

Labor Migration in China

Shan Jin

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

Doctor of Philosophy

in

Economics

Suqin Ge, Chair

Xu Lin

Melinda Miller

Kowk Ping Tsang

April 27, 2021

Blacksburg, Virginia

Keywords: Migration, Wage, Network, Risk Preferences

Copyright 2021, Shan Jin

Labor Migration in China

Shan Jin

(ABSTRACT)

With the transition of the economy in China, migrants start holding a more and more important position in the labor market. Therefore, from this dissertation, we try to explore different topics related to the migrants in China. This dissertation consists of three essays on who chooses to migrate, how networks affect migrants' outcomes, and what is the inter-generational impact of parental migration on children's risk preferences. In the first chapter, we briefly introduce the motivation and contribution, and then we provide the methods and detailed findings in the following chapters.

Chapter 2 examines the impact of the endogeneity of the decision to migrate on the wage differentials between migrants and non-migrants in China. We find that migrants are self-selected from the upper tail of the income distribution in their home location. Consistent with a theoretical model of migration choice, we show that the size of the selection effect on wage depends on the wage differences between the prefectures of origin and destination as well as migration cost. The selection effect also differs among workers with different education and in different cities.

Chapter 3 studies how networks affect migrants' migration decisions, employment, and wage levels by using 2005 China's mini-census. Different from existing studies, this paper takes into account the existence of self-selection in the labor market. With the help of a theoretical model, we have a better understanding of the mechanism of networks as well as the different network effects on rural and urban migrants. We find out that networks affect both rural

and urban migrants' migration decisions positively. In terms of employment, networks exert positive impacts on rural migrants but not on urban ones, which is due to the different quality drops between rural and urban migrants when the networks increase. Such employment effects also lead rural migrants to face a more severe negative wage impact than urban migrants.

Chapter 4 investigates how parental migration affects left-behind underage children's risk preferences. By focusing on migrant parent groups, we are able to estimate whether the influence of nurture could also affect children's risk preference levels or not. The findings suggest that besides the intergenerational transmission of risk preferences between parents and children, parental migrations do have an influence on girls' risk preference levels. In addition, in terms of adult children's risk-related outcomes, we are able to find a positive parental migration impact on daughters' self-employment decisions. Findings help us have a better understanding of the relevant factors of risk preferences, and also confront the impact of the separation of parents and children.

Labor Migration in China

Shan Jin

(GENERAL AUDIENCE ABSTRACT)

Nowadays, migration becomes increasingly common and migrants take a large proportion of the labor market. With the economic development and the closer connections among regions, people are more likely to study or work outside their home locations than before. Even though there is still a strict household registration system in China, we can find the migration supporting systems are becoming established, and facilitate easier migration for more and more people. Having a better understanding of migrants helps us make better policies as well as have the labor market and society develop better.

This dissertation explores who choose to migrate (Chapter 2), how the social connections or networks affect migrants' employment and wages (Chapter 3), and whether there is any intergenerational impact on migrants' children in terms of risk preferences (Chapter 4) using methods from labor economics and economic theory.

We conclude that the best works tend to migrate first in the labor market, and social connections have different impacts on rural and urban migrants in terms of employment and wages. Moreover, we notice that migrant parents affect the left-behind children's risk preferences by both influences of nature and nurture. In sum, we study different topics on migrants and have a deeper understanding of how migrants are affected in the labor market as well as how migrants affect their next generations.

Acknowledgments

Finally, it is the time to express my gratitude to people who help me a lot in this research. It has been a long journey to become an economist. During this journey, I realized the preciousness of perseverance, I learned that life is full of unknowns and challenges, and I got more than I expected. All these small gains and achievements are inseparable from the supports and help of my advisor, committee members, parents, husband, and friends.

First and foremost, I would like to thank my respected advisor, Dr. Suqin Ge, for guiding and supporting me to complete this research. Her precious comments always inspire me to think deeply and make me grow quickly. Thank you!

I sincerely thank my committee members, Dr. Xu Lin, Dr. Melinda Miller, and Dr. Kowk Ping Tsang, for their valuable guidance and suggestion. I feel tremendously fortunate to have you as my committee members during the period of my research.

I am deeply indebted to my parents, Enfu Jin and Chunhua Yin. They are the most important and the most special people in my life. They let me grow in love, and always give their unconditional support and encouragement to me.

I would like to thank my husband, Xiaomeng Zhang. His companionship and support are the source of joy for me during this journey. Lastly, many thanks to my friends, especially to Yuanmei and Caidie, who give me a lot of supports to finish this research work. All of you made a big part of the research as well as my life. Thank you!

Table of Contents

List of Figures	ix
List of Tables	x
1 Introduction	1
2 Self-Selection and the Wage of Migrants in China	4
2.1 Introduction	5
2.2 Related Literature	7
2.3 The Model	11
2.4 Data and Descriptive Statistics	16
2.5 Empirical Implications	19
2.5.1 Empirical Specification	19
2.5.2 Positive Selection and Origins' Wage	24
2.5.3 Positive Selection and Moving Cost	25
2.5.4 Positive Selection and Education Attainment Level	28
2.6 Selection Story in Different Types of Cities	32
2.7 Conclusion	34

3	Network Effects on Migrants' Labor Market Outcomes in China	37
3.1	Introduction	38
3.2	Related Literature	41
3.3	The Model	44
3.3.1	Network Effects on Migrants' Employment	44
3.3.2	Network Effects on Migrants' Wage	48
3.4	Data and Descriptive Statistics	50
3.5	Empirical Implications	53
3.5.1	Type of Selection in the Labor Market	54
3.5.2	Network Effects on Labor Market Outcome	57
3.6	Conclusion	66
4	Parental Migration and Children's Risk Preferences	68
4.1	Introduction	69
4.2	Related Literature	72
4.2.1	Risk Preferences and Determinants	72
4.2.2	Parental Migration and Impacts	74
4.3	Data and Descriptive Statistics	76
4.4	Empirical Analysis	79
4.4.1	Parental Migration and Risk Preference Levels	79

4.4.2	Parental Migration and Risk Related Outcome	85
4.5	Conclusion and Future Work	87
	Bibliography	89
	Appendices	100
	Appendix A Appendix to Chapter 3	101
A.1	Relationship between network and the percentage of early migrants $e(net_c)$.	101
A.2	Proof of Propositions	102
A.2.1	Case 1: when there is no self-selection	102
A.2.2	Case 2: when there is a positive self-selection	103
A.2.3	Case 3: when there is a negative self-selection	103
A.2.4	Case 1: when there is no self-selection	104
A.2.5	Case 2: when there is a positive self-selection	106
A.2.6	Case 3: when there is a negative self-selection	107
	Appendix B Appendix to Chapter 4	110

List of Figures

2.1	Urban Migrants	23
2.2	Rural Migrants	23
2.3	Scatter Plot of City Wage Standard Deviation	24
2.4	Relationship between Positive Selection and Origins' Wage	26
2.5	Relationship between Positive Selection and Origins' Distance	27
2.6	Relationship between Education Return and Migrants Proportions	31
3.1	Migration Status (Data from 2005 1% Population Survey)	38
3.2	Distributions for Early Migrants and All Employers' Log Wage	64
4.1	Numbers of Rural Migrants in China (unit: ten thousand people)	71
4.2	Survey Questions and According Outcomes	78
4.3	Risk Preference Level t - test for Parents with Different Migration Experience	82

List of Tables

2.1	Characteristics of Locals and Migrants with Different <i>Hukou</i> Status	18
2.2	Baseline Results on <i>Hukou</i> Effects	22
2.3	Effect of Origins' Wage and Distances on Positive Selection	27
2.4	<i>Hukou</i> Effects by Education Level	30
2.5	<i>Hukou</i> Effect by City Type	34
3.1	Different Impacts on Migrants	48
3.2	Characteristics of Locals and Migrants	53
3.3	Self-Selection on Rural & Urban Migrants	55
3.4	Relationship between Networks and the Proportion of Employed Early Migrants	56
3.5	Network Effect on New Migrants' Migration Decisions	57
3.6	Network Effect on Migrants' Employment Status (Probit)	59
3.7	Relationship between Networks and Average Education Years for Migrants .	61
3.8	Network Effect on Migrants' Wage	63
3.9	Network Effect on Wage for Rural and Urban Migrants	65
4.1	Characteristics of Parents and Rural Local children	79
4.2	Having a Migrant Parent and Rural Local Underage Children's Risk Prefer- ence Levels	81

4.3	Relationship between Parents' Risk Preference Levels and Migration Status .	82
4.4	Estimates of the ATT with Local Parents as the Control Group (Migration Impact on Risk Levels)	83
4.5	Estimates of the ATT with Migrant Parents as the Control Group (Return Impact on Risk Levels)	83
4.6	Parental Migration and Rural Local Underage Children's Risk Preference Levels	84
4.7	Parental Migration and Rural Local Adult Children's Self-Employment (Probit)	86
B.1	Parental Migration and Rural Local Underage Children's Risk Preference Levels by Family Type	111
B.2	Parental Migration and Local Adult Children's Self-Employment by Family Type (Probit)	112

Chapter 1

Introduction

This dissertation mainly aims to contribute a better understanding on the migrants in the labor market of China using methods from labor economics and economic theory. According to the 2019 China Statistical Yearbook, there are 241 million floating population in China which accounts for almost twenty percent of the entire population. Migrants take such a great important component in the labor market of China, so how they perform is worthy of being studied and discussed. This study not only investigates who choose to migrate (Chapter 2), but also shows how networks (or choosing different types of destinations) affect migrants' labor market outcomes (Chapter 3). Besides these, we are also interested in discussing the intergenerational impact of parental migration on children's risk preferences (Chapter 4). For exploring the impact of having a migrant parent on children's risk preferences, we find China is suited to conduct the study. In China, we have a rich sample of migrant parents and left-behind children. Except for the large sample size, most of the migrant parents in China move within a country, so we can minimize the impact of culture shock on parents' risk preferences or on their children's as well.

Existing studies discuss the wage difference between urban locals and rural migrants (Song, 2016, Whalley and Zhang, 2007), but most of the results pay attention to the impact from rural *Hukou* status. From Chapter 2 of this dissertation, we explore how migrants earn differently from locals under the same condition, and explain why there is a wage difference between them. This study examines the impact of the endogeneity of the decision to migrate

on the wage differentials between migrants and non-migrants in China. By using both theoretical and empirical methods, we find that migrants are positively self-selected from the origins, and such selection effect changes by the wage difference between migrants' origins and destinations as well as migrants' moving costs. Besides this, we also contribute to the literature by showing the heterogeneity effect of selection among workers with different education and in different cities. The results indicate that there is more selection among higher educated groups, which is due to a greater educational return and greater competition for the according group in destinations. Moreover, the magnitude of positive selection increases as the city type changes from the first to the third-tier. The reason for such selection trend is that better cities have many other amenities than wage to attract best workers than the third-tier cities.

Chapter 3 investigates the network effect on migrants various labor market outcomes. Researchers suggest that networks affect migrants various labor market outcomes, such as migration decisions (Kerr and Mandorff, 2015, Patel and Vella, 2013), employment status (Giulietti et al., 2018, Munshi, 2003), as well as wage levels (Chen et al., 2018, Long et al., 2017, Patel and Vella, 2013, Zhang et al., 2012). In this chapter, we attempt to explain the network impact on various migrants' outcomes by using a theoretical model as well as the self-selection pattern among the migrants. Based on Beaman (2012) and Calvo-Armengol and Jackson (2004)'s framework, we study the network effects on migrants' employment and wage levels theoretically. With the theoretical model, we could contribute to the literature by explaining the different network effect on rural and urban migrants. The results indicate that networks exert positive impacts on rural migrants in terms of employment, but not on urban ones, which is due to the different quality drops (due to a positive selection) between rural and urban migrants when the networks increase. Such employment effects also lead rural migrants to face a more severe negative wage impact than urban migrants.

In Chapter 4, we explore the relationship between parental migration and left-behind children's risk preferences. As migrant parents, migration may not only affect migrant parents themselves, but could also have some influence on their left-behind children. Risk preferences are associated with various outcomes or decisions, such as smoking ([Anderson and Mellor, 2008](#)), drinking ([Anderson and Mellor, 2008](#), [Dave and Saffer, 2008](#)), health outcomes ([Anderson and Mellor, 2008](#), [Samek et al., 2021](#), [Sutter et al., 2013](#)), and risky food consumption ([Lusk and Coble, 2005](#)). To a certain extent, such findings prompt us to investigate what determines people's risk preferences and how they are determined. Various studies point out that there is an intergenerational transmission of risk preferences between parents and their children ([Alan et al., 2017](#), [Brown and Van der Pol, 2015](#), [Dohmen et al., 2012](#)). Based on this, we pay attention to the impact from having a migrant parents in this chapter, and point out that there is not only an intergenerational transmission of risk preferences between migrant parents and children, but migration action could also be influential in shaping left-behind children's risk preferences. Moreover, our findings also suggest that there is a positive relationship between parental migration and adult daughters' self-employment decisions.

Chapter 2

Self-Selection and the Wage of Migrants in China

(ABSTRACT)

This paper examines the impact of the endogeneity of the decision to migrate on the wage differentials between migrants and non-migrants in China. We find that migrants are self-selected from the upper tail of the income distribution in their home location. Consistent with a theoretical model of migration choice, we show that the size of the selection effect on wage depends on the wage differences between the prefectures of origin and destination as well as migration cost. The selection effect also differs among workers with different education and in different cities.

2.1 Introduction

The wage difference between migrants and urban locals in China has sparked heated discussion (Meng and Zhang, 2001, Song, 2016, Whalley and Zhang, 2007). Most of the studies indicate that rural-to-urban migrants have been discriminated against in the labor market of China either in terms of wage or job attainment. From Song (2016), we learn that workers with urban *hukou* earn about 50% more than rural *hukou* holders in state-owned enterprises. Meng and Zhang (2001) also find that more than 100% of the intraoccupational wage difference between urban locals and rural migrants is because of the different treatment against the two groups of workers. In fact, when we consider the discrimination against rural migrants, spotlight is always on the rural *hukou* status. However, there are two different ways of classifying *hukou*. Besides the rural/urban status, workers can also be classified into local (residents) or non-local (migrants) according to their place of household registration. If migrants and locals have different qualities, this could be another reason to generate wage difference between rural migrants and urban locals. In the labor market, who choose to migrate and who choose to stay is a self-selection question. Therefore, in this paper, we explore the feature difference between migrants and locals. Additionally, when we mention migrants, that always reminds us of rural-to-urban migrants. But in recent years, the number of urban-to-urban migrants is increasing gradually, which means migrant workers cannot be restricted to only rural-to-urban migrants.¹ Although a number of studies point out the wage differences between rural migrants and urban locals, we find only few studies pay attention to the quality difference between migrants and locals. Thus, we try to fill this gap, and study the who tend to be more likely to migrate in the labor market.

In this paper, we aim to investigate who choose to become migrants, and what is the return

¹Colas and Ge (2018) point out that there are only 21.4% of migrants with urban *hukou* status in 1990, but it has grown to 29.8% in 2005.

of migration. First of all, having a support from theory, we adopt a variant of the Roy model. Based on the theoretical framework, we study the different types of migration decisions and find out some relevant factors that could affect the decisions. The model suggests that, if there is a self-selection on migrants, the magnitude of the selection increases when the wage level in the origin rises. Also, migration cost has a positive impact on the magnitude of self-selection for migrant workers.

In the next section, we use China's 2005 mini-census to conduct an empirical study. The data set contains a rich set of information on individual level, such as people's characteristics, *hukou* status, and job outcomes, which allows us to explore the difference between migrants and locals in various aspects. Using a regression method, we find that urban *hukou* always has a positive impact on individual wage level. But to our surprise, migrants also have a positive wage premium, which indicates there should be a positive selection on migrants. By using the DFL approach (DiNardo et al., 1996), we take one more step to verify that the "best" workers migrate first. We then look over the impact of factors which we acquire from the theoretical model, and the results are consistent with our theoretical prediction by using Beijing and Shanghai sample. In addition, according to the education attainment level information, we are able to categorize the subjects into three groups in total. The findings suggest that the magnitude of positive selection diminishes as the education level decreases. Similarly, we also investigate the selection trend in three different types of cities. The results show that migrants have the lowest magnitude of positive selection in the first-tier cities, and selection becomes greater as the city type changes from the first to the third-tier.

In this study, we are able to supplement the existing literature in a sense. Most existing studies indicate that urban locals have a higher wage level than rural migrants (Fu and Ren, 2010, Liu, 2005, Meng and Zhang, 2001, Song, 2016, Zhang, 2010), but do not pay attention to the different qualities between migrants and locals. Although some researchers realize

that migrants and locals earn differently (Kuhn and Shen, 2015, Zhang and Wang, 2011), we still have unclear evidence on the root cause in the labor market of China. In this paper, we propose to study what kind of people choose to become migrants, and we are also interested in exploring the heterogeneous effect in various groups. This not only helps us have a better understanding on the discrimination against rural migrants, but also gives us a direction of *hukou* reformation. If the negative impact of rural *hukou* status on rural migrants still exists even after we control the quality discrepancy of locals and migrants, such a problem has to be resolved by a reformation of the *hukou* system.

In the next section, we introduce some related literature about the *hukou* system and migrants. Section 2.3 presents the theoretical frame work. Section 2.4 illustrates the data and some descriptive statistics. Section 2.5 and 2.6 provide some empirical implications and empirical results. The last section draws a conclusion.

2.2 Related Literature

The *Hukou* system is a kind of household registration system in China which requires every people to register right after their birth. After founding of the People’s Republic of China, in the 1950s, the *hukou* system starts to give effect. This household registration system was first mandated by law in 1958 (Fan, 2008), and its purpose was to control population mobility, stabilize society, and make food supply easier. However, due to the land reform and Great Leap Forward (*dayuejin*) in the mid-1950s, numbers of peasants lost their lands, became migrants in big cities, and food supply got much less than before. The Great Chinese Famine in 1959 also pushed the central government to send those migrants back to the rural area in order to balance the number of urban-rural population. Right after those big events, since 1963, the system started to classify people’s *hukou* into two statuses: agricultural

(rural) *hukou* or non-agricultural (urban) *hukou* (Chan and Zhang, 1999, Wang, 2005). This classification helps the government balance the population of rural and urban area more easily. From then on, the *hukou* system in China matured gradually.

Newborn children always inherited their mother's *hukou* status before 1998. However, the strict control on children's *hukou* status choice in mixed-*hukou* families has been loosened after 1998 (Colas and Ge, 2018, Wu and Treiman, 2004), which means children can inherit one of their parents' *hukou* status.² The way for the children in mixed-*hukou* families to get urban *hukou* becomes much easier now, but it is still a difficult thing for those rural-*hukou* families. Especially, getting an urban *hukou* in big cities is much more difficult (Zhang and Tao, 2012). Although, recently, the *hukou* system in some small cities has loosened the barriers for entering, it does not benefit most migrants at all. Since most migrants still want to settle down in big cities (Song, 2014), the *hukou* conversion problem is always a problem for those big cities. This policy change still cannot tackle the root cause of difference between rural and urban *hukou* in the labor market.

As mentioned above, one's *hukou* can be classified into a rural *hukou* or an urban *hukou*. However, besides the *hukou* status, there is always a *hukou* location which indicates one's place of residence. Therefore, in a specific place, one can also be categorized into local or non-local. After the State Council raises a new announcement in 2004, the classification of rural/urban *hukou* status is slowly fading out.³ Instead, nowadays, more and more people are classified into local or non-local by their locations of origin. To an extent, it alleviates the unfair treatment against rural-to-urban migrants.

The inequality between rural migrants and urban locals has been reflected in many aspects. *Hukou* mobility is mainly from rural status to urban status in China (Wu and Treiman, 2004),

²Mixed-*hukou* families are the families which mother's *hukou* status and father's *hukou* status are different.

³"Opinions of the State Council on Further Promotion of Reform of the Household Registration System" by the State Council.

but there is a strict restriction from the government on such conversion. The reason why people want to convert their *hukou* status is that there do exist a lot of difference between rural and urban *hukou* status. For example, education resource are unequally distributed between urban and rural area. Proportion of primary school teachers with junior college degrees and above is almost 20% lower in rural area when comparing with urban area (Dewen, 2003). But migrants' children can hardly receive compulsory education outside their *hukou* location (Song, 2014).

Fu and Ren (2010) examine the impact of people's *hukou* status on their income, and find that people with urban *hukou* status have higher income level. The reason is that education return is different for people with different *hukou* status. Based on Liu (2005)'s research, the author dose a pairwise comparison between rural and urban people in Beijing, and also concludes that the *hukou* policy limits rural people to access good quality education, convenient lives and more employment opportunities. The education return for people with rural *hukou* status is low, so that those people with rural *hukou* have low incentives to invest more on education. Liu presents that the *hukou* system indeed causes the inequality in the society.

In addition, public welfare in urban and rural areas also has difference. Based on Song (2014), the rural-to-urban migrants who cannot covered by the employment-based social insurance programs will receive the *hukou*-based social insurance, but the *hukou*-based social insurance program only allow those migrants get reimbursement in their *hukou* locations, and it is inconvenient for migrants to receive medical care. That can also be one of the reasons why people want to convert their *hukou* status instead of just staying as migrants.

Since the *hukou* system prevent labor mobility in China (Whalley and Zhang, 2007), workers with rural *hukou* can hardly have an equal treatment as urban workers both in the respect of wage and opportunity. Meng and Zhang (2001) mention that rural migrants are more

likely to take dirty works than urban residents. Those migrants are always considered to be unstable. However, [Zhang \(2010\)](#) finds that job mobility of temporary migrants with rural *hukou* is significantly lower than local residents. It is mainly due to the institutional discrimination against the migrants without urban *hukou*.⁴ As long as those migrants have enough assets, and stay longer in the city, the job changing rates increase.

Although rural-urban comparison is the mainstream, we are also able to look over some difference between migrants and locals in the labor market. For instance, migrant workers have difficulties to enjoy complete social insurance, and migrants' children could hardly have free compulsory education in the destinations. However, employers can barely solve such problems when hiring migrant workers. Therefore, to some extent, migrant workers are considered to be unstable comparing to local workers. We may expect that migrant workers are not competitive, but [Kuhn and Shen \(2015\)](#) mention that as long as migrants match the job requirements, employers are more likely to hire migrants instead of local residents in low-wage sectors. [Zhang and Wang \(2011\)](#) also find that there is no obvious prejudice against migrant workers with Shanghai sample in China. In fact, there exists more and more well-educated and urban migrants in the labor market, and migrants tend to be employed in better-paid industries than before ([Colas and Ge, 2018](#)). Thus we cannot simply regard migrants as rural-to-urban migrants as previous studies. Even though migrant workers are not shown cold shoulder in some cases, we still do not have a clear idea about the fundamental causes and labor market treatment against migrants in variant groups. Therefore, in this paper, we would like to provide some answer to the questions to meet this need.

If the *hukou* system generates an inequality as the studies suggest, we indeed have to minimize such impact by adopting a better policy. In fact, from 2005, the central government tries

⁴[Song \(2016\)](#) finds that the institutional discrimination mainly exists in those high-wage state owned enterprises, which means that SOEs are more likely to hire local residents when comparing with private enterprises.

to implement a new *hukou* system in order to weaken the function of the *hukou* in China and eliminate rural/urban *hukou* gradually. [Bosker et al. \(2012\)](#) study the consequence of relaxing the *hukou* restrictions and conclude that the relaxation is positively related to the economic development in China. But [Wing Chan and Buckingham \(2008\)](#) mention that those reformation is not actually abolishing the *hukou* system. The central government tries to pass the rights to the local governments, and eventually makes migrants become even harder to settle down in the destination cities. Therefore, how to relax the *hukou* restriction is also worth discussing. It has been more than half a century since the *hukou* system was first set up and played a role in China. It is undeniable that the *hukou* system is important in keeping the order and peace in the society, and providing public services. As the nation keeps developing, the *hukou* need reform and development as well. Nowadays, more and more people choose to migrate, the *hukou* system have to adjust itself to serve the whole nation better. From this study, we can also provide a more accurate rural *hukou* impact by considering and controlling the quality difference between migrants and locals. After having the rural *hukou* impact on migrants' labor market outcomes, the government would also have a better idea on resolving inequality problems which generated by the *hukou* in the labor market.

