advance. For example, Panel A of Figure 4.1 assumes that $\mu_1/\mu_0 = 1$ so that the change of rural inequality caused by intersectoral wage gap can be eliminated and it assumes further that $\mu_0/\sigma_0 = 10$, which is supported by empirics in Chapter 6. Moreover, another parameter ρ_{01} is controlled for explicitly in Panel A: Different curves in Panel A correspond to different values of $\rho_{01} \in [-1,1]$. These iso-ρ_{01} curves are denoted henceforth by $R_V(\sigma_1 / \sigma_0; \rho_{01})$.

In general, Panel A shows clearly that the values of R_V could be either smaller or larger than unity, which suggests that out-migration could decrease or increase the redefined rural inequality. Since the iso-ρ_{01} curves are usually increasing in σ_1 / σ_0, a rising rural inequality ($R_V > 1$) requires "sufficiently" large σ_1 / σ_0, which depend in turn on values of other parameters. For instance, holding others constant, a "sufficiently" large σ_1 / σ_0 tends to be smaller for cases in which ρ_{01} is positive and large than other cases where ρ_{01} is small or even negative in its value.

Next, I provide evidences about comparative statics concerning μ_1/μ_0 when

σ_1/σ_0 takes three somewhat representative values, namely σ_1/σ_0 = 0.5, 1 and 2. Again, I assume that $\mu_0/\sigma_0 = 10$.

Panels B, C and D of Figure 4.1 confirm that out-migration could either decrease or increase rural inequality. Furthermore, an increase in intersectoral wage gap μ_1/μ_0 raises the overall rural inequality when $\sigma_1/\sigma_0 > 1$ (See Panel D), but it is not the case when $\sigma_1/\sigma_0 \leq 1$. In particular, as illustrated in Panel B, if $\sigma_1/\sigma_0 < 1$, an increase in μ_1/μ_0 tends to decrease the overall rural inequality, that is, a sharp decrease followed by a mild increase in the overall rural inequality.



Panel A: $\mu_1/\mu_0 = 1$, $\mu_0/\sigma_0 = 10$

Panel B: $\sigma_1/\sigma_0 = 0.5$, $\mu_0/\sigma_0 = 10$

Panel C: $\sigma_1/\sigma_0 = 1$, $\mu_0/\sigma_0 = 10$

Panel D: $\sigma_1/\sigma_0 = 2$, $\mu_0/\sigma_0 = 10$

Figure 4.1: Comparative Statics of R_V with respect to σ_1/σ_0 and μ_1/μ_0

In sum, all comparative statics in Figure 4.1 give a negative answer to the first question: Under the new definition of rural inequality, out-migration could not only decrease, but increase rural inequality. As for the second question, though the graphical method provides insights into the conditions under which rural inequality could decrease or increase, this method may not suffice to determine these conditions precisely. Furthermore, even if we could determine them by applying this method repeatedly, it would be still difficult to identify underlying parameters in the conditions from the data and then to test these conditions against the data and facts.

The difficulties in answering the second question lead to further exploration of the third question, because comparing to the change of rural inequality, we are usually more knowledgeable about the pattern of migration selection, and also because the identification of the pattern is usually more straightforward than that of the change of rural inequality.[42] If the pattern of migration selection and the change of rural inequality are in fact interrelated, just as what Lipton (1980) and Li (2003) suggests, we can predict the change of rural inequality based on already known or relatively easily identified pattern of migration selection.

Therefore, at the end of this section, I will explore in detail the relationships between the pattern of migration selection and (direction of) change of rural inequality using the analytical toolkit developed so far, which includes the formal definitions of key concepts and the graphic method based on numerical evidences. The following exploration provides also a re-evaluation of the well-known conjecture.

By inspecting the conditions for different patterns of migration selection given

[42] This is mainly because the identification of the pattern of migration selection associates only with comparisons of conditional means (or expectations) of wages earned by different subgroups of rural workers. Most conventional regression techniques serve well for this purpose.

by equations (2.3) - (2.5) and the formula of R_V given by equations (4.21), (4.22), it is easy to find that all of them relate to parameters $\{\sigma_1/\sigma_0, \mu_1/\mu_0, \mu_0/\sigma_0, \rho_{01}\}$. This fact motivates the following exploration of the relationships between the pattern of migration selection and change of rural inequality.

Nevertheless, since the conditions of migration selection contain only two of four parameters, i.e. σ_1/σ_0 and ρ_{01}, in general we cannot predict the change (direction and magnitude) of rural inequality precisely using the pattern of migration selection, unless we assign some specific values to free parameters, i.e. μ_1/μ_0 and μ_0/σ_0 that are excluded from the conditions for migration selection. Considering that the choice of free parameters could have some influences on results and there is no way to know these influences in advance, I will firstly explore the relationships of interest when μ_1/μ_0 and μ_0/σ_0 are set to 1 and 10. Then I will check the robustness of the preliminary findings by assigning alternative values to these free parameters.

Besides, in consistency with previous discussion, the graphic method is used as the tool for the exploration below. If the third question is reformulated in graphic terms, it reads as whether the values of R_V in sub-regions on the σ_1/σ_0 - R_V plane associating with one of the three patterns of migration selection differ considerably. To be more specific, I am particularly interested in whether some of these sub-regions lie entirely above or below the horizontal line $R_V = 1$ on the plane. For simplicity, x and ρ are used to denoted σ_1/σ_0 and ρ_{01}.

I begin by considering cases in which out-migrations exhibit the positive sorting. With new notations, the sufficient and necessary conditions for positive sorting can be written as $\rho > 1/x$ and $x > 1$. These conditions specify one of its boundaries that separates the sub-region in which migrations exhibit the positive sorting from other sub-regions on the x - R_V plane. To see why this is the case, we may think of examples with concrete numbers. For instance, when $\rho = 1$, these conditions imply $x > 1/\rho = 1$, which means along the iso-ρ curve $R_V(x; \rho = 1)$, the section in which migrations exhibit the positive sorting must lie to the right of a threshold $R_V(1; \rho = 1)$; Similarly, when $\rho = 0.5$, we know that along the iso-ρ curve $R_V(x; \rho = 0.5)$, the section in which migrations exhibit the positive sorting must lie to the right of a threshold $R_V(2; \rho = 0.5)$, etc. Connecting these thresholds on all iso-ρ curves with positive ρ gives one of the boundaries of the sub-region within which migrations exhibit the positive sorting denoting by $R_V(x; \rho = 1/x),\ x > 1$. Figure 4.2 suggests that this boundary is likely to be increasing in x.

Another boundary for the same sub-region is often offered by the boundaries of feasible region, which is shorthand for all possible values of (x, R_V), given other parameters. For the special case presented in Figure 4.2 with $\mu_1/\mu_0 = 1$, the iso-ρ curve $R_V(x; \rho = 1)$ offers the upper boundary for the feasible region, and thus for the sub-region with the positive sorting.[43]

The sub-region surrounded by the two boundaries, i.e. $R_V(x; \rho = 1/x)$ and $R_V(x; \rho = 1)$ consists of all cases in which out-migrations exhibit the positive sorting.

[43] For the special case in which $\mu_1 / \mu_0 = 1$, we can prove $R_V(x; \rho)$ is increasing in ρ. Thus, $R_V(x; \rho = 1)$ offers an upper boundary of the feasible region, while $R_V(x; \rho = -1)$ offers a lower boundary.

Likewise, I can show that the sub-region surrounded by another two boundaries, i.e. $R_V(x;\rho = x)$ and $R_V(x;\rho = -1)$ consists of all cases in which out- migrations exhibit the negative sorting. The rest of the feasible region consists of all cases in which migrations exhibit the non-hierarchical sorting.

Furthermore, I find that the boundaries defined by the conditions for the positive and the negative sorting as well as the iso-ρ curve $R_V(x;\rho = 1)$ intersect at the point $(1,1)$. According to Figure 4.2, all of these curves are likely to be increasing in *x*. Therefore, the sub-region in which migrations exhibit the positive sorting lies entirely above the line $R_V = 1$; the sub-region in which migrations exhibit the negative sorting lies entirely below the line $R_V = 1$; while the sub-region in which migrations exhibit the non-hierarchical sorting lies on the both sides of $R_V = 1$. These findings of the graphic analysis imply that migrations exhibiting the positive sorting tend to coexist with increasing rural inequality; migrations exhibiting the negative sorting tend to coexist with decreasing rural inequality, while the relationship between migrations exhibiting the non-hierarchical sorting and the direction of change of rural inequality is underdetermined.

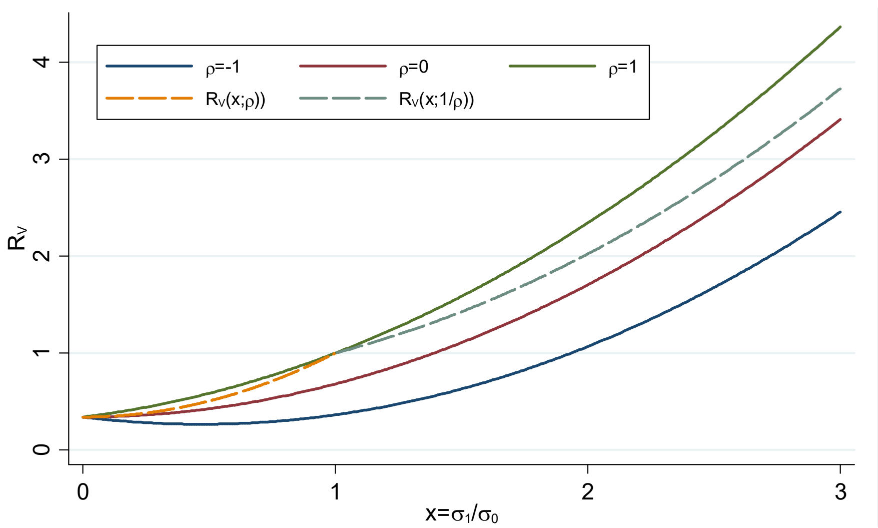


Figure 4.2: Migration Selection and Change of Rural Inequality

Note: The figure is drawn using Stata when free parameter $\mu_1 / \mu_0 = 1$ and $\mu_0 / \sigma_0 = 10$.

By assigning alternative values to parameters μ_1/μ_0 and μ_0 / σ_0, for instance, various $\mu_1/\mu_0 \in [1, 2]$, $\mu_0 / \sigma_0 = 1, 5$ and repeating the graphic analyses above, I conclude that the findings based on specific values of free parameters are likely to be robust to different choices of free parameters.

Until now, I have established the theoretical relationships between the pattern of out-migration selection and the change of the rural inequality. As argued before, these relationships shed also light on empirical studies. Particularly, based on the knowledge of the pattern of migration selection, one can predict the direction of change of the redefined rural inequality. Before concluding this section, it seems necessary to make two further remarks about these relationships: Firstly, unlike the conventional view expressed in Lipton (1980) and Li (2003), these relationships are non-causal. They appear to be interrelated only because they relate to the same group of distributional parameters and hence to the same group of conditions of migration selection. These relationships should be thus labeled as spurious. Secondly, the exploration ignores the random term ς that captures the first category of intra-household economic linkages. In this regard, the exploration is illuminating yet by no means definitive.

CHAPTER 5

THE IMPACTS OF OUT-MIGRATION: THE MODEL WITH HETEROGENEOUS AND IMPERFECTLY SUBSTITUTIVE LABOR

This chapter provides a preliminary discussion on the impacts that out-migration has on rural inequality in a more general setting, namely different types of labor inputs are imperfect substitutes in household agricultural production.

The role of imperfect substitution among different types of labor inputs receives increasing attentions among economists in the past few decades, mainly because it is proven an essential component for understanding the rising skill premia and earning inequalities observed in many countries during that period.[44] As for the context focused by this thesis, in the presence of imperfect substitution, the often selective out-migration tends to affect the wages earned by rural workers and thus rural inequality not only through changing the effective land-labor ratios of household farms, but also through changing the skill mix of workforces in household farms. The second channel – the skill mix effect is new for this chapter.

5.1 The Constrained Maximization Problem

Like in Chapter 4, I consider firstly the microeconomic problem faced by any rural household who attempts to maximize its income by reallocating the predetermined labor endowments provided by members between their own farms and emerging urban employments.

Chapter 5 inherits most of the assumptions from Chapter 4, except that the substitution among different types of labor inputs is now imperfect, which suggests

[44] See Katz and Autor (1999) for a survey.

the parameter $\rho \in [0,1)$.

Therefore, the constrained income maximization problem faced by any rural household takes the form of

$$\begin{aligned}
& \max_{\{L(s)\}} T(A^0)^\alpha [\int_0^1 q(s)L(s)^\rho ds]^{(1-\alpha)/\rho} + \int_0^1 w_1(s)[L^0(s) - L(s)]ds \\
s.t. \quad & 0 < L(s) \le L^0(s) \\
where \quad & \{A^0, \{L^0(s), \forall s\}\} \textit{ are predetermined,} \\
and \quad & \{w_1(s), \forall s\} \textit{ are exogenously given.}
\end{aligned}$$

Notice that as a byproduct, the nested CES specification used in this chapter requires the amounts of agricultural labor inputs at all skill levels must be strictly positive. This requirement causes no serious problem when we study economic agents of large scale such as the firms or aggregated economies, but given that the size of typical rural household farm is usually small, it seems often too restrictive. Nevertheless, bearing its drawback in mind, the preliminary discussion below will not question about its validity.

Solving the constrained maximization problem faced by rural households gives the complimentary-slackness conditions as follows:

$$\begin{cases} \log MPL(s,...) = \log w(s) \Leftrightarrow L(s) < L^0(s); \\ \log MPL(s,...) \ge \log w(s) \Leftrightarrow L(s) = L^0(s), \end{cases} \tag{5.1}$$

where the log agricultural wage faced by a marginal worker at skill level s is given by

$$\log MPL(s,...) = \log T(s) + \alpha \log a + \frac{\rho - 1}{\rho} \log F(s), \tag{5.2}$$

which means that his log agricultural wage is determined by a skill-specific technical term $T(s) \equiv (1-\alpha)q(s)^{1/\rho}T$, the effective land-labor ratio $a \equiv A^0 / (\int_0^1 q(s)L(s)^\rho ds)^{1/\rho}$ of his household farm and the labor share that his skill cell has in the household total agricultural workforce $F(s) \equiv q(s)L(s)^\rho / \int_0^1 q(t)L(t)^\rho dt$. Evaluating equation (5.2) at $(\log a^0, \{\log F^0(s)\})$ and $(\log a^\#, \{\log F^\#(s)\})$ gives the pre- and post-migratory log agricultural wages such that

$$\log MPL^0(s,...) = \log T(s) + \alpha \log a^0 + \frac{\rho-1}{\rho}\log F^0(s), \tag{5.3}$$

$$\log MPL^\#(s,...) = \log T(s) + \alpha \log a^\# + \frac{\rho-1}{\rho}\log F^\#(s), \tag{5.4}$$

To utilize the insights from previous chapters, the complimentary-slackness conditions given by equation (5.1) is rearranged to a form resembling to equations (3.5) and (4.9). A critical step towards finding the alternative formulation of equation (5.1) is to construct the log reservation wages $\log w^r(s,...)$ of out-migration faced by marginal workers at different skill levels. After several mathematical manipulations, the formula of log reservation wages of out-migration is finally given by[45]

$$\log w^r(s,...) \equiv \log T(s) + \alpha \log a^\# + \frac{\rho-1}{\rho}\log\frac{q(s)L^0(s)^\rho}{\int_0^1 q(t)L^\#(t)^\rho dt}. \tag{5.5}$$

For types of labor inputs that are partially migratory, $\log w^r(s,...)$ in equation (5.5) measures the log opportunity costs of out-migration faced by them, that is, the log agricultural wages that these workers could earn if they choose "counterfactually" to stay on-farm, while workers of other types behave optimally according to equation

[45] See Appendix A for more details.

(5.1). For types of labor inputs that are completely non-migratory, their agricultural labor inputs keep unchanged, i.e. $L^{\#}(s)=L^{0}(s)$, thus $\log w^{r}(s,...)$ in equation (5.5) measures the log agricultural wages that workers of these types actually earn.

Reformulating the complimentary-slackness conditions in equation (5.1) in terms of log reservation wage $\log w^{r}(s,...)$ yields an equation with desirable property as follows

$$\log MPL^{\#}(s,...)=\max\{\log w^{r}(s,...),\log w_{1}(s)\},\ \forall s\in[0,1]. \tag{5.6}$$

Moreover, the relation between the log reservation wage and log pre-migratory agricultural wage is given by

$$\begin{aligned}\log w^{r}(s,...)&=\log MPL^{0}(s,...)+(\alpha+\rho-1)(\log a^{\#}-\log a^{0})\\&\equiv\log MPL^{0}(s,...)+\varsigma'.\end{aligned} \tag{5.7}$$

Analogous to Chapter 4, equation (5.7) shows that $\log w^{r}(s,...)$ deviates from $\log MPL^{0}(s,...)$ by a term $\varsigma'\equiv(\alpha+\rho-1)(\log a^{\#}-\log a^{0})$. The deviation ς' combines two often opposite-in-direction effects that out-migration has on log agricultural wages of staying workers and thus on log reservation wages. The first effect should have been familiar to the readers – Out-migration tends to increase the effective land-labor ratio of household farm and hence the log agricultural wages faced by all types of labor inputs. The second effect is that out-migration of some types of labor inputs tends to decrease the agricultural wages earned by types of labor inputs that are *q*- complimentary to the migratory types (Hamermesh, 1993). Suppose the first effect dominates, which is more likely when ρ is large, the overall effect is to increase the agricultural wages; otherwise, the overall effect is to decrease the agricultural wages. Once again, like in Chapter 4, there exists no explicit solution to the additional term ς' and we know little about the distributional properties of the corresponding random term ς' at rural community level.

Equations (5.6) and (5.7) can be adopted for describing the wage determination of all workers who are members of rural households, regardless of their occupations. With the help of the notation $\log \widetilde{w}(s,...)$ introduced in Chapter 4, the wage determination of all rural workers is given by

$$\begin{aligned} &\log \widetilde{w}_i = \max\{\log w_i^r, \log w_{1i}\},\ \forall i, \\ &where\ \log w_i^r = \log MPL_i^0 + \varsigma_i'. \end{aligned} \tag{5.8}$$

5.2 The Impacts on the Overall Inequality

Having obtained equation (5.8) by solving the constrained maximization problem of individual rural household, it appears that like in Chapters 3 and 4, we could use the extended model as the analytical framework to explore the impacts that out-migration has on rural inequality. At the first sight, the subsequent exploration would not differ much from that in Chapter 4 except that the additional random term ς is now replaced by ς'.

Unfortunately, as will be explained below, the imperfectness of substitution among different types of agricultural labor inputs leads to a serious violation of one of the basic assumptions maintained by the standard Roy model and virtually its various extensions in Chapters 3 and 4. Consequently, the Roy models in general cannot give any prediction on the impacts that out-migration has on the overall rural inequality, even in any approximate sense.

