Discrimination in the Chinese Urban Labour Market

By

Dan Li

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Department of Economics

August 2018

Contents

List of Tables

List of Figures

Acknowledgements

X

Abstract

The household registration system (*hukou*) was established in the 1950s to record families' and individuals' residential information. All populations in China are divided into urban and rural *hukou*. With the transition of the Chinese economy from a planned economy to a market one, the *hukou* system has been reformed and its strict limitation relaxed. However, segregation between urban workers and rural-urban migrants in the Chinese labour market caused by the *hukou* system is an unresolved issue. Motivated by the current *hukou* situation in the Chinese labour market, this thesis empirically investigates wage discrimination, income aspirations and subjective well-being of urban workers and rural-urban migrants.

Utilizing the China Health and Nutrition Survey (CHNS) from 1993 to 2011, wage differentials between urban workers and rural-urban migrants are explored in Chapter 2. Unlike previous Chinese research, which has regarded occupational distribution as an exogenous variable, occupation segregation is considered as an endogenous variable in the analysis. The Brown et al. (1980) decomposition method is utilized to decompose wage differentials, and the estimated results are divided into four parts: within-occupation explained and unexplained portions; and inter-occupational explained and unexplained portions. The purpose of Chapter 2 is to estimate the extent to which wage differentials can be attributed to within-occupation or inter-occupation portions and the extent to which wage differentials can be explained by different observed characteristics. The empirical findings show that the unexplained part is higher than that found using the Oaxaca & Ransom (1994) decomposition approach. The existence of wage differentials between urban workers and rural-urban migrants is mainly ascribed to inter-occupational wage differentials. In addition, among the total unexplained component, a large proportion is also due to an inter-occupation wage gap. In other words, rural-urban migrants do not have equal access to occupational choice.

Chapter 3 investigates the relationship between *hukou* discrimination and income aspirations. Using data from the China General Social Survey (CGSS) 2010, this chapter compares the income aspirations of urban workers and rural-urban migrants and estimates the causal effect of *hukou* discrimination on income aspirations. Propensity

score matching is utilized to match urban workers and rural-urban migrants, based on a range of similar human characteristics. The results from the propensity score matching analysis show that, although actual income aspirations of rural-urban migrants are lower than the counterfactual income aspirations, the difference between actual and counterfactual income aspirations is not significant. The reason for the insignificant effect of *hukou* status on income aspirations may be attributed to the fact that rural-urban migrants adapt to their disadvantaged status. In addition, it may be ascribed to unobserved occupational segregation.

Motivated by the previous two chapters, Chapter 4 compares subjective well-being (SWB) between rural-urban migrants and urban workers, and also examines how the income aspirations shape their SWB. Using the same dataset as Chapter 3, rural-urban migrants report similar SWB to urban *hukou* workers. An ordered probit model is applied to explore the impact of income aspirations on subjective well-being. The results show that the income aspirations have no significant impact on SWB. The similar SWB between rural-urban migrants and urban residents cannot be explained by income aspirations, no matter which measurement of income aspirations is utilized in the analysis. In order to explain the similar happiness between rural migrants and urban residents, further analysis about the impact of socioeconomic status on SWB is conducted as well. The results show that, in Chinese context, it is social comparison rather than income aspiration that can help to explain why rural-urban migrants have similar levels of SWB.

This thesis concludes that, although *hukou* limitations were relaxed gradually, individuals with different *hukou* status continue to experience different treatment in terms of income and occupation. However, this different treatment of rural-urban migrants does not have a significant effect on restricting migrants' aspirations. In addition, although rural-urban migrants are discriminated against, they report similar subjective well-being to their urban counterparts and similar perceptions of happiness. Thus, labour market policy should mainly promote an integrated labour market, which can provide equal access to occupational choice and wages.

Key words: wage differentials; occupation segregation; income aspiration; subjective well-being; urban workers; rural-urban migrants; Chinese labour market.

Chapter One

Introduction

1.1 Research Background

1.1.1 Origin and Development of the *Hukou* System

This thesis investigates the dual structure of the Chinese labour market, which is caused by a household registration system known as the *hukou* system. How the household registration status influences the group members in the Chinese urban labour market will be explored.

Hukou (household registration) was established by The Public Security Organ Household Management Institutions, and has long been a legal instrument utilized to record and retain basic household information, including date of birth, relationships between householders, sex, marital status, education levels, and legal home address of all household members. The Chinese household registration system not only provides evidence to prove individuals' official status and identity but is also significant in their daily lives (Cheng & Selden, 1994). For example, individuals cannot go to school outside of their *hukou* areas unless they pay extra fees (Zhao, 2004). In addition, adults cannot marry, utilize medical care, or obtain employment without household registration. The present stage of China's household registration system divides *hukou* into two groups, urban and rural, based on blood relationships and geographical location. The dualistic household system, which builds a barrier between urban and rural residence, is unfair in its treatment of citizenship and involves a certain degree of discrimination.

The reason why the household registration system has long remained is attributed to China's development strategy, which involves agriculture supporting heavy industry. Chan (2010) indicated that China, taking the Stalinist growth strategy as an example, chose rapid industrialization, centred on developing heavy industrial cities; meanwhile, replenishment of an agricultural surplus from the peasantry was achieved after the

Communist Revolution of 1949. The Chinese government, therefore, tried to prevent outflows of rural workers into cities in order to guarantee a sufficient agriculture surplus to feed urban employers engaged in industrial development. In 1958, the government decreed formal laws to restrict population mobility and all migration had to be approved by the local government (Chan, 2010). From 1958 on, each Chinese citizen was classified as either "urban" or "rural" *hukou* and lost the freedom to migrate. The household classification led to two different societies in China. Urban class members could obtain jobs in industrial factories and had access to social welfare benefits, like unemployment entitlement and a retirement salary. The other class was rural peasants, whose members produced the agricultural surplus for industrialization and made a living for themselves. As a consequence, the *hukou* system not only limited labour mobility but also excluded access to state-provided welfare and welfare entitlements for the rural population.

However, at the end of the 1970s, China gradually carried out reforms of the *hukou* system and relaxed the limitations on labour migration. Cai (2011) demonstrated the motivation behind this government behaviour in relation to the household registration reforms and indicated that the Chinese economy had been bound up with two transitions that divided household registration reform into two phases: the transition from a dual economic structure to an integrated economic structure and the transition from a planned economy to a market economy. Under the planned economy, the government was responsible for implementing resource allocation and consumer products, and almost all economic behaviour relied on government instructions. On the other hand, in the market economy, production and sales of products entirely depended on the price mechanisms of the free market, not rigorous regulation from government.

In the first phase of reform, roughly between the early 1980s and the mid-1990s, regulations were relaxed gradually. At first, rural labour could migrate from the agriculture sectors to non-agriculture sectors in rural areas, and then, because of the limited job opportunities in rural and village enterprises, central and local governments acquiesced, allowing rural *hukou* workers to migrate from rural to urban areas and to seek non-agricultural jobs. In the second phase, during the mid-1990s, the policy was further reformed so that rural *hukou* individuals not only had the right to seek jobs in urban enterprises but also were allowed to do their own business in urban cities. During

this period, an increasing number of individuals started businesses, including both urban and rural *hukou* individuals. As long as there is enough financial capital, large ambitions, and reasonable/legal entrepreneurial ideas, everyone, including both urban and rural people, is free to apply for a business license from the specific business department to start their own business. Moreover, there is no different treatment for urban or rural people during the process of self-employment[1] because the government encourages individuals to be self-employed in order to provide more job opportunities in urban cities, which can alleviate the employment pressure in Chinese labour market. In the third phase of reform, approximately from the mid-1990s to 2003, China changed to a labour-intensive and export-oriented growth strategy, leading to more job opportunities for rural migrants, mostly in the coastal provinces and non-public sectors. Due to the increasing demand for labour, a number of reform policies were carried out to eliminate the barriers impeding labour mobility. Furthermore, in order to incentivize a greater number of rural workers to enter the urban labour market to meet the huge demand in urban areas, the Ministry of Public Security relaxed the conditions of entry of rural people into urban cities. As a result, a large number of rural *hukou* people have migrated to urban cities to find a job.

With the reform of the *hukou* system, the situation of migration is changing as well. At the beginning of the *hukou* reform, China was under a planned economy. During this period, although the *hukou* system was reformed and all the citizens were free to migrate to different areas, few rural *hukou* people migrated to urban cities due to limited job opportunities. Most departments and companies were public organizations where urban *hukou* individuals were more likely to be employed. In 1992, the Chinese economic strategy transformed from a planned economy to a market one, increasing the number of private companies in urban area. However, the natural growth rate of the local labour force in urban areas is too slow to satisfy the growing labour demand of

---

[1] Yang (2017) stated that the most common motivation for Chinese individuals starting a business is to achieve greater economic benefits, realize larger ambitions, and provide more job opportunities in the Chinese urban labour market. The situation of choosing to start a business passively because of discrimination in the labour market is extremely rare in China (Zhou et al., 2017) because self-employment has a higher requirement for financial assets. It is difficult for discriminated rural migrants to meet the requirements of self-employment in terms of financial assets. In addition to the similar treatments between urban and rural people during the process of self-employment, rural people do not suffer from customer discrimination because it is very difficult for customers to distinguish these two groups from their appearance; therefore, self-employed people who work for themselves do not suffer from employee discrimination, but also do not suffer from customer discrimination. Consequently, I do not take individuals who are self-employed into account as this thesis mainly focuses on the discrimination against rural-urban migrants in the urban labour market.

urban employers. Since rural migrants can fill the labour force gap, some rural male individuals began to look for jobs in urban cities. These migrants left their parents, children, and wives at home and transferred income back to support their families. Moreover, most of the rural workers were engaged in the construction industry and were around 20-30 years old.

China joined the World Trade Organization (WTO) on November 10, 2001. Thereafter, the manufacturing industry, service industry and foreign investment have grown gradually, especially in southeast coastal cities. Over time, these industries were gradually developing in central and western China, as well, thus the demand for labour increased quickly and an increasing number of rural *hukou* individuals migrated to urban cities. In addition, the level of education of migrants is also gradually improved along with the proportion of female migrant workers and married migrants increased, as well. Female migrants occupied approximately 34.5% in 2009 and this proportion increased to 40.3% in 2016[2]. An increasing number of couples worked in urban cities together and left their parents and children at home, thus more and more empty-nesters and left-behind children in rural China have emerged in recent decade. With the increasing number of female migrants, the occupation types have also expanded to the manufacturing industry and service industry, but the percentage of migrants in construction industry has decreased during recent decade.

1.1.2 Determinants of *Hukou* Status

Zhao (2004), Chan (2010), and Meiyan & Fang (2010) investigated the factors that can determine and change *hukou* status. Zhao (2004) explained that the *hukou* status of each newborn depends on their parents' *hukou* status, especially on the status of the head of household (the father, in most cases), i.e., the baby has rural *hukou* status if his/her father (head of household) holds a rural *hukou*. However, when the head of household has greater educational attainment, the probability of migrating to a city increases, which in turn increases the probability of transforming their status into urban *hukou*. The probability of the children being given urban *hukou* will be higher, too (Zhao, 2004); therefore, the family background (i.e., the father's educational attainment)

---

[2] National Bureau of Statistics of the People's Republic of China.

determines the children's *hukou* status, to a large extent.