2.3 The Model

Based on [Borjas \(1987\)](#)'s framework, we build a model with two different types of workers who want to migrate to a developed place. For simplicity, we assume there are two cities 0 and 1. City 0 is the hometown for migrants while city 1 is the place which migrants move to. There are two types of workers: type 1 and type 2 in city 0 which denote workers with rural *hukou* and those with urban *hukou*.

Workers in city 0 and city 1 have earnings as:

$$\ln w_{0i} = \mu_{0i} + \varepsilon_{0i}, \quad (2.1)$$

$$\ln w_1 = \mu_1 + \varepsilon_1, \quad (2.2)$$

where $i = 1, 2$ denotes the type of migrant workers, $\varepsilon_{0i} \sim N(0, \sigma_{0i}^2)$, $\varepsilon_1 \sim N(0, \sigma_1^2)$, and ρ_i is the correlation coefficient between ε_{0i} and ε_1 , and $\rho_i = \frac{\sigma_{1,0i}}{\sigma_1 \sigma_{0i}}$.

Let $d_i = 1$ if type i workers decide to migrate. The probability is defined as:

$$\begin{aligned} \Pr(d_i = 1) &= \Pr(\ln w_1 > \ln w_{0i} + C_{0i}) \\ &= \Pr(\tau_i < \frac{\mu_1 - \mu_{0i} - C_{0i}}{\sigma_{vi}}), \end{aligned} \quad (2.3)$$

where $\tau_i = \frac{\varepsilon_{0i} - \varepsilon_1}{\sigma_{vi}}$, C_{0i} indicates the moving costs for type i workers, σ_{vi} is the standard deviation of $(\varepsilon_{0i} - \varepsilon_1)$, and $\tau_i \sim N(0, 1)$.

Based on the migration probability, the conditional expected wage in city 1 for migrants is

$$\begin{aligned} E(\ln w_1 | d_i = 1) &= \mu_1 + E(\varepsilon_1 | d_i = 1) \\ &= \mu_1 + \frac{\sigma_1 \sigma_{0i}}{\sigma_{vi}} (\rho_i - \frac{\sigma_1}{\sigma_{0i}}) E(\tau_i | \tau_i < \frac{\mu_1 - \mu_{0i} - C_{0i}}{\sigma_{vi}}). \end{aligned} \quad (2.4)$$

Similarly, for migrants, the conditional expected wage in their hometown is

$$E(\ln w_{0i} | d_i = 1) = \mu_{0i} + \frac{\sigma_1 \sigma_{0i}}{\sigma_{vi}} (\frac{\sigma_{0i}}{\sigma_1} - \rho_i) E(\tau_i | \tau_i < \frac{\mu_1 - \mu_{0i} - C_{0i}}{\sigma_{vi}}), \quad (2.5)$$

while the conditional expected wage in their hometown for non-migrants is

$$E(\ln w_{0i}|d_i = 0) = \mu_{0i} + \frac{\sigma_1\sigma_{0i}}{\sigma_{vi}}\left(\frac{\sigma_{0i}}{\sigma_1} - \rho_i\right)E(\tau_i|\tau_i \geq \frac{\mu_1 - \mu_{0i} - C_{0i}}{\sigma_{vi}}). \quad (2.6)$$

According to [Borjas \(1987\)](#)'s paper, there is a positive selection when $E(\ln w_1|d_i = 1) > \mu_1$ and $E(\ln w_{0i}|d_i = 1) > \mu_{0i}$, which means the 'best' workers are selected to the destinations. Negative selection occurs in the opposite situation, which requires $E(\ln w_1|d_i = 1) < \mu_1$ and $E(\ln w_{0i}|d_i = 1) < \mu_{0i}$. In negative selection, the 'worst' workers are most likely to move.

Let us discuss the positive selection first. Since we know that $E(\tau_i|\tau_i < \frac{\mu_1 - \mu_{0i} - C_{0i}}{\sigma_{vi}}) < 0$ and $E(\tau_i|\tau_i \geq \frac{\mu_1 - \mu_{0i} - C_{0i}}{\sigma_{vi}}) > 0$, the necessary conditions for having positive selection are:

$$\begin{aligned} \rho_i - \frac{\sigma_1}{\sigma_{0i}} &< 0, \\ \frac{\sigma_{0i}}{\sigma_1} - \rho_i &< 0, \end{aligned}$$

which indicates $\sigma_{0i} < \sigma_1$. In positive selection, we have $E(\ln w_1|d_i = 1) > \mu_1$, $E(\ln w_{0i}|d_i = 1) > \mu_{0i}$, and $E(\ln w_{0i}|d_i = 0) < \mu_{0i}$.

Let Q denote the difference between migrants' conditional expected wage and expected wage in city 1, and it measures how large is the magnitude of the selection. Since $(-\tau_i) \sim N(0, 1)$, we have

$$\begin{aligned} Q &= \frac{\sigma_1\sigma_{0i}}{\sigma_{vi}}\left(\rho_i - \frac{\sigma_1}{\sigma_{0i}}\right)E(\tau_i|\tau_i < \frac{\mu_1 - \mu_{0i} - C_{0i}}{\sigma_{vi}}) \\ &= \frac{\sigma_1\sigma_{0i}}{\sigma_{vi}}\left(\frac{\sigma_1}{\sigma_{0i}} - \rho_i\right)\lambda(k), \end{aligned}$$

where $k = \frac{\mu_1 - \mu_{0i} - C_{0i}}{\sigma_{vi}}$.

How does the wage difference Q change with the origins' expected wage μ_{0i} is shown as

below:

$$\frac{\partial Q}{\partial \mu_{0i}} = \frac{\sigma_1 \sigma_{0i}}{\sigma_{vi}^2} \left(\frac{\sigma_1}{\sigma_{0i}} - \rho_i \right) \frac{\partial \lambda(k)}{\partial (-k)}.$$

$\frac{\partial \lambda(k)}{\partial (-k)}$ is always larger than zero since we are measuring how does $E(-\tau_i | -\tau_i > -k)$ change with $(-k)$ when $(-\tau_i) \sim N(0, 1)$. If there is a positive selection, which means $(\frac{\sigma_1}{\sigma_{0i}} - \rho_i) > 0$, then we have $\frac{\partial Q}{\partial \mu_{0i}} > 0$.

The result tends to be the same when we differentiate the cost C_{0i} .

$$\frac{\partial Q}{\partial C_{0i}} = \frac{\sigma_1 \sigma_{0i}}{\sigma_{vi}^2} \left(\frac{\sigma_1}{\sigma_{0i}} - \rho_i \right) \frac{\partial \lambda(k)}{\partial (-k)}.$$

If $(\frac{\sigma_1}{\sigma_{0i}} - \rho_i) > 0$, then $\frac{\partial Q}{\partial C_{0i}} > 0$, which means the magnitude of a positive selection increases as the cost of moving increases.

After discussing positive selection, we have a quick look on negative selection. The necessary conditions for negative selection are:

$$\begin{aligned} \rho_i - \frac{\sigma_1}{\sigma_{0i}} &> 0 \\ \frac{\sigma_{0i}}{\sigma_1} - \rho_i &> 0, \end{aligned}$$

which leads to $\sigma_1 < \sigma_{0i}$. We conclude that the as the μ_{0i} increases, there is more negative selection, and the one who has a higher moving cost have more negative selection.

We assume that rural migrants (type 1) have higher or equal moving costs comparing with urban migrants (type 2) both in psychological and physical aspects. In addition, the expected wage for rural migrants (type 1) is less than or equal to urban migrants (type 2), which means $\mu_{01} \leq \mu_{02}$, if they do not move. Therefore, we know $C_{01} \geq C_{02}$, and $(\mu_1 - \mu_{01}) \geq (\mu_1 - \mu_{02})$

from the assumptions.

Under the equal costs assumption, the type who has smaller wage difference (urban) have more selection. When the wage difference between city 1 and city 0 for both types are equal, $(\mu_1 - \mu_{01}) = (\mu_1 - \mu_{02})$, then the one who has higher moving costs C has a smaller value for $(\frac{\mu_1 - \mu_{0i} - C_{0i}}{\sigma_{vi}})$, which means the difference between conditional expected wage and expected wage in city 1 is larger. Therefore, magnitude of selection becomes larger for the type who has higher moving costs (rural).

We tend to interpret the implication of the positive selection as follows. If there is a positive selection, that means destination cities always select the best workers from other cities. When the mean wage level difference between origin and destination cities is small, fewer less than perfect workers would like to migrate, since there is no higher wage level for them. As a result, the magnitude of selection (difference between migrants wage and destination's mean wage) is greater for the migrants from well-paid cities. Moreover, if the moving cost is high, such less skilled workers would less likely to move, since they are hard to get compensated. So it is also reasonable that the magnitude of positive selection is greater for the migrants with higher moving costs.

In the context of negative selection, if the destination cities always select the worst workers from other cities, migrants from a similar-wage-level city become the worst of worst migrants. Migrants from much more ill-paid origins are more likely to have skilled workers in it. Also, there are fewer skilled workers in a high moving cost group. Therefore, if there is a negative selection, the magnitude of negative selection is higher on the migrants from similar-wage-level cities, or from high moving cost cities.

2.4 Data and Descriptive Statistics

The data which we are using in this paper comes from China's 2005 mini-census. This 1% national population survey was conducted by the National Bureau of Statistics of China, and does not include the population outside the mainland of China. The data set covers information such as gender, age, marital status, education attainment level, working hours, monthly wage, ownership and industry. What's more, the *hukou* status, which is important to this study, is also included. With the various information, the data set allows us to conduct a study on the difference between local and migrant workers in the labor market.

In this paper, we use the people with ages between 16 and 60 in the data set as our sample. Those who are physically unable to work, do not have information on *hukou* registration place, self-employed, and family business helpers are dropped from the set. We also exclude local rural workers from the sample. The reason is that we find over 60% of local rural workers live by farming. We all know that income from land tillage is somewhat different from ordinary income, so such a large scale of local rural workers may not be our object of interest. What's more, migrants are more likely to move to urban areas, and the main competitors for them are local urban workers. Therefore, we would like to mainly study the difference between local urban workers and two different kinds of migrants in this paper. We define a worker as a migrant if she works in a different county from the place of registration. Education years are also assigned manually: 0 year for no education, 3 years for primary school drop-outs, 6 years for primary school graduates, 7.5 years for middle school drop-outs, 9 years for middle school graduates, 10.5 years for high school students and drop-outs, 12 years for high school graduates, 13.5 years for junior college students and drop-outs, 16 years for junior college graduates, 14 years for college students and drop-outs, 16 years for college graduates, 18 years for graduate school students and drop-outs, 20 years for graduate school

graduates. In addition, based on the information of working hours and monthly wage, we are able to get hourly wage from the data set.

Ownership is categorized into state-owned enterprises, private enterprises, and other ownership in this paper. Note that foreign owned enterprises are not defined separately in the data set, so we have to classify such firms to other ownership. In terms of industry, besides the primary industry, manufacturing industry, and construction industry, we separate the service industry into advanced and basic according to the previous study (Ge and Yang, 2009).

Table 2.1 shows the locals and migrants' basic characteristics by *hukou* status. This table provides the information that migrants are younger and less likely to be married for both urban and rural subjects. In addition, migrant subjects tend to have higher working hours. Meanwhile, urban migrant workers have slightly higher average education years, which probably makes a contribution to their higher hourly mean wage. Such finding satisfies our conjecture that migrant workers would like to promote their competitiveness by enhancing education attainment level. Therefore, whether there is any discrimination against migrants or not in the labor market becomes more and more ambiguous.

Table 2.1 also presents the information of the subjects distribution in different ownership enterprises, industries, and various welfare program participation. From the table we know that local workers are mainly working in state-owned enterprises, and migrants work mostly in private enterprises. Meanwhile, 41% of urban local workers choose to work in the advanced service industry. As for social welfare program participation status, urban subjects are much more likely to have necessary insurance than rural subjects, which indicates urban workers probably have more stable jobs. For all migrants, no matter with urban *hukou* or rural *hukou*, working in the manufacturing and basic service industries is the mainstream.

Table 2.1: Characteristics of Locals and Migrants with Different *Hukou* Status

	Local	Non-local	
	Urban	Urban	Rural
Male	0.58	0.57	0.55
Age	37.75	31.41	29.19
Married	0.84	0.60	0.59
Years of Education	12.36	12.53	8.76
Weekly Working Hours	44.83	47.93	54.74
Hourly Wage	7.12	9.89	4.65
<i>Ownership</i>			
State-Owned Enterprise	0.67	0.28	0.11
Private Enterprise	0.26	0.54	0.67
Other Ownership	0.08	0.17	0.22
<i>Industry</i>			
Primary Industry	0.06	0.02	0.06
Manufacturing Industry	0.27	0.33	0.58
Construction Industry	0.04	0.06	0.08
Advanced Service Industry	0.41	0.24	0.05
Basic Service Industry	0.22	0.35	0.22
<i>Social Welfare Program Participation</i>			
Having Unemployment Insurance	0.50	0.35	0.08
Having Pension Insurance	0.68	0.49	0.16
Having Employment Injury Insurance	0.70	0.48	0.20
Number of Respondent	246,766	21,984	92,444
Data comes from 2005 1% national population survey			

From the descriptive statistics above, we notice that even though migrant workers are less competitive in some aspects, general migrant workers are not absolutely ill-paid. Thus, in the next section, we provide some empirical evidence about the existence of wage difference between local and migrant workers, and discuss where the difference comes from. Furthermore, we present how the wage difference between local and migrant workers varies among different educational attainment level groups.

2.5 Empirical Implications

2.5.1 Empirical Specification

First of all, we have to explore whether the wage difference between local and migrant workers still exists when we control for other possible influencing factors. We adopt the classical Mincerian equation (Mincer, 1958) to study the factors which affect the wage of workers in this section. Distinguished from the previous studies, we introduce two more dummies, *migrant* term and *urban* term, to indicate subjects' *hukou* status in the specification. The basic specification can be given as

$$\ln W_i = \beta_0 + \mathbf{X}_i\delta + \beta_1 \times \text{MIGRANT}_i + \beta_2 \times \text{URBAN}_i + \varepsilon_i, \quad (2.7)$$

where W_i is individual i 's hourly wage, $\mathbf{X}_i$ is a vector which represents personal characteristics such as gender, experience, education level, sectors, and so on, MIGRANT_i is a dummy variable that indicates if i is a migrant or not, URBAN_i is also a dummy variable for i 's urban *hukou* status, and ε_i is an error term with zero mean.

The results for equation (2.7) are shown in Table 2.2. We present the baseline results in the first column, then control for industries and ownership in the next two columns respectively. For eliminating the effect from unequal education resources, we use the pupil-teacher ratio and education funding data from China Statistical Yearbook to control for the education quality in the last column.⁵

From Table 2.2 we find that both male and married status have positive effects on individual wage level, and China's non-*Han* minorities tend to earn less when comparing to *Han* people.

⁵Following the Card and Krueger (1992a)'s measurement, we use pupil-teacher ratio and education funding as proxies for education quality. For controlling provinces' general education quality, pupil-teacher ratio in primary school level, middle school level, and high school level are all included in the specification.

We categorize the educational attainment levels into three groups in total. The compulsory education group represents people with middle school educational attainment level or below, the high school group is for people receiving high school education, while the college group includes the subjects with education level above high school (such as junior college, university, and graduate school). As expected, the more educated, the higher the wage obtained. Although the difference between the compulsory education and the high school group is only around 30%, education return for the college group is significantly higher than the other two.

Similar to the previous mainstream studies, we find that urban workers are more likely to have higher wages than rural ones. In order to avoid the effect from type of jobs, we control for the industries and ownership as well. Although the coefficient of urban *hukou* changes slightly after including industries and ownership, we still observe a positive effect on wage. In addition, we use primary industry as a reference group, and the results indicate that subjects who are in the advanced service industry earn the most. In terms of ownership, the private enterprises are treated as a reference group. Compared to other enterprises, wage of state-owned enterprises is the highest.

Another important effect on individual wage level is from migrant identity. Wage differences between migrants and locals do exist. But to our surprise, migrants are not less competitive, instead, they have higher wages than average. The coefficient of the *migrant* term in column (1) is positive, which means that migrants are probably a positively selected group, which exactly follows our positive selection story in section 3. When we control for industries and ownership, we find the positive selection even increases. From column (3) of the table we find that urban migrants have the highest wage premium of 0.482, urban locals get 0.277, while rural migrants enjoy the lowest wage premium of 0.205.

In column (4), we find that when pupil-teacher ratio goes up, which means education quality

decreases, it has a negative effect on individual wage level. However, the impact is a million to one. Similarly, the funding of education cannot influence the wage level very much as well. After controlling for pupil-teacher ratio and education funding, we still have a positive selection on migrant workers. That once again confirms the results in columns (1) to (3) in Table 2.2⁶.

According to Borjas (1987)'s paper, there is a positive selection when $E(\ln w_1|d_i = 1) > \mu_1$ and $E(\ln w_{0i}|d_i = 1) > \mu_{0i}$, which means the “best” workers are selected to the destinations. Negative selection occurs in the opposite situation, which requires $E(\ln w_1|d_i = 1) < \mu_1$ and $E(\ln w_{0i}|d_i = 1) < \mu_{0i}$. In negative selection, the “worst” workers are most likely to work in the destination cities.

Currently, we have $E(\ln w_1|d_i = 1) > \mu_1$ based on Table 2.2. Therefore, at this point, what we try to explore is: $E(\ln w_{0i}|d_i = 1) > \mu_{0i}$. According to DiNardo et al. (1996)'s DFL approach, we show the counterfactual density of wages for both rural and urban migrants. The blue line in the first panel of Figure 2.1 is the wage density of urban locals. The red line is the counterfactual wage density of urban migrants who were paid as urban locals' skill prices. For a more clear comparison, we show the difference⁷ between urban locals' wage density and urban migrants who were paid as urban locals' wage density in the second panel of Figure 2.1. Figure 2.2 shows the similar graph for rural locals and the according counterfactual density. We realize that if migrants did not migrate, their wages would be higher than current local workers, which indicates migrants are the “best workers”. Therefore, we conclude that migrants are self-selected from the upper tail of the income distribution in their origins.

Now that we find positive selection exists in migrants, we are interested in verifying whether

⁶Although it is possible that a migrant choose a ill-paid job which provide a chance to covert to a local *hukou*, this is a nationwide rare case. Moreover, converting *hukou* should be much more difficult in 2005.

⁷Difference is defined as locals' wage minus migrants' wage which is paid as locals' skill prices.

Table 2.2: Baseline Results on *Hukou* Effects

	(1)	(2)	(3)	(4)
Male	0.187*** (0.002)	0.178*** (0.002)	0.176*** (0.002)	0.176*** (0.002)
Working Experience	0.022*** (0.000)	0.019*** (0.000)	0.018*** (0.000)	0.018*** (0.000)
Working Experience Square/100	-0.042*** (0.001)	-0.034*** (0.001)	-0.033*** (0.001)	-0.034*** (0.001)
Married	0.010*** (0.003)	0.028*** (0.003)	0.025*** (0.003)	0.026*** (0.003)
Ethnicity- <i>Han</i>	0.069*** (0.004)	0.055*** (0.004)	0.056*** (0.004)	0.043*** (0.004)
Migrant	0.154*** (0.004)	0.167*** (0.004)	0.205*** (0.005)	0.243*** (0.005)
Urban <i>Hukou</i>	0.270*** (0.005)	0.278*** (0.005)	0.277*** (0.005)	0.257*** (0.005)
High School	0.359*** (0.003)	0.300*** (0.003)	0.279*** (0.003)	0.273*** (0.003)
College	0.873*** (0.003)	0.757*** (0.003)	0.723*** (0.003)	0.716*** (0.003)
Manufacturing Industry		0.555*** (0.005)	0.508*** (0.005)	0.518*** (0.005)
Construction Industry		0.623*** (0.007)	0.572*** (0.007)	0.580*** (0.007)
Advanced Service Industry		0.666*** (0.006)	0.586*** (0.006)	0.593*** (0.006)
Basic Service Industry		0.521*** (0.005)	0.496*** (0.005)	0.498*** (0.005)
State-Owned Enterprises			0.164*** (0.003)	0.162*** (0.003)
Other Ownership			0.091*** (0.003)	0.089*** (0.003)
Province Dummies	Yes	Yes	Yes	Yes
Education Quality	No	No	No	Yes
Constant	1.011*** (0.008)	0.503*** (0.010)	0.478*** (0.010)	0.596*** (0.018)
Observations	361194	361194	361194	361194
r2_a	0.376	0.412	0.419	0.422

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

we can meet the necessary conditions that we have in the model by using the data in hand.

From the model in section 3, we know that there exists positive selection only if $\sigma_{0i} < \sigma_1$,

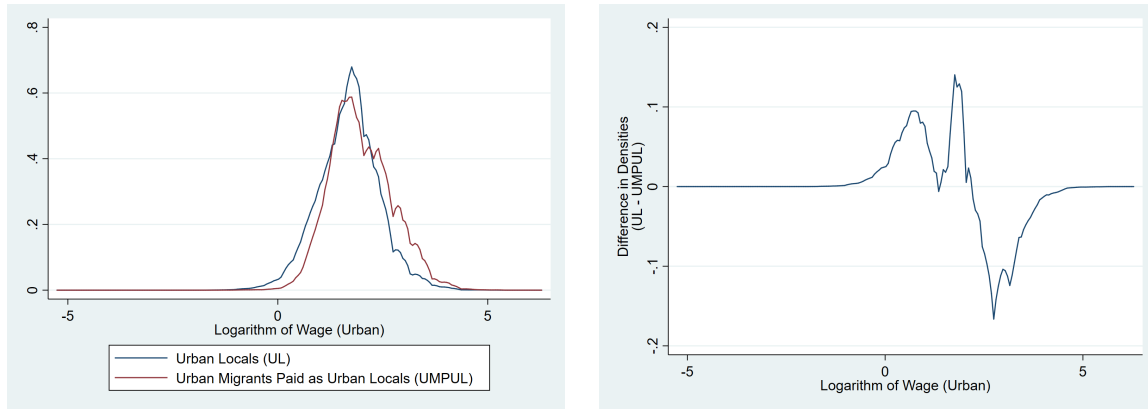


Figure 2.1: Urban Migrants

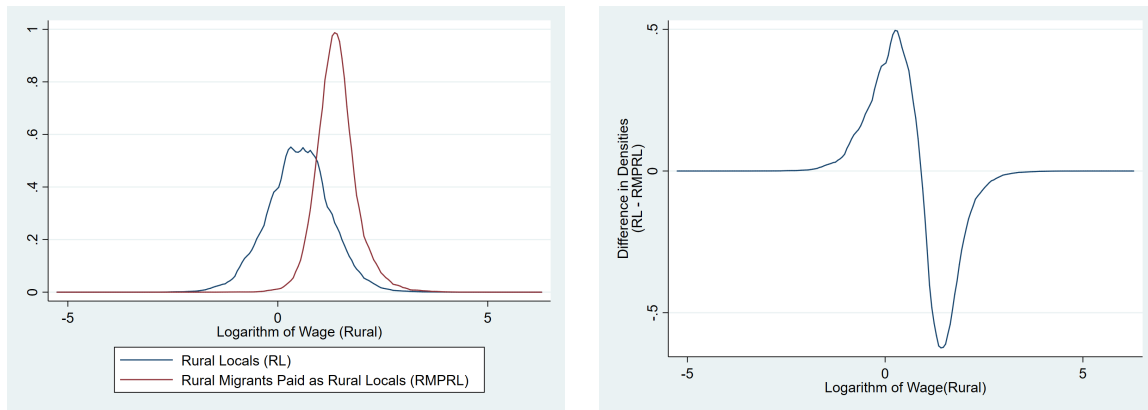


Figure 2.2: Rural Migrants

which means the wage standard deviation of destination cities will be greater than origins. Therefore, cities with a higher proportion of migrants would have higher wage standard deviation. Figure 2.3 presents the relationship between city-level migrant percentage and wage standard deviations. The positive coefficient in the figure gives us the same picture as the model indicates.