For the reason that the distributional properties of ς' are largely unknown, we have to ignore it from this preliminary exploration altogether. As a result, the log reservation wage $\log w^r(s,...)$ equals to the log pre-migratory agricultural wage $\log MPL^0(s,...)$. Because of the imperfectness in labor substitution, observations of log pre-migratory agricultural wages $\log MPL^0(s,...)$ in any rural community cannot be viewed as realizations of a univariate random variable, unless a rare occasion when the initial skill mix characterized by $\{\log F^0(s)\}$ happen to be the same among all households. For the majority cases of interest, the skill mix and thus $\log MPL^0(s,...)$ earned by rural workers should be represented by a continuous-time stochastic process, which can be loosely understood as a collection of possibly interrelated univariate random variables. This observation is responsible for the incapability of the Roy models continuing to serve as an analytical framework when the labor substitution is imperfect.

To illustrate this key observation, let us revisit the expression of $\log MPL^0(s,...)$ given by equation (5.3)

$$\log MPL^0(s,...) = \log T(s) + \alpha \log a^0 + \frac{\rho-1}{\rho}\log F^0(s).$$

By looking at the definition, it is found that at community level the first two right-hand side terms can be in principle represented by univariate random variables. To the contrary, for most cases of interest, the third right-hand side term cannot be represented by a univariate random variable. This is because at every skill level, the pre-migratory labor share $F^0(s)\,|\,s$ is likely to vary across households, which suggests that conditional on the skill level, $F^0(s)$ and its function $\log F^0(s)$ can be often represented by univariate random variables. If $\log F^0(s) \mid s$ at all skill levels in the continuum $s \in [0,1]$ are stacked successively, they form a stochastic process. As suggested by equation (5.3), for the heterogeneous and imperfectly substitutive labor setting with $\rho \in [0,1)$, if $\{\log F^0(s)\}$ is a stochastic process, hence $\{\log MPL^0\}$ must be a stochastic process as well. A graphical illustration why $\{\log F^0(s)\}$ forms a stochastic process is given in Figure 5.1 below.

Since $\log MPL^0(s,...)$ in any community cannot be represented by a univariate random variable, the distribution of $(\log MPL^0(s,...), \log w_1(s))$ at all skill levels cannot be jointly normal. Thus, one of the basic assumptions maintained by all of the Roy models, including the standard Roy model and its extensions in Chapters 3 and 4 are violated. Therefore, in the presence of imperfectness in labor substitution, the Roy models provide no longer proper framework for the theoretical exploration of the overall distributional impacts of out-migration on inequality in source regions.

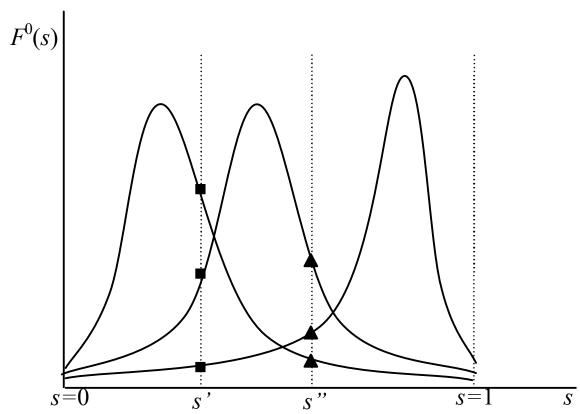


Figure 5.1: An Illustration Why $\{\log F^0(s)\}$ Forms a Stochastic Process

Note that bell-shape curves describe skill mixes of different households in the rural community. At all skill levels, the labor shares $F^0(s)$ are likely to vary across households.

5.3 The Impacts on the Within- and Between-Group Inequalities

Nevertheless, if we could be satisfied with answering narrower and less ambitious questions, for example, how out-migration affects the wage inequalities within skill-specific subgroups, Roy models may still provide insights on these issues.

Conditional on rural workers' skill levels, as already explained, $\log F^0(s) \mid s$ and $\log MPL^0(s,\ldots) \mid s$ can be represented by univariate random variables. Assuming further that the conditional distribution of $(\log MPL^0(s,\ldots), \log w(s))$ is also jointly normal, the discussions in Chapters 3 and 4 have shown clearly that

$$Var[\log \tilde{w}(s,\ldots) \mid s] \leq Var[\log MPL^0(s,\ldots) \mid s], \forall s, \tag{5.9}$$

which suggests that out-migration tends to decrease the wage inequalities within all skill-specific subgroups of rural workers. However, it is worthy to notice that this result remains valid in strict sense only when the existence of the random term ς' does not reverse the signs of inequalities given by equation (5.9).

Until now, I have not explored the impacts that out-migration has on between-group wage inequalities, partly because that question is foremost an empirical one and hence the Roy models can offer limited help in answering it.[46]

In literature on the Chinese labor markets, it is well established that the rate of return to schooling in urban regions exceeds that in rural regions. By assuming that schooling or say, formal education is the only or the dominating contributor to skill of rural workers, such empirical regularity may be stretched as the return to skill in urban regions exceeds its rural counterpart. It follows that between-group wage inequality

[46] The change of between-group wage inequalities resulting from immigration is the very question that Borjas (2003) and Ottaviano and Peri (2008, 2011) attempt to answer.

among skill-specific subgroups tends to increase, since rural workers' post-migratory return to skill would resemble more closely to that in urban regions rather than that in rural regions.

However, there are some doubts whether the skill of rural workers can be chiefly determined by formal education. In fact, empirical studies in Chapter 6 show that in period around the year 2000, the pattern of migration selection is likely to be the non-hierarchical sorting, which provides a rejection to the one-dimensional skill assumption.[47] Besides education, other characteristics, especially working experience could be also important in determining the skill and wage earned by rural workers.

Before leaving for the empirical exploration, I return briefly to the discussion on the microeconomic contents, or more precisely, on the source of pre-migratory wage inequality. Following the practices in Chapters 2 and 3, I calculate firstly the mean of log agricultural wages $\log MPL^0(s,...)$ faced by all rural workers denoting by $\mu_0 \equiv E[\log MPL^0(s,...)]$. I subtract then μ_0 from $\log MPL^0(s,...)$ to obtain the demeaned error term $\varepsilon_0^0(s,...)$. Finally, I attach the microeconomic contents to each of the decomposition components of $\varepsilon_0^0(s,...)$.

Applying the law of iterated expectation gives

$$\begin{aligned}\mu_0 &= E\{E[\log MPL^0(s,...) \mid s]\} \\ &= E\{\log T(s) + \frac{\rho-1}{\rho} E[\log F^0(s) \mid s] + \alpha E[\log a^0 \mid s]\}.\end{aligned}$$

[47] The non-hierarchical sorting is impossible when skill is one-dimensional. See also footnote 11.

By definition of $\log \mathcal{F}^0(s) \equiv E[\log F^0(s) \mid s]$, the formula can be simplified as

$$\begin{aligned}\mu_0 &\equiv E\{\log T(s) + \frac{\rho-1}{\rho}\log \mathcal{F}^0(s)\} + \alpha E\{\log a^0\} \\ &= \log T(\bar{s}) + \frac{\rho-1}{\rho}\log \mathcal{F}^0(\bar{s}) + \alpha \log \bar{a},\end{aligned} \tag{5.10}$$

where the new notation $\mathcal{F}^0(s)$ turns out to be the geometric average of $F^0(s)$ among all households in the rural community. $\bar{s}$ and $\bar{a}$ refer to s and a at which $\log T(s) + \frac{\rho-1}{\rho}\log \mathcal{F}^0(s)$ and $\log a^0$ equal to their community averages.

Hence, the demeaned error term $\varepsilon_0^0(s,...)$ is given by

$$\begin{aligned}\varepsilon_0^0(s,...) &\equiv \log MPL^0(s,...) - \mu_0 \\ &= \frac{\rho-1}{\rho}\log[q(s)/q(\bar{s})] + \alpha \log[a^0/\bar{a}] + \frac{\rho-1}{\rho}\log[F^0(s)/\mathcal{F}^0(\bar{s})],\end{aligned} \tag{5.11}$$

where the first right-hand side term corresponds to the individual-specific term φ_i in equation (2.8), the second term corresponds to the sector-specific term ξ_{ji} in the same equation. Although the third term is also household-specific and difficult to be transferred across sectors, the demeaned term cannot be represented by a univariate random variable. Thus, it does not belong to ξ_{ji} in equation (2.8).

By expanding both $\log q(s)$ and $\log \mathcal{F}^0(s)$ as first-order Taylor's series around $s = \bar{s}$, equation (5.11) can be approximated as follows

$$\varepsilon_0^0(s,...) \approx r_{total}[s-\bar{s}] + \alpha \log[a^0/\bar{a}] + \frac{\rho-1}{\rho}\log[F^0(s)/\mathcal{F}^0(s)], \tag{5.12}$$

where the total rate of return to skill in the agricultural sector r_{total} is given by

$$r_{total} \equiv d\log MPL^{0}(\bar{s})/ds = \frac{1}{\rho} d\log q(\bar{s})/ds + \frac{\rho - 1}{\rho} d\log \mathcal{F}^{0}(\bar{s})/ds. \quad (5.13)$$

The rate of return in equation (5.13) is labeled as the total rate because it consists of two parts. The first right-hand side term refers to the rate of return to skill when labor inputs in rural community of interest are balanced in quality so that the return relates only to the intrinsic agricultural productivity of labor input. The second term refers to the rate of return to skill owing to the relative scarcity of skill in rural community.

Notice that for the special case where heterogeneous labor inputs are perfect substitutes in household agricultural production, i.e. $\rho = 1$, the third right-hand side term in equation (5.12) disappears. Therefore, the demeaned error term $\varepsilon_{0}^{0}(s,...)$ and hence $\log MPL^{0}(s,...)$ faced by all rural workers can be represented by univariate random variables. The problem considered in this chapter reduces to a simpler one that has been explored in detail in Chapter 4. According to equation (5.13), for this special case when $\rho = 1$, the rate of return to skill in agricultural sector depends only on intrinsic skills of rural workers such that

$$r_{partial} \equiv \partial \log MPL^{0}(\bar{s})/\partial s = \frac{1}{\rho} d\log q(\bar{s})/ds. \quad (5.14)$$

CHAPTER 6

EMPIRICAL EXPLORATIONS: PUZZLE AND EXPLANATIONS

This chapter provides empirical explorations supplementing previous theoretical exploration of the impacts that out-migration has on rural inequality. In Chapters 3, 4 and 5, I have analyzed the impacts using the extended Roy models based on various settings where labor inputs are homogenous (Chapter 3), heterogeneous and perfectly substitutive (Chapter 4), heterogeneous and imperfectly substitutive (Chapter 5) in household agricultural production. These analyses show that the predictions of the extended Roy models differ in three settings. The differences remind us the necessity of examining the relevance of these settings to the data before linking these theoretical predictions with the data.

6.1 The Agricultural Technology, Labor Heterogeneity and Substitution

To determine the setting of the highest relevance, I estimate in this section the agricultural production function of household farms in rural China. The results enable us to answer the following questions in sequence, namely (1) whether different types of labor inputs are perfectly substitutes in household agricultural production, and (2) provided that they are perfect substitutes, whether and to what extent they are heterogeneous in terms of intrinsic agricultural productivity. The discussion below centers on these two empirical questions.

6.1.1 Data, Empirical Method and Specification

In this subsection, I specify the data and empirical methods used in estimation, the definitions and measurements of inputs and output in household agricultural production as well as the functional forms of agricultural technology.

Data and Empirical Methods

The following empirical exploration employs the Chinese Health and Nutrition Survey (CHNS) dataset, 1991-2009 as main data source.[48] In each of the seven surveys, information on health, demographic and socioeconomic (such as employment and earning) factors are collected for individuals and their households residing in communities in up to nine provinces. All communities surveyed in the CHNS can be sorted into four regional categories: urban and suburban neighborhoods, county town and village. This section uses only the village subsample in estimation of the agricultural technology, which accounts for slightly more than half of all observations.

Despite the CHNS data are essentially unbalanced panel data, I make no serious effort below to exploit its panel structure to facilitate precise identifications. This is because given sample size and data quality, and particularly the adoption of nonlinear specifications of technology the estimation using panel data involves often great difficulty in numerical optimization. Instead, in explorations below, the CHNS data is treated as if they were repeated cross-sectional and hence the production functions are estimated separately at each of the cross-sections.[49]

The failure of exploiting the panel structure of the data makes it impossible to apply the fixed-effects strategy to mitigate the endogeneity problem of potential importance in estimation. Alternatively, I address this problem using the Generalized Method of Moment (GMM) with instruments.

---

48 Nine panels of the CHNS were collected in years 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009 and 2011. In this thesis, only the data collected between 1991 and 2009 are used. The panel 1989 is not used mainly because the dummy of migration status cannot be constructed in that year. Panel 2011 is only available recently and thus cannot be used in this study. For more details on the data, reader may refer to Appendix D and CHNS documents.

49 By treating panel data as repeated cross-sectional data, the precision of estimations reduces. Nevertheless, in comparison with estimations using the panel data, the present treatment allows for parameters of interest to change more freely over time. In a period of dramatic transitions, the advantage of the present treatment could be obvious.

Inputs and Output

In what follows, the output of household agricultural production is defined as the gross incomes from sale and self-consumption of the final products of household farming, gardening and livestock raising. The agricultural incomes are inflated to the currency values in 2009.

The inputs of household agricultural production include land, material inputs and various types of labor services.[50] The land input is defined as the amount of land cultivated by household in year before the survey. The material inputs consist of all intermediate inputs such as seeds, fertilizer and energy. Since the CHNS provides no detailed account on material inputs, they are approximated by the agricultural production cost inflated to the currency values in 2009.

To answer the empirical questions focused by this section, all agricultural workers and their services are sorted into skill cells in terms of selected characteristics. In literature, educational attainment is one of the most used indicators for skill. Age or working experience is often regarded as another important indicator for skill. Other characteristics such as gender, marital status are sometimes considered.

Because of the relatively small sample size (about 1500 household farms per year), and large amounts of measurement errors in the data, it seems impossible to divide the workforce into fine skill cells and meanwhile, to obtain precise estimates. I consider thus only two dimensions of skill, i.e. education and age.[51] The agricultural

[50] I do not include the capital input, since it contains a large amount of measurement errors and could lead to serious attenuation biases in estimations. For more details, reader may see Appendix D.

[51] Gender could be another dimension of workers' skill. However, given the data, defining skill cells using all of three dimensions of skill would make estimations very difficult. Thus, in empirical studies below, only education and age are of primary concern. This decision could be further justified by the fact that even if functional forms allowing for differences in productivity between male and female

workers are classified into two groups with similar education, namely those who have no Junior High School ("JHS-") diploma and those who have JHS or a higher diploma ("JHS" or "JHS+"). Besides, agricultural workers are also classified into two age groups, namely young workers ("YG" aged below 35), and middle or old aged workers ("MA" aged between 35 and 55; "OA" aged above 55). Hence, there are four broadly defined education-age cells in the workforce.[52] Labor services provided by workers in different cells are viewed as different types of labor inputs.

Moreover, in literature labor inputs can be measured either in head counts or in working hours. Labor economists prefer generally the latter. However, previous experiments suggest that the data of working hours constructed from the CHNS data could be contaminated by measurement errors. Thus, measuring the labor inputs using the hours could lead to severe attenuation bias. For instance, it is not uncommon to find observations of workers who reported more than 11 working hours per day in the year before the survey. Even after excluding these obvious outliers, regressions using

---

workers with similar education and age are adopted, results suggest that gender productivity gaps may not be as large as what originally expected.

[52] In the U.S. literature, skilled and unskilled workers are usually defined in terms of education. Skilled workers are defined as college graduates or their equivalences, while unskilled workers are defined as high school graduates or their equivalences. In rural China, however, I doubt the usefulness of these definitions. This is because the average education in rural China is remarkably lower than that in the U.S. (see Section 6.2). Suppose that the U.S. definition were adopted, the vast majority of agricultural workers in rural China would have to be classified as less skilled workers. Recognizing this difficulty, I proposed to sort all rural workers into three educational groups, i.e. "JHS-", "JHS" and "JHS+", which correspond roughly to the less, medium and highly educated workers. Workers with exact "JHS" diploma are defined separately from others, partly because the law of compulsory education in China requires children keeping enrolled in the school until the completion of JHS education. Consequently, a considerable share (often around 50%) of rural workers has completed their JHS education before they enter the labor markets. Moreover, the urban employers often use the JHS diploma as filter in recruitment of rural workers. Besides, workers are often sorted into age or experience groups. For example, Borjas (2003) defines several five- year experience intervals and sort workers into these experience groups. In this thesis, due to the small sample size of data and estimation methods, I proposed to define age groups more broadly in 20 years age intervals, namely 35-, 35-55 and 55+. According to the scheme, there are up to nine education-age cells.

Unfortunately, for surveys in the 1990s, regressions of agricultural production function using the nine-skill-cell scheme fail to give precise estimates on parameters. This suggests that I am asking questions that may be too subtle for the available data to answer with any precision. Therefore, I turn to the present specification in the main text by combing finer skill cells into broader ones to obtain precise estimates of key parameters. For the surveys 2000 onwards, however, regressions based on the nine-skill-cell scheme are able to give relatively precise estimates of all parameters of interest.

the annual agricultural working hours as an explanatory variable give often estimates of the returns to scale at the magnitude around 0.7 or even lower, which suggests the technology adopted by Chinese household farms exhibits strong decreasing return to scale. This finding is inconsistent with the empirical evidences presented in Xu et al. (2001). Therefore, throughout the empirical explorations below, I decide to measure the agricultural labor inputs in head counts.

Technology

Keeping the two empirical questions in mind, the functional forms that impose overly strong restrictions on the substitution and productivity heterogeneity among different types of labor inputs should be avoided. For the reason that the sizes of households are usually small and skills of household members are often correlated, the widely used Translog functional form may not be appropriate for estimating the agricultural technology of household farms.[53] Instead, like previous chapters, empirical explorations below use the nested CES as main specification. A noteworthy advantage of the nested CES is that it is compatible to zero inputs, while it preserves much of the virtue of the Translog for it places weak ex ante restrictions on the substitution and heterogeneity among labor inputs.