Although the *hukou* status is determined by the family background and the location of birthplace, the status could be changed with subsequent human capital investment and experience. Firstly, current *hukou* status can be changed by marriage. For a couple, if one person has urban *hukou* and the other one has rural *hukou,* one would change his/her *hukou* status to that of the other. Specifically, if a female holding rural *hukou* married a male with urban *hukou*, she would be able to acquire urban *hukou* status. Secondly, as age and work experience are accumulated and length of residence in urban cities is longer, the probability of obtaining local urban *hukou* status increases, as several cities utilize an integrating system of obtaining local urban *hukou*. The integrating system in some cities means that rural *hukou* individuals can earn credits with their length of residence in urban cities, educational attainment, higher standard occupations, contributions to the urban city, amount of tax payment, payment for social security in the local urban city, and so forth. When the credits reach a specific level, those migrants have the opportunity of setting their *hukou* status in the urban city.

Lastly, Meiyan & Fang (2010) found that occupation can partly determine obtaining urban *hukou* status, and this correlation is indirectly caused by educational attainment. Specifically, the professional doctors, rural teachers, and civil servants of township government belong to the staff of government-affiliated institutions and are given urban *hukou* status regardless of the workplace[3]. Similar to other urban *hukou* workers, these workers with urban *hukou* status but working in rural area are paid directly by the state and are directly influenced by their position and administrative level. In addition, all staff of government-affiliated institutions are provided with economically affordable housing in local urban cities, regardless of their work place. All government units and hospitals build economically affordable housing for their staff. If the length of service is not long enough to enjoy economically affordable housing, these civil servants are provided a staff dormitory while waiting to obtain economically

---

[3] As most hospitals and schools in China belong to national system, *hukou* status of people in these public institutions will be registered in urban area regardless of their working place. Their wages are paid directly by the state. In addition, the vast majority of hospitals and schools as public organizations are located in urban area of China which is quite different with the situation in western countries. Moreover, these urban people who work in rural areas are most likely to live in urban cities. Therefore, all the previous Chinese research, in general, considered all the people who hold urban *hukou* status as urban people no matter where they worked. In this thesis, following most previous Chinese research on the *hukou* problem (e.g., urban people and rural migrants), I also regard all these urban *hukou* individuals as urban people regardless of their working place.

affordable housing. In general, there are two typical accommodations for these urban *hukou* workers in rural areas. Firstly, most of them live in urban cities and take the specific bus provided by the working organizations. Secondly, some people who are not required to work every day such as doctors and teachers, choose to live in the staff dormitory during the 2-3 working days and live in the urban cities when they are off-duty. Therefore, although their work place is located in a rural area, their regular wages, welfare, benefits, living conditions, living area, and consumption conditions are no different with other urban workers in urban cities.

With the reform of the *hukou* system, since the strict limitations were relaxed, an increasing number of rural *hukou* individuals with higher education levels can work in cities and change their *hukou* status by living in the urban city for a period of time. However, most are still discriminated against. Rural migrants are more likely to engage in low-quality jobs and have no access to social welfare. Although the *hukou* limitations were reformed, the differential treatment between urban and rural *hukou* individuals has not disappeared. The current situation of *hukou* discrimination between urban residents and rural migrants still exists in the Chinese urban labour market; therefore, Meiyan & Fang (2010) considered the *hukou* system reform to be no more than lip service.

1.2 The Framework of Thesis

According to previous description of *hukou* background, it is clear that the inequality between rural-urban migrants and urban *hukou* individuals still exists even if the *hukou* system has been reformed gradually. Therefore, the entire thesis focuses on the discrimination associated with *hukou* and its negative impacts on the workers employed in the urban labour market.

First of all, I am concerned about the wage differential between urban *hukou* individuals and rural-urban migrants in Chapter 2. The human capital theory points out that the stock of knowledge, skill, capabilities, talents, social and personality attributes contribute to one's productivity (Becker, 1962). From the view of labour economics, the economic value created by individuals is likely to be reflected in their own wages. Based on human capital theory as a theoretical basis, education, training and experience are likely to be the main factors that determine personal income and lead to wage

differences. However, it seems that some unobserved factors may also have influence on the determinant of wages since individuals who have the same human capital characteristics may be paid differently. Moreover, Oaxaca (1973) and Blinder (1973) decomposed the wage differentials and confirmed that the wage gap does not only depend on the differences in the observed human capital investment, it also depends on some unobservable factors, which may be linked to discrimination. Following the viewpoint of human capital theory and previous studies, Chapter 2 analyses the relation between discrimination and wage difference between urban people and rural-urban migrants in urban China.

In the following two chapters I focus on the negative impacts of discrimination on personal aspirations and subjective well-being. Regarding the potential relationship between aspirations and discrimination, the social cognitive career theory (SCCT[4]) suggests that racial-ethnic background may directly influence one's career development and aspirations via confidence in one's capability to perform, potential outcome expectations and determination (Lent, Brown & Hackett, 1994). Discriminated people may underestimate their ability, which leads to a lower level of aspiration. Although racial discrimination does not exist in China, *hukou* separates social classes which bring urban people more opportunities and more social welfare than rural people. Taking into account SCCT, I investigate whether rural-urban migrants' self-judgement of their capabilities is affected by *hukou* status, thereby reducing aspiration levels. Thus, following the relationship between racial discrimination and aspirations which was proposed by SCCT, Chapter 3 focuses on the relationship between *hukou* discrimination and rural-urban migrants' aspirations.

In addition, discrimination may also influence people's subjective well-being. In general, the determinants on subjective well-being involve economic factors (e.g., income) as well as some non-economic factors (e.g., age, health status, gender, education level). From an economic point of view, the most important factor affecting subjective well-being is personal income (Veenhoven, 1991; Diner et al., 1993). Moreover, some economists have proposed that an individual's income aspirations may also have an impact on their wellbeing (Stutzer, 2004; Knight & Gunatilaka, 2010a), which can be seen in Figure 1-1. In China, rural-urban migrants are treated unequally

---

[4] The detailed explanation of SCCT is shown in Chapter 3.

since they are discriminated against in urban labour market. Thus, their actual income as well as their aspiration income may be negatively affected by *hukou* discrimination, as discussed in the previous two chapters. In addition, following previous studies, actual income and aspiration income are likely to have different impacts on SWB, where actual income is likely to positively influence wellbeing while aspiration income has negative effects. I consider whether rural-urban migrants who are discriminated against in urban China are unhappier than their urban counterparts. Based on the relationship between discrimination and income and aspiration income discussed before, Chapter 4 investigates the difference in SWB between rural-urban migrants and urban people from these two economic factors, i.e., actual income and aspiration income.

Figure 1-1 The Framework of the Whole Thesis
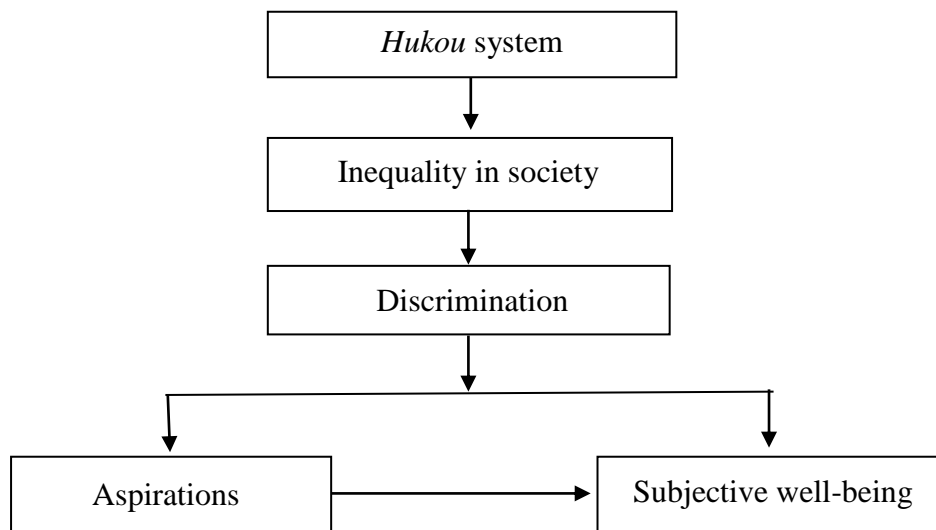
1.3 Contributions and Structure of the Thesis

As discussed in the previous section, an increasing number of individuals with rural *hukou* status have migrated to urban cities to look jobs. Considering the effect of *hukou* status on dual group members in the Chinese urban labour market, the first part of this thesis (Chapter 2) focuses on rural-urban migrants' current working conditions, including income and occupation, in comparison to urban *hukou* individuals. At the beginning, Chapter 2 compares the occupation distribution, working hours and wages by occupation of urban people and rural migrants. In addition, different personal endowment factors are compared between urban workers and rural-urban migrants. The

purpose of Chapter 2 is to investigate whether the different treatments of urban *hukou* workers and rural-urban migrants can be attributed to *hukou* discrimination. Most previous Chinese research has employed the traditional Oaxaca (1973), Blinder (1973) and Oaxaca & Ransom (1994) decomposition to decompose the wage differential between urban *hukou* and rural-urban migrants and ascribe the unexplained proportion to *hukou* discrimination. However, this traditional decomposition method ignores occupational segregation in the Chinese urban labour market and underestimates the discrimination which rural-urban migrants face in the Chinese urban labour market. In this chapter, not only is the wage gap between urban and rural-urban migrants considered, but also occupational segregation. The China Health and Nutrition Survey (CHNS) 1993-2011, a longitudinal dataset which contains detailed occupation information, is utilized. The significant contribution of Chapter 2 is to regard occupation choice as an endogenous variable. The Brown et al. (1980) decomposition is used to decompose the wage differential into two parts: within-occupation and inter-occupation. It is found that the proportion of explained part decreases to 69% and the unexplained portion increases to 31%, when taking occupational segregation into account.

Motivated by existing discrimination against rural migrants and an increasingly significant role of rural-urban migrants in the Chinese urban labour market, Chapter 3 investigates whether long term discrimination has other negative impacts on rural-urban migrants. Psychological studies in western countries have found that racial discrimination restricts the disadvantaged group's career ambition or aspiration. The limited ambitions can exacerbate the gap between the preferred group and the devalued group, which as the potential to intensify social instability. Therefore, considering the possible negative impacts on the disadvantaged group, Chapter 3 aims to investigate empirically whether discrimination against rural migrants also negatively influences rural-urban migrants' aspirations. The 2010 China General Social Survey (CGSS) which contains data on income aspirations of rural-urban migrants and urban workers is utilized. Chapter 3 contributes to the existing literature by comparing income aspirations of urban *hukou* individuals and rural-urban migrants, and examining the impact of *hukou* discrimination on rural-urban migrants' aspirations. To my knowledge, this chapter is the first study to explore whether *hukou* discrimination restricts rural-urban migrants' income aspirations. Propensity score matching is utilized to match urban *hukou* workers and rural-urban migrants, based on a range of characteristics,

which in turn enables us to observe whether urban workers and rural-urban migrants with these characteristics report different income aspirations.

Chapter 4 focuses on the subjective well-being of urban residents and rural-urban migrants, and explores what factors influence subjective well-being. Unlike previous Chinese studies which mainly focuses on the impact of income, personal endowment and social comparisons, this chapter investigates the impact on subjective well-being from the perspective of income aspirations. An ordered probit model is applied to the same dataset as in Chapter 3 (CGSS 2010). The chapter contributes to the literature by discussing how income aspirations play a role to determining the subjective well-being of urban *hukou* workers and rural-urban migrants. It is found that aspiration income has no impact on subjective well-being, no matter which measurement of income aspirations are utilized in the analysis.