After meeting the necessary condition, we would like to see which factors could affect the magnitude of positive selection with our data. In the following sections, we try to provide some empirical evidence on the factors which have impacts on the magnitude of positive

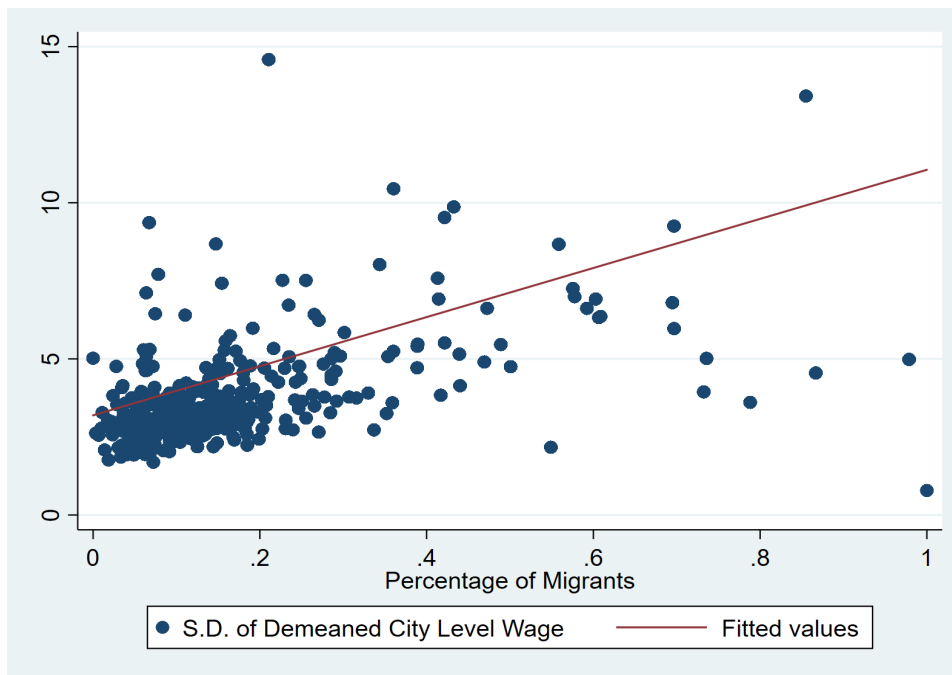


Figure 2.3: Scatter Plot of City Wage Standard Deviation

selection based on the theoretical framework.

2.5.2 Positive Selection and Origins' Wage

From the model we know that, in a specific destination, the magnitude of positive selection is greater for migrants from high wage level cities. In real life, if an origin has a similar wage level to the destination city, workers from that origin would not have much incentive to change their lives. Since we know that destination cities always select best workers from the origins, what are the migration decisions of those less than perfect workers is the crux of the matter. What kind of workers are more likely to be attracted by a high wage level? Are the workers from similar wage level cities or from low pay cities? We would like to suspect that there are more less skilled migrants from low wage level cities which are attracted by the obvious wage difference. In the following text, we find an evidence from the data in hand.

Using the data, we first propose to explore the relationship between wage difference and the magnitude of positive selection. We use the specification as

$$\ln W_i = \beta_0 + \mathbf{X}_i\delta + \beta_1 \times URBAN_i + \sum_j \gamma_j D_{origin_{ij}} + \varepsilon_i, \quad (2.8)$$

where all the variables have the same setting as in equation (7), and $D_{origin_{ij}} = 1$ indicates individual i 's origin is $city_j$. Origin dummies are employed to replace the *migrant* term in the previous tables, so that we get coefficients for each origin city separately from the regression. γ_j represents the wage premium of different origins, and we would like to define it as the magnitude of positive selection.

The following plots indicate the relationship between the magnitude of positive selection and origins' wages. We take the sample from 2 destination cities: Beijing and Shanghai as examples. In Figure 2.4(a), we regress with Beijing sample to obtain γ_j by using equation (8). Regression is done with Shanghai sample in the same way, and the according plot can be found in Figure 2.4(b). Since the sample has a specific destination, the origin with higher wage level has a smaller wage difference from the destination. The positive coefficients for origin dummies first give us an evidence of positive selection. Then we observe positive relationships in both panels, which means that the magnitude of positive selection rises as the wage differences between origins and the destination decrease. We use Beijing and Shanghai sample to find exactly the same result as shown in the theory.

2.5.3 Positive Selection and Moving Cost

In a destination city, there is a common phenomenon that a migrant could get a job with a higher wage level than in the origin but the wage he or she obtains cannot afford the living expense at all. Thus, as a result, the migrant has to give up surviving in the destination.

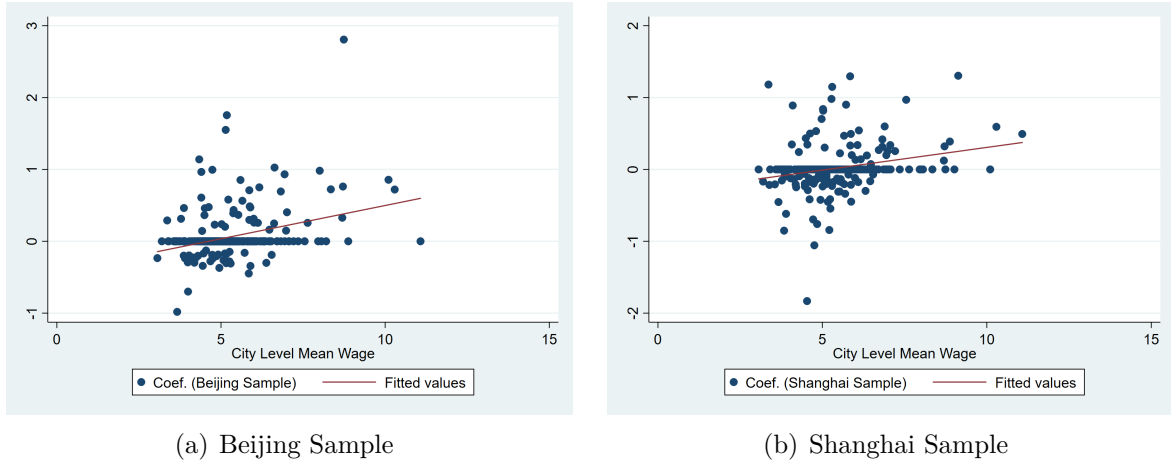


Figure 2.4: Relationship between Positive Selection and Origins' Wage

Moving cost is an important factor that may affect workers' migration choices. Even though there is enough job opportunities in destination cities, moving cost sets the bar too high. Unless the wage covers the moving cost, living as a migrant worker is not attractive.

The theoretical model also indicates that the magnitude of positive selection is greater for migrants with higher moving costs. Then can we find any evidence from the data? What to be the proxy for moving cost is our primary task. According to previous studies, we use the distance between an origin and a destination as the moving cost. We plot the γ_j , which we acquire from equation (8), and distance to the destination in Figure 2.5. X-axis in Figure 2.5 are distances from the origin to the destination. Figure 2.5(a) and 2.5(b) are for Beijing and Shanghai sample respectively.

The positive γ_j in Figure 2.5(a) gives an evidence of positive selection, and the positive trend here indicates the magnitude of positive selection gets greater as the moving cost increases. However, positive trend is not significant in Figure 2.5(b), which means we cannot conclude whether moving cost affects the magnitude of positive selection in Shanghai sample.

In order to study the effects of wage difference and moving cost more precisely, we regress

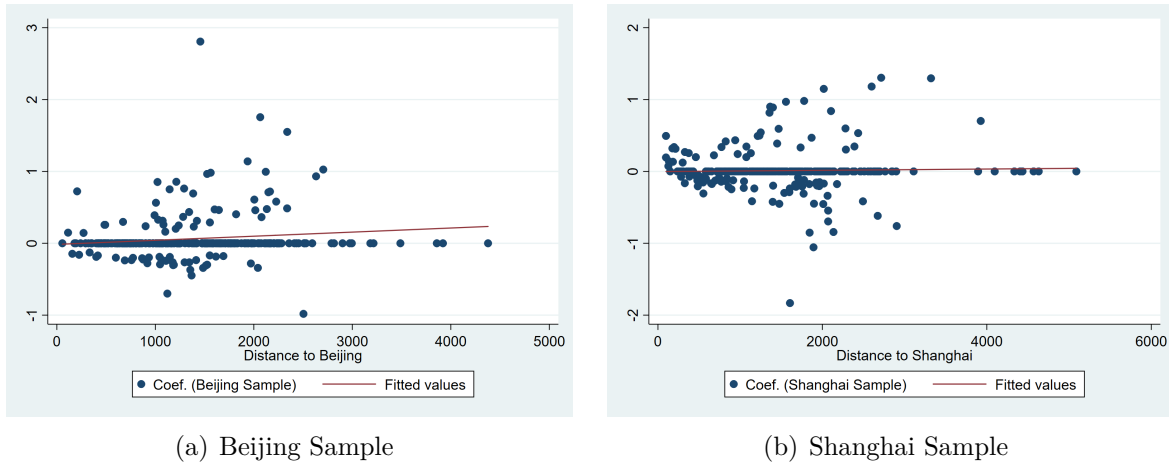


Figure 2.5: Relationship between Positive Selection and Origins' Distance

γ_j on origins' wage and origins' distance in Table 2.3. From Table 2.3 we find that wage difference has a significant effect for both Beijing and Shanghai Sample, but the moving cost seems not so influential. The results indicate that as the wage difference between origins and destinations decreases, the magnitude of positive selection goes up. Moreover, the wage difference effect is a little bit larger in Beijing than in Shanghai. Although the results are not indicating there is no moving cost effect at all in the labor market of China, we find the wage difference effect dominates at least in the two main destination cities.

Table 2.3: Effect of Origins' Wage and Distances on Positive Selection

	(1)	(2)
	Beijing Sample	Shanghai Sample
Origins' Wage	0.102*** (0.016)	0.069*** (0.014)
Distance to Beijing	0.000** (0.000)	
Distance to Shanghai		0.000 (0.000)
Constant	-0.544*** (0.093)	-0.387*** (0.087)
Observations	291	304
r2_a	0.126	0.065

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

2.5.4 Positive Selection and Education Attainment Level

Besides wage difference and moving cost, another possible determinant of the magnitude of positive selection could be education attainment level. Workers with different education attainment levels will apply for different jobs. Some jobs are attractive for migrants with high education level while others may not. We all know that high education attainment level somehow represents the ability, and workers with high ability always look for jobs with high wage level. Such job hunters are less likely to be restricted to the jobs in their origins. We expect that the magnitude of positive selection is high for jobs with tremendous competition. Then, in effect, what kind of jobs are having most selected workers? In the meantime, workers in which education level group are facing the highest magnitude of positive selection?

For accuracy, we use the data in hand to run a regression. The corresponding specification can be expressed as

$$\ln W_i = \beta_0 + \mathbf{X}_i\delta + \beta_1 \times \text{MIGRANT}_i \times \text{EDU}_i + \beta_2 \times \text{URBAN}_i \times \text{EDU}_i + \varepsilon_i, \quad (2.9)$$

where most of the variables have similar meanings to the previous equations, $\mathbf{X}_i$ still represents personal characteristics (*migrant* and *urban hukou* are also included), and EDU_i stands for i 's education attainment level.

The according results are shown in Table 2.4. From Table 2.4 we notice that effects of gender, working experience, marital status, and ethnicity are similar to the baseline results. Wage premium of different industries and ownership still remain similar. Construction and advanced service industries have higher wage premium, while manufacturing and basic service industries are relatively ill-paid. In terms of ownership, state-owned enterprises are probably keenly sought by workers due to the high wage premium.

Interaction terms of education attainment level and *hukou* status have been controlled in equation (9), which means we are able to find different magnitude of migrant or urban *hukou* premium among various education level groups. *Migrant* term and the first two interaction terms ($Education\ Level \times Migrant$) in Table 2.4 show that there exists positive selection, and the magnitude of positive selection becomes greater when education level increases. In the compulsory education level group, migrants have a wage premium of 0.100 comparing to local workers. For the high school group, migrant sample earn 0.176 more than local sample, while in the college group, migrant sample gets 0.317 wage premium comparing to local one. Such phenomenon means that highly-educated groups are more likely to move, and they are selected groups in the labor market.

As shown in Table 4, urban *hukou* has a positive effect, and such effect also goes up as the education level increases. Workers with urban *hukou* enjoy a 0.193 wage premium in the compulsory education group, and has a wage premium of 0.224 in the high school group, while the wage premium increase to 0.266 in the college group. We also find that, from the compulsory education group to the college group, the wage premium of migrant identity increases faster than that of urban *hukou*. In the most educated group, rural migrants even have higher wage premium than urban locals, which somewhat suggests knowledge removes the barrier of *hukou* status⁸.

For both rural and urban workers, we realize that there always exists higher magnitude of positive selection for the higher education attainment level group. Why there is more positive selection in the high education level group? To answer this question, we would like to know whether the education returns vary in different destination cities. Education returns

⁸Wage difference between rural and urban *hukou* holders somewhat also reflects the education quality difference between rural and urban areas.

Table 2.4: *Hukou* Effects by Education Level

	(1)
Male	0.175*** (0.002)
State-Owned Enterprises	0.166*** (0.002)
Other Ownership	0.091*** (0.003)
High School	0.231*** (0.011)
College	0.620*** (0.016)
Migrant	0.100*** (0.007)
High School $\times$ Migrant	0.076*** (0.010)
College $\times$ Migrant	0.217*** (0.010)
Urban <i>Hukou</i>	0.193*** (0.007)
High School $\times$ Urban <i>Hukou</i>	0.031*** (0.011)
College $\times$ Urban <i>Hukou</i>	0.073*** (0.016)
Constant	0.569*** (0.010)
Observations	361194
r2_a	0.420

Standard errors in parentheses

Working experience, working experience square

marital status, ethnicity, industry, province are also controlled

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

are acquired from the equation

$$\ln W_i = \beta_0 + \mathbf{X}_i \delta + \beta_2 \times URBAN_i + \sum_j \gamma_j D_{city_{ij}} + \beta_3 EDU_i + \sum_k \phi_k D_{city_{ik}} \times EDU_{ik} + \varepsilon_i, \quad (2.10)$$

where $D_{city_{ij}}$ is a dummy variable which indicates individual i 's destination is city j , and other variables remain the same settings as previous specifications.

The plot in Figure 2.6 shows the relationship between education returns and migrants pro-

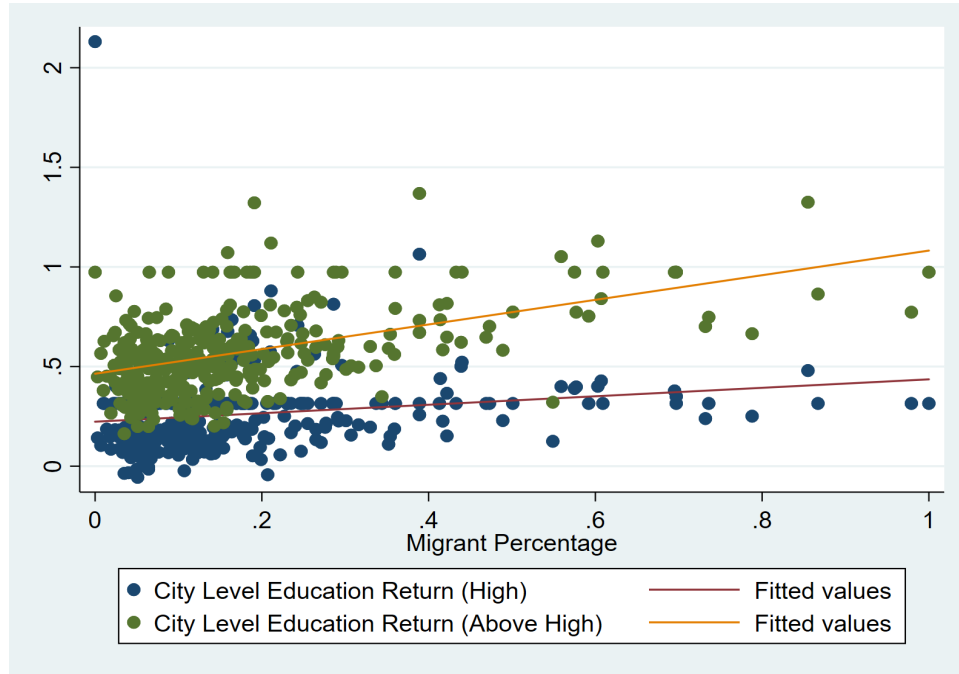


Figure 2.6: Relationship between Education Return and Migrants Proportions

portions. The blue dots are for the high school group, and the green dots are for the college group. As shown in the figure, for both high school and college groups, the education returns go up as the city level migrants' proportion increases. Therefore, according to the information from Figure 2.6, we are able to conclude that destination cities are more likely to have higher education returns than origin cities. Such education returns are even higher for higher education level group. In addition, as can be seen from Figure 2.6, the difference of education return among cities is higher in the college group. As a result, more and more highly-educated workers prefer to move to destination cities to enjoy high education return rather than settle down in their origins. Especially, workers in the college group face more serious competition when migrate to destination cities. Therefore, such phenomenon led to the magnitude of positive selection becomes higher as the education attainment level increases.

2.6 Selection Story in Different Types of Cities

Except the wage difference effect, moving cost effect, and the education attainment level effect, there could also be other factors which influence the migration choices. Different cities have different economic environment. Some cities may attract and prefer migrants, while others probably are more suitable for local workers to work in. Then what kind of cities are more attractive to migrants? If there still exists positive selection, how is the magnitude of selection different among cities? Whether the magnitude of positive selection differs among cities generates a heated discussion.

In this section, we first categorize the cities into three types in total.⁹ The first-tier cities represent Beijing, Shanghai, Guangzhou, Shenzhen; the second-tier cities are provincial capitals; while the third-tier stands for the rest cities.

Then for acquiring different city type effects, individual log wage is decomposed as

$$\ln W_i = \beta_0 + \mathbf{X}_i\delta + \beta_1 \times \text{MIGRANT}_i \times \text{CITYTYPE}_i + \beta_2 \times \text{URBAN}_i \times \text{CITYTYPE}_i + \varepsilon_i, \quad (2.11)$$

where CITYTYPE_i stands for the type of the destination city for individual i , $\mathbf{X}_i$ and all other terms have the same setting as in equation (2.9).

The results are shown in Table 2.5. From Table 2.5 we find that male, working experience, marital status, and *Han* ethnicity still have positive effect on individual wage level. Education returns are also positive for both high school and college groups, and that becomes higher as the education attainment level increases. Comparing to the primary industry, advanced service industry has the highest wage premium among various industries. In terms of ownership, the results suggest that the wage premium of state-owned enterprises and other

⁹Based on the classification of National Bureau of Statistics of China.

ownership becomes not significantly different from each other.

As expected, there is also a wage premium for workers with urban *hukou* status. As can be seen from Table 2.5, urban *hukou* has a wage premium of 0.412 in the first-tier cities. In the second-tier cities, urban *hukou* premium falls to 0.142, and that becomes the lowest in the third-tier cities. Such wage premium decrease means that the wage difference between rural and urban workers or the discrimination against rural workers gets smaller as the city type changes from the first to the third-tier.

The *migrant* term and *city type* $\times$ *migrant* terms suggest that migrants enjoy positive selection in the labor market. The results in Table 2.5 suggest that, in the first-tier cities, the wage difference between migrants and local workers is 0.073. However, as the city type changes from the first to the third-tier, wage difference gets larger. In the second-tier cities, migrant workers have a positive effect of 0.167. In the third-tier cities, wage difference between migrants and local workers becomes 0.213 which is the highest magnitude of positive selection.

Then what is the reason for having such selection trend among cities? The possible reason is that the first-tier cities and provincial capital cities probably have many other amenities than wage for migrants. Such as better education resources for migrants' next generations, more convenient life, more better job opportunities, and so on. Since the cities attract best workers not only by a high wage level, migrants' wage becomes not significantly higher than the average. As a result, the magnitude of positive selection is highest in the third-tier cities.

What's more, based on our theoretical finding, the magnitude of positive selection becomes greater when the wage difference between origin and destination cities decreases. Thus, for a specific migrant, positive selection is more likely to become the highest in the third-tier destination cities. This could also be a theoretical support on the phenomenon in Table 2.5.

Table 2.5: *Hukou* Effect by City Type

	(1)
Male	0.173*** (0.002)
High School	0.260*** (0.002)
College	0.678*** (0.003)
State-Owned Enterprises	0.124*** (0.002)
Other Ownership	0.079*** (0.003)
First-Tier City	0.404*** (0.010)
Second-Tier City	0.121*** (0.011)
Migrant	0.213*** (0.006)
First-Tier City $\times$ Migrant	-0.140*** (0.009)
Second-Tier City $\times$ Migrant	-0.046*** (0.010)
Urban <i>Hukou</i>	0.104*** (0.007)
First-Tier City $\times$ Urban <i>Hukou</i>	0.308*** (0.009)
Second-Tier City $\times$ Urban <i>Hukou</i>	0.038*** (0.011)
Constant	0.110*** (0.009)
Observations	361194
r2_a	0.419

Standard errors in parentheses

Working experience, working experience square
marital status, ethnicity, industry are also controlled* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

2.7 Conclusion

Although previous studies point out rural migrants earn less than urban locals, the quality difference between migrants and locals has not been discussed much. In this paper, we use China's 2005 mini-census to find what kind of people choose to migrate, and what is the

return for migration. If individual wage level can be affected by both rural/urban *hukou* and local/non-local status, how do they come into play is also worth discussing.

Based on the theoretical framework, wage difference between origin and destination cities, and the migration cost cause a fluctuation in the magnitude of self-selection on migrants. In addition, from the empirical findings we notice that migrants earn more than local workers in the labor market, which indicates an existence of positive selection. Using the sample from Beijing and Shanghai, we find such positive selection becomes stronger as the wage of origins increases, but the effect of moving cost is not significant. Moreover, the results present that magnitude of positive selection is larger in higher education level group. That is mainly due to greater education returns for higher education level group in destination cities. Besides education levels, the magnitude of positive selection also varies slightly in different types of cities. We find that the magnitude of positive selection increases as the city type changes from the first to the third-tier. The possible reason for such selection trend is that better cities have many other amenities than wage to attract best workers. If the migrants could be attracted by other amenities, a significant wage difference from average wage level seems to be not so important. Therefore, selected migrants in the third-tier cities have much higher wage than average when comparing to other types of cities.

Since our findings indicate that migrants tend to be self-selected from the upper tail of the income distribution in the origins, we conclude that wage inequality between rural migrants and urban locals is mainly due to the rural *hukou* status. Recently, *hukou* reform is on the rise, whether to abolish the system generates a heated discussion. However, in our opinion, *hukou* is not a problematic system, it helps keep order and provide better public services in the society. Only the classification of rural/urban status indeed causes an inequality, in terms of both opportunities and resources. Such inequality could even be more severe than our expectation since migrants tend to have a higher quality than locals. Therefore, the

reformation of *hukou* has to focus upon eliminating the discrepancies between workers with rural and urban *hukou* status.

Chapter 3

Network Effects on Migrants' Labor Market Outcomes in China

(ABSTRACT)

This paper uses 2005 China's mini-census to study how networks affect migrants' migration decisions, employment, and wage levels. Different from existing studies, this paper takes into account the existence of self-selection in the labor market. With the help of a theoretical model, we have a better understanding of the mechanism of networks as well as the different network effects on rural and urban migrants. We find out that networks affect both rural and urban migrants' migration decisions positively. In terms of employment, networks exert positive impacts on rural migrants but not on urban ones, which is due to the different quality drops between rural and urban migrants when the networks increase. Such employment effects also lead rural migrants to face a more severe negative wage impact than urban migrants.