Following a large literature (Card and Lemieux, 2001; Borjas, 2003; Ottaviano and Peri, 2008 and 2012), this section assumes that the aggregate labor and non-labor inputs are separable in household agricultural production and the technology that

[53] Jacoby (1993) adopts nevertheless the Translog and its special case - the Cobb-Douglas forms in estimating the agricultural production functions in Peru. His strategy is to add one to all inputs to avoid a large number of zero inputs that are incompatible to both functional forms. However, his strategy is not convincing for the reason that although adding one to all inputs could have only marginal effects on the magnitude and direction of input vector, it would be not the case for those of input vector after logarithm-transformation, especially when many inputs are exactly zero.

combines them takes the form of Cobb-Douglas.[54]

The aggregate labor input can be further viewed as the outcome of different types of labor inputs combined by certain CES or nested CES technology. Ideally, one should rely on the knowledge on the ease of substitution to guide the labor aggregation. However, reliable estimates on this issue are often unavailable. As a result, I stick to two widely accepted ways of labor aggregation as follows:

a. Two-Level Nested CES

At the higher level of the two-level nested CES specification, a Cobb-Douglas technology combines the aggregate labor input (*QL*) and non-labor inputs such as land (*A*) and material (*M*). At its lower level, the aggregate labor input is expressed as outcome of a CES function with four education-age-specific labor entries.[55, 56]

$$\begin{aligned} \log y &= \tau + \alpha_A \log A + \alpha_M \log M + \alpha_{QL} \log QL + u \\ where\ QL &= \left(\sum\nolimits_1^4 \beta_s L_s^{\rho}\right)^{1/\rho},\ with \sum\nolimits_1^4 \beta_s = 1 \end{aligned} \tag{6.1}$$

b. Three-Level Nested CES

At the highest level of the three-level nested CES specification, a Cobb-Douglas technology combines the aggregate labor inputs and non-labor inputs. The aggregate

[54] Another literature that focuses on the skill-capital complementarity (see for example, Grilliches, 1969) does not assume this kind of separablity. As pointed out in Borjas (2014), the Cobb-Douglas specification is not entirely innocuous when we analyze the impacts of immigration. Nevertheless, in the literature mentioned in the text, virtually all studies adopt it as the functional form for the highest level of the nested CES.

[55] Strictly speaking, in theoretical chapters I consider mostly the value-added functions instead of the production functions.

[56] This specification is often criticized for it assumes the same elasticity of substitution between any pair of labor inputs, which is often unrealistic for cases with more than two labor inputs. Nevertheless, this specification offers a simple and empirically convenient way to capture both the productivity heterogeneity and the imperfect substitution among different labor inputs. Note that parameters $\beta_1 \sim \beta_4$ have been normalized so that $\Sigma_s \beta_s = 1$. The normalization has no effect on outcomes as long as the constant term τ is included.

labor input is firstly divided into two educational groups at intermediate level. Each of the educational groups is then divided into two age subgroups at its lowest level. This functional form appears repeatedly in the labor economic literature.[57, 58]

$$\begin{aligned}
&\log y = \tau + \alpha_A \log A + \alpha_M \log M + \alpha_{QL} \log QL + u \\
\textit{where } & QL = (\textstyle\sum_i \gamma_i L_i^{\eta})^{1/\eta}, \textit{ with } \textstyle\sum_i \gamma_i = 1 \\
\textit{and } \quad & L_i = (\textstyle\sum_j \delta_{ij} L_{ij}^{\theta})^{1/\theta}, \textit{ with } \textstyle\sum_j \delta_{ij} = 1 \\
& \textit{i: educational groups; j: age subgroups.}
\end{aligned} \tag{6.2}$$

Furthermore, for each of the specifications above, I consider two of its variants in which (i) all parameters in the production functions keep the same for all provinces, and (ii) the hicksian technical factor τ is allowed to vary across provinces.[59]

6.1.2 NLS Estimations

After specifying the data, inputs and outputs, and agricultural technology, I present key results of the Nonlinear Least Square (NLS) regressions. Luckily, these regressions provide relatively clear-cut evidences to give preliminary though still preliminary answers to two empirical questions focused by this section and thus to help us to discriminate three theoretical settings discussed in Chapters 3, 4 and 5.

---

[57] See particularly, Card and Lemieux (2001), Borjas (2003).

[58] Alternatively, the aggregate labor input can be firstly decomposed as two age groups. Each of these age groups can be then decomposed as two educational subgroups. Although this alternative three-level nested CES can be rarely seen in the literature, I see no a priori reason why this specification must be inferior to the one used in the main text. Given the data at hand, results of estimations based on this alternative specification resemble to those based on the two-level nested CES specification. Thus, in the text I do not report the results based on it.

[59] As experiments, I estimated using a more flexible variant of the specification in which α_A, α_M, α_{QL} can vary across provinces. Answers to the two central empirical questions are largely unaffected. Given the data and specification, it turns out difficult to allow further parameters in the production functions to vary across provinces.

For the first empirical question, namely whether different types of labor inputs are perfect substitutes in household agricultural production, estimations suggest that it is likely to be the case. In all regressions presented in Table 6.1, the original hypotheses that the elasticity-of-substitution-related parameters, i.e., $\{\rho, \eta, \theta\}$ take the value of 1 cannot be rejected at 5% significance level, which means that the hypotheses that different labor inputs are perfect substitutes cannot be rejected.[60]

Year	Two-Level		Three-Level	
	i	ii	i	ii
1991	$\rho=0.90^{*\#}$	$\rho=0.81^{*\#}$	$\eta=0.82^{*\#}$; $\theta=1.21^{*\#}$	N.A.
1993	$\rho=1.49^{\#}$	$\rho=1.62^{\#}$	$\eta=1.46^{\#}$; $\theta=1.52^{\#}$	N.A.
1997	$\rho=0.73^{*\#}$	$\rho=0.74^{\#}$	$\eta=0.66^{*\#}$; $\theta=1.00\#$	N.A.
2000	$\rho=0.98^{*\#}$	$\rho=1.00^{*\#}$	$\eta=0.94^{*\#}$; $\theta=1.12^{*\#}$	N.A.
2004	$\rho=1.28^{*\#}$	$\rho=1.31^{\#}$	$\eta=13.98^{\#}$; $\theta=0.68^{\#}$	N.A.
2006	$\rho=0.84^{*\#}$	$\rho=0.90^{*\#}$	$\eta=0.98^{*\#}$; $\theta=0.62^{\#}$	N.A.
2009	$\rho=0.95^{*\#}$	$\rho=0.96^{*\#}$	$\eta=0.91^{*\#}$; $\theta=1.51^{\#}$	N.A.

Table 6.1: NLS Estimates of the Elasticity of Substitution

Note that this table reports only estimates of the elasticity-of-substitution-related parameters. Estimates of other parameters are not reported for the reason of space. [1] Variant i: The same parameters for all provinces; Variant ii: Allowing for τ varying across provinces. [2] Dependent variable: Log agricultural gross income in 2009 RMB. [3] Independent variables: Skill-specific Labor in head counts; Land in Mu (1 Mu=6.67 Ares); Material approximated by the production costs in 2009 RMB. [4] Standard errors are robust and clustered at community level. * indicates significance at 5% level. [5] #: H_0: $\rho=1$ ($\eta=1/\theta=1$) cannot be rejected at 5% significance level. [6] N.A.: Result is not available owing to difficulties in the numerical optimizations.

Next, I turn to the second empirical question. Given the fact that the hypothesis that different types of labor inputs are perfect substitutes in household agricultural production cannot be rejected, I wonder to know whether and to what extent they are heterogeneous in agricultural productivity. To improve the precision of estimation, a

[60] Inspection of Table 6.1 shows also that the estimates of $\{\rho, \eta, \theta\}$ may not be very precise. Thus, I could make sometimes the type II errors. Nevertheless, I find that all estimates of these parameters tend to cluster around unity.

simpler functional form of the agricultural technology is adopted hereafter:

c. Heterogeneous and Perfectly Substitutive

$$\begin{aligned} &\log y = \tau + \alpha_A \log A + \alpha_M \log M + \alpha_{QL} \log QL + u \\ &where\ QL = \sum\nolimits_1^4 \beta_s L_s,\ with\ \sum\nolimits_1^4 \beta_s = 1 \end{aligned} \tag{6.3}$$

Note that the functional form above relates closely to that used in Chapter 4. Suppose different types of labor inputs are homogeneous in intrinsic agricultural productivity, i.e. $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 1/4$, it reduces to a functional form similar to that used in Chapter 3.

Running the NLS regressions with functional form given by equation (6.3) produce a new set of estimates on $\beta_1 \sim \beta_4$, among others. A summary of the key results obtained from these regressions is given in Table 6.2 below.

By inspection of Table 6.2, it is found that the coefficients of $\beta_1 \sim \beta_4$ are usually different from each other, but the differences tend to be small. To answer the question of interest formally, the Wald tests are used. The outputs of the Wald tests show that the original hypothesis H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4$ must be rejected at the 5% significance level in the CHNS 2000 and 2009, and it must be rejected at 10% significance level in the CHNS 2006. By contrast, for the rest surveys, especially for those in the early 1990s, the original hypothesis cannot be rejected at conventional significance levels. Therefore, for the whole period considered, equation (6.3) provides a unified specification for estimating of agricultural technology adopted by household farms.

Parameter	Year	Heterogeneous and Perfect Substitutive	
		i	ii
β_1	1991	0.27*	0.24*
	1993	0.14	0.11
	1997	0.20*	0.26*
	2000	0.23*	0.24*
	2004	0.51*	0.46*
	2006	0.27*	0.30*
	2009	0.29*	0.28*
β_2	1991	0.26*	0.28*
	1993	0.17	0.16
	1997	0.23*	0.22*
	2000	0.23*	0.22*
	2004	0.12	0.13
	2006	0.21*	0.19*
	2009	0.28*	0.26*
β_3	1991	0.20*	0.19*
	1993	0.18*	0.20*
	1997	0.21*	0.22*
	2000	0.16*	0.18*
	2004	0.15	0.19
	2006	0.29*	0.30*
	2009	0.12*	0.12*
β_4	1991	0.27*	0.30*
	1993	0.52*	0.53*
	1997	0.35*	0.30*
	2000	0.38*	0.36*
	2004	0.22	0.22*
	2006	0.23*	0.21*
	2009	0.31*	0.33*

Table 6.2: Estimates on the Labor Heterogeneity

Note that β_4 is calculated using β_4=1-β_1-β_2-β_3. Standard errors of β_4 are obtained by using Stata's *lincom* command. * indicates significance at 5% level. Like Table 6.1, for the variant i, all parameters are the same across provinces, while for the variant ii, parameter τ varies across provinces.

In sum, Tables 6.1 and 6.2 demonstrate that (1) different labor inputs are likely to be perfect substitutes in household agricultural production and (2) the productivity heterogeneity among agricultural labor inputs becomes increasingly important over time. Therefore, based on the NLS estimations, I draw a preliminary conclusion that the theoretical setting analyzed in Chapter 4 in which agricultural labor inputs are heterogeneous and perfectly substitutive could be of the highest relevance to the CHNS data.

6.1.3 Endogeneity and GMM

In all of the regressions above, I pay no attention to the fact that the agricultural inputs, especially the labor and material inputs could be endogenous in the sense that (1) they could correlate with the unobserved productivity of household farms, and (2) they could correlate with shocks anticipated by households.[61] It is well known that in the presence of endogenous variables, the NLS estimator is inconsistent. In other words, the endogeneity problem challenges the validity of previous NLS regressions and thus the validity of the conclusion drawn from them. As a response, this subsection sets out to re-estimate the production functions using the GMM in the hope of mitigating the endogeneity problem.

A prerequisite of applying the GMM is to find enough valid instrumental variables. For the GMM estimations below, I propose to use the log household labor forces at all four skill levels, their community averages, the one-period lag of log material input, together with the log land input as instrumental variables.[62]

Given the data and nonlinear specifications of the form equations (6.1) and (6.2), GMM estimations cannot be used to answer the first empirical question. Hence, this

---

[61] The former is often called as the omitted variable bias, while the latter is related to the simultaneity bias in the econometric literature.

[62] The numbers of household labor forces at four skill levels are likely to be valid instruments because (1) they are largely determined by households' past decisions on fertility and human capital investment, and (2) in the presence of imperfect rural labor market, households' agricultural labor inputs depend on their labor endowments. For similar reasons, community average of households' labor forces at four skill levels could serve as valid instruments as well. The validity of lagged log material input is questionable: It may not correlate with current anticipated shocks, but if the unobserved farm productivity is time-invariant or highly persistent, then the lagged log material input could correlate with unobserved productivity in the error term.

subsection addresses only the second empirical question using the GMM. In all regressions below, household agricultural production functions take the form of equation (6.3). Moreover, to reduce the difficulties in numerical optimization, the production function is assumed to be the same across provinces in each year. Table 6.3 compares the NLS and GMM estimates of $\beta_1 \sim \beta_4$.[63]

Year	Parameter	NLS	GMM
1991	β_1	0.26*	0.23*
	β_2	0.26*	0.26*
	β_3	0.20*	0.20*
	β_4	0.27*	0.31
1993	β_1	0.13	0.07
	β_2	0.17	0.25*
	β_3	0.19*	0.29*
	β_4	0.52*	0.39
1997	β_1	0.20*	0.19*
	β_2	0.23*	0.27*
	β_3	0.21*	0.21*
	β_4	0.35*	0.32*
2000	β_1	0.23*	0.26*
	β_2	0.23*	0.28*
	β_3	0.17*	0.14*
	β_4	0.38*	0.32*
2004	β_1	0.51*	0.25*
	β_2	0.12	0.21*
	β_3	0.15	0.22
	β_4	0.22	0.32*
2006	β_1	0.27*	0.19*
	β_2	0.21*	0.20*
	β_3	0.29*	0.39*
	β_4	0.23*	0.23*
2009	β_1	0.29*	0.14
	β_2	0.29*	0.20
	β_3	0.12*	0.26
	β_4	0.31*	0.40

Table 6.3: Comparison of the NLS and GMM Estimates

[63] The sample used in the NLS and GMM estimations in Table 6.3 differs from that used in Tables 6.1 and 6.2 in a minor aspect, namely in regressions reported in Table 6.3, I exclude the observations of agricultural workers whose annual working hours exceed 4000 hours. This difference could explain the small and often ignorable differences between results of the NLS estimations in Tables 6.2 and 6.3.

Note that for the NLS estimations above, the standard errors are robust and clustered at community level. For the GMM estimations above, the weight matrix is adjusted for correlation within the community. * indicates statistical significance at 5% level.

According to the outputs of Hansen's *J* test of overidentifying, it is found that in all surveys except the CHNS 2000, the original hypothesis of the test stating that all instrumental variables are valid cannot be rejected at 5% significance level, which suggests in absence from other complications, the GMM regressions offer consistent estimates. The Wald test is employed to determine whether different types of labor inputs are heterogeneous. The results suggest that in later surveys, especially in the CHNS 2000 and 2006, there are weak evidences to support the existence of labor heterogeneity in agriculture. Therefore, GMM estimations confirm largely the preliminary conclusion based on NLS estimations.

6.2 The Pattern of Migration Selection

Section 6.1 suggests the setting in Chapter 4 has the highest empirical relevance. One of the implications of the extended Roy model based on that setting are that there exist some approximate relationships between the pattern of migration selection and change of rural inequality. Such relationships can be used to predict the direction of change of rural inequality if the pattern of migration selection can be known in advance. Therefore, the present section attempts to identify the pattern of migration selection from the CHNS data. As byproducts, I obtain predicted log shadow wages faced by rural workers when they were employed on farms and in urban non-agricultural sector. These predicted shadow wages will be useful in next section.

6.2.1 Descriptive Statistics

To get some intuitive feeling on the pattern of migration selection, I present firstly descriptive statistics about characteristics of the migratory and staying rural workers.[64] Preliminary results on the pattern of migration selection can be obtained by comparing these characteristics of both subgroups.

Numerous studies have shown that migratory workers do not make up a random sample of the working age population from source regions. As demonstrated in Table 6.4 below, migratory and staying rural workers differ in several observable individual-specific characteristics as well as in their household backgrounds.[65]

Year	Characteristic	Migratory	Staying
1991	schooling	7.47	5.21
	age	27.19	37.42
	female	0.52	0.52
	schooling (household head)	4.97	5.55
	age (household head)	48.03	44.18
	female (household head)	0.16	0.10
	lagged (*A*/*L*)	2.21	2.09
1993	schooling	7.20	5.40
	age	26.42	38.50
	female	0.50	0.51
	schooling (household head)	5.14	5.67
	age (household head)	49.28	45.78
	female (household head)	0.15	0.10
	lagged (*A*/*L*)	1.88	2.03
1997	schooling	7.62	5.68
	age	25.73	39.47
	female	0.55	0.50
	schooling (household head)	5.06	5.86
	age (household head)	51.30	47.31
	female (household head)	0.15	0.11
	lagged (*A*/*L*)	1.90	2.23
2000	schooling	8.49	6.18
	age	26.38	40.83
	female	0.52	0.50
	schooling (household head)	6.09	6.39
	age (household head)	50.01	48.41
	female (household head)	0.10	0.09
	lagged (*A*/*L*)	2.58	3.19

[64] The concrete definitions of the migratory and staying rural workers can be found in Appendix D.

[65] The sample used to generate the descriptive statistics is restricted to village residents (1) who are aged between 15 and 75 and do not enroll at school at the time of survey and (2) who are members of households who are engaged in agricultural production.

(Continuation)

Year	Characteristic	Migratory	Staying
2004	schooling	8.61	6.50
	age	27.38	43.72
	female	0.49	0.50
	schooling (household head)	6.75	6.89
	age (household head)	51.52	50.69
	female (household head)	0.10	0.08
	lagged (*A*/*L*)	2.40	3.14
2006	schooling	8.45	6.51
	age	29.92	43.79
	female	0.57	0.51
	schooling (household head)	6.01	6.54
	age (household head)	55.48	52.83
	female (household head)	0.12	0.10
	lagged (*A*/*L*)	2.52	3.72
2009	schooling	8.56	6.66
	age	31.99	46.01
	female	0.59	0.50
	schooling (household head)	6.22	6.89
	age (household head)	57.42	54.21
	female (household head)	0.12	0.08
	lagged (*A*/*L*)	3.40	4.13

Table 6.4: Migration Selection on Observable Characteristics

Note that "schooling" denotes the years of completed formal education. "Female" is a dummy variable. "Lagged (*A*/*L*)" denotes the one-period-lagged land-labor ratio of household farms.

According to Table 6.4, migratory rural workers are on average more than 10 years younger than staying rural workers are. The average education of migratory rural workers is also higher than that of staying rural workers.[66] The difference between years of education of both subgroups is about two years.[67] These findings are roughly consistent with existing empirical studies. Besides, the sex ratios in both

[66] Among all young migratory rural workers, those who have exact junior high school diploma take always the highest share. On average, the share for the skill cell "JHS×YG" is about 40 - 50% among migratory rural workers. By contrast, the share for the same skill cell is only about 20% among staying rural workers.

[67] Considering that the standard deviations of years of schooling for migratory and staying rural workers are about 3 and 4, a two-year-difference could be large. However, we cannot conclude that out-migration is strongly positively sorting in terms of education. In fact, migratory rural workers are also much younger than staying rural workers are. Since the educational attainment rises rapidly across age cohorts in past few decades in China, I suspect that a substantial fraction of the educational differential should be attributed to the cohort effect. To control for the cohort effect, years of schooling should be compared within the age cohorts. In particular, given that most out-migrations occur before 35 years old, I pay special attention on that cohort. It turns out that after controlling for the cohort effect, migratory rural workers have still higher education than staying rural workers do. Yet, the educational differential between two subgroups within that cohort becomes smaller: On average, the educational differential within that cohort is about 1/3 as large as the overall differential.

subgroups are close to balanced.