Lastly, Chapter 5 concludes the findings of all parts of this study, puts forward the corresponding limitations and possible future research, based on the results of this study. In addition, I give a comprehensive view of the overall policy implication.

Chapter Two

Wage Differentials between Urban Workers and Rural-Urban Migrants in the Chinese Urban Labour Market

2.1 Introduction

As discussed in Chapter 1, the Chinese economic development strategy has changed from a planned economic system to a labour-intensive and export-oriented growth strategy, which has increased the number of private companies in coastal provinces. The transition in the country's economic development strategy and the corresponding rapid economic development has led to an increased demand for labour. At first, an increasing number of job opportunities were provided to surplus rural labour. Initially, these rural migrants' wages were much lower than the wages of urban residents and could only satisfy their basic needs. With the surplus rural labour in traditional agriculture absorbed into the industrial organizations, a financial zero point was eventually reached. In the second phase, the demand for labour exceeded the supply at the current wage and China experienced a labour shortage, making it difficult for firms to hire rural-urban migrants at a lower cost. After 2003, the labour shortage became widespread and wage rates of rural migrants increased, referred to as the Lewis turning point (Zhang et al., 2011). The government increased the minimum wage and enforced legal employment contracts for all employees, regardless of their household registration type. Although the rural migrants have more access to jobs in the urban sectors and more access to social welfare benefits than before, the differential treatment of rural-urban migrants has not disappeared everywhere in China. Urban people are still more likely to be employed in high-skilled occupations with high wages. Few rural-urban migrants have the same access to these kinds of jobs, and most are engaged in temporary and low-skilled 3-D jobs (dangerous, dirty and demeaning) with lower incomes.

The unfair distribution of income and occupation between urban residents and

rural-urban migrants may impede economic development and generate social instability (Meng & Zhang, 2001). Previous research (e.g., Démurger et al., 2009; Lee, 2012; Messinis, 2013; Liu et al., 2004; Meng & Zhang, 2001) analysed the different wages between urban residents and rural-urban migrants in China. However, most previous studies treat occupational differences as exogenous, including dummy variables for occupations, which ignore the potential discrimination affecting the difference in occupational attainment (Brown et al., 1980). There are factors that determine occupational attainment but have no direct impact on wages, which such as family structure, and people with similar human capital may have different occupations, which earn different wages. Occupation segregation between urban *hukou* individuals and rural-urban migrants is neglected in these studies. Few of these studies analyse the extent to which wage differentials between urban workers and rural-urban migrants is a consequence of intra-occupational or inter-occupational wage differentials.

Although Meng & Zhang (2001) regard the choice of occupation as an endogenous variable and investigate the extent to which the income differential between rural-urban migrants and urban residents can be attributed to inter- and intra-occupational wage differentials, their research only focuses on individuals working in the Shanghai area based on the Shanghai Floating Population Survey 1995. As discussed in the first chapter, China changed their economic development strategy during the mid-1990s and 2003 and joined the World Trade Organization (WTO) in 2001. There have been gradual increases in the manufacturing and service industries and in foreign investment, especially in southeast coastal cities. In 1992, although the Chinese economy had started to transform from a planned to a market economy, there were no diversified jobs available for rural-urban migrants in this early stage of economic restructuring. In addition, the Chinese government aimed at building Shanghai into an economic, financial, and trade center. Therefore, compared with the rest of China, the job opportunities in Shanghai are limited to commodity Trading and Processing, and the textile and construction industry. Most workers were engaged in those few industries and occupation segregation was not widespread in Shanghai at that time. From this data, Meng & Zhang (2001) found that the inter-occupational wage gap is much lower than the intra-occupational wage gap.

This study utilizes the China Health and Nutrition Survey (CHNS), which covers 9

provinces including coastal provinces (where Shanghai is located) and inland provinces such as Henan and Heilongjiang, where the focus is mainly on heavy and manufacturing industries. With the economic transition and the expansion of survey coverage, more diverse occupations appear so the labour force can be dispersed to many kinds of occupations all over the China. It is unknown whether the contribution of the within-occupational wage gap and the inter-occupational wage gap as part of the total wage differential between urban people and rural migrants has variation over time when the coverage of the investigation is not limited to Shanghai. In this chapter, I adopt the Brown et al. (1980) decomposition and utilize the China Health and Nutrition Survey (CHNS), to consider the extent to which the wage gap between urban *hukou* and rural-urban migrants can be attributed to intra-occupational and inter-occupational wage differentials. The next section describes the theory of discrimination, an empirical test for discrimination, and the relevant literature on discrimination in Western countries and China. Section 2.3 outlines the methodology and a discussion of the data is given in Section 2.4. Section 2.5 presents the results of this chapter, followed by a conclusion.

2. 2 Literature Review

2.2.1. The Theory of Discrimination

This section gives a brief introduction to the theory of discrimination, including a definition of discrimination and the identification of different types of discrimination. Discrimination implies that individuals who provide identical services and have equal productivity are treated differently. In other words, individuals with similar observed human characteristics experience unequal outcomes. Two main types of discrimination have been described in the economic literature. Firstly, Becker (1957) defined taste-based discrimination as differential treatment arising from individual prejudices or animosity to members from disliked target groups. This form of discrimination, which relies on individual attitudes, implies that a certain number of individuals would spend more opportunity cost, rather than have any interaction with members from their disliked groups. In Becker's words, if one person is a discriminator, he will prefer to hire members of a favourite group to substitute the members of other disliked groups, even at a higher cost.

In the *Economics of Discrimination*, Becker (1957) delineated three drivers: employer, employee, and consumer taste-based discrimination. Just as their names imply, employer discrimination means that discriminatory employers would rather give up a fraction of revenue in order to hire their favourite applicants. Meanwhile employee discrimination (co-worker discrimination) exists where several members of the majority group have animus towards another other minority group and are not willing to collaborate with its members. Employee discrimination may have an indirect effect on employers, leading them to compensate for the discriminatory employees or reduce the wages of minorities. In addition, Becker (1957) also discussed the consequence of consumer discrimination, when customers prefer to be served by certain specific groups. For example, discriminatory consumers may prefer to be served by white males as doctors and drivers, and white females as nurses, flight attendants or tour guides. Occupational segregation is more likely to be suffered in the social service occupations and involves the separation of people into different occupations on the basis of their characteristics (Lang, 2011). Segregation and discrimination might go together but they are not identical, unless one group's circumstances, such as wages, welfare and the working environment, are obviously worse than that of the other group.

In the early 1970s, Arrow (1973) criticized the Becker's taste-based discrimination model, on the grounds that those discriminatory employers who would rather spend more to employ their favourite workers would be expelled from the market under perfect competition, in the long run. Nevertheless, Arrow (1973) and many other authors have also demonstrated that the relationship between wages and racial prejudice based on Becker's conclusion might survive in an imperfectly competitive market in the long run, because in reality, it is difficult for firms to fill vacancies instantly, and workers also need enough time to find new jobs. Firms cannot know which applicants are seeking jobs and workers are not aware of which companies have vacancies (Lang, 2011). Therefore, although Arrow criticized Becker's model to some extent, it does not mean that Becker's conclusion should be replaced by that of Arrow (1973).

In light of the criticisms of Becker's taste-based discrimination model, Arrow (1973) put forward a discrimination model based on limited information which is referred to as statistical discrimination. Statistical discrimination assumes that firms attempt to maximize profits and are not prejudiced against minorities, and that

individuals are treated according to the group's average situation, rather than the individuals' personal characteristics. In the statistical discrimination model, employers cannot observe all the characteristics of the applicants: for example, it is difficult for employers to assess the productivity of applicants by interviewing and reading their resumes. However, employers can observe the applicant's group membership in terms of e.g. gender and race, and can be aware of the applicants' educational levels, based on their resumes. Because of limited information, Arrow (1973) explained that employers' optimal prediction of productivity was the combination of the average of the individual characteristics and employees' average productivity in the same group. Therefore, in the absence of information, employers will treat individuals as if they have the average features of one group. For example, in general, black people do have more disadvantaged backgrounds, such as lower educational levels or less trust from consumers, so employers may affirm that all applicants from the black membership group will perform worse than those from the white group. Employers will judge one black applicant based on the average productivity of this group, even if this job hunter is better than most white applicants. As a consequence, black people are more likely to earn lower wages and experience a lower probability of being employed in jobs requiring higher professional skills.

2.2.2. Discrimination in Western Countries

At the beginning of research on discrimination, it was apparent that an increasing number of research studies have aimed to investigate racial discrimination between whites and black people in America. Racial discrimination and household registration discrimination may be similar in the sense that two different group members with identical characteristics are treated differently in some respects, such as wages, occupations and employment opportunities. However, western racism is more a question of perception, whereas China's household registration discrimination is a kind of institutional discrimination. There are no decrees in Western countries to confine blacks to working in specific districts or professions, while China's government strictly restricted labour mobility before the 1980s. In spite of the relaxation of the restriction on individuals seeking jobs everywhere, many employers may still regard rural-urban migrants as of lower-status because of their long-term segregation from urban workers and because they may consider rural workers to have lower educational levels, inferior

working skills and lower productivity. Therefore, most employers in China, like those in Western countries, are likely to prefer one group when recruiting. Therefore, it is useful to review research about racial discrimination for reference when analysing household registration discrimination in China.

2.2.2.1. Determinants of Different Treatment in Western Countries

Ritter & Taylor (2011) investigated racial discrimination from two aspects, wage gaps and different rates of unemployment, as they indicated that disparities in both wages and unemployment were the critical features in the U.S. labour market. Focusing on wages alone was insufficient to reflect the achievement differential of the labour market, and employment rates played a significant role in affecting racial disparity as well. They found that the wage differential is more likely to be ascribed to personal characteristics, such as social demographic factors, educational attainment, and cognitive skills but not to racial discrimination. However, the results of unemployment rates demonstrated that a sizeable part of the unexplained unemployment differential rates could be ascribed to the statistical discrimination model. Therefore, they concluded that discrimination plays an important role in determining an individuals' employment or unemployment situation. Thus, I intend to review previous studies from two aspects: personal characteristics and occupation. In terms of human characteristics, I mainly focus on the impact of educational attainment on wage differentials.

A. Influence of Education

The basic approach to testing whether differential treatments are attributable to discrimination is to examine how much of the disparity cannot be explained by human characteristics and then to ascribe the unexplained component to discrimination. It is significant for the success of the test to choose suitable characteristics as the human capital variables in the model. Among a large number of previous studies, demographic factors such as age, gender, education, health and marital status are the ones most commonly examined for their influence on differential treatments. In addition, individuals with different educational levels tend to be employed in different professions, and their wages vary between professions. Therefore, education level is highly correlated to occupations and wages. In this chapter, I intend to include highest

education levels as the independent variable to test the extent to which wages differentials should be attributed to different education levels or discrimination.

Lang & Manove (2011) studied the different returns to education in the US and utilized the statistical discrimination model to explain the phenomenon that blacks were required to have more years of education or higher education levels than whites with similar skills to have wage parity with whites. As explained, the statistical discrimination assumption implies that the model is based on the employers' limited information about the applicants. Employers infer applicants' productivity from observed characteristics, such as education levels or time spent in education in their resumes. The literature about statistical discrimination suggests that companies and employers observe the productivity of white people more accurately than for black people, based on relative educational achievement (Lang & Manove, 2011). Therefore, workers have an incentive to use educational levels as a signal of productivity; and furthermore, education is seen to be more important in relation to black people, as a signal of productivity, than for white people, leading to black people having to obtain more education relative to white people with the same ability. However, they found that black people with higher educational levels earned similar wages to their white counterparts. In other words, with the same ability, the returns to education for black workers were less than for white workers.