3.1 Introduction

Labor migration starts rising since the late 1980s in China. Nowadays, with the perfection of migration supporting system, more and more people try to get job opportunities outside their hometowns. For such people, network in the destination labor market exerts an indispensable role. Networks not only affect migrants' migration decisions, but also influence various labor market outcomes in unspectacular ways. Figure 3.1 shows us the percentages of move-in and move-out in each province, which indicates an unbalanced migration situation in China. Migrants in popular destinations (the eastern coast region) or from central regions are more likely to have greater networks in the destinations when migrating. Some people say that a large network size in the destination helps them settle down more easily, while others are concern about a high competition along with large networks in destinations. Thus, how to choose a proper destination is worth discussing.

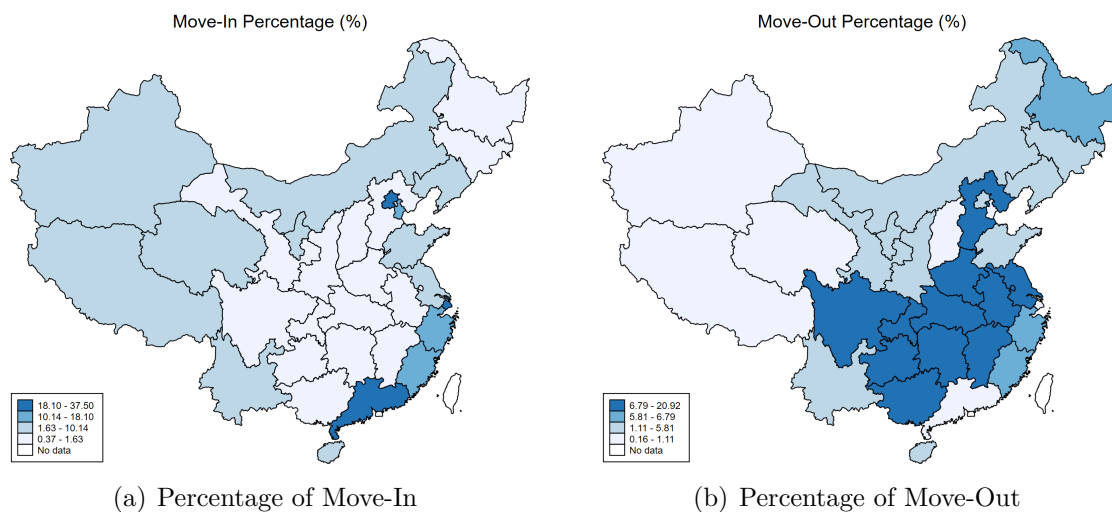


Figure 3.1: Migration Status (Data from 2005 1% Population Survey)

Studies about how networks affect labor market outcomes are many and varied, such as network effects on migrants' migration decisions ([Kerr and Mandorff, 2015](#), [Patel and Vella, 2013](#)), network effects on migrants' employment status ([Giulietti et al., 2018](#), [Munshi, 2003](#)),

and some studies even discuss wage impacts (Chen et al., 2018, Long et al., 2017, Patel and Vella, 2013, Zhang et al., 2012). Obviously, networks play a role that cannot be ignored in the labor market. Most of the existing studies focus on a single aspect of labor market outcomes, which may present only a one-sided picture of the network effects. If we put the pieces together, we would have a more complete story, and have a better understanding of network effects on migrants. So, in this paper, we try to not only find out how networks affect migration decisions, but also examine the employment as well as the wage impacts together. Especially, in terms of wage impact, we pay attention to a general network impact on all migrants instead of the network impact on network users. Different from existing studies, this paper takes into account the existence of self-selection in the labor market. As a result, a change in network size no longer indicates an increase or decrease of identical migrants.

Except for this, another point of this paper is that we explore the difference of network effects between rural and urban migrants. As we know, when we mention migrants, spotlights are only on the rural migrants. However, as the migrants with urban *hukou* status keep rising from 1990, there has already been a sizeable urban migrants group in the labor market. Thus, we try to fill the gap, and find out how networks give rural and urban migrants different impacts on the various labor market outcomes.

First of all, we build up a theoretical foundation in this paper. Based on Beaman (2012) and Calvo-Armengol and Jackson (2004)'s framework, we study the network effects on migrants' employment and wage levels theoretically. Taking the case when there is a positive selection, and when there is a positive relationship between networks and the percentage of employed early migrants as an example¹, we conclude that employment effects of networks depends on the type of network indirect impact, and also depends on the comparison between networks'

¹Other cases are also discussed in the theory model section.

direct and indirect impact on employment². If there is a negative network indirect impact, employment rate decreases as networks increase. If there is a positive network indirect impact, then employment rate decreases when the size of direct impact exceeds indirect impact. In terms of network effects on wage, if the early employed migrants' wage levels are higher than the average workers' wage levels in the market, networks expose a positive wage impact.

In the following part of this paper, we use 2005 China's mini-census to explore network effects on labor market outcomes in China. First of all, we check the type of self-selection as well as the relationship between networks and the percentage of employed early migrants, so that corresponding theoretical propositions could help us in explaining the empirical findings. We then turn to the network effects on various labor market outcomes. From the regression results, we realize that networks affect migration decisions for both rural and urban migrants positively. What's more, networks positively affect overall employment. Nevertheless, when we look over the difference between rural and urban migrants, the results indicate that networks only help rural migrants to be employed, but have negative effect on urban ones. With the help of the theory model, we attribute such difference to the different quality drops of rural and urban migrants when networks increase. In terms of the wage effects of networks, generally speaking, networks affect migrants' wage levels negatively. Although both rural and urban migrants face negative network wage effects, the magnitudes are different. Rural migrants have a more severe negative effects than urban ones, which is because there are more unqualified rural workers who are able to get jobs via networks when networks increase. As a result, negative wage effect for urban migrants mainly reflects the decreasing quality of urban migrants (due to the positive selection), but the wage decreases for rural migrants has already exceeded quality drops from selection.

²Direct impact is how network affects the probability of being employed directly, while indirect impact represents the network impact on the probability of being employed through social network.

In this paper, we fill the gap of existing literature to some extent. As we mentioned, most studies ignore the existence of self-selection in the labor market, and focus on a single aspect of labor market outcomes, which is very likely to read the results incompletely. Also, if we do not consider urban migrants in the study, we are not able to understand migrants group very well. With the help of theoretical model, we try to study a complete mechanism behind the networks in the labor market.

In the next section, we present some related literature about network effects on various labor market outcomes. Section 3.3 shows the theoretical foundation. Section 3.4 introduces the data and the sample. Section 3.5 gives us empirical evidence and according explanations. The last section concludes.

3.2 Related Literature

There are numbers of factors prompt people to migrate, such as higher wages, more amenities, or better education quality for next generations in destinations. Besides such factors, one's networks in the destination could also be a crucial one to affect their migration decisions. Some people think a higher network in the destination is helpful, since employed fellows can provide additional job opportunities and help new migrants settle down as soon as possible. Such people would like to migrate to the place where there is a large size of network. However, others take a higher network for a fierce competition. People from the same origin may have similar geographical characteristics, so migrants from the same origin would be more likely to be good at or to choose to work in the same industry or occupations ([Hashimzade et al., 2014](#), [Kerr and Mandorff, 2015](#), [Patel and Vella, 2013](#)). For such individuals, they prefer to migrate to a destination with fewer origin fellows. It seems like people's opposite choices on migration destinations are due to their different expectations on networks. Some people

believe that networks have positive effects on their labor market outcomes, while others believe not.

According to the existing literature, effects of network on labor market outcomes are varied. [Dolfin and Genicot \(2010\)](#) explore the mechanism of how networks affect migration decisions, and point out networks help migrants in many aspects. The most crucial point to promote people to migrate is that larger networks provide more job information. Similarly, in China's labor market, researchers also find out that as networks increase, people are more likely to migrate ([Chen et al., 2010](#), [Du et al., 2005](#), [Zhao, 2003](#)). Migrants acquire more information on migration experience, on job searching, or even on having financial support by networks. After all, networks are used to reduce migrants' migration costs ([McKenzie and Rapoport, 2007](#)) when people are making migration decisions.

In terms of network effects in destination labor markets, one of the findings is that networks offer help for raising employment rate of migrants. [Munshi \(2003\)](#) explores the network effects on Mexican migrants' employment status, and points out that networks positively affect migrants in the U.S. labor market. *Guanxi*, a Chinese word which indicates networks, increases the probability of being employed in non-farm positions for migrants as well ([Zhang and Li, 2003](#)). [Calvo-Armengol and Jackson \(2004\)](#), [Calvo-Armengol and Jackson \(2007\)](#), [Ioannides and Soetevent \(2006\)](#) provide some theoretical support on such positive network effects. Moreover, there is also a longer tenure for the migrants who get jobs via networks ([Loury, 2006](#)). However, when we focus on different types of networks, there may have some other conclusions. For example, [Beaman \(2012\)](#) mentions that only tenured network members increase migrants' employment, but the new arrivals cannot. In addition, [Giulietti et al. \(2018\)](#) suggest what help migrants in acquiring jobs are strong ties which play different roles from weak ties among migrants.

There exists some evidence to say people get more job information via networks, and the

crucial reason why the job information passed from early migrants is valuable for new entrants is because migrants from the same origin are more likely to choose similar jobs ([Patel and Vella, 2013](#)). Attitudes and beliefs, such as risk preference, honesty weight, could affect people's occupational choices ([Hashimzade et al., 2014](#)), and migrants who come from the same origin also show similar sociodemographic characteristics. As a result, network referral effects seems to be higher among migrants from the same origin ([Bayer et al., 2008](#)), and the effects are more obvious in self-employment ([Allen, 2000](#), [Zhang and Zhao, 2015](#)). [Kerr and Mandorff \(2015\)](#) also suggest that a smaller group has a higher entrepreneurial concentration, which once again confirms the referral effects among migrants from the same origin.

Network effects are not only found in migrants' employment, but there are also some intuitive effects on migrants' wage levels. [Yogo \(2011\)](#) and [Giulietti et al. \(2010\)](#) both point out that higher networks increase migrants wage levels if the migrants use networks to get jobs. What's more, [Patel and Vella \(2013\)](#) also conclude that if new arrivals choose the most popular jobs for their according early migrants, there exists a wage premium for new entrants. Such network effects could also explain a part of wage inequality between ethnicity ([Arrow and Borzekowski, 2004](#)).

However, conclusions of network effects on wages are not consistent. For example, [Beaman \(2012\)](#) suggests that positive network wage effects only come from early migrants. In the meantime, [Ye et al. \(2012\)](#) focus only on high level networks, and find out that the higher level networks increase migrants' wage in the labor market. Even beyond that, there are also couple of researches put forward some other opinions on network wage effects. [Zhang et al. \(2012\)](#) realizes that using network cannot increase all migrants' wage level, [Long et al. \(2017\)](#) and [Chen et al. \(2018\)](#) even find out a negative network effect on wages in informal job searching.

As mentioned above, effects of networks on labor market outcomes are varied. Migrants may

benefit from a higher network, but may also be adversely affected. Also, when we focus on different migrants groups, there could exist various network effects. How networks affect rural and urban migrants differently is also worth discussing. It is undeniable that relationship plays an important role in the labor market. Networks affect people's employment, people's wage level, or even the population distribution structure of a society. Therefore, a correct understanding of network effects on the labor market is valuable and necessary for all of us.

3.3 The Model

Based on [Beaman \(2012\)](#) and [Calvo-Armengol and Jackson \(2004\)](#)'s framework, we build a model to discuss the network effects on migrants' labor market outcomes.

3.3.1 Network Effects on Migrants' Employment

Assume that, there are $(1 - b)$ of the early migrants have already been employed, where $b \in [0, 1]$. Thus, there are three kinds of migrants in the labor market - employed early migrants, unemployed early migrants and unemployed new arrivals. The unemployed early and new migrants can get job offers from employers directly (direct way). In addition, since there exists employed early migrants, the unemployed early migrants and new migrants are able to have job offers passed from employed early migrants (indirect way).

We use net_c to represent the size of network for migrants from community c in the destination labor market, where $c \in C$. According to [Munshi \(2003\)](#), networks are defined as the proportion of the community located at destination, and $net_c \in [0, 1]$. Each community c has a total population of N_c , so $(net_c \times N_c)$ represents the population of migrants from community c in the destination.

Let $e(net_c)$ denote the percentage of early migrants, so that $(e(net_c) \times net_c \times N_c)$ is the number of early migrants. We use en_c to represent the network that comes from early migrants, while $rn_c(en_c)$ is the network from recent migrants, and the size of recent network is correlated with the size of early network. Therefore, we have

$$\begin{aligned} net_c &= en_c + rn_c(en_c) \\ &= e(net_c) \times net_c + rn_c(en_c), \end{aligned}$$

where $rn'_c(en_c) \geq 0$ ³. To explore how the percentage of early migrants $e(net_c)$ changes with net_c , we take the first derivative of $e(net_c)$, and find out

$$\begin{aligned} \frac{\partial e(net_c)}{\partial net_c} &= \frac{\partial \frac{en_c}{net_c}}{\partial net_c} \\ &= \frac{1}{net_c^2} \times \frac{rn_c(en_c) - rn'_c(en_c)en_c}{1 + rn'_c(en_c)}. \end{aligned}$$

Thus, when $\frac{rn'_c(en_c)}{rn_c(en_c)} \leq \frac{1}{en_c}$, there is a positive relationship between networks and the percentage of early migrants ($e'(net_c) \geq 0$), otherwise we have a negative relationship⁴.

s_c denotes the probability of being employed for unemployed migrants from community c ,

$$s_c = a(net_c) + r_c, \tag{3.1}$$

where $a(net_c)$ is the probability of being employed in a direct way, while r_c represents the probability of being employed through an indirect way, $a(net_c) \in [0, 1]$, $r_c \in [0, 1]$. We assume that employed early migrants pass the offer she receives via networks. As long as individual i is unemployed, i accept the offer that passed from an employed early migrant.

³According to [Chen et al. \(2010\)](#), [Du et al. \(2005\)](#), [Zhao \(2003\)](#), a high existing network in the destination prompts people to migrate.

⁴See Appendix [A.1](#).

How does network affect $a(net_c)$ depends on the average ability change with the network. For example, if there is a positive self-selection on migrants, the average ability of migrants from community c decreases as the network size increases. As a result, the probability of receiving job offers from employers decreases when net_c increases. Therefore, we have $a'(net_c) \leq 0$, $a''(net_c) \geq 0$. If there is a negative self-selection, we have $a'(net_c) \geq 0$, $a''(net_c) \leq 0$. For no selection case, we have $a'(net_c) = 0$.

The probability of being employed through an indirect way (r_c) can be written as

$$r_c = \frac{(1-b)e(net_c)a(net_c)}{1-(1-b)e(net_c)}.$$

Thus, the probability of being employed is

$$\begin{aligned} s_c &= a(net_c) + r_c, \\ &= \frac{a(net_c)}{1-(1-b)e(net_c)}. \end{aligned}$$

As can be seen from above, migrants have a direct impact ($a(net_c)$), and an indirect impact (r_c) from networks when applying for jobs. Thus, how networks affect migrants' employment depends on the size of direct and indirect impact.

Case 1: when there is no self-selection

In this case, we know that the probability of being employed in a direct way does not change as the network size increases ($a'(net_c) = 0$). This is also the case which most existing studies discuss.

Proposition 3.1. *If $e'(net_c) \geq 0$, s_c (weakly) increases as net_c increases, $\forall net_c \in [0,1]$. If $e'(net_c) < 0$, s_c (weakly) decreases as net_c increases, $\forall net_c \in [0,1]$.*

Proof. See Appendix A.2.1. □

Case 2: when there is a positive self-selection

If there is a positive self-selection on migrants, we have $a'(net_c) \leq 0$.

Proposition 3.2. *If $e'(net_c) \geq 0$, s_c (weakly) increases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \in [-\frac{(1-b)e'(net_c)}{1-(1-b)e(net_c)}, 0]$; s_c (weakly) decreases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \in (-\infty, -\frac{(1-b)e'(net_c)}{1-(1-b)e(net_c)}]$. If $e'(net_c) < 0$, s_c (weakly) decreases as net_c increases, $\forall net_c \in [0, 1]$.*

Proof. See Appendix A.2.2. □

Case 3: when there is a negative self-selection

If there is a negative self-selection on migrants in the destination labor market, we have $a'(net_c) \geq 0$.

Proposition 3.3. *If $e'(net_c) \geq 0$, s_c (weakly) increases as net_c increases, $\forall net_c \in [0, 1]$. If $e'(net_c) < 0$, s_c (weakly) increases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \in [-\frac{(1-b)e'(net_c)}{1-(1-b)e(net_c)}, +\infty)$; s_c (weakly) decreases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \in [0, -\frac{(1-b)e'(net_c)}{1-(1-b)e(net_c)}]$.*

Proof. See Appendix A.2.3. □

For a better understanding, we take the case when $e'(net_c) \geq 0$ and when there is a positive self-selection on migrants as an example. As Table 3.1 shows, in the destination labor market, migrants face both direct and indirect impacts from networks. We know that as the network increases, there is a negative direct impact ($a'(net_c) \leq 0$). If as network increases, indirect impact decreases, we find a negative network impact on migrants' employment. However, if as network increases, indirect impact increases, then we have to discuss. When direct impact

dominates, migrants have a negative network impact, but when indirect impact dominates, migrants have a positive network impact on employment.

Table 3.1: Different Impacts on Migrants

Conditions	Direct & Indirect Impact	Overall Impact
Network $\uparrow$ Indirect Impact $\downarrow$	Negative Direct Impact Negative Indirect Impact	Negative Network Impact
Network $\uparrow$ Indirect Impact $\uparrow$ Direct Impact $\geq$ Indirect Impact	Negative Direct Impact Positive Indirect Impact	Negative Network Impact
Network $\uparrow$ Indirect Impact $\uparrow$ Direct Impact $<$ Indirect Impact	Negative Direct Impact Positive Indirect Impact	Positive Network Impact

3.3.2 Network Effects on Migrants' Wage

In terms of jobs with different wage levels, individual i can receive a job offer with wage w_i in the destination labor market. We assume that if i is unemployed, she takes any offer she gets. If i has already been employed at the beginning of the period with wage w_i^* , she compares the wage she receives, w_i , with her own wage w_i^* . If $w_i \geq w_i^*$, she takes this offer, otherwise, she passes the offer to others in her network. Employed early migrant i earns the same wage as at the beginning of the period if she does not take any other offer.

Wage has an *i.i.d.* distribution from $[w_{low}, w_{high}]$ with a cumulative distribution function $F(x)$. $G(x)$ is a cumulative distribution function of employed early migrants' wage.

Case 1: when there is no self-selection

In this case, the probability of being employed in a direct way does not change with the network size, so $a'(net_c) = 0$.

Proposition 3.4. *If $e'(net_c) \geq 0$, all employed migrants' average wage (weakly) increases as net_c increases when G first-order stochastically dominates F ($G(x) \leq F(x)$); all employed migrants' average wage (weakly) decreases as net_c increases when $G(x) \geq F(x)$. If $e'(net_c) < 0$, all employed migrants' average wage (weakly) increases as net_c increases when $G(x) \geq F(x)$; all employed migrants' average wage (weakly) decreases as net_c increases when $G(x) \leq F(x)$.*

Proof. See Appendix [A.2.4](#). □

Case 2: when there is a positive self-selection

In this case, we know that $a'(net_c) \leq 0$.

Proposition 3.5. *If $e'(net_c) \geq 0$, all employed migrants' average wage (weakly) increases as net_c increases when $G(x) \leq F(x)$; all employed migrants' average wage (weakly) decreases as net_c increases when $G(x) \geq F(x)$. If $e'(net_c) < 0$, all employed migrants' average wage (weakly) increases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \geq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \geq F(x)$, or when $\frac{a'(net_c)}{a(net_c)} \leq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \leq F(x)$; all employed migrants' average wage (weakly) decreases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \geq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \leq F(x)$, or when $\frac{a'(net_c)}{a(net_c)} \leq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \geq F(x)$.*

Proof. See Appendix [A.2.5](#). □

Therefore, taking the case when there is a positive self-selection and $e'(net_c) \geq 0$ as an example, if the quality of early employed migrants is higher than the average quality in the market, networks expose positive wage effects; otherwise, migrants' wage decreases as networks increase.

Case 3: when there is a negative self-selection

In this case, there is a positive relationship between network size and the probability of being employed in a direct way. Therefore, $a'(net_c) \geq 0$.

Proposition 3.6. *If $e'(net_c) \geq 0$, all employed migrants' average wage (weakly) increases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \geq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \geq F(x)$, or when $\frac{a'(net_c)}{a(net_c)} \leq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \leq F(x)$; all employed migrants' average wage (weakly) decreases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \geq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \leq F(x)$, or when $\frac{a'(net_c)}{a(net_c)} \leq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \geq F(x)$. If $e'(net_c) < 0$, all employed migrants' average wage (weakly) increases as net_c increases when $G(x) \geq F(x)$; all employed migrants' average wage (weakly) decreases as net_c increases when $G(x) \leq F(x)$.*

Proof. See Appendix [A.2.6](#). □

From the model, we realize that how network affects migrants employment and wage depends on the relationship between networks and the percentage of early migrants ($e'(net_c)$), as well as the type of self-selection ($a'(net_c)$) in the labor market.

3.4 Data and Descriptive Statistics

In this paper, we use the China's 2005 mini-census data which is published by the National Bureau of Statistics of China. As a national survey data, the survey covers 1% of the national population, and the data set has various information for each subject, such as subjects' gender, age, education status, marital status, household registration information, employment status, and even has wage information, which help us broaden our research.

Since we have to acquire subjects' networks from household registration information and residence information, subjects who do not have household registration or residence information are not included. According to one's household registration information (*hukou*), subjects can be classified into rural *hukou* or urban *hukou* type. However, besides the *hukou* type, there is also a *hukou* location which indicates one's place of residence (origin). For most of the people in China, it is difficult to change the *hukou* location. Therefore, in a specific place, one can also be categorized into a local or a migrant according to the *hukou* location.

We mainly focus on the working age population whose ages are between 16 and 60 in this study. Subjects who are not able to work, self report to be retired or in school are dropped from the data set. Also, people who report to be employed but the income resources are unemployment insurance or property income are also not considered as our objects. Migrants are defined by their household registration and residence information. A subject is a migrant if she works in a different province from their household registration place (province). There are two reasons for why we focus on inter-provincial migrants. First of all, inter-provincial migration is mainstream in the labor market of China. As [Hou et al. \(2005\)](#) shows, the vast majority of migrants choose the east region as their migration destinations in China. [Li et al. \(2014\)](#) also conclude that there is an uptrend in inter-provincial migration in China since 1985. The amount of inter-provincial migrants in 2000-2005 is 38,042,340, which is almost four times of the number in ten years ago. Another reason is that most people are used to distinguish themselves by province in social settings. Thus, inter-provincial migration is more proper in network analysing.

According to the education status, we are able to acquire subjects' education years, and categorize the sample into three educational levels: compulsory and below education level (middle and primary school or below), high school level, college and above level. What's more, we have three sectors in the data set which are state-owned enterprises, private en-

terprises, and other ownership. Based on previous research, industries are classified into primary industry, manufacturing industry, construction industry, basic service industry, and advanced service industry (Ge and Yang, 2009). Hourly wage is acquired by monthly wage and working hours from the data set.

Table 3.2 shows the basic characteristics of the sample. From Table 3.2, we find out that migrants are much younger than locals, and there are significant fewer married subjects among migrants. When we separate the migrants into urban and rural sample, we notice that urban migrants even have more education years than locals. Employed rates are slightly higher for migrants, since migrants are more eager to make money to survive in destinations.