As for the household backgrounds of both worker subgroups, migratory rural workers tend to come from relatively disadvantageous households: For instance, in all surveys, heads of migratory rural workers' household are slightly older, less educated than heads of staying rural workers' households.[68] The percentage of female-headed households is usually higher for migratory rural workers than staying rural workers. Furthermore, in most surveys, household farms from which migratory rural workers come have smaller ex ante land-labor ratios.[69] The shortage of cultivated land of household farms could push rural workers to search for off-farm employments.

6.2.2 Migration Selection in terms of Log Agricultural Wages

To determine the pattern of migration selection with higher precision, I return to its definition given in Chapter 2.[70,71] By definition, the pattern is jointly determined by the relative sizes of average log wages earned by migratory and staying rural workers when they were employed in household farms and in urban sector. The present subsection focuses on comparison of average log agricultural wages, while the next subsection focuses on comparison of average log urban wages.

---

[68] On average, the former are about 3 years older than the latter are. The years of schooling for the former are about 0.5 year shorter than the latter are.

[69] The land-labor ratios reported in Table 6.4 have not adjusted for labor heterogeneity. Moreover, all land-labor ratios in the table are the one-period lags for the reason that a resident's current migratory status is usually influenced by the ex ante ratio of his household farm. Unlike other characteristics in the table, out-migration is likely to have direct impacts on the ratio of household farm. Therefore, the one-period lags of the ratios are used here as proxies for the ex ante ratios.

[70] It is not uncommon that the patterns predicted by different characteristics disagree. To solve the problem, we could map all of these characteristics onto the real line and then evaluate the pattern based on the real index. A natural choice for such an index is the (log) wage itself. This observation provides a rationale for returning to the form definitions.

[71] See also Appendix D.

To begin with, let us consider the average log agricultural wage of staying rural workers, i.e. $\overline{\log w_0(I=0)}$, which is the sample counterpart of $E(\log w_0 \mid I=0)$.

As assumed in Chapter 4, all rural households of interest employ the same agricultural technology in the form of

$$y = \tau NL^{\alpha_{NL}} QL^{\alpha_{QL}}, \tag{6.4}$$

where *NL* and *QL* denote the aggregate non-labor and labor inputs in agriculture. Since different types of labor inputs are by assumption heterogeneous and perfectly substitutive, the aggregate household labor input in agriculture is given by

$$QL = \sum_s \beta(s)L(s) = \sum_s \beta(s)\sum_i L(s;i). \tag{6.5}$$

Applying the chain rule and taking logarithm yield the log agricultural wage that any staying rural worker whose skill level is s could earn as follows

$$\log w_0 = \log \beta(s) + \log \partial y / \partial QL.$$

Therefore, we obtain

$$E(\log w_0 \mid I=0) = E(\log \beta(s) \mid I=0) + E(\log \partial y / \partial QL \mid I=0),$$

and its sample counterpart

$$\overline{\log w_0(I=0)} = \sum_s \lambda(s;I=0)\log \beta(s) + \sum_i \log \partial y / \partial QL(I=0), \tag{6.6}$$

where $\lambda(s;I=0)$ denotes the share of s-type workers in all staying rural workers.[72] Thus, the first right-hand side term in equation (6.6) refers to the weighted average of

[72] See Appendix D, Table D.1.

log individual-specific agricultural productivity among staying rural workers. The second right-hand side term refers to the average of log agricultural productivity relating to their household farms' production conditions.

With modifications, equation (6.6) can be adapted for evaluating the average log agricultural wage earned by rural migratory workers, i.e. $\overline{\log w_0(I=1)}$. However, I need to address two difficult issues beforehand, both of which relate to the issues of counterfactual inference. The first difficulty lies in the fact that for migratory rural workers, especially for those who have withdrawn completely from agriculture, their agricultural wages cannot be observed. The second difficulty lies in the fact that to infer the agricultural wages earned by migratory rural workers, we need to think of a hypothetical scenario in which migratory rural workers would stay on-farm. The counterfactual stay of migratory workers could affect the agricultural wages earned by staying rural workers or even migration decisions of household members at the margin of migration and stay.

As a partial solution to the first issue, the exploration below employs a largely untested assumption to help the identification, namely within each broadly defined education-age cells, migratory and staying rural workers have the same individual-specific agricultural productivity. Under this assumption, the estimates of $\{\beta(s)\}$ obtained in Section 6.1 can be used to evaluate the log agricultural wages of both subgroups. To avoid the complication caused by the second issue, one needs to calculate the log agricultural wages of both subgroups on a comparable basis. Since in the CHNS, a large fraction of migrants have only recently moved to urban regions, I use the one-period-lags of households' agricultural inputs $X_{t-1} \equiv \{A_{t-1}, M_{t-1}, QL_{t-1}\}$, and current technical parameters $\Theta_t \equiv \{\tau_t, \alpha_{A,t}, \alpha_{M,t}, \alpha_{QL,t}, \{\beta_t(s)\}\}$ to evaluate the log agricultural wages of both subgroups.

By comparing the average log agricultural wages of migratory and staying rural workers evaluated on the common basis of X_{t-1} and Θ_t, I obtain the pattern of migration selection in terms of average log agricultural wages.

Year	NLS			GMM		
	Migratory	Staying	Δ	Migratory	Staying	Δ
1991	6.29	6.35	-0.06	6.27	6.35	-0.08
1993	5.45	5.51	-0.06	5.87	6.03	-0.16
1997	6.41	6.53	-0.12	6.24	6.44	-0.20
2000	6.84	7.03	-0.19	6.90	7.21	-0.31
2004	6.04	5.90	0.14	6.84	6.92	-0.08
2006	7.08	6.99	0.09	8.39	8.19	0.20
2009	7.41	7.71	-0.30	7.63	7.59	0.04

Table 6.5: Migration Selection in terms of the Log Agricultural Wages

Note that the difference of average log agricultural wages faced by migratory and staying rural workers is defined by $\Delta \equiv \Sigma_s \lambda(s; I=1)\log\beta(s) - \Sigma_s \lambda(s; I=0)\log\beta(s)$. Δ denotes the growth rate. In the left panel of Table 6.5, the log agricultural wages are evaluated using the technical factors Θ_t obtained from the NLS regressions in Section 6.1, whereas in the right panel of the table, the log agricultural wages are evaluated using the technical factors Θ_t obtained from the GMM regressions in Section 6.1.

Inspecting of Table 6.5 shows that the difference of average log agricultural wages faced by both rural worker groups denoting by Δ changes its sign from negative to positive around the early or middle 2000s, which means on average migratory workers may have caught up with and even overtaken staying workers in terms of the agricultural wages. Further analysis suggests that the overtaking is driven primarily by the increase in individual-specific agricultural productivity of migratory workers comparing with their staying counterparts.

6.2.3 Migration Selection in terms of Log Urban Wages

According to the definition of pattern of migration selection, I proceed to determine the pattern in terms of log urban non-agricultural wages.

This task is challenging for two reasons: The log urban wages faced by staying rural workers are unobservable; More importantly, partly owing to difficulties in data collection, the CHNS offer limited information on employment and wages of migratory rural workers.[73] The limited data availability rules out the possibility of using existing structural approaches to determine the pattern more rigorously.[74]

As a preliminary attempt, I propose to predict the log urban wages faced by both migratory and staying rural workers using the earning functions of urban blue-collar workers.[75] This proposition is not self-evident, some justifications are thus offered for it below.

Migration has usually profound impacts on the destination: It causes often adjustments in multiple good and factor markets. Under certain circumstances, it affects the wage determination at the source and destination. Various effects of migration on the wage determination are called as the general equilibrium effects.[76] Unfortunately, in the literature that uses the Roy model as analytical framework, the general equilibrium effects are difficult to be incorporated and are thus usually ignored for simplicity. I follow the literature hereafter by treating the wage

---

[73] Up to the CHNS 2000, only about 30% of working age migrants report positive wages. From the CHNS 2004 onwards, owing to changes of the questionnaire design, virtually no working age migrant reports positive wage. Instead, only the remittances are reported. The empirical study below shows that the reported and predicted log hourly wages earned by rural migrants are usually similar. The former is only about 10% lower than the latter. The correlation coefficients between rural migrants' reported and predicted log hourly wages are also reasonably high. They are often at the level around 0.6.

[74] See for example the empirical methods surveyed by French and Taber (2011).

[75] Blue-collar workers include skilled and unskilled labor, service workers, drivers and athlete, since the earning functions of workers with these occupations differ often significantly from those with professional occupations, while their earning functions resemble to that of unskilled labor.

[76] The general equilibrium effects are ignorable if (1) the size of the newly arrived rural migratory workers is small relative to that of the urban native workers, or (2) urban industries can expand or contract freely to absorb the labor supply shocks, and the good markets adjust accordingly. Notably, if the condition (2) is the case, then the general equilibrium effects disappear entirely. This is known as the Rybczynski theorem. For an engaging discussion on the theorem and its application in the immigration context, reader may refer to Leamer (2009).

determination in urban sector as fixed.

In practice, the term "wage determination" is often understood as an equivalent to the earning function that captures the relationship between wage and various skill components. Thus, I treat the earning function of urban workers also as fixed.

The reasoning above does not guarantee that we can use the earning functions of all urban workers to predict the log urban wages faced by both migratory and staying rural workers. This is largely because in many developing countries like China, the phenomenon "occupational segmentation" exists to varying extent in the urban labor market. Migratory rural workers and native urban workers tend to take different occupations.[77,78] As for the Chinese context, a plenty of evidences show that migratory rural workers have a much higher possibility of being employed as blue-collar workers than urban workers do.[79] Furthermore, as suggested by the descriptive statistics in Table 6.4, staying rural workers are on average less educated and older than migratory rural workers are, thus comparing to their migratory counterparts, staying rural workers would be even more likely to be employed as blue-collar workers if they could work in the urban sector. The empirical exploration below will make use of this key observation that rural workers as a whole tend to take blue-collar urban occupations to give reasonable predictions about the log urban wages faced by all rural workers.

---

[77] In the tradition of development economics, explicit theoretical modeling of the occupational segmentation can be traced back at least to Todaro (1969).

[78] Of course, to conduct study along this line, there must be some overlaps between occupations of both subgroups.

[79] Based on a survey in 1995, Du (1997) reports that the majority of rural migrants are employed in a small number of urban industries such as construction and service. Meng and Zhang (2001) find that in the middle 1990s, only 4% of rural migrants in Shanghai are white-collar workers with professional, technical or managerial occupations, while 35% of urban residents are white-collar workers. Based on the China 1% National Population Sample Survey, 2005, Xing (2008) shows that most rural migrants work in urban manufacturing and service industries, but the industrial and occupational distributions of urban residents are far more disperse.

To justify the proposition, I need to argue further that within blue-collar occupations, the urban wages are primarily determined by their observable skill components instead of their residency so that the earning functions of blue-collar urban workers can be applied to predict the log urban wages faced by all rural workers. Related empirical evidences are somewhat mixed: Meng and Zhang (2001) find more than 100 percent of total wage differentials between rural migrants and urban residents are the "unexplained intra-occupational differentials" or the discrimination within occupations.[80] To the contrary, Xing (2008) finds that about 90% of the hourly wage differentials between rural migrants and urban residents can be explained by skill differences (such as schooling, age and occupation). Altogether, the evidences above suggest that the predictions based on the proposition above offers upper limits for the log urban wages faced by all rural workers. The predictions would resemble more closely to their true values when the empirical findings in Xing (2008) could be of higher relevance to the reality.

For the rest of this subsection, I apply the method proposed above to determining the pattern of migration selection in terms of log urban wages. Like many empirical studies, the earning function of urban blue-collar workers has the Mincerian specification (Mincer, 1974).[81] In keeping with the discussion in Section 6.2.2, the dependent variable of the earning function is the log annual wage in the currency value 2009. The explanatory variables always include a linear term of schooling, a quadratic term of working experience and a gender dummy. In addition, the functions

[80] I have no intention to deny the existence of discrimination in the workplace. Nevertheless, Meng and Zhang (2001) may exaggerate its importance: In inequality decomposition exercises, shares larger than 100% could often imply that some mistakes occur. In addition, it makes little sense to attribute all of the unexplained intra-occupational wage differentials to the discrimination. The unexplained wage differentials may only measure our ignorance about the wage determination.

[81] Readers may see Card (1999) and Lemieux (2006) for interesting and sympathetic reappraisals on the Mincerian function.

may or may not include provincial dummies, depending on the concrete specification. As a baseline, the outputs of the Ordinary Least Square (OLS) are presented in the first and second columns of Table 6.6. Moreover, to mitigate the ability biases in the OLS estimations, regressions in the third and fourth columns include parents' education as additional controls.[82,83]

Year	Variable	(1)	(2)	(3)	(4)
1991	Schooling	0.02	0.02*	0.02	0.02
	Exper	0.04*	0.04*	0.03*	0.04*
	$Exper^2/100$	-0.05*	-0.05*	-0.05*	-0.05*
	Female	-0.19*	-0.20*	-0.20*	-0.20*
1993	Schooling	0.02	0.02	0.01	0.02
	Exper	0.04*	0.04*	0.03*	0.02*
	$Exper^2/100$	-0.05*	-0.05*	-0.04*	-0.03
	Female	-0.15*	-0.15*	-0.15*	-0.16*
1997	Schooling	-0.01	-0.01	-0.02	-0.01
	Exper	0.03*	0.04*	0.03*	0.03*
	$Exper^2/100$	-0.07*	-0.07*	-0.07*	-0.08*
	Female	-0.26*	-0.26*	-0.28*	-0.28*
2000	Schooling	0.01	0.01	0.01	0.01
	Exper	0.02*	0.02*	0.02*	0.03*
	$Exper^2/100$	-0.04*	-0.04*	-0.05*	-0.05*
	Female	-0.18*	-0.18*	-0.17*	-0.18*
2004	Schooling	0.03	0.03*	0.03*	0.03*
	Exper	0.03*	0.03*	0.02	0.03*
	$Exper^2/100$	-0.04*	-0.05*	-0.03	-0.05*
	Female	-0.33*	-0.33*	-0.34*	-0.33*
2006	Schooling	0.05*	0.06*	0.05*	0.05*
	Exper	0.03*	0.03*	0.03*	0.03*
	$Exper^2/100$	-0.06*	-0.06*	-0.06*	-0.06*
	Female	-0.44*	-0.44*	-0.44*	-0.45*
2009	Schooling	0.04*	0.03*	0.04*	0.03*
	Exper	0.04*	0.04*	0.02	0.03*
	$Exper^2/100$	-0.07*	-0.08*	-0.05*	-0.07*
	Female	-0.32*	-0.34*	-0.35*	-0.37*

Table 6.6: Earning Functions of the Urban Blue-Collar Workers

Note the specification (1) does not include provincial dummies and father's education; the specification (2) includes provincial dummies, but it excludes father's education; the specification (3) excludes provincial dummies, but includes father's education; the specification (4) includes both provincial

[82] The Instrumental Variable (IV) method is one of the leading methods used to mitigate the ability bias. However, after trying various instruments proposed in Card (1999), I find that most of these instruments are too weak to provide any reliable result for the data at hand. For example, schooling-proximity-related variables, such as the accessibility of schools to communities, distances between community and schools are proven to be weak instruments. The quarter of birth instrument advocated by Angrist and Krueger (1992) may suffer the weak instrument problem more severely in the Chinese context, mainly because the compulsory education law in China requires all students remain enrolled at school until they complete junior high school education, rather than until they are 16 years old.

[83] Missing parental education is imputed using the community level average whenever possible. Considering that the educational attainments vary across age cohorts, the community averages of parental education are calculated within each of the 10-year-age cohorts. In all of the regressions controlling for parental educations, imputation flags and interactions between parental educations and imputation flags are also included.

dummies and father's education. The standard errors are robust and clustered at community level. * indicates significance at 5% level. Father's education refers to father's schooling, imputation flags, and interaction between father's schooling and flags. I have also experimented to use mother's schooling or both parents' schooling as controls. Results are in general robust to these controls.

Among all regression outputs above, two regularities are of particular interest for further empirical explorations: (1) the rate of return to schooling rises abruptly in the period between 2000 and 2004. The rate before this short period is extremely low in size and is often not statistically different from zero, while the rate after this period stabilizes. (2) The explanatory power of the Mincerian earning functions is limited in the sense that the adjusted R^2 never exceeds 0.20. The average adjusted R^2 is around 0.10.[84] The low goodness of fit could cause downwards bias in estimating of the key parameter (σ_1/σ_0).

Given the earning functions of urban blue-collar workers, I can predict the log urban wages faced by all rural workers. By comparing the average predicted values of log urban wages faced by migratory and staying rural workers, I can finally determine the pattern of migration selection in terms of log urban wages.

Both panels of Table 6.7 below suggest that on average migratory rural workers have caught up with and then overtaken their staying counterpart in terms of log urban wages. However, the timing of overtaking cannot be uniquely determined and it depends on concrete specification in estimation. If the timing of overtaking is defined as the year when sign of Δ changes for the first time to non-negative, then for regressions without controlling for parents' education, the overtaking occurs around the middle 1990s, while for regressions controlling for parents' education, the overtaking occurs around 2000. In addition, the intensity of migration selection in terms of log urban wage is low in general. For instance, the most negative entry in Table 6.7 is only -0.13, whereas the most positive entry in the same table is 0.17.

[84] In the U.S. literature, the adjusted R^2 of the standard Mincerian earning function is often about 1/3.

	(1)			(2)		
Year	Migratory	Staying	Δ	Migratory	Staying	Δ
1991	7.95	8.05	-0.10	7.94	8.04	-0.10
1993	8.13	8.21	-0.08	8.13	8.19	-0.06
1997	8.53	8.52	0.01	8.47	8.46	0.01
2000	8.83	8.80	0.03	8.77	8.74	0.03
2004	9.03	9.00	0.03	8.96	8.92	0.04
2006	9.03	8.90	0.13	8.98	8.85	0.13
2009	9.34	9.18	0.16	8.33	8.16	0.17

Panel A: Before controlling for Fathers' Education

	(3)			(4)		
Year	Migratory	Staying	Δ	Migratory	Staying	Δ
1991	7.93	8.06	-0.13	7.91	8.04	-0.13
1993	8.13	8.26	-0.13	8.12	8.23	-0.11
1997	8.50	8.57	-0.07	8.44	8.50	-0.06
2000	8.82	8.82	0.00	8.76	8.77	-0.01
2004	9.04	9.06	-0.02	8.97	8.97	0.00
2006	9.08	8.94	0.14	9.03	8.90	0.13
2009	9.32	9.26	0.06	9.31	9.23	0.08

Panel B: After controlling for Father' Education

Table 6.7: Migration Selection in terms of Log Urban Wages

Note that the details on specifications (1) - (4) are given in the notes below Table 6.6. The definition of Δ is given in notes below Table 6.5.