O'Neill (1990) researched the impact of human capital on earnings differences between black and white people, based on data from the National Longitudinal Survey of Youth. The author found that, since 1940, as education levels and the quality of black people's performance in school improved, the racial difference of human capital tended to narrow, so that black people now earn much more than previously. Thus, the wage differential between black and white people has narrowed, gradually. The author explained the increase in the relative wages of black workers from two significant factors as resulting from a rise in the cognitive skills of blacks and a decline in labour market discrimination. Firstly, the estimated results indicated that the narrowing of educational distinction between black and white people made a great contribution to convergence in the earnings gap, which meant that educational achievement and working skills account for most of the earning gaps between black and white people. However, O'Neill also stated that human capital factors were not the only cause of

narrowing wage differentials. The reduction in discrimination may also have been influenced by the relative rise in black earnings, as better job opportunities and training became available to them, providing cumulative career benefits. The author therefore concludes that the rise of the Black-White wage ratio is related to interaction between human capital and the reduction in discrimination.

Card & Krueger (1991) studied the relationship between school quality and the returns to education for black and white workers and found similar evidence to O'Neill (1990) that the narrowing of schooling quality between black and white people had contributed to the convergence of black-white relative wages. Nevertheless, their estimates indicated that school quality can account for only 15% to 25% of the narrowing of the black-white wage ratio. The rest of the convergence could be attributed to the higher requirements of black people's educational levels by employers. In addition, government policy and judicial pressure incentivizes the demand for skilled black employees, which can be explained by the reduction of discrimination, similar to O'Neill's second viewpoint.

In addition, Maxwell (1994) also researched the effects of the quality and quantity of education on the black-white wage differential using the National Longitudinal Survey of Youth. He tested the extent to which the black/white wage difference might be explained by the quality and quantity of educational attainment. The analysis demonstrated that the critical cause of the wage differential was the quality of the education levels of the two groups, rather than the quantity of schooling (Maxwell, 1994). A number of firms preferred to gauge unobserved variables, like working skills and workers' productivity as markers of their workers' educational quality. Therefore, the conclusion suggested that attempts to close the black-white wage gap should concentrate on the closing of the educational quality differential by offering higher quality schooling to blacks. However, as previous research such as Card & Krueger (1991) has shown, the narrowing of the racial gap in education quality and basic skills learned from school only contributes to the convergence of wages by two-thirds rather than the whole differential.

B. Impacts of Occupation

Apart from the impact of human characteristics on wage differential, few studies (Brown et al., 1980; Gabriel & Schmitz, 1989; De Beyer & Knight, 1989; Fearon & Wald, 2011) consider that the wage differential between whites and blacks may be attributed to their inequality opportunity of occupation. Blacks may suffer from racial discrimination when applying for jobs, therefore their wages are restricted by their lower quality jobs. In this section, those studies which regard occupation segregation as an endogenous factor influencing wage differentials will be reviewed.

Both Gabriel & Schmitz (1989) and De Beyer & Knight (1989) investigated how occupational segregation influences wage differentials using Blinder-Oaxaca (1973) decomposition[5]. Gabriel & Schmitz (1989) found that blacks were less likely to be hired in higher-paid occupations. In light of their estimated results, if blacks had experienced identical occupational opportunities, the probability of blacks being employed as managers, and in professional jobs with higher incomes, would have increased, and blacks would have been less likely to work in sales and service jobs. Therefore, the researchers argued that parts of the black-white wage differential could be attributed to this occupational disparity. However, the results from Gabriel & Schmitz (1989) showed that while earnings-related characteristics such as type of occupation, had an impact on wage differentials, they could just account for about 33 percent of the black-white wage differential and the rest should be attributed to unexplained factors, such as differential treatment of blacks in the labour market, i.e. discrimination.

De Beyer & Knight (1989) emphasized the importance of both human capital and job-related characteristics on the determination of wages, and both factors were likely to be interconnected. For example, the type of occupation in which workers are engaged in is related to their educational levels, sex, race and working experience. In addition, De Beyer & Knight (1989) also considered the different productivity among different occupations, concluding that the returns to productivity in non-manual occupations were higher than in manual occupations, and the returns were lowest for unskilled jobs. Thus, like Gabriel & Schmitz (1989), they deemed that occupation had no independent impact on wages, and only played a proxy role for unmeasured personal characteristics, such as unobserved ability or productivity (De Beyer & Knight, 1989). The occupation had a

[5] This is a traditional decomposition method for the wage gap. The Blinder-Oaxaca (1973) decomposition method is described in detail in the methodology section of this chapter.

positive effect on wages, but the role of characteristics which determine occupational access suggests the existence of discrimination. Therefore, the wage differential was partly attributed to occupation segregation as much as the presence of discrimination. In addition, Fearon & Wald (2011) utilized Brown et al.'s decomposition method, which takes occupational segregation into account, to research the wage differential between blacks and whites in Canada. The estimated results suggested that wage discrimination and occupational segregation account for a large proportion of the wage gap, while human characteristics account for a small part (Fearon & Wald, 2011).

2.2.3 Discrimination in China

According to the previous research on racial discrimination in Western countries, an increasing number of researchers have begun to use the former principles and methods to analyse household registration discrimination as well. Lee (2012) researched wage differentials between rural migrants and urban *hukou* employees using the 2005 China Urban Labour Survey (CULS). Unlike most previous research on wage differentials, Lee (2012) took other welfare aspects, such as the annual bonus and insurance cover, into account when carrying out the decomposition of the wages using the Blinder-Oaxaca (1973) decomposition. The decomposition approach divided the wage differential between rural migrants and urban *hukou* workers into an explained and unexplained component. The explained component of the wage differentials was ascribed to human characteristics such as age, gender, job tenure, marital status, health, years of education and highest education levels. The unexplained part of the wage gap was attributed to discrimination. The results demonstrated that a majority of rural migrant workers did not receive full insurance entitlements; and, moreover, migrant workers earned less than their counterpart urban *hukou* employees, holding their similar characteristics constant. A considerable income gap between rural migrants and urban *hukou* workers was unexplained, and the reason for the unexplained component may be that employers are statistically discriminatory towards the minorities, as they were not able to observe the applicant's capacity.

Messinis (2013) proposes a different conclusion to Lee (2012) on the issue of the wage gap between urban *hukou* individuals and rural-urban migrants, where China Household Income Project (CHIP 2002) data is used to research the productivity of rural

migrant and urban *hukou* individuals in the Chinese labour market. Messinis (2013) utilizes matching and the Instrumental Variable method to assess different conditional and unconditional quantile treatment impacts in terms of different levels of education and different migrant status. By dividing the education into different levels and defining the duration of being non-local migrants, Messinis (2013) compares the returns of education between local urban workers and long-term migrants within the same education levels. The dataset (CHIP) covers local urban households and rural migrant households samples. Messinis (2013) regards these rural migrant households in CHIP as long-term migrants living with their family, which is more directly comparable to urban households. The estimated results show that urban *hukou* people are advantaged and better treated, in comparison to rural-urban migrants. The different treatment can be completely explained by different personal factors and the lower human capital investment of rural migrants (Messinis, 2013).

In terms of the occupational segregation, Démurger et al. (2009) analysed the different treatments of rural migrants and urban *hukou* workers, from segregation to different sectors and different incomes within in the same sector. Using the China Household Income Project (CHIP), they utilized Blinder-Oaxaca (1973) decomposition to divide the difference into four parts: (1) difference in sector allocation; (2) different working time (working time effect); (3) different wages within sectors; (4) different population structure[6]. In addition, the authors separated the sectors into self-employment, private and public sectors, in order to evaluate the impact of difference in earnings and opportunities for urban people and rural *hukou* applicants hired into the different sectors. Their findings showed that rural *hukou* individuals were more likely to be self-employed or to be hired in the private sector; in contrast, urban *hukou* applicants were more likely to work in the public sector, receiving higher wages and better benefits. What is more, annual wages of urban workers were 1.3 times higher than those of rural migrant ones. However, their analysis suggests that the differential treatments toward urban and rural workers did not contribute to segregation in terms of access to jobs but the different population characteristics, which indicated that rural migrants were much less experienced and educated than urban *hukou* applicants.

---

[6] This aspect refers to the effect of the distribution of human characteristics in the populations.

Regarding the income inequality and occupation segregation, Meng & Zhang (2001) were the first to utilize the Brown et al. (1980)[7] decomposition and to regard the occupation choice as an endogenous variable. They regarded the migrants as a disadvantaged group with lower wages than the majority, less benefits and insurance. They decomposed the wage gaps into four parts: intra-occupation explained; intra-occupation unexplained; inter-occupation explained and inter-occupation unexplained portions. In addition, Meng & Zhang (2001) also utilized the Blinder-Oaxaca (1973) decomposition and compared the results with Brown et al. (1980) decomposition results. They found that a large proportion of wage differential was attributed to the intra-occupational unexplained part, which are more likely to be due to discrimination. In addition, the total unexplained portion in Brown et al. (1980) decomposition was higher than that in the traditional Blinder-Oaxaca (1973) decomposition method. Therefore, these results suggest that the Blinder-Oaxaca (1973) approach underestimates the discrimination against disadvantaged group because occupation segregation is ignored. However, Meng & Zhang (2001) only aimed at exploring the income inequality and occupational segregation between urban people and rural migrants in Shanghai area. They did not consider whether this situation existed in the whole of China. Therefore, this chapter will fill the gap in previous Chinese literature and explore the wage gap between urban workers and rural-urban migrants in China as a whole and investigate the extent to which wage gaps between urban workers and rural-urban migrants can be ascribed to inter- or intra-occupational wage differentials.

2.3 Methodology

2.3.1 Human Capital Theory and Earning Function

Human capital theory corresponds to the stock of knowledge, skills, social and personality attributes that it represents the ability to perform labour thus contribute to productivity (Becker, 1962). In other word, human capital is a collection of personalities which includes knowledge, competencies, skills talents, experience, training that are possessed by individuals. It can be defined as a measurement of economic value of

---

[7] An extension of the traditional Blinder-Oaxaca (1973) decomposition, will be explained in detail in this section.

one's capabilities. From the labour economics point of view, human capital refers to a set of people's characteristics which contributes a worker' productivity. In the current labour market, the economic value of the production created by a person is often reflected by their wage. Human capital has increasingly been seen as a unidimensional object such as knowledge or skill that it directly contributes to the production or wage function (Becker, 1962; Becker, 1993). Therefore, wage difference amongst individuals are often reflected by the diverse elements of human capital such as knowledge, skills, ability, institutions, etc.

In terms of the difference in human capital investment, it is useful to explore the diversity of human capital sources including innate and acquired ability. In the biology and social biology point of view, innate ability refers to the genetic IQ. For example, even though two individuals are in the same situation (e.g. education background, economic conditions, social environment) and they have the same access to the investment opportunities, they still can create different human capital and act differently in their work. According to Blundell (1999), wages difference is caused by the difference in human capital investment mainly focus on schooling and on the job training. The purposes of schooling and on the job training both aim to improve individual's knowledge, skills and abilities that to create further productivity or economic value.