In terms of the employed sample in our data set, migrants' mean working hours are greater than that of locals. For rural migrants, the mean hourly wage is not higher than urban locals' which indicates rural migrants are more likely to work in low-paid positions. Self-employment rate is similar for both locals and migrants, and rural migrants have a slightly higher self-employment rate. From the table, we also notice that urban locals are mainly working in state-owned enterprises, while migrants are more likely to choose to work in private enterprises. Besides sectors, we also have an industry distribution in Table 3.2. The most popular industries for migrants are basic service industry and manufacturing industry, while urban locals are employed most in advanced service industry.

From the above summary statistics, we learn about our sample distribution, and also have some idea about the subjects' labor market performance. It is undeniable that networks play an important role in our culture as well as in the society. Thus, how migrants' labor market outcomes are affected by networks is worth a special study and discussion. In the next section, we provide some empirical evidence to show how the networks affect various labor market outcomes, and also present different network effects on rural and urban migrants.

Table 3.2: Characteristics of Locals and Migrants

	Local		Migrant	
	Urban	Rural	Urban	Rural
<i>All</i>				
Male	0.53	0.49	0.52	0.52
Age	37.96	38.16	31.85	30.04
Married	0.83	0.82	0.63	0.65
Years of Education	11.47	7.19	12.18	8.52
Employed	0.76	0.89	0.85	0.89
Number of Respondent	386,746	1,017,144	18,233	91,184
<i>Employed</i>				
Weekly Working Hours	45.88	45.70	48.58	55.18
Hourly Wage	6.82	2.45	10.61	4.71
Self-Employed	0.12	0.09	0.10	0.13
<i>Ownership</i>				
State-Owned Enterprise	0.58	0.03	0.20	0.09
Private Enterprise	0.33	0.90	0.59	0.67
Other Ownership	0.09	0.06	0.20	0.24
<i>Industry</i>				
Primary Industry	0.06	0.76	0.01	0.04
Manufacturing Industry	0.25	0.10	0.34	0.57
Construction Industry	0.04	0.03	0.06	0.09
Advanced Service Industry	0.36	0.02	0.21	0.05
Basic Service Industry	0.29	0.09	0.38	0.26
Number of Respondent	292,455	905,957	15,525	81,317
Data comes from 2005 1% national population survey				

3.5 Empirical Implications

In this section, we explore the network effects on migrants' migration decisions, employment status, and wages. Before turning to the network effects on the various labor market outcomes, we have to check several theoretical conditions, such as the type of self-selection, and the relationship between networks and the percentage of employed early migrants ($e'(net_c)$).

3.5.1 Type of Selection in the Labor Market

First of all, we propose to explore the type of self-selection on migrants in the labor market. We use the Mincerian equation ([Mincer, 1958](#)) as the basic structure. The specification can be shown as

$$\ln W_i = \beta_0 + \mathbf{X}_i\delta + \beta_1 MIG_i + \varepsilon_i, \quad (3.2)$$

where W_i represents i 's hourly wage, $\mathbf{X}_i$ is a vector which indicates one's gender, working experience, marital status, ethnicity, educational status, industries and so on, MIG_i is a dummy variable which indicates i 's migration status, and ε_i is an error term with zero mean.

Table [3.3](#) shows the basic OLS regression results. In the regression, subjects are separated into three educational level - compulsory and below level, high school level, college and above level. The compulsory and below level includes the subjects who have no educational experience or attend up to middle school; high school level stands for the subjects with high school educational experience; college and above level represents the subjects with more than high school education (both college and graduates school). We also categorize industries into 5 groups ([Ge and Yang, 2009](#)) - primary industry, manufacturing industry, construction industry, advanced service industry, and basic service industry; and categorize sectors into 3 groups - state-owned enterprises, private enterprises, and other ownership. For eliminating the effect from unequal educational resources, we use the pupil-teacher ratios in primary school level, middle school level, high school level, and educational funding data of migrants' origin communities to control for education quality in each province.⁵ Destination fixed effects are also controlled in the regression.

From Table [3.3](#) we know that male and married status both provide a wage premium. Also, as can be seen from the table, wage becomes higher for the one with a higher education

⁵Similar to [Card and Krueger \(1992b\)](#), we use pupil-teacher ratios and education funding as proxies for education quality. Data comes from China Statistical Yearbooks.

Table 3.3: Self-Selection on Rural & Urban Migrants

	(1) All Sample	(2) Rural Sample	(3) Urban Sample
Male	0.203*** (0.001)	0.207*** (0.001)	0.186*** (0.002)
Working Experience	0.009*** (0.000)	0.005*** (0.000)	0.019*** (0.000)
Married	0.073*** (0.002)	0.092*** (0.002)	0.029*** (0.003)
Migrant	0.085*** (0.004)	0.105*** (0.004)	0.265*** (0.007)
Urban <i>Hukou</i>	0.105*** (0.002)		
Migrant $\times$ Urban <i>Hukou</i>	0.214*** (0.006)		
High School	0.173*** (0.002)	0.128*** (0.002)	0.260*** (0.003)
College	0.640*** (0.003)	0.538*** (0.010)	0.707*** (0.003)
Province Dummies	Yes	Yes	Yes
Constant	0.740*** (0.015)	0.626*** (0.017)	1.223*** (0.039)
Observations	1295254	987274	307980
r2_a	0.522	0.391	0.411

Standard errors in parentheses

Working experience square, ethnicity, industry, ownership type, and education qualities are also controlled

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

level. In the first column of Table 3.3, subjects with high school educational experience earn 0.173 more than those with lower education level, and people who have been in college earn almost 0.467 more than high school graduates. Additionally, from the regression results we learn that, the one who works in the advanced service industry is more likely to earn the highest wage, and state-owned enterprises are better-paid than other sectors.

One of the most important findings in Table 3.3 is that both rural and urban migrants earn more than their according locals when we control for other related factors. Consistent with Ge and Jin (2020), we are able to say there is a positive selection on both rural and urban migrants in the labor market. This indicates that the average quality of migrants decreases

as the proportion of the community located at destination (network) increases.

In Table 3.4, we show an evidence of another theoretical condition that we have in the propositions of section 3. Relationship between the percentage of (employed) early migrants and networks are shown in the following table. Based on the model settings, we define networks as the proportion of the community located at destination⁶. Moreover, we have different employed early migrants in the two columns based on the migration years⁷. As can be seen from the table, both give us significant positive correlations.

Table 3.4: Relationship between Networks and the Proportion of Employed Early Migrants

	(1) Proportion of Employed Early Migrants (Define More than 4 Years as Early Migrants)	(2) Proportion of Employed Early Migrants (Define More than 3 Years as Early Migrants)
Network	0.231*** (0.011)	0.266*** (0.010)
Province Dummies	Yes	Yes
Constant	0.338*** (0.001)	0.414*** (0.001)
Observations	32263	41182
r2_a	0.345	0.285

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

So far we have the empirical evidence on the type of self-selection as well as the relationship between networks and the percentage of employed early migrants. So we currently expect that corresponding theoretical propositions would help us in explaining the empirical findings in the next section. In the following section, we find out how networks affect migrants' various labor market outcomes, and also pay attention to the difference of network effects between rural and urban migrants in the labor market of China.

⁶According to Munshi (2003), network is calculated by the number of migrants from origin community c (province level) who work in destination m (province level) divided by the population of origin community c .

⁷According to Zhao (2003), the one who has migrated more than 48 months can be regarded as "early migrants." For robustness checking, we also regress 3+ year migrants on networks in column (2).

3.5.2 Network Effects on Labor Market Outcome

Network Effect on Migration Decisions

Networks could affect numbers of decisions made from migrants. One of the most basic decisions for migrants is whether to become migrants or not. Therefore, in Table 3.5, we first explore the network effect on migrants' migration decisions.

We use the specification as

$$NEWMIG_{icm} = \beta_0 + \beta_1 EARLYNET_{icm} + \delta PRO_{im} + \varepsilon_i, \quad (3.3)$$

where $NEWMIG_{icm}$ is normalized new migrant i who come from origin community c work in destination m , $EARLYNET_{icm}$ indicates i 's early networks which is the proportion of employed early migrants within an origin community c located at destination m , PRO_{im} is a destination province indicator, and ε_i is an error term with zero mean.

Table 3.5: Network Effect on New Migrants' Migration Decisions

	Normalized New Migrants (≤ 4 Years)			Normalized New Migrants (≤ 3 Years)		
	All (1)	Rural (2)	Urban (3)	All (4)	Rural (5)	Urban (6)
Early Network (from 4+ Years Migrants)	57.454*** (0.132)	56.983*** (0.145)	59.659*** (0.350)			
Early Network (from 3+ Years Migrants)				44.850*** (0.111)	44.499*** (0.122)	46.562*** (0.297)
Province Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.987*** (0.001)	-0.984*** (0.001)	-0.996*** (0.002)	-0.978*** (0.001)	-0.974*** (0.001)	-0.990*** (0.002)
Observations	71608	61134	10474	61687	52915	8772
r2_a	0.904	0.904	0.902	0.908	0.907	0.910

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5 shows the according regression results. Column (1) to (3) is the regression results of a case when we define new migrants as the ones who migrate to current destination within

4 years, while the following three columns is for new migrants who migrate no more than 3 years. Column (2) and (5) are for rural sample, while (3) and (6) are for urban sample only. As can be seen from the table, numbers of new migrants increase when the early networks raise. In other words, if there exists a high proportion of the community located at destination, that promotes the migrants' fellows to migrate. This result is true for both rural and urban migrants.

Network Effect on Employment

Apart from migration decisions, we are also interest in the network effect on migrants' employment. We use the data in hand to find out whether networks help migrants to be employed more or not. Moreover, we are able to explore whether the network effect is different for rural and urban migrants. Table 3.6 and 3.7 show us empirical evidence to answer these questions.

In Table 3.6, we first do a probit regression to explore how networks affect migrants' employment status. The specification can be written as

$$EMPLOYED_i = \beta_0 + \mathbf{X}_i\delta + \beta_1 NET_i + \varepsilon_i, \quad (3.4)$$

where $EMPLOYED_i$ indicates whether a migrant is employed or not, $\mathbf{X}_i$ is the same vector as we use in equation (2), NET_i represents migrants' network size, and ε_i is an error term with zero mean.

Table 3.6 shows us the regression results of equation (4). We use all migrants as our sample in the first two columns, and then restrict the sample to recent migrants⁸ in column (3) and (4). As the table shows, networks help all migrants generally in being employed. No

⁸We define migrants who migrate no more than 4 years as recent migrants in Table 3.6.

Table 3.6: Network Effect on Migrants' Employment Status (Probit)

	All Migrants		Recent Migrants	
	General (1)	Rural/Urban (2)	General (3)	Rural/Urban (4)
Network	0.380** (0.062)**	0.609*** (0.054)*	0.504** (0.076)**	0.683*** (0.070)**
Urban <i>Hukou</i> $\times$ Network		-1.374*** (-0.245)**		-1.202*** (-0.195)*
Urban <i>Hukou</i>	-0.236*** (-0.038)**	-0.182*** (-0.043)**	-0.201*** (-0.030)**	-0.153*** (-0.034)**
High School	0.048*** (0.008)**	0.049*** (0.008)**	0.027 (0.004)	0.027 (0.004)
College	0.176*** (0.027)**	0.170*** (0.026)**	0.066* (0.010)*	0.059 (0.009)
Male	1.038*** (0.168)**	1.038*** (0.168)**	0.950*** (0.144)**	0.950*** (0.144)**
Age	0.091*** (0.015)**	0.090*** (0.015)**	0.103*** (0.016)**	0.103*** (0.016)**
Married	-0.647*** (-0.105)**	-0.646*** (-0.105)**	-0.748*** (-0.113)**	-0.747*** (-0.113)**
Province Dummies	Yes	Yes	Yes	Yes
Constant	-0.802***	-0.832***	-1.001***	-1.022***
Observations	109,417	109,417	72,053	72,053

Margins in parentheses

age square, ethnicity, and education qualities are also controlled

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

matter using all migrants or recent migrants as our sample, we are able to find out that networks are helpful to rural migrants, but cannot help urban migrants to be employed. Among recent migrants, the marginal effect of network for rural migrants is 0.070, while that for urban migrants is -0.125. Additionally, urban workers are less likely to be employed when comparing with rural ones. Perhaps rural migrants have higher moving cost, so they are more eager to make money in the destination. In terms of education attainment level, the more educated, the more likely to be employed. Moreover, male or single migrants are relatively favorable in the labor market.

From the above regression, we find out networks help rural migrants but cannot help urban migrants in being employed. Let us recall the theoretical settings in Section 3.3. In equation

(3.1), we assume that there is only one kind of migrants in the market. However, in reality, we are able to categorize migrants into rural and urban ones. Therefore, for migrants in group k , the probability of being employed can be expressed as:

$$s_{kc} = a_k(net_c) + \frac{(1-b)e(net_c)a(net_c)}{1-(1-b)e(net_c)},$$

where $k \in \{rural, urban\}$. Network indirect impact will be the same for both rural and urban migrants. Since migrants could have job offers passed from both rural and urban employed early migrants. Thus, we suppose that, for current rural and urban job seekers (migrants), the difference in network effect is mainly due to the difference in direct impact ($a'_{rural}(net_c)$ vs. $a'_{urban}(net_c)$). In other words, different network employment effect for current rural and urban job seekers (migrants) comes from the different quality drops for the two groups of migrants.

Different from the probability of being employed for job seekers, let w_c denote the probability of being employed for all the migrants in the labor market:

$$\begin{aligned} w_c &= (1-b)e(net_c) + [1-(1-b)e(net_c)]s_c, \\ &= (1-b)e(net_c) + a(net_c). \end{aligned}$$

Similar to the s_{kc} ,

$$w'_{kc} = (1-b)e'(net_c) + a'_k(net_c),$$

where $k \in \{rural, urban\}$. $e'(net_c)$ is the same for both rural and urban migrants. Thus, the different network employment effect for all rural and urban migrants also comes from the different quality drops for the two groups of migrants ($a'_{rural}(net_c)$ vs. $a'_{urban}(net_c)$).

Table 3.7 not only supports the existence of positive selection in the labor market, but also shows the different quality drops for rural and urban migrants as networks increase. We regress overall, rural, urban migrants' average education years on networks, and the results indicate that when the network size increases, migrants' average education years go down significantly. No matter for the rural or the urban migrants, migrants' average ability decreases as the migration rate increases, and the average quality drop for urban migrants is greater than rural migrants ($a'_{urban}(net_c) < a'_{rural}(net_c) < 0$.)

Table 3.7: Relationship between Networks and Average Education Years for Migrants

	(1) Average Education Years All	(2) Average Education Years Rural	(3) Average Education Years Urban
Network	-8.885*** (0.076)	-3.163*** (0.054)	-10.765*** (0.206)
Province Dummies	Yes	Yes	Yes
Constant	10.423*** (0.009)	9.082*** (0.007)	13.508*** (0.020)
Observations	109417	91184	18233
r2_a	0.402	0.498	0.409

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Since overall network employment effect is positive (Table 3.6), we conclude that, for current job seekers (migrants), indirect impact is positive (because direct impact $a'(net_c) \leq 0$). When the size of indirect impact is the same for both rural and urban migrants, the one with smaller direct impact has a smaller (even negative) network effect on employment. What's more, for all of the rural and urban migrants in the labor market, we know that $(1 - b)e'(net_c) \geq 0$. Thus, the one with smaller $a'_k(net_c)$ has a smaller (even negative) network effect on employment.

Up to now, we explore network effects on migration decision and employment status. Besides this, how networks affect migrants' wage remains a mystery. In the following subsection, we will show empirical evidence on the network effects on migrants' wage, as well as how rural and urban migrants affected differently by networks.

General Network Effect on Wage

In this section, we discuss the network effect on migrants' log hourly wage. From the following regressions, we learn about whether migrants with a higher network have a wage premium or not in the destination labor market. Based on the classic Mincerian equation (Mincer, 1958), the log wage is composed as

$$\ln W_i = \beta_0 + \mathbf{X}_i\delta + \beta_1 NET_i + \varepsilon_i, \quad (3.5)$$

where $\ln W_i$ is a log hourly wage of individual i , $\mathbf{X}_i$ a vector which stands for personal characteristics, NET_i is migrants' network size, and ε_i is an error term with zero mean.

According regression results are shown in Table 3.8. In the first column of Table 3.8, we do the basic OLS regression. Nevertheless, there still exists some factors which affect both network size and wage levels. For dealing with the endogeneity problem, we adopt a lagged network size as an instrument variable (IV), and the 2SLS regression results are shown in the second column of Table 3.8. The instrument variable is from China's 2000 Census data, and the data coverage is similar to the 2005 mini-census. Additionally, the lagged network size is correlated with the current migration proportion, but cannot affect other aspects, such as economic conditions, in 2005⁹.

From Table 3.8, we are able to realize that networks decrease migrants' wage levels. Even after we do the 2SLS regression, the negative coefficient remains similarly. As we know that there is a positive selection on migrants in the labor market¹⁰, which means as networks increase, the average quality of migrants falls. As a result, when networks increase, migrants'

⁹Hausman test result strongly rejects the null hypothesis. The first-stage results of 2SLS regression indicate that there is a strong correlation between network and its instrument variable. Cragg-Donald Wald F statistic shows there is no weak instrument problem.

¹⁰From Table 3.3 and 3.7.

Table 3.8: Network Effect on Migrants' Wage

	OLS Results (1)	2SLS Results (2)
Network	-0.865*** (0.051)	-0.819*** (0.053)
Male	0.155*** (0.003)	0.155*** (0.003)
Working Experience	0.015*** (0.001)	0.015*** (0.001)
Married	0.039*** (0.005)	0.039*** (0.005)
Urban <i>Hukou</i>	0.217*** (0.006)	0.218*** (0.006)
High School	0.259*** (0.005)	0.260*** (0.005)
College	0.879*** (0.010)	0.881*** (0.008)
Province Dummies	Yes	Yes
Constant	0.874*** (0.027)	0.869*** (0.025)
Observations	96842	96275
r2_a	0.362	0.361

Standard errors in parentheses

Working experience square, ethnicity, industry, ownership type, and education qualities are also controlled

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

wage drops by 0.819. Besides the network impact, we find out that, among migrants, male, married or urban *Hukou* holders are more likely to have a relatively higher wage. Also, migrants with more working experience or higher educational attainment level are also better paid. As shown in Table 3.8, migrants with high school education earn 0.260 more than those with compulsory education or lower, and college graduates have 0.621 wage premium when comparing to high school graduates. From the regression results we are also able to conclude that the best-paid industry is the advanced service industry, while the worst-paid one is, as we expected, the primary industry. In terms of sectors, the largest wage gap exists between state-owned enterprises and private enterprises.

So far, we notice that there is a negative network effect on migrants' wage. Then let us recall

what we have in the theoretical model. In Proposition 5, we mention that all employed migrants' average wage (weakly) decreases as network increases when: $G(x) \geq F(x)$ ¹¹. Therefore, in Figure 3.2(a) and 3.2(b), we check whether we have such condition in our data¹².

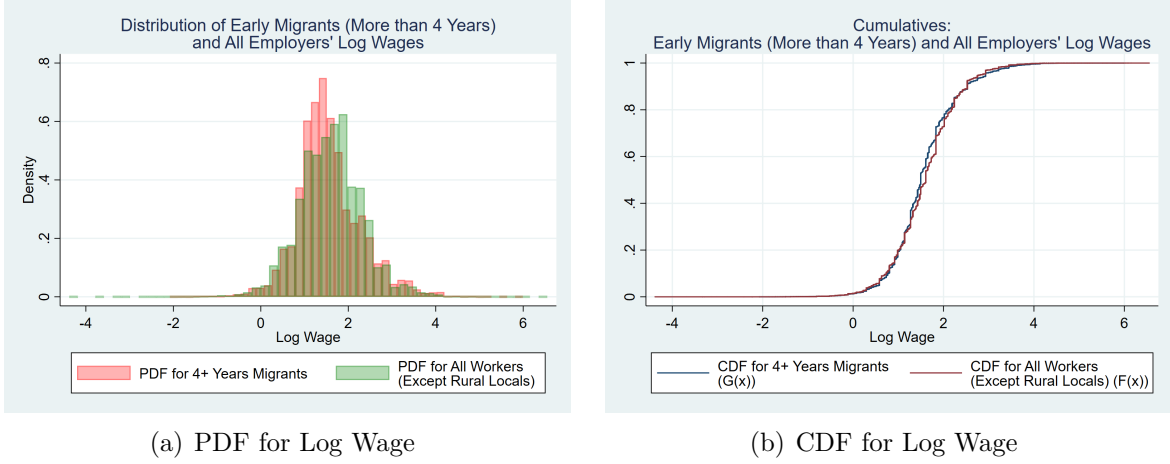


Figure 3.2: Distributions for Early Migrants and All Employers' Log Wage

Figure 3.2 gives us a PDF graph and an according CDF graph for log wage. In both figures, we treat people who migrate to current destination more than 4 years as early migrants. Figure 3.2(a) shows that early migrants' wage distributions skew leftward when comparing to the wage distribution in the labor market. Accordingly, we also find out a slightly leftward moved CDF curves for employed early migrants¹³. Therefore, after having a negative network wage effect empirically, we observe its theoretical condition from the data as well.

¹¹Wage has an *i.i.d.* distribution from $[w_{low}, w_{high}]$ with a cumulative distribution function $F(x)$. $G(x)$ is a cumulative distribution function of employed early migrants' wage.

¹²The reason why we exclude rural locals in both figures is that the vast majority of migrants migrate out for non-farming jobs, but from the data we realize that more than 70% of local rural workers live by farming. Therefore, we suppose that migrants' main competitors are urban locals.

¹³Kolmogorov-Smirnov tests statistic for panel (b) is 0.0687 [0.000].

Network Effect on Wage for Rural and Urban Migrants

From the previous subsection, we realize that networks have a negative effect on migrants' wage. In this subsection, we pay further attention to whether there is any difference between rural and urban migrants, and try to figure out the reason why migrants are facing different network effects.

Table 3.9: Network Effect on Wage for Rural and Urban Migrants

	(1)	(2)
	OLS Results	2SLS Results
Network	-0.935*** (0.051)	-0.900*** (0.055)
Urban <i>Hukou</i> $\times$ Network	0.489*** (0.112)	0.544*** (0.101)
Urban <i>Hukou</i>	0.195*** (0.008)	0.193*** (0.007)
Male	0.155*** (0.003)	0.155*** (0.003)
Working Experience	0.015*** (0.001)	0.015*** (0.001)
Married	0.039*** (0.005)	0.038*** (0.005)
High School	0.260*** (0.005)	0.260*** (0.005)
College	0.881*** (0.010)	0.883*** (0.008)
Province Dummies	Yes	Yes
Constant	0.885*** (0.027)	0.882*** (0.025)
Observations	96842	96275
r2_a	0.362	0.361

Standard errors in parentheses
Working experience square, ethnicity, industry,
ownership type, and education qualities are
also controlled

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We adopt an interaction term of *urban* $\times$ *network* in Table 3.9. Similar to the previous table, we do a simple OLS regression in column (1), and do a 2SLS regression in the second column. From the table, we notice that networks still have negative effects on both rural and urban migrants. What's more, as we would expect, networks affect rural and urban

migrants differently. When networks increase, rural migrants' wage drops by 0.900, while that of urban migrants only decreases by 0.356. The results indicate that networks have a more severe impact on rural migrants. From Section 3.5.2 we know that networks help rural migrants to be employed. As a result, as networks increase, more rural recent migrants who were not able to get jobs by themselves now are working. Therefore, as networks increase, the wage decrease for urban migrants mainly reflects the decreasing quality of urban migrants¹⁴, but the wage decreases for rural migrants exceed the quality drops from positive selection. That is because there are more unqualified rural migrants get jobs via networks as the networks increase.

3.6 Conclusion

By using 2005 China's mini-census, this paper studies how networks affect migrants' migration decisions, employment, as well as wage levels. Moreover, this paper also pays attention to the existence of self-selection and the difference between rural and urban migrants, so that we consider the heterogeneity of migrants, and have a better understanding of the roles that networks play in the labor market of China.