Up to this point, I always maintain the assumption that all skill components are perfectly transferable across sectors. For skill components like working experience, however, this assumption may not be realistic. To be specific, experience collected from agricultural activities may depreciate in urban sector.[85] As a result, the conventional measure of working experience, i.e. years since the entry in the labor market after leaving school could be inaccurate in estimating of the log urban wages faced by rural workers. Because staying rural workers tend to be older, less educated, and thus more experienced in agriculture than migratory rural workers, previous estimates of log urban wages faced by staying rural workers could suffer more from inaccurate measurement of experience. To account for the imperfect transferability of experience, it is ideally to adjust the estimated log urban wages of staying rural workers and, to a lesser extent for those of rural migratory workers downwards.

[85] For a highly relevant discussion in the international migration context, reader may see Borjas (2003), Section VI.A.

Nevertheless, as long as the adjustments are not too big, the pattern of selection in terms of log urban wages would remain qualitatively unchanged, though the time point of overtaking could shift to an earlier year.

To conclude Section 6.2, we need to combine empirical evidences presented in Sections 6.2.2 and 6.2.3 to determine the overall pattern of migration selection. Based on these evidences, it is found that in the period 1991-2009, the pattern of migration selection in rural China may undergo two transitions. The overall pattern of migration selection in the early and middle 1990s is the negative sorting, because in this period, rural migratory workers tend to be outperformed on average by staying rural workers in both sectors. In the period around 2000, the overall pattern of migration selection changes to the non-hierarchical sorting, because in this period, migratory rural workers are still outperformed by staying rural workers in agriculture, but rural migratory workers start to outperform rural staying workers in the urban sector. In the middle and late 2000s, the overall pattern of migration selection changes to the positive sorting, because in this period, rural migratory workers tend to outperform staying rural workers in both sectors. Though the timing of transitions depend on concrete specification and cannot be determined precisely, these findings are likely to be qualitatively robust.

6.3 Testing Theory with the CHNS Data: Puzzle and Explanations

Based on the knowledge on the patterns of migration selection in three sub-periods mentioned above, the extended Roy model in Chapter 4 predicts that in the first and third sub-periods, out-migration tends to decrease and increase rural inequality, while in the second sub-period, the direction of impacts of out-migration on rural inequality is unclear. These theoretical predictions are exactly what this section aims at testing for using the CHNS data. The tests and associating discussions below demonstrate that the extended Roy model and CHNS data could disagree about the role of intersectoral wage gap in determining the rural inequality. I offer thus several explanations for the disagreement near the end of this section.

6.3.1 Empirically Testable Implications

Before testing the predictions, it is necessary to realize a tension between theory prediction and data: The extended Roy model in Chapter 4 and virtually all other theoretical models in this thesis are static or comparative static in the sense that the time factor does not enter them explicitly. By contrast, the data with which the theory is tested are collecting from a highly dynamic world. To facilitate appropriate empirical tests, we need to modify these theoretical predictions so that they are in principle empirically testable. To be concrete, as in Section 6.2, it is straightforward to determine the pattern of migration selection from the data. However, the empirical counterpart for another key theoretical concept, i.e. the change of rural inequality R_V remains unclear and thus requires additional efforts.

Further analysis makes clear that in order to preserve the relationships between the pattern of migration selection and the direction of change of rural inequality in Chapter 4, the empirical counterpart of R_V must take the form as follows

$$R_{V,t} \equiv Var(\log \widetilde{w} \mid \Theta_t) / Var(\log w_0 \mid \Theta_t), \ where \ \log \widetilde{w} \approx \max\{\log w_0, \log w_1\}. \quad (6.7)$$

The numerator in equation (6.7) measures the actual rural inequality at the time t, where the conditions Θ_t denote the vector of distributional parameters at that time. The denominator measures the rural inequality in a counterfactual scenario in which all rural workers were engaged full-time in agriculture even when the opportunities of urban employments were already available for them. Therefore, $R_{V,t}$ refers to the change of rural inequality in comparison with that in the counterfactual scenario.

6.3.2 Empirical Tests and Some Disagreements

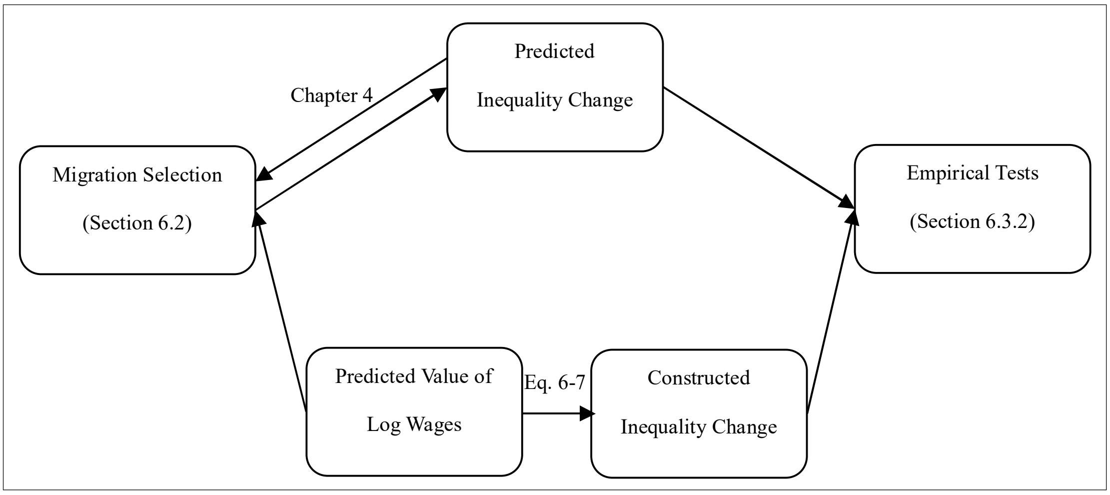


Figure 6.1: Roadmap of Empirical Tests

As preparation for the empirical tests, I construct measures of the change of rural inequality $R_{V,t}$ defined by equation (6.7) using the data of log shadow wages faced by rural workers in both sectors.[86] The signs of $R_{V,t}$ -1 suggest the directions of change of rural inequality: Positive sign indicates increase in rural inequality.

	(1)	(2)	(3)	(4)
1991	1.43	1.33	1.40	1.31
1993	2.09	1.41	2.06	1.39
1997	2.52	2.06	2.44	2.01
2000	2.17	1.68	2.15	1.67
2004	3.18	2.08	3.18	2.09
2006	2.58	0.97	2.68	1.01
2009	2.12	1.89	2.08	1.87

Table 6.8: Constructed $R_{V,t}$

Note that in constructing $R_{V,t}$ in the first column, the log agricultural wages are predicted using the specification in Table 6.5, the left panel; the log urban wages are predicted using the same empirical specification in Table 6.7, Panel A, the second column. For $R_{V,t}$ in the second column, the log

[86] The denominator of $R_{V,t}$ measures the counterfactual rural inequality when all rural workers would work full-time in agriculture, which provides a reference to the actual rural inequality in the same year. Tabulation of predicted log agricultural wages faced by rural workers does not show any clear trend in it. The quantile regressions that are robust to outliers confirm this finding.

agricultural wages are predicted using the specification in Table 6.5, the right panel; the log urban wages are predicted using the specification in Table 6.7, Panel A, the second column. For $R_{V,t}$ in the third column, the log agricultural wages are predicted using the specification in Table 6.5, the left panel; the log urban wages are predicted using the specification in Table 6.7, Panel B, the second column. For $R_{V,t}$ in the fourth column, the log agricultural wages are predicted using the specification in Table 6.5, the right panel; the log urban wages are predicted using the specification in Table 6.7, Panel B, the second column.

Unlike the theoretical predictions based on knowledge of pattern of selection, Table 6.8 points out that out-migration nearly always increases the rural inequality. Therefore, for the first sub-period in the early and middle 1990s, there is an obvious disagreement between theoretical predictions and empirical results in Table 6.8, since the pattern in that sub-period is negative sorting, which implies that rural inequality should decrease instead.

To understand the reasons why theory and data sometimes disagree, I decompose the difference between the numerator and denominator of $R_{V,t}$ using the law of total variance such that

$$\begin{aligned}
&\Delta Var(total) \\
&= \Pr(I=1)[Var(\log w_1 \mid I=1) - Var(\log w_0 \mid I=1] \\
&+ \Pr(I=0)[Var(\log w_0 \mid I=0) - Var(\log w_0 \mid I=0] \\
&+ \Pr(I=1)\Pr(I=0)\{[E(\log w_1 \mid I=1) - E(\log w_0 \mid I=0)]^2 \\
&-[E(\log w_0 \mid I=1) - E(\log w_0 \mid I=0)]^2\} \\
&\equiv Sh(m)\Delta Var(m) + 0 + \Delta Var(between).
\end{aligned} \tag{6.8}$$

The decomposition results under various specifications are given in Table 6.9.

	$\Delta Var(total)$	$Sh(m)\Delta Var\ (m)$	$Sh\ (s)\Delta Var\ (s)$	$\Delta Var\ (between)$
1991	0.168	-0.025	-0.002	0.196
1993	0.689	-0.073	-0.004	0.766
1997	0.400	-0.024	-0.003	0.427
2000	0.403	-0.035	-0.013	0.451
2004	1.507	-0.136	-0.010	1.653
2006	0.761	-0.124	-0.018	0.904
2009	0.460	-0.126	-0.018	0.604

Panel A: Specification (1)

	ΔVar (*total*)	Sh (m)ΔVar (m)	Sh (s)ΔVar (s)	ΔVar (*between*)
1991	0.162	-0.031	-0.002	0.195
1993	0.358	-0.129	-0.006	0.492
1997	0.425	-0.037	-0.005	0.467
2000	0.298	-0.043	-0.016	0.357
2004	0.630	-0.109	-0.009	0.747
2006	-0.014	-0.142	-0.018	0.145
2009	0.520	-0.170	-0.029	0.719

Panel B: Specification (2)

	ΔVar (*total*)	Sh (m)ΔVar (m)	Sh (s)ΔVar (s)	ΔVar (*between*)
1991	0.156	-0.026	0.000	0.182
1993	0.664	-0.073	-0.001	0.738
1997	0.380	-0.024	-0.003	0.407
2000	0.395	-0.035	-0.013	0.443
2004	1.511	-0.136	-0.009	1.656
2006	0.877	-0.126	-0.017	1.019
2009	0.445	-0.128	-0.017	0.590

Panel C: Specification (3)

	ΔVar (*total*)	Sh (m)ΔVar (m)	Sh (s)ΔVar (s)	ΔVar (*between*)
1991	0.151	-0.031	0.000	0.182
1993	0.343	-0.129	-0.002	0.474
1997	0.404	-0.037	-0.004	0.446
2000	0.291	-0.043	-0.016	0.350
2004	0.636	-0.108	-0.008	0.752
2006	0.006	-0.143	-0.016	0.166
2009	0.504	-0.172	-0.027	0.703

Panel D: Specification (4)

Table 6.9: Decomposition using the Law of Total Variance

The notations Sh (m) and Sh (s) denote shares of migratory and staying rural workers, ΔVar (m) and ΔVar (s) denote differences of variances within subgroups of migratory and staying rural workers, while ΔVar (*between*) denote the difference of between-group variance. Note that owing to missing values and rounding errors, values of Sh (s)ΔVar (s) in the table are not exactly zero.

Without exception, all decompositions in Table 6.9 attribute the majority of the increase in overall rural inequality to the increase in between-group variance captured by the large and positive $\Delta Var(between)$. Furthermore, considering that the intensities of migration selection in terms of log wages in both sectors are usually low, the expression of $\Delta Var(between)$ can be further approximated without much loss of the precision as follows

$$\Delta Var(between) \approx \Pr(I=1)\Pr(I=0)(\mu_1 - \mu_0)^2. \tag{6.9}$$

Therefore, the decompositions above identify the large intersectoral wage gap as the primary source of the increase in rural inequality reported in Table 6.8.

Nevertheless, as we already know,[87] the Roy models in general reject such a possibility entirely for the reason that in that model a large intersectoral wage gap tends to result in a large and instantaneous migration that suffices to narrow the wage gap. Thus, for the Roy models a large intersectoral wage gap (μ_1 / μ_0) cannot be a major determinant of the change of rural inequality $R_{V,t}$. This is exactly the point at which theory and data disagree.

By examining the data closely, however, there is no evidence that supports the presumption of the extended Roy model that large intersectoral wage gap does result in migration that is large enough to narrow the wage gap effectively. In fact, the shares of migratory workers among all rural workers in the CHNS data are often low. Even in the CHNS 2009, this share is still below 40%. As a result, the remaining intersectoral wage gap continues to affect the overall rural inequality.

To summarize, the discussions above demonstrate that the disagreement between theory and data could lie ultimately in an old puzzle in literature, namely given the large observed intersectoral wage gap, why so little migration occurs.

6.3.3 Possible Explanations

At the end of this section, I offer several possible explanations for the puzzle of missing intersectoral migration mentioned above and hence for the disagreement

[87] See particularly Figure 4.1.

between theory and data. These explanations are largely complimentary to each other.

The first group of explanations associate with data issues that could cause upwards biases in estimation of the intersectoral wage gap (μ_1 / μ_0).[88] Suppose that the true value of (μ_1 / μ_0) is significantly lower than its current estimate, the magnitude of observed migration can be legitimate. The over-estimates of (μ_1 / μ_0) could also contribute to the increasing rural inequality reported in Table 6.8.

Firstly, as argued before, because of the large amount of measurement errors in constructed working hours, current estimates of (μ_1 / μ_0) have to rely on predicted log annual wages of rural workers when they were employed in both sectors. Suppose that typical workers in both sectors differ significantly in their annual working hours, current estimates of (μ_1 / μ_0) would offer misleading measures for the intersectoral hourly wage gap, which relates more directly to the migration decision. Indeed, as demonstrated by Table 6.10 below, a typical urban employee in the CHNS data works per year about twice as long as a typical agricultural worker does.

[88] Other data issues could lead to imprecise estimates of parameter (σ_1 / σ_0). Particularly, given the low goodness of fit of the Mincerian regressions of urban blue-collar workers, inferring the true σ_1^* by using the standard deviation of the predicted log urban wages faced by rural workers σ_1 could lead to downwards bias. For example, suppose the adjusted $R^2 \approx R^2 \equiv SSE / SST = (\sigma_1 / \sigma_1^*)^2 \approx 0.1$ - the value that most adjusted R^2 take, the true standard deviation $\sigma_1^* \approx \sigma_1 / \sqrt{0.1} \approx 3\sigma_1$. In other words, if similar biases in inferring σ_0 is ignorable, values of (σ_1 / σ_0) need multiplying by a factor 3 to correct for the downwards bias caused by the low goodness of fit in Mincerian regressions. The corrected (σ_1 / σ_0) between 1991 and 2004 (except 1997) would be still smaller than or equal to 1. However, the values of true (σ_1 / σ_0) in years 2006 and 2009 would exceed 1. In both years, the values of true (σ_1 / σ_0) do not violate the sufficient and necessary conditions of the patterns of migration selection.

	Urban	Rural	Urban/Rural
1991	2455.48	1520.34	1.62
1993	2419.25	1291.87	1.87
1997	2189.58	1209.22	1.81
2000	2174.89	1023.47	2.13
2004	2227.63	1017.77	2.19
2006	2271.97	914.92	2.48
2009	2237.08	812.42	2.75

Table 6.10: Differences in Annual Working Hours

Note that the first column entitled "Urban" reports the average annual working hours of urban workers. The second column entitled "Rural" reports the average annual working hours of agricultural workers. The third column reports the ratios between annual working hours of urban and agricultural workers in each year. To avoid outliers, only working hours of employees who work less than 4000 hours per year are included in the tabulation.

Secondly, if discrimination plays a role in the wage determination within urban occupations, as asserted by some empirical literature, the predicted log urban wages faced by rural workers obtained from Section 6.2.3 offer only upper limits for their true values. Consequently, the current estimates of (μ_1 / μ_0) tend to be upwards biased as well.

After correcting for the upwards biases introduced by data issues, the new estimates of (μ_1 / μ_0) should be lower than their current values and thus the puzzle of missing migration could be in part resolved. Nevertheless, the corrected wage gap would remain substantial: Prior to any correction, the average level of urban wages faced by all rural workers (e^{μ_1}) is about eight to ten times as large as that of the agricultural wages faced by the same population (e^{μ_0}).[89] After accounting for the

[89] In Lewis (1954), he mentions once to use the Average Product of Labor (*APL*) in agriculture as the opportunity cost for rural-to-urban migration. In his words, "in economies where the majority of the people are peasant farmers, working on their own land, we have a more objective index, for the minimum at which labour can be had is now set by the average product of the farmer (pp. 148-9) ". However, Lewis does not work out the model. Hu (1994) argues that in the post-reform China where rural households are de facto owners of the land they cultivate and thus own (the majority of) the agricultural incomes attributable to both land and labor inputs, the *APL* is likely to be a superior measure for their members' incomes. *APL* could also be a superior measure for the opportunity cost, if rural migrants give up their rights to share the rentals with staying members. To see this point, assuming that the agricultural technology takes the Cobb-Douglas form such that $Y=TA(0)^{\alpha}L^{1-\alpha}$. We

difference of annual working hours between workers in both sectors, the ratio would be still around four to five. It is hard to believe that discrimination within urban occupations alone can explain the rest of intersectoral wage gap.

For the large corrected intersectoral wage gap, the puzzle of missing migration and thus the disagreement between theory and data continue to exist. Therefore, more efforts are required to explain them. In what follows, I provide three further possible explanations for them.

The first explanation emphasizes the role of large yet unobserved migration costs. A typical reaction of economists to the puzzle of missing migration would be to assume the existence of unobserved migration costs that prevents rural-to-urban migration in response to the large intersectoral wage gap, although the economists know little about the sources, magnitude and its functional form of the migration costs. In the presence of migration costs, the effective inter-sectoral wage gap faced by rural workers can be reduced. Hence, fewer rural workers find profitable to out-migrate. Furthermore, even if migration can narrow the effective intersectoral wage gap

---

obtain $APL=Y/L=MPL+MPA\times A(0)/L=MPL/(1-\alpha)$. Estimations on the agricultural value-added functions suggest that α – the share of land input in the total value-added is often smaller than 1/2. Hence, we have $MPL<APL<2MPL$. Particularly, when $\alpha=1/2$, $APL=2MPL$; when $\alpha=1/3$, $APL=1.5MPL$.

If I follow Hu's suggestion to use *APL* to measure the income of agricultural workers, then the intersectoral wage gap can be further narrowed. After adjusting for the sectoral difference in annual working hours and then using *APL* to measure the agricultural income, the average wage in the urban sector would be only about three times as large as the average *APL*. The corrected wage gap is roughly comparable to the ratio between the average disposable income of urban residents and average per capita net household income of rural residents routinely reported in the official statistics of China (see for example, China Compendium of Statistics, 1949-2008, Table 1-23, "Per Capita Annual Income, Expenditure and Engle Coefficient of Urban and Rural Households"). Note that the net household income is obtained by subtracting the monetary value of all material inputs from the gross agricultural outputs and thus consists of both labor income and land rental. Therefore, the official statistics provide some supports to my estimates on the annual shadow wages in both sectors.