Schooling has been the most commonly utilized component of human capital investment model. Schooling is often regarded as an institution that it dedicates to specializing in the training for people (Becker, 1962). Although one cannot make money directly from attending school, the training and taught knowledge received at school has a close relationship to their future working wage or income. The amount of investment in schooling one receives can have great impact on their wage after graduation. The one who gets more training in school can often have better income and wage. Besides the time spent on school education, the quality of schooling is deemed as another important component that can determine the human capital accumulation. For example, previous research found when two different students holding urban and rural *hukou* status separately, even the two students have similar innate intelligence and receives the same years of education, they still develop different human capital due to the different schooling quality in urban and rural cities (Afridi et al., 2015). This can be explained by

different quality of schooling education. However, even when two people receive the same quality of education, they still develop different human capital that may be caused by their innate capability and different effort they put into their study.

Training, especially on the job training, is other important component of human capital investment that is mostly received after leaving school during their work. Unlike the campus training (schooling) that specializes in teaching, on-the-job training specially focuses on the job relevant training that aims to enhance one's productivity and create more economic value in the future work (Becker, 1962; Becker, 1993; Konings & Vanormelingen, 2015). However, it is difficult for workers to make investment on the occupational training for themselves. Most on-the-job training is provided by firms who require their employees to have specific skills and would like to spend a large proportion of benefit on the training investment. The expenditure and the length of training rely on the specific skills required in each firm. Therefore, on-the-job training is not a purely individual investment in human capital rather than a joint investment by both companies and workers.

There are two types of on-the-job training which are general training and specific training (Becker, 1962). General training refers to the process of skills training that can be applicable to many other firms. For example, the skills learned by an interned doctor in one hospital can also be applied to other hospitals. Thus, the general training a firm provided not only increases their marginal productivity, it also has the same productivity function in other firms. Unlike the general training, there are some kinds of training that only increases the productivity of workers in the firms providing the training. The specific training that brings more productivity to the firm which provide training is called specific training. Therefore, the participation of on-the-job training, investment on training, different types and length of training of each worker received will have a great contribution to the difference in human capital.

Regarding to the utility of training for human capital, Mincer (1958) and Mincer (1974) argued that the working experience accumulated through learning from previous colleagues and summing up their projects at work is also an essential part of training. Although working experience differs from the formal training either general training or specific training measured by the amount of investment and length of time spent, it is

also crucial to influence one's human capital accumulation. For example, besides the professional training a doctor received for the job, the number of surgeries he/she participated also underpin the skills they have generated. Therefore, previous research proved that one's working performance is determined by both on-the-job training and experience accumulated from their work (Mincer, 1974). Individuals become more productive as they get more experienced and their wages will also increase with the accumulation of experience.

However, some labour economists have criticized the human capital theory that tries to explain all the difference in wages by the difference in human capital. Wages can be higher due to some factors such as gender and occupational status instead of some traditional human capital elements. Regarding to the occupational status, Mincer (1958) argued the positive association between the level of occupation status and earnings. Mincer (1958) proposed that occupational status is a meaningful indicator to quantify the "formal training". Occupation difference determines difference in the total amount of training. The professional or specialized skills that is required by a specific job can directly affect the amount of investment and time spending on the formal training. Occupational status can be divided into different levels, from the lower-skilled occupations that required basic training (e.g., waiter) to those highly professional occupations which require a large amount of specific training (e.g., doctor). Thus, it is more likely that occupations with higher wages require more specific formal training, i.e., either general or specific training. In addition, the concept of human capital also includes unobserved heterogeneity such as individual's social connection, personal characteristics and discrimination against the minority groups, which can also lead to different wages amongst workers. Therefore, when exploring the wage difference in labour market, it is very important to make a right balance between assigning wage differences to unobserved heterogeneity and difference in human capital investment. Based on the general description of human capital theory above, a relevant human capital model that is relevant to earnings function and investment in human capital can be established.

$$\ln wage = \beta_0 cons \tan t + \underbrace{\beta_1 schooling + \beta_2 training + \beta_3 \exp erience}_{traditional \ human \ capital \ investment} + \underset{unobserved \ \text{heterogeneity}}{\mu} \qquad (2\text{-}1)$$

where μ refers to a set of unmeasurable indicators other than human capital, such as discrimination against minority group.

2.3.2 Decomposition

Based on human capital model, Oaxaca (1973) and Blinder (1973) proposed a decomposition method to disentangle the observed from the unobserved factors affecting wage differential. The starting point of the decomposition method is to specify two linear Mincer wage functions based on human capital model for different groups A and B:

$$Y_A = X'_A \beta_A + \varepsilon_A, \qquad E(\varepsilon_A) = 0 \qquad (2\text{-}2)$$

$$Y_B = X'_B \beta_B + \varepsilon_B, \qquad E(\varepsilon_B) = 0 \qquad (2\text{-}3)$$

$Y_A(Y_B)$ is the outcome variable (earnings), $X_A(X_B)$, which is a set of human capital indicators, such as education, work experience and training; $\varepsilon_A (\varepsilon_B)$ is a stochastic error term. The key issue with the Blinder-Oaxaca (1973) decomposition is how much of the mean outcome variable difference,

$$R = E(Y_A) - E(Y_B) \qquad (2\text{-}4)$$

is accounted for by group differences in the predictors (X_t). $E(Y)$ denotes the expected value of the outcome variable.

In this chapter, I intend to estimate the wage differential between urban *hukou* individuals (A) and rural-urban migrants (B); thus, for the given two groups "urban" and "rural" *hukou*, I specify two wage equations for urban *hukou* workers and rural-urban migrants, respectively.

$$Y_{it,u} = X'_{it,u} \beta_u + \varepsilon_{it,u}, \quad \text{urban } hukou \text{ workers} \qquad (2\text{-}5)$$

$$Y_{it,r} = X'_{it,r} \beta_r + \varepsilon_{it,r}, \quad \text{rural-urban migrants} \qquad (2\text{-}6)$$

where Y_{it} refers to the log of monthly wages of the *i-th* employee in year *t*, and the subscripts *u* and *r* indicate that the sample is restricted to urban *hukou* people and rural-urban migrants, respectively. *X* is a vector of explanatory variables, including basic socio-demographic factors such as age, marital status, gender, education level, occupation type, working unit, and a constant. *β* is the vector of coefficients to be

estimated, containing slope parameters and the intercept. The average difference in outcome (Y) can be expressed as the difference in the linear model between urban workers' wage equation (2-5) and rural-urban migrants' wage model (2-6), giving

$$
\begin{aligned}
R &= E(Y_u) - E(Y_r) \\
&= E(X_u)'\beta_u - E(X_r)'\beta_r
\end{aligned}
\tag{2-7}
$$

since $E\ (\varepsilon_u) = 0$ and $E\ (\varepsilon_r) = 0$ by assumption. The equation of raw wage differential can be transformed as follows:

$$
R = \{E(X_u) - E(X_r)\}'\beta^* + \{E(X_u)'(\beta_u - \beta^*) + E(X_r)'(\beta^* - \beta_r)\}
\tag{2-8}
$$

We now have a twofold decomposition, where β^* is the non-discriminatory coefficient vector. The first component $\{E(X_u) - E(X_r)\}'\beta^*$ represents the portion of the wage gap that is explained by group differences in the characteristics, and the second component, $E(X_u)'(\beta_u - \beta^*) + E(X_r)'(\beta^* - \beta_r)$ is the unexplained portion, which is more likely to be ascribed to discrimination. When dealing with this equation, the reference group should be specified. In general, most researchers assume that the discrimination is only directed against one group. Following the classical Blinder-Oaxaca (1973) decomposition, since discrimination is assumed to be directed only against rural migrants, the coefficient of urban people can be regarded as the non-discriminatory vector; or the urban people are chosen as a reference vector that is $\beta^*=\beta_u$. The wage differential can be expressed as

$$
R = \overline{Y}_u - \overline{Y}_r = \underbrace{(\overline{X'_u} - \overline{X'_r})\hat{\beta}_u}_{\mathrm{exp}lained} + \underbrace{\overline{X'_r}(\hat{\beta}_u - \hat{\beta}_r)}_{un\mathrm{exp}lained}
\tag{2-9}
$$

However, in terms of calculating the reference vector or non-discriminatory vector, Neumark (1988) suggested using the coefficients from a pooled estimation over both groups, which is utilized more commonly in the literature. As pointed out by Oaxaca & Ransom (1994), the wage structure (2-8) can be transformed in the following equation:

$$
\begin{aligned}
R &= (\overline{X}_u - \overline{X}_r)'(W\beta_u + (I - W)\beta_r) \\
&\quad + \{(I - W)'\overline{X}_u + W\overline{X}_r\}'(\beta_u - \beta_r)
\end{aligned}
\tag{2-10}
$$

where W indicates a matrix of observed weights containing the urban and rural migrant samples, and I is the identity matrix. Choosing $W=I$ in Eq. (2-10) is equivalent to setting $\beta^*=\beta_u$ in the classical Blinder-Oaxaca (1973) decomposition. The first part on the right-hand side $(\overline{X}_u - \overline{X}_r)'(W\beta_u + (I - W)\beta_r)$ is the explained portion, which can be

explained by human capital or individuals' characteristics, and the second part $\{(I - W)'\overline{X}_u + W'\overline{X}_r\}'(\beta_u - \beta_r)$ is the unexplained portion that may be due to discrimination. Oaxaca & Ransom (1994) also utilized Ω to represent the fitted value of the weighted matrix: $\hat{W} = \Omega = (X_u'X_u + X_r'X_r)^{-1}X_u'X_u$.

However, even though we can know whether the wage differential has been affected by the unobserved indicator (e.g., discrimination) in human capital model through Oaxaca & Ransom (1994) decomposition, this decomposition approach is likely to underestimate the actual discrimination against rural-urban migrants. In this model, if wage differential were attributed to different occupation types, it would be difficult to decide whether the wage gap results from observed human capital indicators leading to difference in occupational status or discrimination constraints on occupational entry, since applicants might be confronted with discrimination while applying for jobs. In general, it is well-known that senior professional and technical workers (doctors, professors, lawyers and engineers) earn higher wages than service workers, such as waiters, doorkeepers and counter salespeople. But rural-urban migrants might be faced with discrimination while seeking jobs and they may have limited access to these senior professional or technical occupations. As Brown et al. (1980) indicated, persons with similar characteristics may attain different occupations and earn different wages. In the context of China, rural-urban migrants are more likely to be engaged in less skilled jobs with lower wages and worse working conditions than urban workers, even if they possess similar characteristics. *Hukou* discrimination against rural-urban migrants not only influences the wage gap but also generate constraints on rural migrant's occupational choice or entry, to some extent. Therefore, if the dummy variables of occupational status are included in the human capital model as justified and exogenous indicators, the types of occupational status may be affected by the potential discrimination in unobserved disturbance. This approach ignores the impact of potential discrimination among the differences in occupational distribution on wage differential. The discrimination faced by rural-urban migrants in the urban labour market will be underestimated. Therefore, I follow Brown et al. (1980) in regarding occupational status as an endogenous variable and take the potential discrimination in occupational attainment into account. A separate model of occupational attainment is proposed in the estimation of the income gap to estimate the wage differential between urban *hukou*

workers and rural-urban migrants.