The empirical findings indicate that networks give both rural and urban migrants a positive impact on their migration decisions. For employment, networks only exert positive impacts on rural migrants, but not on urban ones. As networks increase, urban migrants even face a lower employment rate. With the help of theoretical framework, we are able to find out the reason for such rural-urban difference. This is due to the different quality drops between rural and urban migrants when the networks increase. Such employment effect has also led to different negative network wage effects for rural and urban migrants. Since, among rural

¹⁴Due to the positive selection.

migrants, there are more unqualified workers get jobs via networks, the wage decreases for rural migrants has already exceeded quality drops from positive selection.

Chapter 4

Parental Migration and Children's Risk Preferences

(ABSTRACT)

This paper investigates how parental migration affects left-behind underage children's risk preferences. By focusing on migrant parent groups, we are able to estimate whether the influence of nurture could also affect children's risk preference levels or not. The findings suggest that besides the intergenerational transmission of risk preferences between parents and children, parental migrations do have an influence on girls' risk preference levels. In addition, in terms of adult children's risk-related outcomes, we are able to find a positive parental migration impact on daughters' self-employment decisions. Findings help us have a better understanding of the relevant factors of risk preferences, and also confront the impact of the separation of parents and children.

4.1 Introduction

Risk preferences are used to explain many of people's behaviors and outcomes. According to [Tanaka et al. \(2010\)](#), if people are extremely risk averse, entrepreneurial intentions with large amount of risky cash flows would be avoided, so ultimately, risk aversion could be one of the factors which make people stay poor. Not only in terms of investment behavior, risk preferences are also associated with many other outcomes or decisions, such as smoking ([Anderson and Mellor, 2008](#)), drinking ([Anderson and Mellor, 2008](#), [Dave and Saffer, 2008](#)), health outcomes ([Anderson and Mellor, 2008](#), [Samek et al., 2021](#), [Sutter et al., 2013](#)), and risky food consumption ([Lusk and Coble, 2005](#)). To a certain extent, such findings prompt us to take one more step to investigate what determines people's risk preferences and how they are determined.

Among the relevant determinants, the intergenerational correlation of risk preferences between parents and children has been fiercely discussed. Various studies point out that there is an intergenerational transmission of risk preferences between parents and their children ([Alan et al., 2017](#), [Brown and Van der Pol, 2015](#), [Dohmen et al., 2012](#)), while some other researchers also suggest that environment or parents' academic achievement could affect children's risk levels ([Dohmen et al., 2011](#), [Zyphur et al., 2009](#)). Now that studies suggest environment could be one of the factors that affect people's risk preferences, we therefore would like to explore the intergenerational impact of migrant parents on children's risk preferences. Migrant parents affect children in many aspects, for instance children's educational outcomes ([Antman, 2011, 2012](#), [Chang et al., 2019](#), [Giannelli and Mangiavacchi, 2010](#), [Liu et al., 2018](#), [Lyu and Chen, 2019](#), [Zhao et al., 2014](#)), children's future job choices ([Lee, 2019](#)), children's health outcomes ([Shuangyue, 2018](#), [Zhou et al., 2015](#)), and even children's behavior ([Cadsby et al., 2019a,b](#)). [Luo et al. \(2009\)](#) even mention that according to the 2003-2005 statistics of underage children's theft and robbery cases in China, left-behind children ac-

counted for a great proportion¹. However, we find that there is no study discuss how having a migrant parent affect children's risk preferences. According to the existing literature, we speculate that migrant parents are not only able to transfer the risk preferences to their children genetically, but also could change the family environment of children via migration. Therefore, for exploring whether the intergenerational correlation of risk preferences between parents and children is due to a influence of nature or nurture, we look over the impact of having a migrant parent on children's risk preferences in this paper.

For exploring the impact of having a migrant parent on children's risk preferences, we find China is quite suited to conduct the study. In China, there is a large number of migrant parents and left-behind children, so we have a rich sample. With the transition of the economy and urbanization, China is gradually promoting and developing industrialization. As a result, more and more people from rural areas try to work outside their hometowns instead of being engaged in agriculture. As can be seen from Figure 4.1², the numbers of rural migrants in China keep increasing. In 2018, there are over 288 million rural migrants in the labor market, and the number of 20 to 50-year-old migrants, who may have underage children, is almost 216 million. Such a large number of 20 to 50-year-old rural migrant leads to a mass of children become left behind. According to the data from All-China Women's Federation (ACWF), there are more than 61 million left-behind children in 2010, which accounts for 37.7% of the rural underage children. Except for the sample size, in China, most of the migrant parents move within a country, so we can minimize the impact of culture shock on parents' risk preferences or on their children's as well.

Therefore, in this paper, we conduct an empirical study on the impact of having a migrant parent on children's risk preferences using the 2018 CFPS data. The 2018 CFPS data has rich

¹Luo et al. (2009) also mention that, in some counties, more than 60% of the theft and robbery cases are committed by left-behind children.

²Data from National Bureau of Statistics.

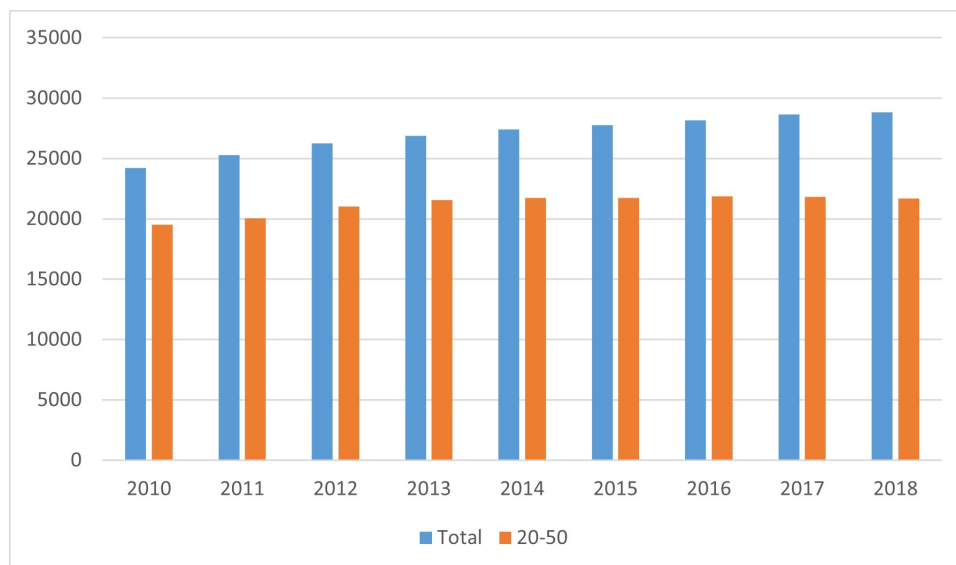


Figure 4.1: Numbers of Rural Migrants in China (unit: ten thousand people)

information in individual level, such as subjects' background information, educational and health outcomes, and also contains a bunch of questions related to one's risk preferences. With such information, we first estimate a linear regression model, and find a positive correlation between having a migrant parent and children's risk preference levels, especially for girls. From existing studies we understand that parents' risk preference levels do affect children's genetically. Despite that, we still wonder, for migrant parents, would there be any other channels that could affect children's risk preference levels besides the influence of nature.

To explore this question, we need to show the stability of parents' risk preference levels first, which means parents' risk preference levels do not change with their migration or return actions. Such stable parents' risk preference levels demonstrate the stability of intergenerational transmission of risk attitudes between parents and their children. After this, we control parents' risk preference levels in the linear regression, and still find a positive correlation between having a migrant parent and girls' risk preference levels. As we expect, the results indicate that even after we control the intergenerational transmission, we still observe

the parental migration impact. In terms of the relationship between parental migration and children's risk related outcomes, we look over the rural local adult children's self-employment decisions. The results suggest that children's educational outcomes and marital status have strong correlation with their self-employment decisions, and parental migration is influential for daughter but not for sons.

In this paper, we mainly pay attention to the impact of parental migration on their left-behind children. Children's risk preferences are not only related to their various outcomes, such as educational and health outcomes, but also affect their future behaviors. Although this is an important topic, few studies discuss it. Therefore, in this study, we try to fill the gap of the existing literature, and further explore such intergenerational impacts is due to the influence of nature or nurture. Beyond that, when we have 2020 CFPS Data, we could even check whether such intergenerational impacts from migrant parents are temporary or permanent impact.

In the following section, we present the existing literature on risk preferences and various intergenerational impacts. Section 4.3 introduce the data that we use in this paper. Section 4.4 shows the empirical results. Lastly, we provide the conclusion in section 4.5.

4.2 Related Literature

4.2.1 Risk Preferences and Determinants

Risk preferences play an important role in a world of options and uncertainty. People with high level of risk tolerance would prefer a risky but tempting offer, while risk averse people choose stability instead of high yields. For instance, [Barbosa et al. \(2007\)](#) use 528 college students from Russian, Norway and Finland as the sample to find that people with high

preference of risk have more entrepreneurial intentions. Existing literature also indicate that risk averse people are less likely to smoke or to forgo using seat belt when driving ([Anderson and Mellor, 2008](#)), and less likely to consume alcohol products ([Dave and Saffer, 2008](#)) or genetically modified food ([Lusk and Coble, 2005](#)). Moreover, [Sutter et al. \(2013\)](#) conduct an experiment on children and adolescents in Austria, and find a positive relationship between risk aversion and lower BMI.

Studies present such a key role of risk preferences in our lives, and therefore the relevant factors of risk preferences also catch a lot of attentions. Existing literature shows the gender difference of risk preferences. For example, [Andreoni et al. \(2020\)](#) use experimental data to conclude that adolescent girls are more risk averse than adolescent boys, but this result is not found with young girls and boys. [Charness and Gneezy \(2012\)](#) also use the data from 15 different studies to find a strong evidence that women invest less than men in investment games, which shows female are financially more risk averse. Additionally, social capital ([Nielsen et al., 2013](#)), economic circumstance of the community ([Tanaka et al., 2010](#)), environment ([Zyphur et al., 2009](#)), traumatic experience ([Malmendier and Nagel, 2011](#)), culture factors ([Rieger et al., 2015](#)) or parents' academic achievement ([Dohmen et al., 2011](#)) are considered to be related to people's risk preferences. Other than the external factors, intergenerational transmission of risk preferences between parents and children is also an important way to affect one's risk preference ([Alan et al., 2017](#), [Brown and Van der Pol, 2015](#), [Dohmen et al., 2012](#)). From this study, we would like to focus on migrant parent group, to see whether parental migration action itself could affect children's risk preferences or not.

4.2.2 Parental Migration and Impacts

Even though there is no study discuss the relationship between parental migration and children's risk preferences, we do learn various parental migration impacts from existing studies. After the founding of China, a great migration starts rising at the late 1980s ([Wu, 1994](#), [Wu and Zhou, 1996](#)). Nowadays, with the perfection of migration supporting system, more and more rural residents realize the feasibility and benefit of migration. Migration not only helps mitigate instability of agricultural income, but significantly improve household income level. Nevertheless, as parents, their migration also comes with reducing involvement with children. Thus, how parental migration affect children is an important question for the potential migrants parents.

Parental migration on children's educational performance has been discussed in many aspects. For example, existing study realizes that father's Mexico to U.S. migration reduces children's study hours and increase work hours ([Antman, 2011](#)). Similarly, [Chang et al. \(2011\)](#) conclude that when there is a migrant within a household, elderly women and girls' time spent on work increases. Besides the impact on left-behind children's time spent, [Giannelli and Mangiavacchi \(2010\)](#) also finds out that parental migration has a negative effect on children's school attendance. However, in rural China, there are mixed results of parental migration impacts on children's educational performance. [Zhao et al. \(2014\)](#) mentions that having migrant parents reduces children's math score significantly, while [Chang et al. \(2019\)](#) conclude that there is no significant relationship between parental migration and math score.

Except for that, there are some studies focus on long-term impacts of parental migration. [Antman \(2012\)](#) concludes that left behind children's educational attainment may be affected positively if parents migrate from Mexico to U.S., but no significant results for domestic parental migration. In terms of occupation choices, children with U.S. legal permanent

residents mothers are less likely to be employed in blue collar jobs (Lee, 2019). In rural China, Lyu and Chen (2019) points out a significant negative relationship between parental domestic migration and children's educational attainment as well as future income.

Parental migration not only affects children's educational or occupational outcomes, but also influences children's health and behavior. For instance, Shuangyue (2018) concludes that mother's migration negatively affects children's health and mental health. Zhou et al. (2015) also suggest that left-behind children perform poorly in most of the health indicators. In addition, researchers also find out that left behind children are more likely to be depressed (He et al., 2012, Wu et al., 2015), to feel anxiety (Chang et al., 2019), to feel lonely (Jia and Tian, 2010), and to have emotional problems (Lahaie et al., 2009). Not only that, but parental migration could affect children's behavior as well. Cadsby et al. (2019b) point out that parental migration has significant impact on children's dishonesty, and from Cadsby et al. (2019a), we learn that left-behind children are more likely to be altruistic.

It is undeniable that parental migration affect children a lot, especially for the left-behind children. Up to now, we notice the parental migration's intergenerational impacts in terms of education, health, behavior, and emotion. Except for that, migrant parents are always correlated with less risk averse or more likely to enter a competition (Akgüç et al., 2016, Dustmann et al., 2020, Hao et al., 2016, Jaeger et al., 2010). According to the existing literature, if people's risk preferences could be determined both genetically and environmentally, how do migrant parents affect left-behind children's risk preferences is worth discussing. Therefore, in this paper, we would like to check whether there is an intergenerational transmission between migrant parents and their left-behind children, and also find out whether migrant parents could affect the left-behind children's risk preferences other than the intergenerational transmission or not.

There are presently a mass of children who cannot study outside their household registration

places due to the policy restrictions. Having a better understanding on the left-behind children helps us confront the impact of separation of parents and children. In doing so, we will make better education and migration policies to help the children grow better, and help the society develop better.

4.3 Data and Descriptive Statistics

The main data set that we use in this study is the 2018 China Family Panel Studies (CFPS) data. The CFPS data is a comprehensive longitudinal survey data of Chinese communities, families, and individuals. According to [Xie and Hu \(2014\)](#), [Xie and Lu \(2015\)](#), the baseline survey of the CFPS was first launched by Peking University in 2010, and the later data sets of 2011, 2012, 2014, 2016, and 2018 are published consecutively. The CFPS data covers 25 provinces in the mainland China³, and in the most recent 2018 wave, the survey visits almost 15,000 households, and collects near 44,000 individuals' data. This nation-wide survey data gives us a rich set of information, such as subjects' background information, educational outcomes, migration information, economic activities, health outcomes, and even answers related to the Big Five personality. With such information in the 2018 CFPS data set, we try to explore the relationship between having a migrant parent and children's risk preference levels in this paper.

For understanding how having a migrant parent be correlated with children's risk preference levels, we restrict the baseline regression sample into local underage (age ≤ 18) children who are alive and live in rural areas⁴. In terms of subjects' migration information, there is a

³[Xie and Lu \(2015\)](#) mention that Hong Kong, Macao, Taiwan, Xinjiang, Tibet, Qinghai, Inner Mongolia, Ningxia, and Hainan are excluded from the data set. The CFPS sampling represents almost 94.5% of the population in mainland China, so that the sample is considered to be nationally representative.

⁴Areas are classified into rural and urban based on the classification rules of the National Bureau of Statistics.

question about people's type of current residence. We define a subject as a migrant if she lives in a different county from the place of household registration (*Hukou*). However, in the 2018 CFPS data set, there is a mass of missing information about the type of current residence. Therefore, we adopt the data from the 2010 to 2016 CFPS data sets, to acquire subjects' migration information more precisely. Based on subjects' recorded resident address and migration information in each wave, we are also able to judge if she ever migrated before and currently stays as a local.

In this study, the main variable of our interest is people's risk preference levels. As Figure 4.2 shows, there are a series of questions related to subjects' risk preferences in the 2018 CFPS survey. By answering 2 or 3 of total 5 questions, subjects get different outcomes. According to the selection process, subjects are assigned different levels of risk preference. If subjects always choose the sure payment, the outcome is 50 at the end, and their risk preference levels will be one. If subjects keep choosing the 0/200 lottery, the final outcome is getting the lottery, and their according risk preference levels will be the highest, six. Therefore, we would like to describe this risk preference levels as the levels of subjects' risk tolerance. The one with a higher risk tolerance level, she is more likely to make a risky choice.

Additionally, by matching personal ID with subjects' relatives' information (parental ID), we are able to have parental information in terms of parents' age, education years, self-employment status, number of children, migration experience, and risk preference levels for each subject. After matching such parental information with subjects, we are able to control family background in this study.

Table 4.1 describes some basic characteristics of parents and rural local children in the data set. The way we define "parents" is based on subjects' child information. If the one who has at least one child, he or she will be defined as a parent. As can be seen from Table 4.1, we realize that migrant parents are younger, more educated than local parents. Local

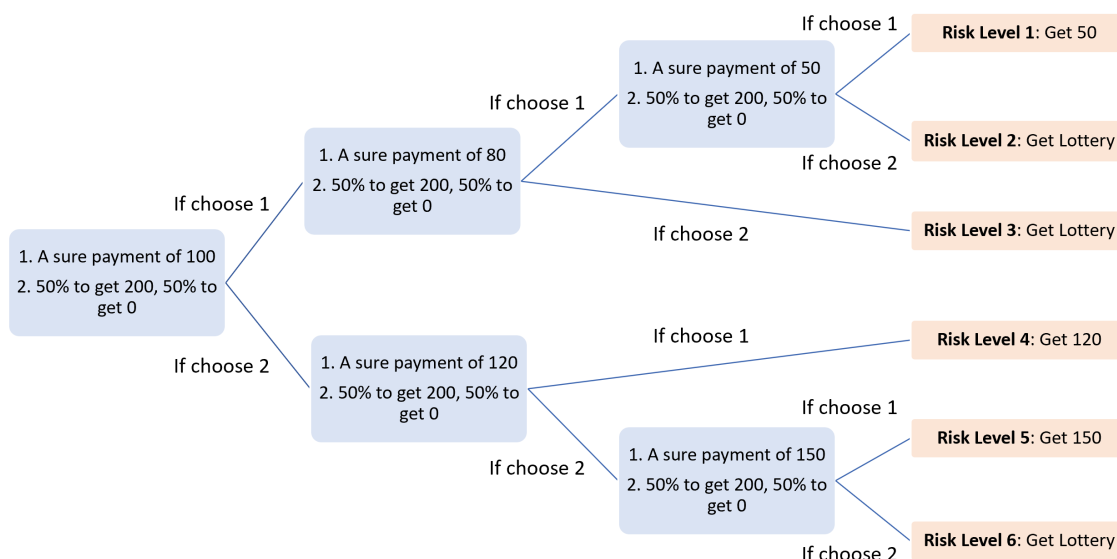


Figure 4.2: Survey Questions and According Outcomes

parents have slightly more child, and are more likely to be self-employed. Since we have subjects' migration experience by adopting the 2010 to 2016 CFPS data sets, we are able to differentiate return parents (parents who ever migrated before) from local parents (no migration experience before). Table 4.1 indicates that there are 10 percent of current local parents have migration experience before.

According to subjects' age, we separate rural local children into two groups - underage rural local children and adult rural local children⁵. From the table we notice that the average age of underage group is 13.79, while that of adult group is 34.27. Adult rural local children seem to have compulsory educational attainment level on average, and almost 93% of them are employed.

The above statistics gives us an idea about the sample. In the following section, we will do some empirical analysis to explore how having a migrant parent affect rural local underage

⁵Underage children are subjects with age below or equal to 18, while adult children are those with age above 18. The way we define adult subjects as children is that if subjects have at least one of their parents' information (family background), we define them as children.

Table 4.1: Characteristics of Parents and Rural Local children

	Parents		Rural Local Children	
	Migrants	Locals	Underage	Adult
Male	0.47	0.47	0.54	0.77
Age	43.69	52.78	13.79	34.27
Years of Education	9.66	7.49	6.11	9.28
Urban <i>Hukou</i> Holder	0.38	0.25	0.03	0.06
Married	0.94	0.90		0.59
Number of Children	1.68	2.00		1.10
Employed	0.73	0.75		0.93
Self-Employed	0.34	0.63		0.52
Ever Migrate		0.10		
Number of Respondent	1,635	18,091	1,349	1,718
Data comes from CFPS				

children's risk preference levels, and also discuss whether parental migration is related to subjects risky outcomes or not.

4.4 Empirical Analysis

4.4.1 Parental Migration and Risk Preference Levels

In this section, we provide some empirical evidence on how having a migrant parent affect children's risk preference levels. As the existing studies indicate, parents' risk preference levels do affect children's risk attitudes. However, most of these studies mainly focus on intergenerational transmission, but ignore other channels that may also have impact on children's risk preferences. In this study, we pay attention to a special parent group - migrant parents, and would like to show that besides intergenerational transmission, parental migration itself could also be influential.

First of all, in order to explore the primary relationship between having a migrant parent

and rural local underage children's risk preference levels, we estimate a linear model with the specification as:

$$RISK_i = \beta_0 + \mathbf{X}_i\delta + \beta_1 HHMIG_i + \varepsilon_i, \quad (4.1)$$

where $RISK_i$ represents i 's risk preference level, $\mathbf{X}_i$ is a vector which indicates one's characteristics, such as gender, education years, household registration information, and family background (parent's education and age information), $HHMIG_i$ is a dummy variable which indicates whether i 's parent migrate or not⁶, and ε_i is an error term with zero mean. The coefficient β_1 represents the relationship between having a migrant parent and children's risk preference levels. Subjects' characteristics and family background $\mathbf{X}_i$ are also included in this model so that the impact from other potential factors would be controlled.

By using the 2018 CFPS data, we estimate the linear model in equation (1), and show the according results in Table 4.2. In the first column of Table 4.2, we only control subjects' background, while in the following columns we also control family background as well. The results indicate that, generally speaking, children who live in a household with migrant parent would have a higher risk tolerance level. In Column (2) of Table 4.2, we find that the risk preference level of rural local underage children with migrant parent is 0.492 higher than the risk level of those with all local parents.

When we separate the sample by gender in Column (3) and (4), we notice that household with migrant parent is more influential among rural local girls. For rural local underage girl sample, girls with migrant parent are more likely to have higher risk levels by 0.956 when comparing to girls with no migrant parent. However, such migrant parent impact on

⁶The reason why we do not differentiate migrant father and migrant mother is that parental migration decision is a kind of family decision within a household. According to [Dustmann et al. \(2020\)](#), one's migration is related to the risk attitude of household members. Although migrant father and migrant mother may have different impact on children, their migration can be regard as a decision of both parents. In the appendix, we also show how different household types affect children's risk preference levels. See Appendix B Table B.1.

children's risk preference level is not significant for rural local underage boy sample.

Table 4.2: Having a Migrant Parent and Rural Local Underage Children's Risk Preference Levels

	(1) All	(2) All	(3) Girl	(4) Boy
Male	0.029 (0.106)	0.101 (0.129)		
Age	-0.119 (0.252)	-0.245 (0.313)	-0.316 (0.452)	-0.130 (0.440)
Age Square/100	0.553 (0.897)	1.238 (1.103)	1.852 (1.583)	0.443 (1.559)
Education Years	-0.052 (0.043)	-0.103** (0.052)	-0.171** (0.074)	-0.019 (0.072)
Urban <i>Hukou</i>	0.087 (0.349)	0.195 (0.382)	-1.434** (0.597)	1.014** (0.514)
Household with Migrant Parent	0.160 (0.190)	0.492** (0.250)	0.956*** (0.368)	0.260 (0.345)
Average Years of Parents' Education		0.029 (0.030)	-0.024 (0.044)	0.074* (0.042)
Average Age of Parents		-0.015 (0.013)	-0.017 (0.019)	-0.008 (0.018)
Province Dummies	Yes	Yes	Yes	Yes
Constant	4.069** (1.845)	5.075** (2.318)	4.443 (3.441)	5.148 (3.192)
Observations	1345	948	434	514
r2_a	0.004	0.007	0.039	0.013

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Up to now we find that there is a positive correlation between having a migrant parent and girls risk preference levels. We would wonder whether this is because migrant parents are more risk tolerant, children with migrant parent would have a higher risk preference level due to the intergenerational transmission. To see whether it is due to an intergenerational transmission, we first compare migrant parents' risk preference levels with that of locals in Figure 4.3. From the t -test result, we conclude that migrant parents are more risk tolerant than local parents in general. For accuracy, in the following Table 4.3, we also control for some other potential factors that could affect one's risk preference levels, and still able to conclude that migrant parents are more risk tolerant.