Unfortunately, although I consider in the thesis the land ownership in China explicitly, similar to the standard Roy model, virtually all theoretical models in Chapters 3, 4 and 5 are still neo-classical. Thus, profit- or income- maximization in these models requires more or less equality between urban wage and agricultural *MPL*, rather than that between urban wage and agricultural *APL*. It would be thus difficult, but important for the economists who are interested in the Chinese labor markets to think about ways to incorporate the *APL* as a variable directly relating to the decision-making of maximizing households into the neo-classical models.

efficiently, it cannot narrow the observed wage gap owing to migration costs. Thus, the intersectoral wage gap owing to migration costs continues to affect the observed rural inequality.

The second explanation stresses the fact that migration may lag well behind the changes of the economic environment, since many activities relating to migration, for example, financing for the migration, transportation, searching jobs and finding lodging at the destination may be time-consuming for rural workers and households. Therefore, different from what the extended Roy model assumes, migration in the reality is unlikely to adjust instantaneously in response to the changes of intersectoral wage gap (μ_1 / μ_0). The existence of convex adjustment costs would further slowdown the out-migration. Consequently, the intersectoral wage gap tends to affect the overall rural inequality in a long period. Under this explanation, the observed staying rural workers may also include those who find out-migration profitable but are not well prepared for out-migration.

The third explanation goes beyond the simple framework in which all households are assumed to be total-income-maximizers, which is maintained in all of the previous discussions. In a more general household utility-maximization framework, we may consider various non-pecuniary benefits and costs relating to migration that could affect the effective intersectoral wage gap perceived by rural workers and households. Notably, as documented in Table 6.10, a typical urban worker tends to work longer than a typical worker on household farms does. This finding implies that to be employed in the urban sector, rural workers have to sacrifice a large number of leisure and possibly, to accept more strict discipline in the workplace, both of which could lead to disutility and hence decrease in the perceived intersectoral wage gap.

More interestingly, by allowing for the income pooling between migratory and staying household members in household income-maximization framework to be imperfect, i.e. $0<\theta<1$, we may be able to resolve to a considerable extent the puzzle of missing migration and hence the disagreement between theory and data. If in making migration decisions, maximizing rural households anticipate and take the imperfectness of income pooling into account, the effective intersectoral wage gap faced by rural households should be adjusted downwards accordingly. Smaller effective inter-sectoral wage gap could help to explain the puzzle of missing migration and hence the disagreement between theory and data. Another obvious advantage of adopting this slightly modified household income-maximization framework to studies of intersectoral migration and its impacts is that there will be no need to assume the equalization of wages earned from both sectors. The modified framework requires only that at the margin the agricultural wage equals to a fraction (θ) of the given urban wage. Thanks to this advantage, this modified framework could have a large potential for the theoretical and empirical explorations of the intersectoral migration in developing countries and its impacts on the source and destination.

CHAPTER 7

CONCLUSIONS AND POLICY IMPLICATIONS

In the last chapter, I provide a brief summary on this thesis. Then I give some preliminary discussions on the policy implications of previous theoretical and empirical explorations.

In theoretical chapters of this thesis, namely Chapters 2 to 5, I extend the standard Roy model by combining several simple microeconomic models capturing decisions of income-maximizing rural households subjective to constraints of factor endowments with the statistical framework of the standard Roy model. The extended Roy models have thus more solid microfoundations than the standard Roy model does. In particular, the extended Roy models are capable of taking the intra-household economic linkages that are totally missing in the standard Roy model into account.

The extended Roy models do shed light on the source of the pre-migratory rural inequality. Nevertheless, the fundamental difference between the standard and extended Roy models lies in the fact that they associate with two interrelated yet different statistical problems: As illustrated in Chapter 3, the former associates with the problem of truncation, whereas the latter associate with the problem of censoring.

The difference in the statistical structure leads to different predictions about the impacts of out-migration on rural inequality. For example, in the setting where all labor inputs are homogeneous in household agricultural production, Chapter 3 shows clearly that although the extended Roy model predicts also a decreasing rural inequality, the predictions derived from both models differ quantitatively. In general, the extended Roy model predicts a smaller decrease in rural inequality than the standard Roy model does.

The difference in prediction becomes more apparent after introducing the labor heterogeneity. In the setting where labor inputs are heterogeneous and perfect substitutive in household agricultural production, Chapter 4 illustrates that out-migration could not only decrease, but also increase the rural inequality measured by the censored variance of log wages. Furthermore, by ignoring the effects that out-migration has on the productivity and wages of staying rural workers, it is found that there are non-causal relations between the pattern of migration selection and the direction of change of rural inequality. Particularly, out-migrations exhibiting positive sorting tend to coexist with increase in rural inequality; out-migrations exhibiting the negative sorting tend to coexist with decrease in rural inequality; the direction of change of rural inequality is unclear if out-migrations exhibit non-hierarchical sorting.

Chapter 5 considers the impacts of out-migration on rural inequality in a more general setting where labor inputs are heterogeneous and imperfectly substitutive in household agricultural production. The discussion suggests that since one of the basic assumptions are violated in that setting, the Roy models in general stop to offer useful analytical framework for studying the impacts.

Chapter 6 presents empirical findings supplementing the theoretical chapters. A few of them should be of interest for broader researches as well.

Firstly, by estimating the production function of household farms, the results suggest that (1) the hypothesis that labor forces from different broadly defined education-age cells are perfect substitutes in household agricultural production cannot be rejected, and (2) heterogeneity in intrinsic productivity in agriculture gains its importance over time. These empirical findings together point out that the setting in

Chapter 4 as the most relevant setting.

Secondly, to my knowledge, Chapter 6 is the first study that reports the finding that the pattern of selection in the Chinese rural-to-urban migration is likely to undergo two transitions during the period 1991-2009. The dominating pattern of migration selection changes from the negative sorting to the non-hierarchical sorting around the late 1990s, and it changes again from the non-hierarchical sorting to the positive sorting around the middle 2000s. The transitions imply that during that period, rural migratory workers on average have caught up with and eventually overtaken their staying counterparts in terms of potential log wages in both sectors.

Because comparing to rural staying workers, the characteristics of migratory rural workers remain roughly stable, the overtaking of migratory rural workers should be chiefly explained by the changes of relative shadow prices of different skill components during the period.[90] Notably, evidences in Chapter 6 makes clearly that schooling plays an increasingly important role in determining the log wages faced by rural workers in both sectors. Moreover, even in agricultural sector, where working experience is highly valued for the production, the importance of schooling may increases relative to working experience.

Thirdly, for the first sub-period in the early and middle 1990s, the theoretical prediction based on knowledge of the pattern of migration selection could disagree with what data of predicted log shadow wages reveal. In that sub-period, the theory predicts a decrease in rural inequality (compared to the counterfactual inequality when

[90] Though there is no convincing evidence, it is still interesting to note that the timing of the two transitions, especially the second one coincides roughly with China's entry to the WTO on December 11, 2001. I suspect a freer access to the world economy owing to the entry to the WTO could response for the changes of the effective endowments of skills and other resources, the shadow prices and hence the pattern of migration selection.

all rural workers were engaged full-time in agriculture). By contrast, like in other sub-periods, the data suggest an increase in the rural inequality.

Further explorations attribute much of the disagreement between theory and data to the puzzle of missing migration, which means that facing with a large intersectoral wage gap, migration is often smaller than what theory predicts and thus cannot narrow the wage gap effectively. To understand and resolve the puzzle, I propose several possible explanations. Aside from the two data issues, the following three reasons could have large potentials as further explanations, which include (1) the existence of unobserved yet sizable migration costs, (2) the time lags between the changes of economic environments and migration, and (3) the imperfectness of the income pooling between migratory and staying rural workers. Nevertheless, all of these explanations are preliminary. There is much to be done to clarify them and to determine the empirical relevance of them to the puzzle of missing migration and thus to the disagreement between theory and data.

At the end of this thesis, I discuss the possible policy implications of previous theoretical and empirical chapters.

The theoretical chapters do not have rich policy implications. This is mainly because most of relationships found in these chapters, especially the one between the pattern of migration selection and the direction of change of the rural inequality in Chapter 4 are not causal. Therefore, there would be little role for public policy.

The empirical chapter, however, does have policy implications. For example, the discussion on the pattern of migration selection shows that on average rural migratory workers' productivities in agriculture overtake those of rural staying workers, which suggests that it makes no longer sense to categorize rural migratory workers as the surplus laborers in Lewis (1954). The out-migration since around the middle 2000s could have thus large negative impacts on the agricultural outputs in China. Therefore, to promote the growth of agriculture, rural households should have better accesses to the factor markets of non-labor inputs such as capital and material as well as to new agricultural technologies. The evidences highlight also the importance of schooling in determining the wages. Hence, public policies that improve educational attainments of rural workers would be highly desirable. Furthermore, Chapter 6 may suggest the existence of obstacles that keep rural workers from out-migration and integration in urban sector. Public policies that aim at removing these obstacles should to be implemented with courage and determination.

APPENDIX A

MATHMATICAL APPENDIX

In this appendix, I present mathematical derivations that are essential for a thorough understanding of the main text in Chapters 2 to 5.[91] This appendix is organized as follows: Firstly, I recall the basic assumptions required for most derivations below. Next, based on these assumptions, I derive in some detail (1) the sufficient and necessary conditions for different patterns of migration selection and (2) the moments of truncated and censored normal distributions. At the end of this appendix, I explain the way in which the reservation wage of out-migration is constructed in Chapter 5.

The basic assumptions are given in Section 2.1. In particular, they assert that the log sectoral wages faced by rural workers are jointly normal distributed, and that all rural workers are wage-income maximizers.

Under these assumptions, I follow the classical treatment in Borjas (1987) to define three patterns of migration selection according to the signs of two selection biases, $Q_0 \equiv E(\log w_0 \mid \log w_0 \leq \log w_1) - \mu_0$ and $Q_1 \equiv E(\log w_1 \mid \log w_0 \leq \log w_1) - \mu_1$. They measure differences in earning capacity between average rural out-migrants and the population of rural workers when they were employed in agricultural and non-agricultural sectors. By definition, the positive sorting occurs when both Q_0 and Q_1 are positive, the negative sorting occurs when both Q_0 and Q_1 are negative, while the non-hierarchical sorting occurs when Q_0 is negative and Q_1 is positive.

To determine the concrete conditions under which each of three patterns of

[91] In preparing the text and the Appendix A, I assume the readers have exposed to probability theory at the level roughly equivalent to Casella and Berger (2001).

migration selection occurs, I express Q_0 and Q_1 in terms of underlying distributional parameters. For example, Q_0 can be reformulated as

$$\begin{aligned} Q_0 &\equiv E(\log w_0 \mid \log w_0 \leq \log w_1) - \mu_0 \\ &= E[\varepsilon_0 \mid \varepsilon_1 - \varepsilon_0 \geq -(\mu_1 - \mu_0)]. \end{aligned}$$

Under the normality of ε_0 and $\varepsilon_1 - \varepsilon_0$, ε_0 can be decomposed as its projection on $\varepsilon_1 - \varepsilon_0$ denoting by $B(\varepsilon_1 - \varepsilon_0)$ and a residual υ that is orthogonal to $\varepsilon_1 - \varepsilon_0$, where B is the regression coefficient with $B \equiv Cov(\varepsilon_1 - \varepsilon_0, \varepsilon_0) / Var(\varepsilon_1 - \varepsilon_0)$.

Thus, the equation above becomes

$$\begin{aligned} Q_0 &= E[B(\varepsilon_1 - \varepsilon_0) + \upsilon \mid \varepsilon_1 - \varepsilon_0 \geq -(\mu_1 - \mu_0)] \\ &= BE[\varepsilon_1 - \varepsilon_0 \mid \varepsilon_1 - \varepsilon_0 \geq -(\mu_1 - \mu_0)]. \end{aligned}$$

Now applying Theorem 24.2 in Greene (2007) gives

$$Q_0 = B\sigma_{\Delta\varepsilon}\lambda(-c) = \sigma_0 \frac{\rho_{01}\sigma_1 - \sigma_0}{\sigma_{\Delta\varepsilon}} \lambda(-c) \equiv \sigma_0 \beta \lambda(-c),$$

where $\sigma_{\Delta\varepsilon}^2 \equiv Var(\varepsilon_1 - \varepsilon_0)$, c is the standardized intersectoral wage gap defined as $c \equiv (\mu_1 - \mu_0) / \sigma_{\Delta\varepsilon}$, $\lambda(.)$ is the inverse Mills ratio defined as $\lambda(.) \equiv \phi(.) / [1 - \Phi(.)]$.

Likewise, another selection bias Q_1 can be reformulated as

$$Q_1 = \sigma_1 \frac{\sigma_1 - \rho_{01}\sigma_0}{\sigma_{\Delta\varepsilon}} \lambda(-c) \equiv \sigma_1 \gamma \lambda(-c).$$

Since the inverse Mills ratio $\lambda(.)$ is always positive, the signs of Q_0 and Q_1 are determined only by the signs of β and γ. For instance, for the positive sorting, both β and γ must be positive, which implies $\sigma_0 / \sigma_1 < \rho_{01} < \sigma_1 / \sigma_0$. Considering that ρ_{01} is a correlation coefficient and thus cannot exceed 1, we have $\sigma_0 < \sigma_1$. Hence, the sufficient and necessary conditions for the positive sorting are given by equation (2.3) in the form of $\rho_{01} > \sigma_0 / \sigma_1$ and $\sigma_0 < \sigma_1$. Similarly, we obtain the conditions for the negative sorting and the non-hierarchical sorting given by equations (2.4) and (2.5).

In the standard Roy model, the post-migratory inequality in source sector is conventionally measured by the truncated variance $Var(\log w_0 \mid \log w_0 \geq \log w_1)$. Under the joint normality assumption, I prove equation (2.6) stating that the truncated variance must be smaller than pre-truncated variance below.

$$\begin{aligned}
& Var(\log w_0 \mid \log w_0 \geq \log w_1) \\
&= Var(\varepsilon_0 \mid \varepsilon_0 - \varepsilon_1 \geq \mu_1 - \mu_0) \\
&= Var[-B(\varepsilon_0 - \varepsilon_1) + \upsilon \mid \varepsilon_0 - \varepsilon_1 \geq \mu_1 - \mu_0] \\
&= B^2 Var[\varepsilon_0 - \varepsilon_1 \mid \varepsilon_0 - \varepsilon_1 \geq \mu_1 - \mu_0] + Var(\upsilon).
\end{aligned}$$

According to Theorem 24.2 in Greene (2007), the first right-hand side term in the equation above equals to $B^2 \sigma_{\Delta\varepsilon}^2 [1 - \delta(c)] = \sigma_0^2 \beta^2 [1 - \delta(c)]$ and the second right-hand term $Var(\upsilon) = Var[\varepsilon_0 + B(\varepsilon_0 - \varepsilon_1)] = \sigma_0^2 (1 - \beta^2)$. Thus

$$Var(\log w_0 \mid \log w_0 \geq \log w_1) = \sigma_0^2 [1 - \beta^2 \delta(c)].$$

Since $0 < \delta(c) < 1$ for all c, β is a correlation coefficient and thus $|\beta| \leq 1$, we obtain equation (2.6) in the form $Var(\log w_0 \mid \log w_0 \geq \log w_1) \leq \sigma_0^2 = Var(\log w_0)$.

Using the same procedure yields the following formulas for the truncated means and variances of normal distribution that are useful in equations (4.19) and (4.20):

$$E(\log w_0 \mid \log w_0 \geq \log w_1) = \sigma_0^2[1-\beta^2\delta(c)];$$

$$E(\log w_1 \mid \log w_0 \leq \log w_1) = \sigma_1^2[1-\gamma^2\delta(-c)];$$

$$Var(\log w_0 \mid \log w_0 \geq \log w_1) = \mu_0 - \sigma_0\beta\lambda(c);$$

$$Var(\log w_1 \mid \log w_0 \leq \log w_1) = \mu_1 + \sigma_1\gamma\lambda(-c).$$

By applying the law of total variance, the censored variance $Var(\log \widetilde{w})$ can be decomposed as the sum of within- and between-group variances, just as what equation (4.19) shows. Inserting formulas for the truncated moments into equation (4.19) gives equation (4.20).

Finally, I show the way in which one can construct the log reservation wage $\log w^r(s,...)$ faced by rural workers when they make migration decisions.

I start by considering the first row of the complimentary-slackness conditions given by equation (5.1), which describes the optimal decisions made by partially migratory types of rural workers in the rural household of interest.

$$\begin{aligned}
& L^{\#}(s) < L^0(s) \\
\Leftrightarrow & \frac{q(s)L^{\#}(s)^{\rho}}{\int_0^1 q(t)L^{\#}(t)^{\rho}\,dt} < \frac{q(s)L^0(s)^{\rho}}{\int_0^1 q(t)L^{\#}(t)^{\rho}\,dt} \\
\Leftrightarrow & \log T(s) + \alpha \log a^{\#} + \frac{\rho-1}{\rho}\log[\frac{q(s)L^{\#}(s)^{\rho}}{\int_0^1 q(t)L^{\#}(t)^{\rho}\,dt}] \\
& > \log T(s) + \alpha \log a^{\#} + \frac{\rho-1}{\rho}\log[\frac{q(s)L^0(s)^{\rho}}{\int_0^1 q(t)L^{\#}(t)^{\rho}\,dt}].
\end{aligned}$$

Note that the left-hand side term in the last line of equation above is just $\log MPL^{\#}(s,...)$. The right-hand side term is the log reservation wage, or say, the log opportunity costs of out-migration that I attempt to construct, because it measures the log agricultural wages that any migratory type of workers could earn if they would stay, while all other types of workers would behave optimally. Since for these partially migratory types of rural workers, we also have $\log MPL^{\#}(s,...) = \log w_1(s)$, the first row of equation (5.1) can be rewritten as

$$\log MPL^{\#}(s,...) = \max\{\log w^r(s,...), \log w_1(s)\},\ if\ L^{\#}(s) < L^0(s).$$

Similarly, starting with the second row of equation (5.1), which describes the optimal decision made by staying types of rural workers, we obtain

$$\begin{aligned} & L^{\#}(s) = L^0(s) \\ \Leftrightarrow\ & \log MPL^{\#}(s,...) = \log T(s) + \alpha \log a^{\#} + \frac{\rho - 1}{\rho} \log[\frac{q(s) L^0(s)^{\rho}}{\int_0^1 q(t) L^{\#}(t)^{\rho}\, dt}], \end{aligned}$$

where the right-hand side term is again the log reservation wage $\log w^r(s,...)$. Since their agricultural labor inputs remain unchanged in the course of out-migration, i.e. $L^{\#}(s) = L^0(s)$, $\log w^r(s,...)$ measures the actual log agricultural wages earned by staying types of rural workers. Moreover, according to equation (5.1), we have for the staying labor types $\log MPL^{\#}(s,...) \geq \log w_1(s)$. Hence, the second row of equation (5.1) can be rewritten as

$$\log MPL^{\#}(s,...) = \max\{\log w^r(s,...), \log w_1(s)\},\ if\ L^{\#}(s) = L^0(s).$$

Collecting the rewritten complimentary-slackness conditions for all types of rural workers gives equation (5.5) such that

$$\log MPL^{\#}(s,...) = \max\{\log w^r(s,...), \log w_1(s)\},\ \forall s.$$

APPENDIX B

INDIRECT EVIDENCE ON THE PREVELENCE OF THE PARTIAL OUT-MIGRATION

This appendix provides indirect evidences on the prevalence of partial out-migration in post-reform Chinese context.