Brown et al. (1980) proposed an approach to combine occupation choice and the wage model, and derive a more accurate decomposition to estimate the gender wage differential. This decomposition is an extension of the conventional Oaxaca decomposition, which considers occupational discrimination. Brown et al. (1980) incorporate the difference in the probability of being engaged in one occupation between urban workers and rural-urban migrants. Based on Brown et al. (1980), the wage gaps in this chapter between urban workers and rural-urban migrants can be estimated as the sum of intercept, within-occupational and inter-occupational portions. Within-occupation wage gaps refer to unequal pay for equal work. It means that urban workers and rural-urban migrants earn different wages, although they have the same occupation and are engaged in the same work. The inter-occupation wage gap captures the difference in occupations between urban workers and rural-urban migrants, which considers that urban people and rural-urban migrants experience different opportunities, holding their characteristics constant. In other words, urban workers and rural-urban migrants with similar characteristics, such as the age, education level, gender and marital status, might be employed in different occupations. Rural-urban migrants might be engaged in a worse job than their urban counterpart. Hence, wages for the workers are likely to be different. Assuming there are j occupations, Brown et al. (1980) decomposed the wage differential into three portions,

$$R = \overline{Y_u} - \overline{Y_r} = \sum_j (p_{uj}\overline{w_{uj}} - p_{rj}\overline{w_{rj}})$$

$$= \sum_j (p_{uj}a_{uj} - p_{rj}a_{rj}) + \sum_j (p_{uj}\overline{X_{uj}}\hat{\beta}_{uj} - p_{rj}\overline{X_{rj}}\hat{\beta}_{rj})$$

$$= \sum_j p_{rj}(a_{uj} - a_{rj}) + \left\{\sum_j p_{rj}(\hat{\beta}_{uj}\overline{X_{uj}} - \hat{\beta}_{rj}\overline{X_{rj}})\right\} + \left\{\sum_j (\hat{\beta}_{uj}\overline{X_{uj}} + a_{uj})(p_{uj} - p_{rj})\right\}$$

$$= \underbrace{\sum_j p_{rj}(a_{uj} - a_{rj})}_{I} + \underbrace{\left\{\sum_j p_{rj}(\hat{\beta}_{uj}\overline{X_{uj}} - \hat{\beta}_{rj}\overline{X_{rj}})\right\}}_{within-occupation} + \underbrace{\left\{\sum_j \overline{w_{uj}}(p_{uj} - p_{rj})\right\}}_{inter-occupation}$$

(2-11)

where p_{uj} (p_{rj}) is the proportion of urban workers and rural-urban migrants employed in the j-th occupation, and $\overline{X}_t$ is a $K \times N_t$ vector of mean personal characteristics, $t=u, r$ which refer to urban individuals and rural-urban migrants. a_{tj} and β_{tj} represent the intercept and the coefficients of the personal characteristics, respectively. I is the

unexplained portion of the intercepts. The within-occupation proportion is similar to the Oaxaca & Ransom (1994) decomposition, which focuses on unequal pay for equal work. This proportion can be decomposed as the Oaxaca & Ransom (1994) decomposition into an explained part (WE), which can be attributed to human characteristics, and an unexplained part, due to the difference in coefficients of the wage equations between urban workers and rural-urban migrants. In general, this portion (WU) can be attributed to discrimination.

$$\underbrace{\sum_j p_{rj}(\hat{\beta}_{uj}\overline{X_{uj}} - \hat{\beta}_{rj}\overline{X_{rj}})}_{within-occupation} = \underbrace{\sum_j p_{rj}\hat{\beta}_{uj}(\overline{X_{uj}} - \overline{X_{rj}})}_{WE} + \underbrace{\sum_j p_{rj}\overline{X_{rj}}(\hat{\beta}_{uj} - \hat{\beta}_{rj})}_{WU} \qquad (2\text{-}12)$$

In addition, the inter-occupation differential can also be further decomposed into an explained part due to different observed characteristic (IE). The qualifications for one occupation might have a close relationship with different education levels, experience, gender or marital status. In addition, the second portion of the inter-occupation differential is an unexplained part, due to differences in the structure of occupational attainment between urban workers and rural-urban migrants (IU) not based on observed characteristics, which mean that urban workers and rural-urban migrants with similar characteristics experience different treatment while looking for jobs.

$$\underbrace{\sum_j \overline{w_{uj}}(p_{uj} - p_{rj})}_{inter-occupation} = \underbrace{\sum_j \overline{w_{uj}}(p_{uj} - \hat{p}_{rj})}_{IE} + \underbrace{\sum_j \overline{w_{uj}}(\hat{p}_{rj} - p_{rj})}_{IU} \qquad (2\text{-}13)$$

The new term ($\hat{p}_{rj}$) in this model is the counterfactual proportion of rural-urban migrants in occupation j, if rural-urban migrants faced the same occupational opportunities as urban workers. The estimation of this new term ($\hat{p}_{rj}$) requires a model to predict rural occupational choices, assuming that rural-urban migrants encounter the same opportunities as urban workers. Brown et al. (1980) used the multinomial logit model to estimate the probability of an individual's employment in a particular occupation ($\hat{p}_{rj}$).

$$p_{ij} = p(y_i = occu_j) = f(x_i) = e^{x_i'\beta_i} / \sum_{j=1}^{J} e^{x_i'\beta_j} \qquad (2\text{-}14)$$

$i=1, 2,..., N$

$j=1, 2,..., J.$

where the dependent variable y is a categorical variable with J outcomes: $occu_1$, $occu_2$,..., $occu_J$. The categories are numbered 1 through J, and $p_{ij} = p(y_i = occu_j)$ is the probability of individual I working in occupation j and x_i refers to variables influencing the choice of occupation. N is the number of observations and β_j represents the vector of coefficients corresponding to the j-th occupation. Using the multinomial logit model, I can estimate the unobserved $\hat{p}_{rj}$ by predicting the occupational choice of urban *hukou* workers using the rural migrants' human capital indicators. This is a measure of the rural occupation choice in the absence of discrimination. The parameters $(\beta_1, \beta_2, ..., \beta_J)$ in the multinomial logit model can be estimated by maximum likelihood, and the log-likelihood function is given by

$$LL = \sum_{i=1}^{N} \sum_{j=1}^{J} y_{ij} \ln(p_{ij}) \tag{2-15}$$

where $y_{ij}=1$, if alternative occupation j is chosen by individual i, and 0 if not, for the J possible outcomes. Therefore, for each individual only one of the y_{ij}'s is 1. The purpose of the estimation of these parameters is to analyse the difference in the structure of occupational attainment decomposed from the inter-occupational wage differentials.

2.4 Data

The China Health and Nutrition Survey (CHNS) used in this chapter is conducted by an international team of researchers and collects information from 12 aspects: population density, traditional markets, modern markets, transportation infrastructure, sanitation, communications, housing, education, diversity, health, and social services. The purpose of the CHNS is to investigate the health effects of nutrition and family planning policies, and to examine the role of China's social and economic changes on the health and nutrition status of the population. The survey began in 1989 with eight provinces (Henan, Guangxi, Guizhou, Hubei, Hunan, Liaoning, Shandong and Jiangsu) and added a ninth, Heilongjiang, in 1997 and three autonomous cities, Beijing, Shanghai, and Chongqing, in 2011, and ran from 1991 to 2011. The Longitudinal Survey covers nine waves 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009 and 2011. By the last wave in 2011, the CHNS survey covered 32,877 individuals from 8619 households; detailed information about this data set is given in the Appendix.

After the combination of different separate datasets and deleting the necessary missing values, the estimation sample includes 12,221 observations and 7,186 individuals.[8] The dependent variable in my analysis is the log of hourly wages, defined as the log of monthly income divided by monthly hours worked. The mean value of log of hourly real wages is about 0.89. The distribution of the log of hourly wages is showed by the Figure 2-1.

As mentioned in the methodology section, Brown et al. (1980) estimated the earning function and occupational choice model separately. In terms of the earnings function, most studies utilize the human capital model as the theoretical basis (Becker, 1962). Based on human capital theory, wage increases with the accumulated skills including education, training and working experience. Further categories of explanatory variables will be taken into account as well, demographic factors (e.g., gender and marital status) and characteristics of job (e.g., working overtime and work unit). Based on Chinese-specific characteristics that transform the planned economy to the market economy, there are two types organizations in China, the public sectors and private sectors each having different mechanism to determine the employees' earnings. Therefore, the type of working unit is also considered in the wage equation. The detailed description of each determinant on earnings is shown in following paragraphs.

Regarding the occupational model, Boskin (1974), Schmid & Strauss (1975) and Brown et al. (1980) consider that individuals' occupational attainment depends on employers' willingness to hire one person and the job-hunters' ambition of a specific occupation. Employers' willingness to hire one person replies on the employees' education attainment, work experience, and training received. Job hunters' desire for a job depends on income, the consumption aspects of working, and whether the career is a stable job, often induced by the family size. Following the argument from Boskin (1974), Schmid & Strauss (1975), and Brown et al. (1980), educational attainment, work experience, number of children, and other demographic factors (gender and marital status) are included in the occupational model.

---

[8] The detailed description of getting this sample from the original sample is shown in Section 2.7.2.2 Data structure, in the appendix. In addition to the literal description, I also give Table 2-15 to show the sample clearance to guarantee it looks much clearer.

Figure 2- 1 Distribution of Lnwage

In terms of each independent variable, "*hukou*" contains the answer to the question "To which type of household registration do you belong?" and is used to divide all the individuals into one of the two *hukou* categories. The question regarding living area can be used to separate the rural *hukou* people into rural residents and rural-urban migrants. In addition, since I am analysing urban labour market discrimination between urban *hukou* workers and rural-urban migrants, it is necessary to choose individuals who are presently working, therefore, I only focus on working people in my sample. Only females between 16 and 55 and males between 16 and 60 are included in the sample since Chinese labour law sets the minimum legal working age at 16 and the retirement age for male and female workers at 60 and 55, respectively.

When measuring work experience, there is no available information in the CHNS. Also, since Heilongjiang province was added in 1997 and three autonomous cities, Beijing, Shanghai, and Chongqing were added in 2011, it is difficult to obtain information of history of actual work experience for individuals in these newly added areas. The most common method to measure work experience is taking the age of the individual minus years spent in education; however, it is less accurate if I estimate the years of schooling based on respondents' highest level of education they obtained. The primary school system is five years in some provinces while it is six years in other areas. Moreover, the length of normal undergraduate degree is four years, whereas it is five years for medical, architectural, and civil engineering domains. I, therefore, intend to follow Drolet & Mumford (2012) and utilize age as a proxy for work experience. When using age instead to interpret the estimated results, it is important to understand that the

relationship between age and earnings may be underestimated if individuals were not employed for a long time period (Chatterji, Mumford, & Smith, 2011; Mumford & Smith, 2007). Finally, there is no available information in the CHNS about the length of time that employees spent in occupational training either.

Education and occupation are categorical variables, and I will divide these two into multiple dummy variables. For education level, the Chinese government decreed the "Nine-year Compulsory Education Law" in 1986, which ruled that primary school and lower middle school were compulsory for all children. According to education law, all school-age children and teenagers must receive nine years of compulsory education, and the government covers all the tuition fees of this period. It is the obligation of the compulsory schools, parents, and the community to let school-age children receive the appropriate education. Those who do not fulfil this obligation may be punished by law. The purpose of this law is to enhance the integral educational levels of all persons in order to avoid illiteracy. After nine years of compulsory education, some students may choose to study further in higher middle school or technical school, while other students may give up studying further and devote their time to working, therefore, the categorical variable "highest education level" is divided into 4 dummy variables: no educational attainment, compulsory education, senior higher school, and university.