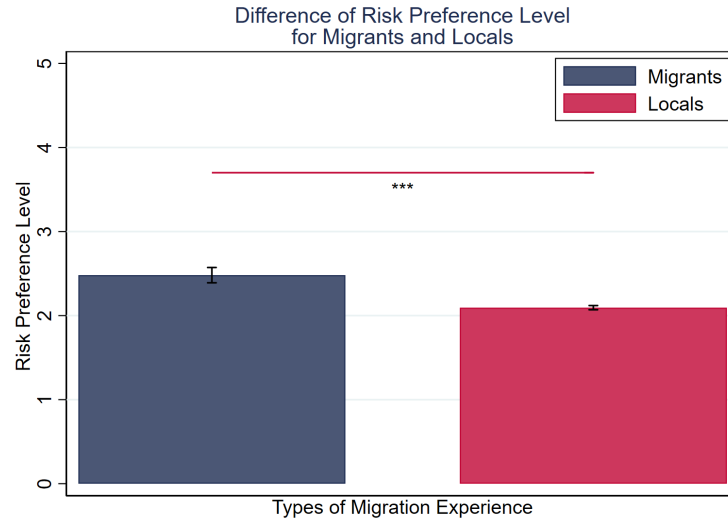
Figure 4.3: Risk Preference Level t - test for Parents with Different Migration Experience

Table 4.3: Relationship between Parents' Risk Preference Levels and Migration Status

	Parent Sample Risk Preference Level
Male	0.239*** (0.025)
Age	-0.046*** (0.006)
Migrant	0.138*** (0.045)
Province Dummies	Yes
Constant	3.835*** (0.223)
Observations	19292
r2_a	0.053

Standard errors in parentheses

Age Square, education years, and *Hukou* status are also controlled* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Although current evidence shows us that migrant parents have higher risk preference levels than local parents, we still wonder whether “migration” changes parents’ risk preference levels. Therefore, in Table 4.4, we use the propensity score matching (PSM) method to explore the average treatment effects on treated group (parents who migrate) with local parents as the according control group. By using different matching methods, we find that

the ATTs are not significant. So we conclude that migration itself cannot significantly change parents' risk preference levels.

Table 4.4: Estimates of the ATT with Local Parents as the Control Group (Migration Impact on Risk Levels)

	Number of Treated (Return Migrants)	Number of Control (Migrants)	ATT	S.E.	t
Nearest Neighbor Matching	348	620	0.036	0.158	0.231
Kernel-Based Matching	348	1823	0.154	0.111	1.384
Stratification Matching	348	1823	0.115	0.114	1.014

Data comes from CFPS
 * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Similarly, we also have to check whether the local parents who ever migrated before have fluctuation in the risk attitudes after their return. So we look over whether “return” causes parents' risk preference levels change. By using the propensity score matching, we explore the average treatment effects on treated group (return migrants) with migrant parents as the according control group. From the results in Table 4.5, we notice that “return” could almost not change parents' risk preference levels, which once again verify that parents' risk preference is stable⁷.

Table 4.5: Estimates of the ATT with Migrant Parents as the Control Group (Return Impact on Risk Levels)

	Number of Treated (Return Migrants)	Number of Control (Migrants)	ATT	S.E.	t
Nearest Neighbor Matching	193	100	-0.215	0.280	-0.765
Kernel-Based Matching	193	155	-0.272	0.209	-1.304
Stratification Matching	193	155	-0.298	0.215	-1.384

Data comes from CFPS
 * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Results in Table 4.2 suggest that there is a positive correlation between having a migrant parent and children's risk preference levels, especially for girls. Now that we find parents'

⁷Existing literature point out that one's risk preference does not change with wealth (Brunnermeier and Nagel, 2008) or migration status (Jaeger et al., 2010). Chuang and Schechter (2015) also conclude that risk preference that measured by survey is stable over time.

risk preference levels are stable, and migrant parents do have higher risk tolerance than local parents, is the positive correlation in Table 4.2 due to an intergenerational transmission of risk preferences between migrant parents and their children? To answer this question, we control parents' risk preference levels in Table 4.6 to find out whether children's risk preference levels are totally affected by parents' risk preference levels or not.

Table 4.6: Parental Migration and Rural Local Underage Children's Risk Preference Levels

	(1) All	(2) All	(3) Girl	(4) Boy	(5) Girl	(6) Boy
Male	0.029 (0.106)	0.079 (0.128)				
Age	-0.119 (0.252)	-0.214 (0.310)	-0.296 (0.448)	-0.101 (0.436)	-0.393 (0.450)	-0.136 (0.437)
Age Square/100	0.553 (0.897)	1.103 (1.093)	1.702 (1.569)	0.343 (1.547)	2.050 (1.573)	0.456 (1.549)
Education Years	-0.052 (0.043)	-0.097* (0.051)	-0.152** (0.074)	-0.022 (0.072)	-0.158** (0.074)	-0.021 (0.072)
Urban <i>Hukou</i>	0.087 (0.349)	0.046 (0.380)	-1.519** (0.592)	0.814 (0.514)	-1.600*** (0.587)	0.839 (0.516)
Household with Migrant Parent	0.160 (0.190)	0.530** (0.248)	1.059*** (0.366)	0.238 (0.342)	1.029*** (0.363)	0.254 (0.343)
Average Years of Parents' Education		0.020 (0.030)	-0.039 (0.043)	0.070* (0.042)	-0.024 (0.044)	0.078* (0.043)
Average Risk Level of Parents		0.197*** (0.047)	0.207*** (0.070)	0.190*** (0.063)		
Average Age of Parents		-0.010 (0.013)	-0.016 (0.019)	0.001 (0.019)	-0.012 (0.019)	0.002 (0.019)
Father's Risk Level (1=Risk Averse)					-0.016 (0.053)	0.105** (0.048)
Mother's Risk Level (1=Risk Averse)					0.227*** (0.055)	0.078 (0.052)
Province Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	4.069** (1.845)	4.309* (2.304)	3.975 (3.412)	4.220 (3.182)	4.567 (3.424)	4.346 (3.185)
Observations	1345	948	434	514	428	510
r2_a	0.004	0.025	0.056	0.029	0.077	0.028

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In Table 4.6, we estimate the linear regression model based on the specification in equation (1), and also add parents' risk preference levels as subjects' family background. We control

subjects' basic characteristics in the first column, and add family background with parents' average risk preference levels in the second column. In Column (3) to (6), we separate the sample into girls and boys to find out the impact of parental migration in different gender groups. We control parents' average risk preference levels in Column (3) and (4), and control father's and mother's risk preference levels separately in Column (5) and (6).

Results in Table 4.6 indicate that parents' risk preference levels have positive significant correlation with rural local underage children's risk preference levels. In Column (5) and (6), the results suggest that boys' risk preference levels are mainly influenced by fathers' risk attitudes, while girls are not only affected by mothers' risk preference levels, but also affected by parental migration. Even after controlling the intergenerational transmission of risk preferences, girl's risk preference level still increases when there is a migrant parent within her household⁸.

4.4.2 Parental Migration and Risk Related Outcome

In this subsection, we further check whether parental migration could affect children's other risk related outcomes or not. Based on the existing literature, one's risk preference can be correlated with various outcomes, such as one's health status, investment behavior, and even career choice. For having a better understanding on subjects' risk related outcomes, we use adult children as our sample, and find out the relationship between their parental migration and self-employment decisions.

⁸One's migration choice seems to be not a random but a self-selected behavior. As parents, their migration choices could be influenced by many factors. For checking whether there is a self-selection problem, we focus on the girl sample only, and use the method of Heckman Two-Step. In the first step of the regression, we use the numbers of children for parents as the exclusive restriction, and also control parents' characteristics. However, we do not find a significant self-selection problem by running this regression, so we currently stick with the results in Table 4.6. We try to also include the log of saving per capita as the exclusive restriction, but there are too many missing values in the data set. Therefore, if more data could be involved in the future, we probably can have better exclusive restrictions in resolving the self-selection problem.

We do a probit regression with the model:

$$SELF_i = \beta_0 + \mathbf{X}_i\delta + \beta_1 HHMIG_i + \varepsilon_i, \quad (4.2)$$

where $SELF_i$ is a dummy variable which indicates whether i is self-employed or not. $\mathbf{X}_i$ is a vector variable which includes one's gender, age, education years, household registration information, and family background. $HHMIG_i$ equals to 1 means i is in a household with at least one migrant parent, while $HHMIG_i$ equals to 0 represents there is no migrant parent within the household of i . Finally, ε_i is an error term with zero mean. β_1 is the main coefficient that we are interested in, which shows the relationship between parental migration and children's self-employment decisions. The according regression results are shown in Table 4.7.

Table 4.7: Parental Migration and Rural Local Adult Children's Self-Employment (Probit)

	(1) All	(2) Daughter	(3) Son
Male	-0.043 (-1.257)		
Age	0.014* (1.720)	-0.007 (-0.295)	0.008 (0.896)
Age Square/100	-0.004 (-0.428)	0.038 (0.977)	0.003 (0.254)
Education Years	-0.028*** (-7.765)	-0.019*** (-3.351)	-0.024*** (-5.533)
Married	0.120*** (4.091)	0.156*** (3.154)	0.083** (2.405)
Household with Migrant Parent	0.126* (1.778)	0.194* (1.704)	0.084 (1.007)
Urban <i>Hukou</i>	-0.144** (-2.750)	-0.478** (-2.141)	-0.139** (-2.339)
At Least One Parent Self-Employed	0.060* (1.658)	0.083 (1.453)	0.059 (1.371)
Province Dummies	Yes	Yes	Yes
Observations	1,364	231	1,119

Report margins in the table

z-statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In the first column of Table 4.7, we use all the rural local adult children as the sample, and separate the sample into daughter sample and son sample in the following two columns. The results suggest that children’s self-employment decisions are correlated with their education years, marital status, and household registration status. Additionally, daughters’ self-employment decisions are also positively correlated with their parental migration⁹. The results indicate that the probability of being self-employed is higher by 0.194 for daughters with migrant parent¹⁰.

4.5 Conclusion and Future Work

This paper mainly investigates the impact of parental migration on rural local underage children’s risk preference levels. Migrant parents are often regard to be risk tolerant since one’s migration action is identified as risky choice. However, children with migrant parent in the household not only have a high risk tolerance genetically, but also have different family environment when comparing to children with local parents. Thus, for the special parent group - migrant parents - whether they only affect children’s risk preference levels by intergenerational transmission or not becomes the question of our interest.

The findings indicate that there is a positive relationship between parental migration and rural local underage girls’ risk preference levels. Even after we control the parents’ risk preference levels, or use a different model to resolve the self-selection problem, we still have such positive correlation between parental migration and girls’ risk preference levels. The according rural local adult daughters’ self-employment decisions are also positively correlated

⁹We control the household type instead of “household with migrant parent” in the later Appendix B Table B.2.

¹⁰However, there is an unbalanced adult daughter and son sample in the regression. Since CFPS is a family survey data, there would be more adult daughters who do not live with parents within a household than adult sons. We also have some reservations about it.

with parental migration. Therefore, we would like to conclude that besides migrant parents' intergenerational transmission of risk preferences, parents' migration action itself could also have an important impact on children's risk preference levels.

In fact, except for current results, we also want to check whether parental migration impact is a temporary or permanent impact, whether such parental migration impact disappear or not after migrant parents' return. However, currently, we only have one wave of risk preference data in the most recent 2018 data set. In the future work, if we can have the next wave of CFPS data, we could look over the consistency of parental migration impact on children's risk preferences over time.

Bibliography

- Akgüç, M., Liu, X., Tani, M., and Zimmermann, K. F. (2016). Risk Attitudes and Migration. *China Economic Review*, 37:166–176.
- Alan, S., Baydar, N., Boneva, T., Crossley, T. F., and Ertac, S. (2017). Transmission of Risk Preferences from Mothers to Daughters. *Journal of Economic Behavior & Organization*, 134:60–77.
- Allen, W. D. (2000). Social Networks and Self-Employment. *The Journal of Socio-Economics*, 29(5):487–501.
- Anderson, L. R. and Mellor, J. M. (2008). Predicting Health Behaviors with An Experimental Measure of Risk Preference. *Journal of health economics*, 27(5):1260–1274.
- Andreoni, J., Di Girolamo, A., List, J. A., Mackevicius, C., and Samek, A. (2020). Risk Preferences of Children and Adolescents in Relation to Gender, Cognitive Skills, Soft Skills, and Executive Functions. *Journal of Economic Behavior & Organization*, 179:729–742.
- Antman, F. M. (2011). The Intergenerational Effects of Paternal Migration on Schooling and Work: What Can We Learn from Children’s Time Allocations? *Journal of Development Economics*, 96(2):200–208.
- Antman, F. M. (2012). Gender, Educational Attainment, and the Impact of Parental Migration on Children Left Behind. *Journal of Population Economics*, 25(4):1187–1214.
- Arrow, K. J. and Borzekowski, R. (2004). Limited Network Connections and the Distribution of Wages.

- Barbosa, S. D., Gerhardt, M. W., and Kickul, J. R. (2007). The Role of Cognitive Style and Risk Preference on Entrepreneurial Self-Efficacy and Entrepreneurial Intentions. *Journal of Leadership & Organizational Studies*, 13(4):86–104.
- Bayer, P., Ross, S. L., and Topa, G. (2008). Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy*, 116(6):1150–1196.
- Beaman, L. A. (2012). Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S. *The Review of Economic Studies*, 79(1):128–161.
- Borjas, G. J. (1987). Self-Selection and the Earnings of Immigrants. *The American Economic Review*, pages 531–553.
- Bosker, M., Brakman, S., Garretsen, H., and Schramm, M. (2012). Relaxing Hukou: Increased Labor Mobility and China’s Economic Geography. *Journal of Urban Economics*, 72(2):252 – 266.
- Brown, H. and Van der Pol, M. (2015). Intergenerational Transfer of Time and Risk Preferences. *Journal of Economic Psychology*, 49:187–204.
- Brunnermeier, M. K. and Nagel, S. (2008). Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-Evidence on Individuals. *American Economic Review*, 98(3):713–36.
- Cadsby, C. B., Song, F., and Yang, X. (2019a). Are “Left-Behind” Children Really Left Behind? A Lab-in-Field Experiment Concerning the Impact of Rural/Urban Status and Parental Migration on Children’s Other-Regarding Preferences. *Journal of Economic Behavior & Organization*.

- Cadsby, C. B., Song, F., and Yang, X. (2019b). Dishonesty among Children: Rural/Urban Status and Parental Migration. In *Dishonesty in Behavioral Economics*, pages 31–52. Elsevier.
- Calvo-Armengol, A. and Jackson, M. O. (2004). The Effects of Social Networks on Employment and Inequality. *American Economic Review*, 94(3):426–454.
- Calvó-Armengol, A. and Jackson, M. O. (2007). Networks in Labor Markets: Wage and Employment Dynamics and Inequality. *Journal of Economic Theory*, 132(1):27–46.
- Card, D. and Krueger, A. B. (1992a). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *Journal of Political Economy*, 100(1):1–40.
- Card, D. and Krueger, A. B. (1992b). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *Journal of Political Economy*, 100(1):1–40.
- Chan, K. W. and Zhang, L. (1999). The Hukou System and Rural-Urban Migration in China: Processes and Changes. *The China Quarterly*, 160:818–855.
- Chang, F., Jiang, Y., Loyalka, P., Chu, J., Shi, Y., Osborn, A., and Rozelle, S. (2019). Parental Migration, Educational Achievement, and Mental Health of Junior High School Students in Rural China. *China Economic Review*, 54:337–349.
- Chang, H., Dong, X.-y., and MacPhail, F. (2011). Labor Migration and Time Use Patterns of the Left-Behind Children and Elderly in Rural China. *World Development*, 39(12):2199–2210.
- Charness, G. and Gneezy, U. (2012). Strong Evidence for Gender Differences in Risk Taking. *Journal of Economic Behavior & Organization*, 83(1):50–58.

- Chen, Y., Jin, G. Z., and Yue, Y. (2010). Peer Migration in China. Technical report, National Bureau of Economic Research.
- Chen, Y., Wang, L., and Zhang, M. (2018). Informal Search, Bad Search?: The Effects of Job Search Method on Wages among Rural Migrants in Urban China. *Journal of Population Economics*, 31(3):837–876.
- Chuang, Y. and Schechter, L. (2015). Stability of Experimental and Survey Measures of Risk, Time, and Social Preferences: A Review and Some New Results. *Journal of development economics*, 117:151–170.
- Colas, M. and Ge, S. (2018). Transformations in China’s Internal Labor Migration and Hukou System. *Available at SSRN*.
- Dave, D. and Saffer, H. (2008). Alcohol Demand and Risk Preference. *Journal of Economic Psychology*, 29(6):810–831.
- Dewen, W. (2003). China’s Rural Compulsory Education: Current Situation, Problems and Policy Alternatives. Technical report, Working paper series.
- DiNardo, J., Fortin, N. M., and Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica: Journal of the Econometric Society*, pages 1001–1044.
- Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2012). The Intergenerational Transmission of Risk and Trust Attitudes. *The Review of Economic Studies*, 79(2):645–677.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the european economic association*, 9(3):522–550.

- Dolfin, S. and Genicot, G. (2010). What Do Networks Do? The Role of Networks on Migration and “Coyote” Use. *Review of Development Economics*, 14(2):343–359.
- Du, Y., Park, A., and Wang, S. (2005). Migration and Rural Poverty in China. *Journal of Comparative Economics*, 33(4):688–709.
- Dustmann, C., Fasani, F., Meng, X., and Minale, L. (2020). Risk Attitudes and Household Migration Decisions. *Journal of Human Resources*, pages 1019–10513R1.
- Fan, C. C. (2008). Migration, Hukou, and the City. *China Urbanizes: Consequences, Strategies, and Policies*, pages 65–89.
- Fu, Q. and Ren, Q. (2010). Educational Inequality under China’s Rural-Urban Divide: The Hukou System and Return to Education. *Environment and Planning A: Economy and Space*, 42(3):592–610.
- Ge, S. and Jin, S. (2020). Self-Selection and the Wage of Migrants in China. *Working Paper*.
- Ge, S. and Yang, D. T. (2009). Accounting for Rising Wages in China. *Manuscript, Virginia Tech and Chinese University of Hong Kong*.
- Giannelli, G. C. and Mangiavacchi, L. (2010). Children’s Schooling and Parental Migration: Empirical Evidence on the ‘Left-Behind’ Generation in Albania. *Labour*, 24:76–92.
- Giulietti, C., Guzi, M., Zhao, Z., Zimmermann, K. F., et al. (2010). Social Networks and the Labour Market Outcomes of Rural to Urban Migrants in China. *Unpublished Manuscript*.
- Giulietti, C., Wahba, J., and Zenou, Y. (2018). Strong versus Weak Ties in Migration. *European Economic Review*, 104:111–137.
- Hao, L., Houser, D., Mao, L., and Villeval, M. C. (2016). Migrations, Risks, and Uncertainty: A Field Experiment in China. *Journal of Economic Behavior & Organization*, 131:126–140.

- Hashimzade, N., Myles, G. D., Page, F., and Rablen, M. D. (2014). Social Networks and Occupational Choice: The Endogenous Formation of Attitudes and Beliefs about Tax Compliance. *Journal of Economic Psychology*, 40:134–146.
- He, B., Fan, J., Liu, N., Li, H., Wang, Y., Williams, J., and Wong, K. (2012). Depression Risk of ‘Left-Behind Children’ in Rural China. *Psychiatry research*, 200(2-3):306–312.
- Hou, W., Bao, S., and Shi, A. (2005). Analysis of the Changing Spatial Patterns of Migration in China. *Population Science of China*, 5.
- Ioannides, Y. M. and Soetevent, A. R. (2006). Wages and Employment in A Random Social Network with Arbitrary Degree Distribution. *American Economic Review*, 96(2):270–274.
- Jaeger, D. A., Dohmen, T., Falk, A., Huffman, D., Sunde, U., and Bonin, H. (2010). Direct Evidence on Risk Attitudes and Migration. *The Review of Economics and Statistics*, 92(3):684–689.
- Jia, Z. and Tian, W. (2010). Loneliness of Left-Behind Children: A Cross-Sectional Survey in A Sample of Rural China. *Child: care, health and development*, 36(6):812–817.
- Kerr, W. R. and Mandorff, M. (2015). Social Networks, Ethnicity, and Entrepreneurship. Technical report, National Bureau of Economic Research.
- Kuhn, P. and Shen, K. (2015). Do Employers Prefer Migrant Workers? Evidence from a Chinese Job Board. *IZA Journal of Labor Economics*, 4(1):22.
- Lahaie, C., Hayes, J. A., Piper, T. M., and Heymann, J. (2009). Work and Family Divided Across Borders: The Impact of Parental Migration on Mexican Children in Transnational Families. *Community, Work & Family*, 12(3):299–312.
- Lee, R. (2019). How do Parental Migration Histories Matter for Children’s Economic Outcomes? *Migration and Development*, 8(1):75–92.

- Li, Y., Liu, H., Tang, Q., Lu, D., and Xiao, N. (2014). Spatial-Temporal Patterns of China's Interprovincial Migration, 1985-2010. *Journal of Geographical Sciences*, 24(5):907–923.
- Liu, Z. (2005). Institution and Inequality: the Hukou System in China. *Journal of Comparative Economics*, 33(1):133 – 157.
- Liu, Z., Yu, L., and Zheng, X. (2018). No Longer Left-Behind: The Impact of Return Migrant Parents on Children's Performance. *China Economic Review*, 49:184–196.
- Long, W., Appleton, S., and Song, L. (2017). The Impact of Job Contact Networks on Wages of Rural-Urban Migrants in China: A Switching Regression Approach. *Journal of Chinese Economic and Business Studies*, 15(1):81–101.
- Loury, L. D. (2006). Some Contacts Are More Equal than Others: Informal Networks, Job tenure, and Wages. *Journal of Labor Economics*, 24(2):299–318.
- Luo, J., Wang, W., and Gao, W.-B. (2009). Review of the Studies on Rural Left-Behind Children in China. *Advances in Psychological Science*, 17(05):990.
- Lusk, J. L. and Coble, K. H. (2005). Risk Perceptions, Risk Preference, and Acceptance of Risky Food. *American Journal of Agricultural Economics*, 87(2):393–405.
- Lyu, L. and Chen, Y. (2019). Parental Migration and Young Migrants' Wages in Urban China: An Exploratory Analysis. *Urban Studies*, 56(10):1968–1987.
- Malmendier, U. and Nagel, S. (2011). Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? *The quarterly journal of economics*, 126(1):373–416.
- McKenzie, D. and Rapoport, H. (2007). Network Effects and the Dynamics of Migration and Inequality: Theory and Evidence from Mexico. *Journal of Development Economics*, 84(1):1–24.