The proof starts with a well-established fact in literature, namely percentages of migratory households, i.e. households with at least one migratory worker are usually higher than percentages of migratory workers in the same samples. Perhaps the best guess of the ratio between both percentages in literature is about two.[92] In what follows, I use the **proof by contradiction** to argue that this fact can be interpreted as indirect evidence for the prevalence of partial out-migration in the Chinese context.

Let us think about a scenario in which the following assumptions are satisfied:

(1) There are *N*-households in the rural community before migration. The number of members in each of the households is given by $h_i,\ i = 1, 2, ..., N$. For simplicity, I assume all household members belong to the labor force.

(2) *M* of *N* households ($0 \le M \le N$) participate in rural-to-urban migration in the sense that at least one member in each of these households out-migrates. They are called henceforth as the migratory households.

(3) Suppose that only complete family migration is possible. In other words, all members of a household make the same migration decision. This will serve as

[92] See for example, Zhao (1997), Zhao (1999b), Du and Park (2003). For a more complete review on this issue, readers may refer to Lu (2008), Appendix Table 3.

the null hypothesis (H_0) to be tested.

Under the scenario above, the percentage of migratory households in the rural community is given by M/N, while the percentage of migratory members is given by $\sum_{j=1}^{M} h_j / \sum_{i=1}^{N} h_i$. The percentage of migratory members can be rewritten as

$$\frac{\sum_{j=1}^{M} h_j}{\sum_{i=1}^{N} h_i} = \frac{M}{N} \frac{\frac{1}{M}\sum_{j=1}^{M} h_j}{\frac{1}{N}\sum_{i=1}^{N} h_i} \equiv \frac{M}{N} \frac{\overline{h}_j}{\overline{h}_i}$$

To ensure that the percentage of migratory households is much higher than that of migratory workers, we have

$$\Leftarrow \frac{M}{N} >> \frac{M}{N} \frac{\overline{h}_j}{\overline{h}_i}, \quad \Leftarrow \overline{h}_i >> \overline{h}_j,$$

which suggests that the average size of migratory households ($\overline{h}_j$) should be sufficiently smaller than the average size of all households in the rural community.

If we accept the best guess on the relation between two percentages, then

$$\Leftarrow \frac{M}{N} \approx 2 \frac{M}{N} \frac{\overline{h}_j}{\overline{h}_i}, \quad \Leftarrow \overline{h}_j \approx \frac{1}{2} \overline{h}_i,$$

which means the average size of migratory households should be about one half as large as that of all households in the community. To my knowledge, no literature has ever documented such a big difference between the average sizes of two types of households in the same community.

Therefore, the fact that percentage of migratory households is significantly larger than corresponding percentage of migratory members is likely to provide evidence against the null hypothesis given by (3). Consequently, alternative hypothesis (H_1) stating that not all the migrations are complete family migration must be accepted.

Nevertheless, accepting H_1 does not mean automatically that the partial migration is the prevailing form of migration in the Chinese context – An alternative hypothesis that is stronger than H_1. Hereafter, I call this stronger alternative hypothesis as $H_1^{'}$. To prove that $H_1^{'}$ should be accepted, I consider a slightly different scenario in which assumptions (1) and (2) listed above are maintained, while the assumption (3) is replaced by a new one as follows:

(3') Not all migrations are family migration. Among *M*-migratory households, *pM* of them migrate partially, while the rest of them engage in family migration. For households in the first subgroup, the numbers of members are given by $h_j, j = 1,..., pM$ and the numbers of migratory members are given by $l_j, j = 1,..., pM$. Obviously, we have $0 < l_j < h_j, \forall j$. For households in the second subgroup, the numbers of members are given by h_k, $k = 1,...,(1-p)M$ and the numbers of migratory members are given by $l_k = h_k, k = 1,...,(1-p)M$.

Under new assumptions (1) (2) and (3'), the percentage of migratory households is still given by M/N, while that of migratory members is now given by

$$\frac{\sum_{j=1}^{pM} l_j + \sum_{k=1}^{(1-p)M} l_k}{\sum_{i=1}^{N} h_i} = \frac{pM\bar{l}_j + (1-p)M\bar{l}_k}{N\bar{h}_i}.$$

In order to have

$$\Leftarrow \frac{M}{N} \approx 2\frac{pM\bar{l}_j + (1-p)M\bar{l}_k}{N\bar{h}_i}, \quad \Leftarrow \frac{1}{2} \approx \frac{p\bar{l}_j + (1-p)\bar{l}_k}{\bar{h}_i}.$$

It is clear that $0 < \bar{l}_j < \bar{h}_j$ and $\bar{l}_k = \bar{h}_k$. To facilitate further discussion, I define $\bar{l}_j \equiv \overline{\theta h_j}$, $\bar{\theta} \in (0,1)$, where $\bar{\theta}$ denotes average fraction of migratory members in the partially migratory households.

Substituting $\bar{l}_j \equiv \overline{\theta h_j}$ and $\bar{l}_k = \bar{h}_k$ into previous equation gives

$$\Leftarrow \frac{1}{2} \approx \frac{p\overline{\theta h_j} + (1-p)\bar{h}_k}{\bar{h}_i}.$$

If the average sizes measured by amount of household members are roughly the same for the partially, family migratory households and staying households, i.e. $\bar{h}_j \approx \bar{h}_k \approx \bar{h}_i$, then the equation above can be further simplified as

$$\Leftarrow \frac{1}{2} \approx p\bar{\theta} + (1-p).$$

Thus, for any given $\bar{\theta} \in (0,1)$, we have $p \approx \frac{1}{2}\frac{1}{1-\bar{\theta}} > \frac{1}{2}$. Moreover, since p is a probability, $p \le 1$ must hold and thus $\bar{\theta} \le \frac{1}{2}$.

To conclude, the results above suggest that within the migratory households, the share of partially migratory households is likely to exceed 1/2, which could be interpreted as a supportive evidence of the alternative hypothesis $H_1^{'}$ stating that the partial out-migration prevails in the Chinese rural-to-urban migration. Furthermore, the last equation may suggest among all partially migratory households, the average fraction of migratory members is likely to be low.

APPENDIX C

ECONOMIC INTERPRETATIONS FOR PARAMETERS $\{q(s)\}$

The main purpose of this appendix is to provide clear economic interpretations for the set of parameters $\{q(s)\}$ associating with different types of labor inputs in the CES production function. As will be shown below, these parameters relate closely to the marginal products of different labor inputs when total labor inputs are **balanced** in quality. With some stretches, these parameters relate also to the **intrinsic** skills of different labor inputs. The reason why I emphasize the term "intrinsic" is that the scarcity of inputs could also affect labor inputs' marginal products. By insisting on evaluating marginal products of labor at balanced labor-input-bundles, the impacts of the scarcity on the marginal products can be partialled out.

To simplify the following discussion, I consider a simple two-level nested CES production function such as[93]

$$y = TA^{\alpha} QL^{1-\alpha},$$
$$where\ QL = (q_1 L_1^{\rho} + q_2 L_2^{\rho})^{1/\rho},\ \text{with}\quad q_1 + q_2 = 1.$$

In what follows, I reconsider the economic interpretation for (q_1, q_2) that has often been vaguely understood as terms reflecting the skill-specific technical changes.

Since the production function is nonlinear, I approximate it using its first- and second order Taylor's series. I begin by considering a simpler case in which the production function is approximated as a first-order Taylor's series around (L_{10}, L_{20}).

[93] The discussion based on the simple two-level nested CES specification could be extended easily to the cases when the two-level nested CES specification with more than two labor inputs involves. After some modifications, this discussion would also shed light on the cases when the more complex nested CES specifications involve.

$$y = TA^{\alpha} Q L^{1-\alpha} = TA^{\alpha}(q_1 L_1^{\rho} + q_2 L_2^{\rho})^{(1-\alpha)/\rho}$$

Taking natural logarithms gives

$$\log y = \log T + \alpha \log A + \frac{(1-\alpha)}{\rho} \log(q_1 L_1^{\rho} + q_2 L_2^{\rho}) \equiv \tau + \alpha \log A + \frac{(1-\alpha)}{\rho} \log g.$$

Approximating this equation using the first-order Taylor's expansion yields

$$\begin{aligned} \log y &\approx \tau + \alpha \log A + \frac{1-\alpha}{\rho}[\log g_0 + \frac{\partial \log g_0}{\partial L_1}(L_1 - L_{10}) + \frac{\partial \log g_0}{\partial L_2}(L_2 - L_{20})] \\ &\approx \tau + \alpha \log A + \frac{1-\alpha}{\rho}[\log g_0 + \frac{\partial \log g_0}{\partial \log L_1} \log(\frac{L_1}{L_{10}}) + \frac{\partial \log g_0}{\partial \log L} \log(\frac{L_2}{L_{20}})], \end{aligned}$$

where $\log g_0 = \log(q_1 L_{10}^{\rho} + q_2 L_{20}^{\rho})$ and $\dfrac{\partial \log g_0}{\partial \log L_s} = \rho \dfrac{q_s L_s^{\rho}}{q_1 L_1^{\rho} + q_2 L_2^{\rho}}$, $\forall s = 1, 2$.

In particular, if the expansion point (L_{10}, L_{20}) is chosen such that it locates along the ray $L_{20} = L_{10} = \ell, \ \forall \ell > 0$, then a greatly simplified expression of $\log y$ can be obtained as follows

$$\log y = \tau + \alpha \log A + \frac{(1-\alpha)}{\rho} \log g \approx \tau + \alpha \log A + (1-\alpha)(q_1 \log L_1 + q_2 \log L_2).$$

Taking partial derivatives gives the expressions for the parameters (q_1, q_2):

$$\begin{aligned} q_1 &= \frac{1}{1-\alpha} \frac{\partial \log y}{\partial \log L_1} \equiv \frac{1}{1-\alpha} \varepsilon_{y1}; \\ q_2 &= \frac{1}{1-\alpha} \frac{\partial \log y}{\partial \log L_2} \equiv \frac{1}{1-\alpha} \varepsilon_{y2}, \end{aligned}$$

where ε_{y1} and ε_{y2} denote the elasticities of output with respect to L_1 and L_2.

Moreover, suppose the labor inputs are balanced in quality in the sense that $L_1 = L_2$, the relationship between parameters q_1 and q_2 have a clear economic interpretation. Without losing generality, if it is found that $q_1 \geq q_2$, then according to the discussions above, we know $\varepsilon_{y1} \geq \varepsilon_{y2}$. Furthermore, since the elasticities are defined as $\varepsilon_{ys} = \frac{\partial \log y}{\partial \log L_s} = \frac{L_s}{y}\frac{\partial y}{\partial L_s} \equiv \frac{L_s}{y} MPL_s$, $q_1 \geq q_2$ implies that $MPL_1 \geq MPL_2$ when labor inputs are balanced in quality. Therefore, provided that the first-order Taylor's series provides an acceptable approximation to the nested CES function, then the following results are proven to be equivalent

$$q_1 \geq q_2 \Leftrightarrow \varepsilon_{y1} \geq \varepsilon_{y2} \Leftrightarrow MPL_1 \geq MPL_2 \text{ when } L_1 = L_2 .$$

Nevertheless, the simplified expression of $\log y$ presented above turns out to take the form of the log Cobb-Douglas, which could be restrictive for my purpose. Particularly, the Cobb-Douglas form imposes a strong restriction asserting that the elasticity of substitution between L_1 and L_2 must be unity. To generalize the key finding above, this restriction has to be relaxed. Thus, I proceed to consider the second-order Taylor's approximation.[94] Besides, the second-order Taylor's series would give a better approximation to the original nested CES function. Hence, I expect the economic interpretation of parameters (q_1, q_2) based on the second-order Taylor's approximation would be more accurate.

Applying the formula of the Taylor's expansion gives

[94] In literature, the CES function is often approximated by the so-called Translog function (see for example, Christensen et al., 1973; Chambers, 1988), which could be viewed as the second-order Taylor's approximation of the CES function.

$$
\begin{aligned}
\log y &= \tau + \alpha \log A + \frac{(1-\alpha)}{\rho} \log g \\
&\approx \tau + \alpha \log A + \frac{(1-\alpha)}{\rho} [\log g_0 + \frac{\partial \log g_0}{\partial L_1}(L_1 - L_{10}) + \frac{\partial \log g_0}{\partial L_2}(L_2 - L_{20}) \\
&\quad + \frac{1}{2}\frac{\partial^2 \log g_0}{\partial L_1^2}(L_1 - L_{10})^2 + \frac{1}{2}\frac{\partial^2 \log g_0}{\partial L_2^2}(L_2 - L_{20})^2 + \frac{\partial^2 \log g_0}{\partial L_1 \partial L_2}(L_1 - L_{10})(L_2 - L_{20})]
\end{aligned}
$$

Rearranging and simplifying the right-hand side term gives

$$
\begin{aligned}
\log y &\approx \tau + \alpha \log A + \log(q_1 L_{10}^\rho + q_2 L_{20}^\rho) \\
&\quad + \rho \frac{q_1 L_{10}^\rho}{q_1 L_{10}^\rho + q_2 L_{20}^\rho}(\log L_1 - \log L_{10}) + \rho \frac{q_2 L_{20}^\rho}{q_1 L_{10}^\rho + q_2 L_{20}^\rho}(\log L_2 - \log L_{20}) \\
&\quad + \frac{1}{2}\rho^2 \frac{(q_1 L_{10}^\rho)(q_2 L_{20}^\rho)}{(q_1 L_{10}^\rho + \beta_2 L_{20}^\rho)^2}[(\log L_1 - \log L_{10}) - (\log L_2 - \log L_{20})]^2 \\
&\quad - \frac{1}{2}\rho \frac{1}{q_1 L_{10}^\rho + q_2 L_{20}^\rho}\{q_1 L_{10}^\rho (\log L_1 - \log L_{10})^2 + q_2 L_{20}^\rho (\log L_2 - \log L_{20})^2\}.
\end{aligned}
$$

Hence, the input-output elasticities are given by

$$
\begin{aligned}
\varepsilon_{y1} &= \frac{\partial \log y}{\partial \log L_1} \\
&\approx (1-\alpha)\{\frac{q_1 L_{10}^\rho}{\beta_1 L_{10}^\rho + q_2 L_{20}^\rho} + \rho \frac{(q_1 L_{10}^\rho)(q_2 L_{20}^\rho)}{(\beta_1 L_{10}^\rho + q_2 L_{20}^\rho)^2}[(\log L_1 - \log L_{10}) - (\log L_2 - \log L_{20})] \\
&\quad - \frac{q_1 L_{10}^\rho}{q_1 L_{10}^\rho + q_2 L_{20}^\rho}(\log L_1 - \log L_{10})\}
\end{aligned}
$$

$$
\begin{aligned}
\varepsilon_{y2} &= \frac{\partial \log y}{\partial \log L_2} \\
&\approx (1-\alpha)\{\frac{q_2 L_{20}^\rho}{q_1 L_{10}^\rho + q_2 L_{20}^\rho} - \rho \frac{(q_1 L_{10}^\rho)(q_2 L_{20}^\rho)}{(q_1 L_{10}^\rho + q_2 L_{20}^\rho)^2}[(\log L_1 - \log L_{10}) - (\log L_2 - \log L_{20})] \\
&\quad - \frac{q_2 L_{20}^\rho}{q_1 L_{10}^\rho + q_2 L_{20}^\rho}(\log L_2 - \log L_{20})\}
\end{aligned}
$$

If the expansion point is again chosen at a point along the ray $L_{10} = L_{20} = \ell$, $\forall \ell > 0$, the expressions of elasticities can be simplified as

$$\varepsilon_{y1} = (1-\alpha)[q_1 + \rho q_1 q_2(\log L_1 - \log L_2) - q_1(\log L_1 - \log \ell)];$$
$$\varepsilon_{y2} = (1-\alpha)[q_2 - \rho q_1 q_2(\log L_1 - \log L_2) - q_2(\log L_2 - \log \ell)].$$

Furthermore, when labor inputs are balanced, that is, $L_1 = L_2 = L$, we obtain

$$\varepsilon_{y1}\,|_{L_1=L_2=L} = (1-\alpha)q_1(1-\log L + \log \ell)$$
$$\varepsilon_{y2}\,|_{L_1=L_2=L} = (1-\alpha)q_2(1-\log L + \log \ell)$$

Thus, the expressions for the parameters (q_1, q_2) are given by

$$q_1 = \frac{\varepsilon_{y1}\,|_{L_1=L_2=L}}{(1-\alpha)(1-\log L + \log \ell)};$$
$$q_2 = \frac{\varepsilon_{y2}\,|_{L_1=L_2=L}}{(1-\alpha)(1-\log L + \log \ell)}.$$

Consequently, $q_1 \geq q_2$ implies $\varepsilon_{y1} \geq \varepsilon_{y2}$ when labor inputs are balanced in quality. Moreover, it implies also that $MPL_1 \geq MPL_2$ under the same condition.

However, unlike the previous discussion based on the first-order Taylor's series, these results are not always equivalent when labor inputs are balanced. This is because ε_{y1} and ε_{y2} are input-output elasticities and thus their values should be non-negative. Hence, these results are equivalent if these labor inputs are balanced and if the additional restriction $L_1 = L_2 = L \leq e\ell$ is satisfied.

To conclude, by approximating the two-level nested CES function by a second-order Taylor's series around $(L_{10}, L_{20}) = (\ell, \ell)$, the following results are proven to be equivalent.

$$q_1 \geq q_2 \Leftrightarrow \varepsilon_{y1} \geq \varepsilon_{y2} \Leftrightarrow MPL_1 \geq MPL_2, \text{ when } L_1 = L_2 = L < e\ell\,.$$

This conclusion is no doubt the most important finding in this appendix. With this finding, I make the economic interpretation for parameters (q_1, q_2) explicitly, just as what has been argued at the beginning of this appendix. In addition, this finding could also serve as a useful guidance to order different types of labor inputs whose characteristics such as education, age or working experiences and possibly gender are different in terms of their "intrinsic" skills.