The variable stemming from the question "What type of work unit is this?" is a categorical variable including government, state service institute, state-owned enterprise, small collective enterprise (township-owned), large collective enterprise (such as owned by county, city, province), family contract farming, private enterprise, and three-capital enterprise (owned by foreigners, overseas Chinese and joint ventures). In the Chinese economic system, all the organizations and institutions can be divided into public ownership and private ownership. In general, public ownership means that all the production materials and capital goods are owned by all workers together. These workers devote themselves to working jointly and are in possession of the products jointly as well. In CHNS, government departments, state service institutes, state-owned enterprises, small collective enterprises, and large collective enterprises belong to the public ownership category. Private ownership, as opposed to the public ownership economy, refers to the situation where the means of production are possessed by individuals or some collectives exclusively, which means that individuals have

exclusive possession right and do not need to share the production material with others. In CHNS, private ownership includes family contract farming, private enterprise and three-capital enterprises. Therefore, I will transform this categorical variable into a dummy variable, where 1 refers to public ownership while 0 stands for private ownership. In addition, working time for different occupations is also taken into account. The dummy variable of overtime describes whether an individual always works overtime, where 1 refers to working overtime.

The categorical variable occupation type, which is the dependent variable in the occupational choice model in the Brown decomposition, stems from the question "What is your primary occupation?", and it includes senior professional workers (i.e., doctor, lawyer, professor and architect), junior professional workers (including midwife, nurse, editor, teacher, and photographer), administrator/executive/managers (i.e., government official, working proprietor, department director, section chief, and village leader), office staff (secretary, office helper), skilled worker (i.e., foreman, craftsman, and group leader), non-skilled worker (ordinary labourer, logger), service worker (i.e., housekeeper, counter staff, cook, sale person, and waiter), army officer, ordinary soldier, athlete or actor, and other. I have left out three categories, army officer, ordinary soldier, and athlete or actor. For the army officer and ordinary soldier, this is because their wages are decided directly to the national government, and are also related to their military rank. All army officers and soldiers who are at the same military level will receive the same wages from the national government, no matter to which company they belong. In addition, the athlete and actor are very particular occupations, as the athlete could earn significant amounts of money winning one competition and the actor could earn a very large remuneration for one movie, therefore, they cannot be compared with workers who earn regular wages. Hence, I kept only 8 categories from the original occupation-type framework, senior professional, junior professional, administrator/executive/manager, office staff, skilled worker, non-skilled worker, service worker, and other. I combined skilled and non-skilled workers as blue-collar workers, and drivers, service workers and other categories are merged as service workers. Therefore, there are 6 occupation types left in the models. Table 2-1 shows a description of all the factors in this chapter.

Table 2- 1 Description of Variables

Variables	Description
Dependent variable	
Hourly wages	Individuals' regular hourly real wages
Treatment dummy	
Rural	1- All respondents with rural *hukou* status but work in urban city
	0- All respondents with urban *hukou* status regardless of their working place
Independent variables	
Male	1- Male
	0- Female
Age	Age of respondent in years
Married	1- Married
	0- Unmarried
Work unit	1- Public sector
	0- Private sector
Working overtime	The legal working hours are 8 hours per day and 5 days per week.
	1- Respondents working hours per week exceeds 40 hours
	0- Otherwise
Respondents' highest education level	
Non-education	1- No education
	0- Otherwise
Compulsory level	1- Highest education level is compulsory school
	0- Otherwise
Senior high school	1- Highest education level is senior level (professional high school, senior high school and technical school)
	0- Otherwise
University	1- Highest education level is university
	0- Otherwise
Type of occupation	China as a dual structural country, rural area is mainly engaged in the production of agricultural activities. The diversity of non-agricultural occupations is mainly concentrated in urban cities. Individuals who are engaged in these occupations are urban *hukou* people and rural-urban migrants. The occupational categories given in the dataset is shown below:
Senior professional	1- Respondents who are engaged in senior professional jobs (doctors, lawyer, professor, architect and engineer)
	0- Otherwise
Junior professional	1- Respondents who are engaged in junior professional jobs (midwife, nurse, editor, teacher, photographer and journalist)
	0- Otherwise
Administration/executive/manager	1- Respondents who are engaged in administration job (administrator/ executive/ manager, government official, department director, section chief and leader)
	0- Otherwise
Office staff	1- Respondents who are office staff (secretary, office helper, typist)
	0- Otherwise
Blue collar worker	1- Respondents who are engaged in blue collar jobs (foreman, craftsman, group leader, ordinary labourer, logger)
	0- Otherwise
Service worker	1- Respondents who are engaged in service jobs (housekeeper, counter staff, cook, sale person and waiter/waitress)
	0- Otherwise

Note: China Health and Nutrition Survey is utilized.

Table 2-2 lists the summary statistics of the independent and dependent variables for all respondents. According to Table 2-2, the average value of the log of hourly wages for all the respondents is about 0.89. In addition, the results from Table 2-2 show the mean value of human characteristics and the probability of being in one occupational category for both urban workers and rural-urban migrants. Tables 2-3 and 2-4 present information on the differences in occupation distributions, hours worked, wages, and different individual characteristics between urban workers and rural-urban migrants.

Table 2- 2 Summary Statistics

Variables	Obs.	Mean	Std. Dev	Min	Max
Dependent variable					
Lnwage	12221	0.889	0.917	-2.079	5.187
Hourly wages	12221	3.821	6.122	0.125	178.947
Independent variables					
Age	12221	38.858	10.132	16	60
Age^2	12221	1612.602	792.097	256	3600
Rural *hukou*	12221	0.208	0.406	0	1
Married	12221	0.842	0.364	0	1
Male	12221	0.574	0.494	0	1
Public sector	12221	0.789	0.408	0	1
Overtime	12221	0.468	0.499	0	1
Highest education level					
Non-education	12221	0.028	0.165	0	1
Compulsory school	12221	0.398	0.489	0	1
Senior high school	12221	0.397	0.489	0	1
University	12221	0.176	0.381	0	1
Occupation type					
1. Senior professional	12221	0.107	0.309	0	1
2. Junior professional	12221	0.105	0.306	0	1
3.Admin./executive/manager	12221	0.114	0.318	0	1
4. Office staff	12221	0.131	0.338	0	1
5. Blue collar worker	12221	0.357	0.479	0	1
6.Service worker	12221	0.186	0.389	0	1

Note: China Health and Nutrition Survey is utilized.

Firstly, as can be seen from Table 2-3, the difference between occupational choice between urban workers and rural-urban migrants is obvious from the 6 occupation categories. Senior/junior professional, administration, and office staff are considered "white-collar", and only 23.16% of rural-urban migrants have white-collar jobs, while 51.63% of urban workers are in this category. However, about 76.84% of rural-urban

migrants are engaged in blue collar and service jobs, whilst 48.37% of urban workers are employed in these categories. Hence, according to the descriptive statistics, it is necessary to take occupation segregation into account and regard occupation attainment as an endogenous variable, since rural-urban migrants and urban workers are faced with different occupational distributions.

Table 2- 3 Occupational Distribution, Hours Worked, Earnings of Rural and Urban Workers

Occupation	Occupational distribution				Hours worked by occupation			
	Rural		Urban		Rural		Urban	
	Freq.	%	Freq.	%	Mean	SD	Mean	SD
Senior professional	56	2.20	1252	12.94	48.57	10.13	40.74	6.98
Junior professional	159	6.25	1122	11.59	42.91	9.88	41.10	6.97
Administration/ executive/manager	214	8.42	1176	12.15	44.19	13.89	42.01	6.95
Office staff	160	6.29	1447	14.95	44.18	9.36	41.38	7.12
Blue collar worker	1397	54.94	2963	30.62	51.77	12.56	45.22	9.74
Service worker	557	21.90	1718	17.75	49.43	14.54	47.17	12.58
Total	2543	100	9678	100	49.52	13.13	43.54	9.40
Occupation	Monthly wages by occupation				Hourly wages by occupation (ln)			
	Rural		Urban		Rural		Urban	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Senior professional	571.23	452.51	944.89	1239.07	0.79	0.86	1.37	0.94
Junior professional	431.85	319.45	721.95	785.68	0.65	0.86	1.12	0.93
Administration/ executive/manager	485.22	815.58	825.77	1300.51	0.58	0.93	1.16	0.93
Office staff	526.38	370.42	732.19	1072.84	0.89	0.75	1.08	0.92
Blue collar worker	490.84	480.65	472.36	699.22	0.61	0.77	0.61	0.85
Service worker	589.55	835.24	581.12	702.60	0.82	0.81	0.81	0.87
Total	512.30	597.22	663.62	954.73	0.68	0.81	0.94	0.94

Note: China Health and Nutrition Survey is utilized.

Secondly, rural-urban migrants are not only more likely to be employed in less prestigious jobs, they also work longer hours and earn lower wages than urban *hukou* workers. On average, rural-urban migrants work approximately 8 hours per week more than the urban *hukou* individuals, and the difference is particularly evident in the senior professional. Thirdly, rural-urban migrants earn lower wages than urban workers. The average monthly wages of rural-urban migrants (512.30) is 22.8% lower than that of urban *hukou* workers (663.62). The difference in working hours between rural migrants

and urban workers is taken into account and the dependent variable is defined as the difference in log of hourly real wages (the log (monthly real wages/ working hours)). Since rural-urban migrants work longer shifts than their urban counterparts, the hourly wage gap is wider than the monthly wage gap. The average log of hourly wages of rural-urban migrants (0.68) is 27.6% lower than the wages of urban *hukou* workers (0.94), which is similar findings by Meng & Zhang (2001).

According to Becker (1957), different wages and different occupational attainment should be attributed to productivity-related characteristics, such as educational levels, age, work experience, and non-productive related characteristics, such as gender, race, and the *hukou* status. If individuals with the same productivity related characteristics and other human capital indicators are treated differently in terms of occupation and wages because of race, gender and *hukou* status, it will be defined as discrimination. Therefore, before analysing the extent to which the characteristics influence the different wages and occupational choices, it is necessary to know the difference in individual characteristics for the wage model and occupation choice model, respectively.

Table 2-4 shows the comparison in personal characteristics between urban workers and rural-urban migrants, and presents information of human capital endowment and personal characteristics between the two groups. Rural-urban migrants are 3 years younger than their urban counterparts, on average. The descriptive statistics in Table 2-4 show the overtime variable is consistent with the statistics from Table 2-3. In Table 2-3, on average, rural-urban migrants work 8 hours per week are more than urban workers. Table 2-4 shows that 70% of rural-urban migrants work overtime, while only 41% of urban residents do. Lee (2012) using Chinese data also finds that the weekly working hours for rural-urban migrants is 13 hours more than for urban workers, on average, which is similar to this chapter.

As expected, the highest education level of rural-urban migrants is lower than for urban workers. Most rural-urban migrants only have compulsory education level (64%), while only 26% and 4% of rural migrants have graduated from senior higher education and university, respectively. Therefore, these rural-urban migrants may not be qualified for jobs with a high skill requirement. In comparison, 22% and 43% of the urban

workers graduated from university and senior high school. Hence, it is clear that the education levels of rural-urban migrants are lower than urban workers, in this context. In addition, the urban residents are more likely to be hired in the public sector (83% of urban workers), including governmental departments, state service institutes, state-owned enterprises, and small and large collective enterprises.