- Meng, X. and Zhang, J. (2001). 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(3):485 – 504.
- Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 66(4):281–302.
- Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the US Labor Market. *The Quarterly Journal of Economics*, 118(2):549–599.
- Nielsen, T., Keil, A., and Zeller, M. (2013). Assessing Farmers’ Risk Preferences and Their Determinants in A Marginal Upland Area of Vietnam: A Comparison of Multiple Elicitation Techniques. *Agricultural Economics*, 44(3):255–273.
- Patel, K. and Vella, F. (2013). Immigrant Networks and Their Implications for Occupational Choice and Wages. *Review of Economics and Statistics*, 95(4):1249–1277.
- Rieger, M. O., Wang, M., and Hens, T. (2015). Risk Preferences Around the World. *Management Science*, 61(3):637–648.
- Samek, A., Gray, A., Datar, A., and Nicosia, N. (2021). Adolescent Time and Risk Preferences: Measurement, Determinants and Field Consequences. *Journal of Economic Behavior & Organization*, 184:460–488.
- Shuangyue, S. (2018). Mother’s Migration and Children’s Health - Evidence from CFPS Data. In *West Forum on Economy and Management*, number 1, page 10.
- Song, Y. (2014). What Should Economists Know about the Current Chinese Hukou System? *China Economic Review*, 29:200 – 212.
- Song, Y. (2016). Hukou-Based Labour Market Discrimination and Ownership Structure in Urban China. *Urban Studies*, 53(8):1657–1673.

- Sutter, M., Kocher, M. G., Glätzle-Rützler, D., and Trautmann, S. T. (2013). Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior. *American Economic Review*, 103(1):510–31.
- Tanaka, T., Camerer, C. F., and Nguyen, Q. (2010). Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam. *American economic review*, 100(1):557–71.
- Wang, F.-L. (2005). *Organizing through Division and Exclusion: China's Hukou System*. Stanford University Press.
- Whalley, J. and Zhang, S. (2007). A Numerical Simulation Analysis of (Hukou) Labour Mobility Restrictions in China. *Journal of Development Economics*, 83(2):392 – 410. Papers from a Symposium: The Social Dimensions of Microeconomic Behaviour in Low-Income Communities.
- Wing Chan, K. and Buckingham, W. (2008). Is China Abolishing the Hukou System? *The China Quarterly*, 195:582–606.
- Wu, H. X. (1994). Rural to Urban Migration in the People's Republic of China. *The China Quarterly*, 139:669–698.
- Wu, H. X. and Zhou, L. (1996). Rural-to-Urban Migration in China. *Asian-Pacific Economic Literature*, 10(2):54–67.
- Wu, Q., Lu, D., and Kang, M. (2015). Social Capital and the Mental Health of Children in Rural China with Different Experiences of Parental Migration. *Social science & medicine*, 132:270–277.
- Wu, X. and Treiman, D. J. (2004). The Household Registration System and Social Stratification in China: 1955–1996. *Demography*, 41(2):363–384.

- Xie, Y. and Hu, J. (2014). An Introduction to the China Family Panel Studies (CFPS). *Chinese sociological review*, 47(1):3–29.
- Xie, Y. and Lu, P. (2015). The Sampling Design of the China Family Panel Studies (CFPS). *Chinese journal of sociology*, 1(4):471–484.
- Ye, J., Bo, S., Cong, L., and Zhou, Y. (2012). The Levels of Social Network and Wage of Rural-Urban Migrants: From a Perspective of Identity Model. *Economic Review*, 4.
- Yogo, U. T. (2011). Social Network and Wage: Evidence from Cameroon. *Labour*, 25(4):528–543.
- Zhang, H. (2010). The Hukou System’s Constraints on Migrant Workers’ Job Mobility in Chinese Cities. *China Economic Review*, 21(1):51 – 64.
- Zhang, J. and Zhao, Z. (2015). Social-Family Network and Self-Employment: Evidence from Temporary Rural-Urban Migrants in China. *IZA Journal of Labor & Development*, 4(1):4.
- Zhang, L. and Tao, L. (2012). Barriers to the Acquisition of Urban Hukou in Chinese Cities. *Environment and Planning A*, 44(12):2883–2900.
- Zhang, X. and Li, G. (2003). Does Guanxi Matter to Nonfarm Employment? *Journal of Comparative Economics*, 31(2):315–331.
- Zhang, Y., Mouhoud, E., and Fan, Y. (2012). Heterogeneous Social Network and Rural Migrant Worker’s Wage: Evidence from China. *South China Journal of Economics*, 2.
- Zhang, Y. and Wang, H. (2011). *Hukou* Discrimination and Geographical Discrimination in Urban Labor Market: A Study Based on Census Data. *Guan Li Shi Jie (Management World)*, (7):42–51.

- Zhao, Q., Yu, X., Wang, X., and Glauben, T. (2014). The Impact of Parental Migration on Children's School Performance in Rural China. *China Economic Review*, 31:43–54.
- Zhao, Y. (2003). The Role of Migrant Networks in Labor Migration: The Case of China. *Contemporary Economic Policy*, 21(4):500–511.
- Zhou, C., Sylvia, S., Zhang, L., Luo, R., Yi, H., Liu, C., Shi, Y., Loyalka, P., Chu, J., Medina, A., et al. (2015). China's Left-Behind Children: Impact of Parental Migration on Health, Nutrition, and Educational Outcomes. *Health affairs*, 34(11):1964–1971.
- Zyphur, M. J., Narayanan, J., Arvey, R. D., and Alexander, G. J. (2009). The Genetics of Economic Risk Preferences. *Journal of Behavioral Decision Making*, 22(4):367–377.

Appendices

Appendix A

Appendix to Chapter 3

A.1 Relationship between network and the percentage of early migrants $e(net_c)$

How does the percentage of early migrants $e(net_c)$ change with net_c :

$$\begin{aligned}\frac{\partial e(net_c)}{\partial net_c} &= \frac{\partial \frac{en_c}{net_c}}{\partial net_c} = \frac{\frac{\partial en_c}{\partial net_c} net_c - en_c}{net_c^2} \\ &= \frac{1}{net_c^2} \left[\frac{\partial [net_c - rn_c(en_c)]}{\partial net_c} net_c - en_c \right] \\ &= \frac{1}{net_c^2} \left[net_c - \frac{\partial rn_c(en_c)}{\partial net_c} net_c - net_c + rn_c(en_c) \right] \\ &= \frac{1}{net_c^2} \left[rn_c(en_c) - \frac{\partial rn_c(en_c)}{\partial [en_c + rn_c(en_c)]} net_c \right].\end{aligned}$$

Let $en_c + rn_c(en_c) = k(en_c) = y$, $en_c = k^{-1}(y)$. Therefore,

$$\begin{aligned}\frac{\partial rn_c(en_c)}{\partial [en_c + rn_c(en_c)]} &= \frac{\partial [y - k^{-1}(y)]}{\partial y} = 1 - \frac{1}{k'[k^{-1}(y)]} \\ &= 1 - \frac{1}{k'(en_c)} = 1 - \frac{1}{1 + rn'_c(en_c)}.\end{aligned}$$

$$\begin{aligned}
\frac{\partial e(net_c)}{\partial net_c} &= \frac{1}{net_c^2} [rn_c(en_c) - (1 - \frac{1}{1 + rn'_c(en_c)}) \times net_c] \\
&= \frac{1}{net_c^2} \times \frac{rn_c(en_c) - rn'_c(en_c)en_c}{1 + rn'_c(en_c)}.
\end{aligned}$$

A.2 Proof of Propositions

A.2.1 Case 1: when there is no self-selection

Proposition 3.1. *If $e'(net_c) \geq 0$, s_c (weakly) increases as net_c increases, $\forall net_c \in [0, 1]$. If $e'(net_c) < 0$, s_c (weakly) decreases as net_c increases, $\forall net_c \in [0, 1]$.*

Proof.

$$\frac{\partial s_c}{\partial net_c} = \frac{a'(net_c)[1 - (1 - b)e(net_c)] + (1 - b)e'(net_c)a(net_c)}{[1 - (1 - b)e(net_c)]^2}, \quad (\text{A.1})$$

$$= \frac{a'(net_c) - (1 - b)[e(net_c)a'(net_c) - e'(net_c)a(net_c)]}{[1 - (1 - b)e(net_c)]^2}. \quad (\text{A.2})$$

We know that there is no self-selection on migrants ($a'(net_c) = 0$).

(1) If $e'(net_c) \geq 0$, we have $a'(net_c) - (1 - b)[e(net_c)a'(net_c) - e'(net_c)a(net_c)] = (1 - b)e'(net_c)a(net_c) \geq 0$, $\frac{\partial s_c}{\partial net_c} \geq 0$.

(2) If $e'(net_c) < 0$, $a'(net_c) - (1 - b)[e(net_c)a'(net_c) - e'(net_c)a(net_c)] = (1 - b)e'(net_c)a(net_c) \leq 0$, $\frac{\partial s_c}{\partial net_c} \leq 0$.

□

A.2.2 Case 2: when there is a positive self-selection

Proposition 3.2. *If $e'(net_c) \geq 0$, s_c (weakly) increases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \in [-\frac{(1-b)e'(net_c)}{1-(1-b)e(net_c)}, 0]$; s_c (weakly) decreases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \in (-\infty, -\frac{(1-b)e'(net_c)}{1-(1-b)e(net_c)}]$. If $e'(net_c) < 0$, s_c (weakly) decreases as net_c increases, $\forall net_c \in [0, 1]$.*

Proof. In this case, the probability of being employed in a direct way decreases as the network increases, so $a'(net_c) \leq 0$.

(1) According to equation (8), if $e'(net_c) \geq 0$:

when $a'(net_c) - (1-b)[e(net_c)a'(net_c) - e'(net_c)a(net_c)] \geq 0 \Leftrightarrow a'(net_c) \geq -\frac{(1-b)e'(net_c)a(net_c)}{1-(1-b)e(net_c)}$, $\frac{\partial s_c}{\partial net_c} \geq 0$. Therefore, we have $\frac{\partial s_c}{\partial net_c} \geq 0$ when $\frac{a'(net_c)}{a(net_c)} \in [-\frac{(1-b)e'(net_c)}{1-(1-b)e(net_c)}, 0]$;

when $a'(net_c) - (1-b)[e(net_c)a'(net_c) - e'(net_c)a(net_c)] \leq 0 \Leftrightarrow a'(net_c) \leq -\frac{(1-b)e'(net_c)a(net_c)}{1-(1-b)e(net_c)}$, $\frac{\partial s_c}{\partial net_c} \leq 0$. Therefore, we have $\frac{\partial s_c}{\partial net_c} \leq 0$ when $\frac{a'(net_c)}{a(net_c)} \in (-\infty, -\frac{(1-b)e'(net_c)}{1-(1-b)e(net_c)}]$.

(2) If $e'(net_c) < 0$, $a'(net_c) - (1-b)[e(net_c)a'(net_c) - e'(net_c)a(net_c)] \leq 0$, $\frac{\partial s_c}{\partial net_c} \leq 0$.

□

A.2.3 Case 3: when there is a negative self-selection

Proposition 3.3. *If $e'(net_c) \geq 0$, s_c (weakly) increases as net_c increases, $\forall net_c \in [0, 1]$. If $e'(net_c) < 0$, s_c (weakly) increases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \in [-\frac{(1-b)e'(net_c)}{1-(1-b)e(net_c)}, +\infty)$; s_c*

(weakly) decreases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \in [0, -\frac{(1-b)e'(net_c)}{1-(1-b)e(net_c)}]$.

Proof. In this case, we have $a'(net_c) \geq 0$.

(1) If $e'(net_c) \geq 0$, we have $a'(net_c) - (1-b)[e(net_c)a'(net_c) - e'(net_c)a(net_c)] \geq 0$, $\frac{\partial s_c}{\partial net_c} \geq 0$.

(2) According to equation (8), if $e'(net_c) < 0$:

when $a'(net_c) - (1-b)[e(net_c)a'(net_c) - e'(net_c)a(net_c)] \geq 0 \Leftrightarrow a'(net_c) \geq -\frac{(1-b)e'(net_c)a(net_c)}{1-(1-b)e(net_c)}$, $\frac{\partial s_c}{\partial net_c} \geq 0$. Since $a'(net_c) \geq 0$, we have $\frac{\partial s_c}{\partial net_c} \geq 0$ when $\frac{a'(net_c)}{a(net_c)} \in [-\frac{(1-b)e'(net_c)}{1-(1-b)e(net_c)}, +\infty)$;

when $a'(net_c) - (1-b)[e(net_c)a'(net_c) - e'(net_c)a(net_c)] \leq 0 \Leftrightarrow a'(net_c) \leq -\frac{(1-b)e'(net_c)a(net_c)}{1-(1-b)e(net_c)}$, $\frac{\partial s_c}{\partial net_c} \leq 0$. Therefore, we have $\frac{\partial s_c}{\partial net_c} \leq 0$ when $\frac{a'(net_c)}{a(net_c)} \in [0, -\frac{(1-b)e'(net_c)}{1-(1-b)e(net_c)}]$.

□

A.2.4 Case 1: when there is no self-selection

Proposition 3.4. *If $e'(net_c) \geq 0$, all employed migrants' average wage (weakly) increases as net_c increases when G first-order stochastically dominates F ($G(x) \leq F(x)$); all employed migrants' average wage (weakly) decreases as net_c increases when $G(x) \geq F(x)$. If $e'(net_c) < 0$, all employed migrants' average wage (weakly) increases as net_c increases when $G(x) \geq F(x)$; all employed migrants' average wage (weakly) decreases as net_c increases when $G(x) \leq F(x)$.*

Proof. Let us discuss the proportion of employed individuals whose wages are greater than x for all job seekers, $\forall x$. For all job seekers,

$$\begin{aligned}
\% \text{ of } \frac{\text{wage} > x}{\text{employed}} &= \frac{\% \text{ early employed wage} > x}{\% \text{ early employed} + (\text{early unemployed} + \text{new}) \times \% \text{ of being employed}} \\
&+ \frac{[\text{early employed wage} < x + (\text{early unemployed} + \text{new})] \times \% \text{ of current wage} > x}{\% \text{ early employed} + (\text{early unemployed} + \text{new}) \times \% \text{ of being employed}}, \\
&= \frac{(1-b)e(\text{net}_c)(1-G(x))}{(1-b)e(\text{net}_c) + [1 - (1-b)e(\text{net}_c)] \times (a(\text{net}_c) + r_c)} \\
&+ \frac{[1 - (1-b)e(\text{net}_c)(1-G(x))] \times [a(\text{net}_c)(1-F(x)) + \frac{a(\text{net}_c)(1-b)e(\text{net}_c)(1-G(x))}{1-(1-b)e(\text{net}_c)(1-G(x))}(1-F(x))]}{(1-b)e(\text{net}_c) + [1 - (1-b)e(\text{net}_c)] \times (a(\text{net}_c) + r_c)}, \\
&= \frac{(1-b)e(\text{net}_c)(1-G(x)) + a(\text{net}_c)(1-F(x))}{(1-b)e(\text{net}_c) + a(\text{net}_c)}.
\end{aligned}$$

$$\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial \text{net}_c} = \frac{[(1-b)(1-G(x))e'(\text{net}_c) + (1-F(x))a'(\text{net}_c)][(1-b)e(\text{net}_c) + a(\text{net}_c)]}{[(1-b)e(\text{net}_c) + a(\text{net}_c)]^2} \quad (\text{A.3})$$

$$- \frac{[(1-b)e'(\text{net}_c) + a'(\text{net}_c)][(1-b)e(\text{net}_c)(1-G(x)) + a(\text{net}_c)(1-F(x))]}{[(1-b)e(\text{net}_c) + a(\text{net}_c)]^2}, \quad (\text{A.4})$$

$$= \frac{(1-b)(1-F(x))[e(\text{net}_c)a'(\text{net}_c) - e'(\text{net}_c)a(\text{net}_c)]}{[(1-b)e(\text{net}_c) + a(\text{net}_c)]^2} \quad (\text{A.5})$$

$$+ \frac{(1-b)(1-G(x))[a(\text{net}_c)e'(\text{net}_c) - e(\text{net}_c)a'(\text{net}_c)]}{[(1-b)e(\text{net}_c) + a(\text{net}_c)]^2}, \quad (\text{A.6})$$

$$= \frac{(1-b)}{[(1-b)e(\text{net}_c) + a(\text{net}_c)]^2} [e(\text{net}_c)a'(\text{net}_c) - e'(\text{net}_c)a(\text{net}_c)][G(x) - F(x)]. \quad (\text{A.7})$$

we know that, in this case, $a'(\text{net}_c) = 0$.

(1) According to equation (13), if $e'(net_c) \geq 0$,

when $G(x) - F(x) \geq 0 \Leftrightarrow G(x) \geq F(x)$, we have $\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial net_c} \leq 0$;

when $G(x) - F(x) \leq 0 \Leftrightarrow G(x) \leq F(x)$, we have $\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial net_c} \geq 0$.

(2) If $e'(net_c) < 0$,

when $G(x) - F(x) \geq 0 \Leftrightarrow G(x) \geq F(x)$, we have $\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial net_c} \geq 0$;

when $G(x) - F(x) \leq 0 \Leftrightarrow G(x) \leq F(x)$ or, we have $\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial net_c} \leq 0$.

□

A.2.5 Case 2: when there is a positive self-selection

Proposition 3.5. *If $e'(net_c) \geq 0$, all employed migrants' average wage (weakly) increases as net_c increases when $G(x) \leq F(x)$; all employed migrants' average wage (weakly) decreases as net_c increases when $G(x) \geq F(x)$. If $e'(net_c) < 0$, all employed migrants' average wage (weakly) increases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \geq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \geq F(x)$, or when $\frac{a'(net_c)}{a(net_c)} \leq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \leq F(x)$; all employed migrants' average wage (weakly) decreases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \geq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \leq F(x)$, or when $\frac{a'(net_c)}{a(net_c)} \leq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \geq F(x)$.*

Proof. In this case, $a'(net_c) \leq 0$. Based on equation (13),

(1) if $e'(net_c) \geq 0$,

when $G(x) - F(x) \geq 0 \Leftrightarrow G(x) \geq F(x)$, we have $\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial net_c} \leq 0$;

when $G(x) - F(x) \leq 0 \Leftrightarrow G(x) \leq F(x)$ or, we have $\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial net_c} \geq 0$.

(2) If $e'(net_c) < 0$,

when $\frac{a'(net_c)}{a(net_c)} \geq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \geq F(x)$, or when $\frac{a'(net_c)}{a(net_c)} \leq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \leq F(x)$, we have $\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial net_c} \geq 0$;

when $\frac{a'(net_c)}{a(net_c)} \geq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \leq F(x)$, or when $\frac{a'(net_c)}{a(net_c)} \leq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \geq F(x)$, we have $\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial net_c} \leq 0$.

□

A.2.6 Case 3: when there is a negative self-selection

Proposition 3.6. *If $e'(net_c) \geq 0$, all employed migrants' average wage (weakly) increases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \geq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \geq F(x)$, or when $\frac{a'(net_c)}{a(net_c)} \leq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \leq F(x)$; all employed migrants' average wage (weakly) decreases as net_c increases when $\frac{a'(net_c)}{a(net_c)} \geq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \leq F(x)$, or when $\frac{a'(net_c)}{a(net_c)} \leq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \geq F(x)$. If $e'(net_c)$*

< 0 , all employed migrants' average wage (weakly) increases as net_c increases when $G(x) \geq F(x)$; all employed migrants' average wage (weakly) decreases as net_c increases when $G(x) \leq F(x)$.

Proof. In this case, $a'(net_c) \geq 0$. Based on equation (13),

(1) if $e'(net_c) \geq 0$,

when $\frac{a'(net_c)}{a(net_c)} \geq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \geq F(x)$, or when $\frac{a'(net_c)}{a(net_c)} \leq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \leq F(x)$, we have $\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial net_c} \geq 0$;

when $\frac{a'(net_c)}{a(net_c)} \geq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \leq F(x)$, or when $\frac{a'(net_c)}{a(net_c)} \leq \frac{e'(net_c)}{e(net_c)}$ and $G(x) \geq F(x)$, we have $\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial net_c} \leq 0$.

(2) If $e'(net_c) < 0$,

when $G(x) - F(x) \geq 0 \Leftrightarrow G(x) \geq F(x)$, we have $\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial net_c} \geq 0$;

when $G(x) - F(x) \leq 0 \Leftrightarrow G(x) \leq F(x)$ or, we have $\frac{\partial \% \text{ of } \frac{\text{wage} > x}{\text{employed}}}{\partial net_c} \leq 0$.

□

So, with different conditions, the proportion of employed workers whose wages are greater than x (weakly) increases or decreases with network increases for $\forall x$. Since x can be any non

negative numbers, we conclude that, with different conditions, the average wage of employed workers (weakly) increases or decreases with network increases.

Appendix B

Appendix to Chapter 4

Table B.1: Parental Migration and Rural Local Underage Children's Risk Preference Levels by Family Type

	(1) All	(2) All	(3) Girl	(4) Boy	(5) Girl	(6) Boy
Male	0.106 (0.125)	0.140 (0.138)				
Age	-0.045 (0.299)	-0.052 (0.334)	0.020 (0.478)	0.052 (0.472)	0.079 (0.478)	0.107 (0.472)
Age Square/100	0.421 (1.061)	0.578 (1.176)	0.701 (1.662)	-0.198 (1.668)	0.485 (1.666)	-0.376 (1.669)
Education Years	-0.080 (0.050)	-0.118** (0.056)	-0.192** (0.076)	-0.035 (0.081)	-0.182** (0.077)	-0.041 (0.080)
Urban <i>Hukou</i>	-0.111 (0.372)	-0.078 (0.399)	-1.581** (0.620)	1.069** (0.536)	-1.434** (0.623)	0.983* (0.533)
Local Father and Migrant Mother	-0.022 (0.376)	0.352 (0.462)	-0.043 (0.535)	2.209** (0.887)	-0.058 (0.542)	1.966** (0.881)
Migrant Father and Local Mother	0.020 (0.313)	0.113 (0.387)	1.563** (0.627)	-0.398 (0.494)	1.620** (0.635)	-0.421 (0.492)
Migrant Father and Migrant Mother	0.874* (0.516)	0.782 (0.527)	2.431** (0.940)	-0.323 (0.640)	2.534*** (0.946)	-0.337 (0.638)
Father's Education Years		-0.004 (0.027)	0.007 (0.039)	-0.011 (0.037)		
Mother's Education Years		0.042 (0.031)	-0.004 (0.045)	0.105** (0.043)		
Father's Risk Level (1=Risk Averse)		0.044 (0.038)	-0.062 (0.056)	0.151*** (0.052)		
Mother's Risk Level (1=Risk Averse)		0.166*** (0.040)	0.249*** (0.059)	0.086 (0.055)		
Father's Age		0.021 (0.023)	0.041 (0.036)	0.008 (0.030)		
Mother's Age		-0.020 (0.023)	-0.046 (0.036)	0.005 (0.031)		
Average Years of Parents' Education					-0.019 (0.045)	0.071 (0.045)
Average Risk Level of Parents					0.173** (0.073)	0.243*** (0.068)
Average Age of Parents					-0.009 (0.020)	0.012 (0.020)
Province Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	3.482 (2.142)	2.693 (2.474)	1.228 (3.634)	2.456 (3.424)	0.965 (3.625)	2.275 (3.425)
Observations	1010	820	384	436	390	440
r2_a	0.008	0.028	0.091	0.050	0.058	0.046

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.2: Parental Migration and Local Adult Children's Self-Employment by Family Type (Probit)

	(1) All	(2) Daughter	(3) Son
Male	0.001 (0.029)		
Age	0.032** (2.094)	0.013 (0.405)	0.038** (2.068)
Age Square/100	-0.032 (-1.510)	0.003 (0.059)	-0.041 (-1.593)
Education Years	-0.027*** (-5.703)	-0.021** (-2.198)	-0.024*** (-4.154)
Married	0.159*** (4.040)	0.213*** (3.065)	0.130*** (2.755)
Local Father and Migrant Mother	0.015 (0.103)	-0.039 (-0.161)	-0.027 (-0.146)
Migrant Father and Local Mother	0.124 (1.213)	0.373** (2.230)	0.013 (0.102)
Migrant Father and Migrant Mother	- -	- -	- -
Urban <i>Hukou</i>	-0.146* (-1.951)	-0.305 (-0.990)	-0.167* (-1.913)
At Least One Parent Self-Employed	0.035 (0.836)	0.108 (1.293)	0.037 (0.728)
Province Dummies	Yes	Yes	Yes
Observations	811	138	641

Report margins in the table

z-statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$