APPENDIX D

DATA APPENDIX

The data are drawn from the Chinese Health and Nutrition Survey (CHNS) longitudinal data in years 1991, 1993, 1997, 2000, 2004, 2006 and 2009.[95] In each year, the survey covers up to nine provinces that vary substantially in geographic, economic and health indicators: They are Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi and Guizhou.[96] The CHNS data provide detailed information on health, demographic and socioeconomic factors (such as employment and earnings) at individual, household and community levels. Partly because of its accessibility, the CHNS data is one of the most widely used microdata from China.

(1) Skill Components and Imputations

The CHNS data offer two alternative measures of individual agents' educational attainments, namely the years of schooling ($a11$) and the highest degree of education ($a12$). The former is useful in preparing descriptive statistics (Section 6.2.1) and in estimating the Mincerian earning functions (Section 6.2.3), while the latter is useful in subdividing total labor force into skill groups and thus in estimating agricultural production functions of household farms (Section 6.1). In original CHNS data, educations are sometimes missing. I attempt to recover some of them by making use of the panel structure of the data. For instance, if a person's education is missing at one survey, then I will search for his education in previous and subsequent surveys. Suppose his educations in both surveys are available and they happen to be the same,

[95] For more details, interested readers should refer to the CHNS documents, questionnaires in all years and the working manuals in 1993 and 2006.

[96] The rankings of these provinces in terms of the per capita GDP, 2000 are 8, 10, 6, 9, 18, 13, 17, 29 and 31 among the 31 provinces reported in the Chinese Statistical Yearbook, 2000.

the missing education is replaced by one from adjacent surveys. Also, if a person's education is missing, but his education is available in the previous survey and it is known that he left school at the time of previous survey ($a13=0$), then the missing education is replaced by education in previous survey. Besides, if one of the measures of education is missing, then I impute the missing one using the non- missing alternative measure whenever possible.

The information on age is also missing for a small number of individual observations. I impute missing age of a person using his ages available from other surveys and the interview dates of two surveys. The potential working experience is defined for workers as the years since they left school and entered the labor market, that is, *experience = age - schooling - 6*.

(2) Agricultural Employments

The agricultural workers are defined as rural workers who are engaged in one of the following agricultural activities: household farming, gardening and livestock-raising. For each of these activities, I construct a dummy variable that takes value 1 if the worker reports his participation or if he reports positive working hours in it, and 0 otherwise. A household is identified as an agricultural household if the respondent reports the participation of this household in agriculture or if one of its members is identified as an agricultural worker. Notice that agricultural workers may also be employed in non-agricultural sector and earn wages in that sector.

Suppose the labor inputs is homogeneous and perfectly substitutive in household agricultural production, the total labor input of a household farm can be obtained by simply aggregating labor services provided by all household members who are identified as agricultural workers. To be concrete, the amount of agricultural labor inputs can be measured either in persons or in hours. A worker's annual agricultural working hours are constructed by multiplying months worked per year, a constant 4.35 – average weeks per month, days worked per month and hours worked per day for all three agricultural activities and then adding them together.

Alternatively, suppose like in Chapters 4 and 5 the labor inputs are in fact heterogeneous, labor inputs endowed with different skill components should be treated as separate entries of agricultural production. In this thesis, I categorize all agricultural workers and their labor services into two educational groups: workers without junior high school diploma and those with junior high school diploma or higher degrees. Meanwhile, I categorize the same population into two age groups: workers aged below 35 and those above 35 years old. Altogether, there are four broadly defined education-age cells. The amounts of persons and annual hours worked are calculated within all of four education-age cells.

Year	Skill Cell	Migratory	Staying
1991	JHS-×YG	0.31	0.22
	JHS-×MA	0.07	0.31
	JHS-×OA	0.05	0.13
	JHS×YG	0.40	0.20
	JHS×MA	0.04	0.07
	JHS×OA	0.01	0.01
	JHS+×YG	0.11	0.05
	JHS+×MA	0.01	0.02
	JHS+×OA	0.00	0.00
1993	JHS-×YG	0.35	0.18
	JHS-×MA	0.08	0.31
	JHS-×OA	0.03	0.14
	JHS×YG	0.40	0.21
	JHS×MA	0.03	0.09
	JHS×OA	0.00	0.01
	JHS+×YG	0.10	0.04
	JHS+×MA	0.01	0.03
	JHS+×OA	0.02	0.00

(Continuation)

1997	JHS-×YG	0.36	0.18
	JHS-×MA	0.02	0.28
	JHS-×OA	0.02	0.15
	JHS×YG	0.51	0.21
	JHS×MA	0.01	0.09
	JHS×OA	0.00	0.01
	JHS+×YG	0.08	0.04
	JHS+×MA	0.01	0.04
	JHS+×OA	0.00	0.00
2000	JHS-×YG	0.22	0.13
	JHS-×MA	0.03	0.27
	JHS-×OA	0.01	0.15
	JHS×YG	0.56	0.21
	JHS×MA	0.05	0.13
	JHS×OA	0.00	0.02
	JHS+×YG	0.13	0.04
	JHS+×MA	0.02	0.05
	JHS+×OA	0.00	0.00
2004	JHS-×YG	0.18	0.09
	JHS-×MA	0.06	0.24
	JHS-×OA	0.02	0.20
	JHS×YG	0.51	0.16
	JHS×MA	0.07	0.18
	JHS×OA	0.00	0.03
	JHS+×YG	0.16	0.03
	JHS+×MA	0.01	0.06
	JHS+×OA	0.00	0.01
2006	JHS-×YG	0.18	0.07
	JHS-×MA	0.09	0.22
	JHS-×OA	0.01	0.21
	JHS×YG	0.43	0.17
	JHS×MA	0.11	0.18
	JHS×OA	0.00	0.04
	JHS+×YG	0.14	0.04
	JHS+×MA	0.03	0.06
	JHS+×OA	0.00	0.01
2009	JHS-×YG	0.12	0.04
	JHS-×MA	0.14	0.20
	JHS-×OA	0.02	0.25
	JHS×YG	0.35	0.14
	JHS×MA	0.17	0.20
	JHS×OA	0.01	0.06
	JHS+×YG	0.15	0.04
	JHS+×MA	0.04	0.05
	JHS+×OA	0.00	0.01

Table D.1: Skill Structures of Migratory and Staying Rural Workers

Note: The notations "JHS-", "JHS" and "JHS+" refer to worker groups without junior high school diploma, with exact junior high school diploma, and with senior high school, or college diploma. "YG", "MA" and "OA" refer to worker groups aged between 15 and 34, 35-54, and 55-75. The notation "×" means "and".

(3) Non-Labor Agricultural Inputs

The non-labor agricultural inputs include land, material and capital. The land cultivated by rural households is directly available from the CHNS data. When land is missing for some rural households, I impute land using the similar strategy as I did for the missing education. The information on the land redistribution ($e11e$) is used in these imputations whenever feasible. Since the CHNS data provide no information on detailed material inputs such as seeds, fertilizers, energy etc., I have to approximate the material input by the available production cost data. The capital input measures the market values of all farm machinery and draft animals (before CHNS 1993) employed in household agricultural production. I manage to construct the capital input by multiplying the unit price and amount of each of the farm machinery and draft animals and then adding them. Unfortunately, the constructed capital input is likely to be very inaccurate and thus cannot be used in estimations. By inspecting the data, I tend to attribute the inaccuracy of the capital data to the large amount of measurement errors in unit prices of farm machinery and draft animals reported in the original data, since it is not uncommon to find the cases where unit prices of the same machinery owned by the same household vary dramatically across years.

(4) Total Agricultural Income

The total agricultural income equals to the sum of market values of final products (including sold and self-consumed) of household farming, gardening and livestock raising. Unlike the conventional practice of the CHNS-team, if income earned from one of three agricultural activities is found to be missing, I set the total agricultural income to be missing.

(5) Migration Status

In this thesis, rural migrants are defined as those who left the village where his household resides for at least six months. This definition bears some similarity to that adopted in the Chinese Population Census, 2000 onwards, except that the latter puts emphasis on the destination region. To apply such definition to the CHNS data, I need thus two sets of variables indicating where and when rural workers move out. Because of the redesign of the questionnaires, variables used to construct the migration dummy are slightly different before and after CHNS 1993. In CHNS 1991 and 1993, the variables $a7$ ("How many months did [you] not live here?"), $aa12$ ("When moved out?") are used to determine whether a rural worker moved out for at least six months, $aa13$ ("Where lives now?") is used to determine whether he stayed in the same community where his household resides. In CHNS 1997 and later surveys, however, $a5e$ ("Still live your household?") and $aa12$ are employed for the first purpose, while $a5f$ ("How long gone?") and $aa13$ are employed for the second.

The advantage of adopting this definition lies partly in its resemblance to that in the Chinese Population Census, and partly in its applicability to a longer period. By contrast, many scholars propose to use variable *a5e* alone to identify rural migrants. Unfortunately, this variable appears for the first time in CHNS 1997. Admittedly, given that applying the definition preferred by this thesis requires somewhat different sets of variables in two sub-periods, the consistency of the definition could be difficult to check.

Moreover, rural workers who are not identified as migrants are labeled as staying workers. Rural migratory households are defined as rural households with at least one migrant. The rest of rural households are non-migratory.

Considering that the definition above relies little on the occupational information of rural residents, the identified rural migrants may include those who have weak or no attachments to the labor market. Therefore, I will further restrict to rural migratory and staying residents who were not at school and who were in their working ages at the time of survey when necessary.

Further details on the data management, the definitions of samples and variables essential for estimation and prediction that allow an exact reproduction of the results presented in Chapter 6 are documented in a series of Stata do files, which can be available from the author upon request.

REFERENCES

Acemoglu, Daron. "Technical Change, Inequality, and the Labor Market." Journal of Economic Literature 40 (2002): 7-72.

Angrist, Joshua D., and Alan Krueger B. "Does Compulsory School Attendance Affect Schooling and Earnings?" Quarterly Journal of Economics 106 (1991): 979-1014.

Benjamin, Dwayne et al. "Income Inequality during China's Economic Transition." In: Loren Brandt and Thomas G. Rawski, eds., China's Great Economic Transformation: 729-775. Cambridge: Cambridge University Press.

Blau, Francine and Lawrence Kahn. 2015. "Immigration and the Distribution of Incomes." In: Barry R. Chiswick and Paul W. Miller, eds., Handbook of the Economics of International Migration, Volume 1: 793-843. Amsterdam: Elsevier.

Borjas, George J. "Self-Selection and the Earnings of Immigrants." American Economic Review 77 (1987): 531-553.

Borjas, George J. "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market." 118 (2003): 1335-1374.

Borjas, George J. 2014. Immigration Economics. Cambridge: Harvard University Press.

Card, David. "The Impact of the Mariel Boatlift on the Miami Labor Market." Industrial and Labor Relations Review 43 (1990): 245-257.

Card, David. "The Causal Effect of Education on Earnings." In: Orley Ashenfelter and David Card, eds., Handbook of Labor Economics, Volume 3: 1801-1863. Amsterdam: Elsevier.

Card, David and Lemieux, Thomas. "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis." Quarterly Journal of Economics 105 (2001): 705-746.

Casella, George and Roger L. Berger, 2001. Statistical Inference, 2nd edition. Duxbury Press.

Chambers, Robert G. 1988. Applied Production Analysis: A Dual Approach. Cambridge: Cambridge University Press.

Chan, Kam Wing, 2009. "The Chinese Hukou System at 50." Eurasian Geography and Economics, 50 (2): 197-221.

Chayanov, A.V. 1986. The Theory of Peasant Economy. Madison: University of Wisconsin Press.

Cheng, Tiejun and Mark Selden. "The Origin and Social Consequences of China's Hukou System." The China Quarterly 139 (1994): 644–668.

Chiswick, Barry R. “Are Immigrants Favorably Self-Selected?” AEA Papers and Proceedings (1999): 181-185.

Christensen, Laurits R., Dale W. Jorgenson and Lawrence J. Lau. “Transcendental Logarithmic Production Frontiers.” The Review of Economics and Statistics 55 (1973): 28-45.

Du, Ying. “Individual Characteristics and Macroeconomic Background of the Recent Rural-Urban Migration in China.” Chinese Rural Economy 06 (1997): 4-11 (In Chinese).

Food and Agriculture Organization of the United Nation (FAO). 2011. The State of Food and Agriculture. Rome.

French, Eric and Christopher Taber. 2011. “Identification of Models of the Labor Market.” In: Orley Ashenfelter and David Card, eds., Handbook of Labor Economics: 537-617. Amsterdam: Elsevier.

Gould, Eric D. “Rising Wage Inequality, Comparative Advantage, and the Growing Importance of General Skills in the United States.” Journal of Labor Economics 20 (2002): 105-147.

Greene, William H. 2007. Econometric Analysis, 6th edition. Upper Saddle River: Prentice Hall.

Greenwood, Michael J. 1997. “Internal Migration in Developed Countries.” In: M. R. Rosenzweig and O. Stark, eds., Handbook of Population and Family Economics: 647-720. Amsterdam: Elsevier.

Grilliches, Zvi. “Capital-Skill Complementarity.” Review of Economics and Statistics 51 (1969): 465-68.

Hamermesh, Daniel S. 1993. Labor Demand. Princeton: Princeton University Press.

Heckman, James J. and Bo. E. Honoré, “The Empirical Content of the Roy Model.” Econometrica 58 (1990): 1121-1149.

Heckman, James J. and Guilherme Sedlacek. “Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market.” Journal of Political Economics 93 (1985): 1077-1125

Hellerstein, Judith K., David Neumark and Kenneth R. Troske. “Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations.” Journal of Labor Economics 17 (1999): 409-446.

Hu, Jingbei. “An Explanation of Rising Wage in the Process of Economic Development.” Economic Research Journal 03 (1994): 33-43 (In Chinese).

Hu, Jingbei. “Concepts and the Stylized Facts of the Intersectoral Transfer of Agricultural Labor Force.” Working Papers for Economic Development 03 (2010): 1-24 (In Chinese).

Jacoby, Hanan G. “Shadow Wages and Peasant Family Labor Supply: An Econometric Application to the Peruvian Sierra.” Review of Economic Studies 60 (1993): 903-921.

Jin, Songqing and Klaus Deininger, 2004. “Land Rental Markets as an Alternative to Reallocation? Equity and Efficiency Considerations in the Chinese Land Tenure System.” China Economic Quarterly 03 (2004): 1003-1028 (In Chinese).

Katz, Lawrence F. and David H. Autor. 1999. “Changes in the Wage Structure and Earning Inequality.” In: Orley Ashenfelter and David Card eds., Handbook of Labor Economics, Volume 3: 1464-1555. Amsterdam: Elsevier.

Katz, Lawrence F. and Kevin M. Murphy. “Changes in Relative Wages, 1963-1987: Supply and Demand Factors.” Quarterly Journal of Economics 107 (1992): 35-78.

Kuznets, Simon. “Economic Growth and Income Inequality.” American Economic Review 45 (1955): 1-28.

Leamer, Edward E. 2009. The Craft of Economics: Lessons from the Heckscher-Ohlin Framework (Ohlin Lectures). Cambridge: The MIT Press.

Lemieux, Thomas. 2006. “The “Mincer Equation” Thirty Years After Schooling, Experience, and Earnings.” In: Shoshana Grossbard, eds., Jacob Mincer: A Pioneer of Modern Labor Economics: 127-145. New York: Springer.

Lewis, Arthur W. “Economic Development with Unlimited Supplies of Labour.” The Manchester School 22 (1954): 139-191.

Lewis, Ethan. “Local Open Economies Within the US: How Do Industries Respond to Immigration?” Federal Reserve Bank of Philadelphia Working Paper, 2003.

Lewis, Ethan. “Immigration, Skill Mix, and Capital Skill Complementarity.” Quarterly Journal of Economics 126 (2011): 1029-1069.

Li, Shi. “Researches on the Personal Income Distribution in China: Review and Perspective.” China Economic Quarterly 02 (2003) (In Chinese).

Lipton, Michael. “Migration from Rural Areas of Poor Countries: The Impacts on Rural Productivity and Income Distribution.” World Development 8 (1980).

Lu, Yunhe. “Who Leave the Countryside? A Survey on the Characteristics of the Rural-Urban Migrants in China.” Working Papers for Economic Development 04 (2008): 1-55 (In Chinese).

Meng, Xin and Junsen Zhang. “The Two-Tier Labor Market in Urban China: Occupational Segregation and Wage Differentials between Urban Residents and Rural Migrants in Shanghai.” Journal of Comparative Economics 29 (2001): 485-504.

Mincer, Jacob. 1974. Schooling, Experience, and Earnings. New York: National Bureau of Economic Research.

Mishra, Prachi. "Emigration and Wages in Source Countries: Evidence from Mexico." Journal of Development Economics 82 (2007): 180-199.

National Bureau of Statistics of China. 2010. China Compendium of Statistics 1949-2008. Beijing: China Statistics Publishing House.

National Bureau of Statistics of China. Chinese Statistical Yearbooks, 2000 and 2014. Beijing: China Statistics Publishing House.

Neal, Derek and Sherwin Rosen. 2000. "Theories of the Distribution of Earnings." In: A. B. Atkinson and F. Bourguignon, eds., Handbook of Income Distribution, Volume 1: 379-427.

Ottaviano Gianmarco I. P. and Giovanni Peri. 2008, "Immigration and National Wages: Clarifying the Theory and the Empirics." NBER Working Paper No. 14188.

Ottaviano Gianmarco I. P. and Giovanni Peri. "Rethinking the Effect of Immigration on Wages." Journal of the European Economic Association 10 (2012): 152-196.

Ravallion, Martin and Shaohua Chen. "China's (Uneven) Progress Against Poverty." Journal of Development Economics 82 (2007): 1-42.

Roy A. D. "Some Thoughts on the Distribution of Earnings." Oxford Economic Papers 3 (1951): 135- 146.

Todaro, Michael P. "A Model of Labor Migration and Urban Unemployment in Less Developed Countries." American Economic Review 59 (1969): 138-148.

Xing, Chunbing. "Income Differential between Rural Migrants and Urban Workers." Management World 05 (2008): 55-64 (In Chinese).

Xu, Qing, Yin Rongliang and Zhang Hui. "Economies of Scale, Return to Scale and the Problem of Optimal-Scale of Farm Management: An Empirical Study based on Grain Production in China." Economic Research Journal 03 (2011): 59-71 (In Chinese).

Ye, Jianping, Jiang Yan and Feng Lei. "Research on the Land Rental Market in Rural China." China Rural Survey 04 (2006): 48-55 (In Chinese).

Zhao, Yaohui. "Labor Mobility and the Role of Education in Migration in Rural China." Economic Research Journal 02 (1997): 37-42, 73 (In Chinese).

Zhao, Yaohui. "Leaving the Countryside: Rural-To-Urban Migration Decisions in China." AEA Papers and Proceedings (1999): 281-286.

Zhao, Zhong. "Rural-Urban Migration in China: What do We Know and What do We Need to Know?" China Economic Quarterly 03 (2004): 517-36 (In Chinese).

Eidesstattliche Erklärung

Ich erkläre hiermit, dass ich die vorgelegte Dissertation selbstständig und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis niedergelegt sind, eingehalten.