Table 2- 4 Comparison of Personal Characteristic between Rural and Urban Workers

	Rural-urban migrants		Urban residents	
	Mean value	SD	Mean value	SD
Independent variables				
Age	36.18	10.75	39.56	9.84
Number of kid	0.59	0.76	0.48	0.60
Married	0.80	0.40	0.85	0.35
Male	0.72	0.43	0.57	0.49
Overtime	0.70	0.46	0.41	0.49
Public sector (work unit)	0.62	0.46	0.83	0.37
Highest education level				
Non-education	0.06	0.48	0.02	0.37
Compulsory school	0.64	0.48	0.33	0.47
Senior high school	0.26	0.44	0.43	0.49
University	0.04	0.19	0.22	0.41

Note: China Health and Nutrition Survey is utilized.

In comparison, only 62% of the rural-urban migrants are employed in the public sector. Considering gender, about 72% of the rural-urban migrants and 57% of the urban workers are men. The higher proportion of men in rural-urban migrants can be explained by the nature of migration which has been discussed in Chapter 1. In urban China, the great majority of urban women choose to work due to the higher cost of living in urban cities. However, for rural *hukou* migrants, there were limited job opportunities that were concentrated in the construction industries in the early stage of migration. Thus, most of the rural *hukou* individuals who migrated to urban cities were men. Most rural women chose to stay in rural area and were reluctant to work in urban cities. With the accession to the WTO, types of jobs in urban cities are gradually diversifying and more job opportunities are available than that before. The number of rural *hukou* female who migrated to urban cities have increased gradually. About 85% of urban workers are married, while 80% of rural-urban migrants are married, which may be due to the fact that the average age of the urban *hukou* workers is higher (39.56)

than for rural-urban migrants (36.18).

The summary statistics in Table 2-3 and Table 2-4 are informative, but cannot tell us how the difference in the characteristics between urban and rural groups influences wages and occupation choice. In addition, it can also be difficult to ascertain the extent to which the difference in wages and occupational attainment between urban workers and rural-urban migrants can be explained by the difference in individual characteristics. In the next section, I will use the Brown et al. (1980) decomposition to disentangle this issue.

2.5 Results

In this section, the results come from applying the econometric techniques described in the methodology section to the data. Since some researchers have analysed wage differentials in relation to gender, race or *hukou* status by the Oaxaca decomposition approach previously, I will compare both approaches by Brown et al. (1980) and Oaxaca & Ransom (1994), and discuss the advantage of the Brown et al. (1980) decomposition.

Table 2-5 presents results of wage regressions by *hukou* status. The explanatory variables in the wage regressions are age, gender, marital status, type of work unit, a dummy variable for working overtime, a time dummy, and dummies for highest education level. As can be seen, all the explanatory variables have a significant effect on hourly wages for both urban people and rural migrants except the work unit. According to Table 2-5, there was an increase in wages for both rural migrants and urban *hukou* workers from 1997 to 2011. The overtime dummy variable indicates whether working overtime has a negative effect on hourly wages. The variable age has a positive effect on hourly wages: the hourly wages for rural-urban migrants increase by 1.4% for each 1 year increases in age, which is consistent with other research. That is because most researchers have regarded age as a proxy for work experience. The older individuals might have an abundance of work experience and might be more skilful; thus, they may be more likely to be hired in superior positions with higher wages.

The type of work unit has no significant impact on either urban workers or

rural-urban migrants. Married individuals are more likely to obtain higher wages than those unmarried. Married urban people earn 4.9% higher wages than unmarried urban people. In terms of education level, individuals who graduated from senior high school and university are more likely to be in superior occupations and positions that require higher levels of skills and knowledge; therefore, they might earn higher wages than those who do not receive education. In addition, the rate of return of a university education is much higher than that of senior high school. For example, compared to non-education for rural-urban migrants, individuals graduating from university earn 21.9% higher wages, and individuals graduating from senior high school earn 14.5% higher wages, on average. Workers with a bachelor's degree are more likely to be engaged in senior professional jobs, such as lawyer, architect, and doctor, while ones who graduated from senior high school might be occupied as office helpers, craftsmen or group leaders. Thus, workers with a bachelor's degree enjoy a higher rate of return than individuals who graduate from senior school. In addition, it can be seen that the rate of return to university level is lower for rural workers than for urban residents, similar to the findings by Messinis (2013) and Démurger et al. (2009). These authors indicated that the difference in terms of the rate of return of education levels for rural-urban migrants and urban *hukou* workers is most likely to be due to discrimination.

As discussed previously, I took occupational segregation into account and estimated with the multinomial logit model for occupational choice to process the Brown et al. (1980) decomposition. Tables 2-6 and 2-7 present the results of the multinomial logit model, which shows the effect of individual characteristics on occupational choice. Firstly, according to Table 2-6, for rural-urban migrants, in comparison to those who do not have education attainment (the base category), applicants graduating from senior high school and university tend to work in white-collar jobs, which includes senior and junior professional jobs, administrative and manager, and office jobs. The effect of senior high school and university are significant on all the categories for white-collar jobs. There is not a significant difference in the probability of being engaged in white-collar jobs (except for office staff) for respondents who only receive compulsory education. The insignificant effect of compulsory education level should be attributed to senior and junior professional jobs, and administrative and manager, which have higher requirements for employees'

education level. Individuals only having compulsory education level are not competent for most high-level jobs such as doctor and professor. Furthermore, the probability of being in a blue-collar job is lower than for people graduating from compulsory and senior high school and university, compared with individuals without any education attainment.

Table 2- 5 OLS Estimation of Hourly Wages Equations

Variables	Rural migrants		Urban workers	
	Coefficient	Cluster S. E.	Coefficient	Cluster S.E.
Constant	-0.325**	0.165	-0.651***	0.107
Independent variable				
Age	0.014**	0.005	0.021***	0.005
Age^2	-0.0003**	0.0001	-0.0002***	0.00006
Overtime	-0.350***	0.026	-0.423***	0.014
Male	0.278***	0.024	0.163***	0.013
Public (work unit)	-0.016	0.031	0.021	0.020
Married	0.065*	0.032	0.049**	0.021
Highest educational levels				
Compulsory school	-0.028	0.047	-0.029	0.047
Senior high school	0.145**	0.052	0.102**	0.046
University	0.219**	0.072	0.365***	0.048
Year 1997	0.465***	0.039	0.338***	0.020
Year 2000	0.631***	0.043	0.681***	0.023
Year 2004	0.859***	0.048	1.010***	0.023
Year 2006	1.008***	0.046	1.158***	0.024
Year 2009	1.343***	0.045	1.483***	0.025
Year 2011	1.567***	0.042	1.721***	0.022
Number of observations	2543		9678	
Adjusted R^2	0.5701		0.6818	

Note: (1) In highest education level, compulsory (including primary and junior middle school) is the base category.

(2) ***, ** and * indicate 1%, 5%, and 10% significant levels, respectively.

(3) China Health and Nutrition Survey is utilized.

(4) Standard errors adjusted for individual-level clustering in parentheses.

For example, rural-urban migrants who are university graduates are 5.5 percent and 9.8 percent more likely to work in senior professional jobs and junior professional jobs, respectively. Furthermore, they are 54.4 percent and 20.7 percent less likely to work in blue collar and service jobs, in comparison to those without educational attainment. The reason for these results is likely to be that the senior professional jobs, such as professors, doctors, architect and lawyers require more advanced and professional

knowledge or skills, which are based on higher education levels, whereas blue collar and service jobs, such as waiters, counter staff, and hairdressers, do not require the applicants to have an advanced educational background. In this case, companies can save on labour costs by hiring individuals without higher education levels.

For urban *hukou* workers, the results are similar to rural-urban migrants; individuals graduating from senior high school and university are more likely to be hired in white-collar jobs, including senior and junior professional jobs, administrative and manager, and office staff, in comparison to those without education attainment. As with rural-urban migrants, all the educational levels in Table 2-7 have negative and significant effects on being employed in blue collar and service jobs. Individuals graduating from university are 61.3 and 21.7 percentage points less likely to be engaged in blue collar and service jobs, respectively, compared to those with non-education attainment. For the blue collar and service categories that do not have higher educational requirements, the higher the education level is, the lower the probability of individuals being in these two occupational categories. For most white collar occupational categories, having a higher education level leads to a higher probability of being engaged in these occupations, except people with compulsory education attainment. For example, those who graduated from university are 29.3 percent more likely to be engaged in senior professional jobs. And those graduating from senior high school are 17.0 percent more likely to be in this category, in comparison to people with non-education attainment.

Comparing Tables 2-6 and 2-7, we also see that for the highest occupation category (senior professional) and the lowest occupation category (service jobs), urban *hukou* workers have a higher chance of being engaged in the senior professional category and a lower chance of being employed in a service job, in comparison to rural-urban migrants. For example, if individuals graduated from university, the probability of urban *hukou* workers in senior professional occupations would increase by 29.3 percentage points, while the probability of rural-urban migrants being in this category would only increase by 5.5 percentage points. In addition, for the most inferior occupations (service jobs) among those graduating from university, the probability of urban workers decreases by 21.7 percentage points, whereas the probability of rural-urban migrants drops by 20.7 percent. This situation is identical for the senior high school, for example, among those

Table 2- 6 Marginal Effects of Multinomial Logit Model, Rural-Urban Migrants

Occupation	Senior professional	Junior professional	Administrative /executive/manager	Office staff	Blue collar worker	Service worker
Compulsory school	0.021	0.028	0.017	0.025*	-0.163***	-0.072**
	(0.019)	(0.024)	(0.013)	(0.012)	(0.019)	(0.014)
Senior high school	0.038**	0.064***	0.040***	0.053***	-0.313***	-0.116***
	(0.011)	(0.013)	(0.007)	(0.012)	(0.020)	(0.016)
University	0.055***	0.098***	0.103***	0.082***	-0.544***	-0.207***
	(0.012)	(0.014)	(0.011)	(0.013)	(0.046)	(0.039)
Age	-0.001	-0.004***	0.007***	-0.007	-0.007**	-0.003
	(0.001)	(0.001)	(0.002)	(0.005)	(0.003)	(0.003)
Age2	0.0001	0.0005**	-0.0001***	0.0001	-0.0001	-0.0001
	(0.0003)	(0.0001)	(0.00001)	(0.0001)	(0.0002)	(0.0003)
Married	-0.001	0.002	-0.006	-0.003	0.018	-0.022
	(0.004)	(0.004)	(0.008)	(0.004)	(0.015)	(0.016)
Number of children	0.0001	0.005**	0.010***	-0.001	-0.011	-0.002
	(0.001)	(0.002)	(0.002)	(0.002)	(0.007)	(0.004)
Male	0.007**	-0.014**	0.028***	-0.003	0.047***	-0.030***
	(0.002)	(0.003)	(0.004)	(0.002)	(0.008)	(0.007)

Note: (1) ***, ** and * indicate 1%, 5%, and 10% significant levels, respectively.

(2) Highest education level includes non-education attainment (base category), compulsory school, senior high school and university.

(3) Standard errors adjusted for individual-level clustering in parentheses.

(4) China Health and Nutrition Survey is utilized.

Table 2- 7 Marginal Effects of Multinomial Logit Model, Urban *Hukou* Workers

Occupations	Senior professional	Junior professional	Administration /executive/manager	Office staff	Blue collar worker	Service worker
Compulsory school	0.006	0.108	0.024	0.046**	-0.110**	-0.074**
	(0.060)	(0.070)	(0.037)	(0.021)	(0.032)	(0.026)
Senior school	0.170**	0.210**	0.035**	0.092**	-0.328***	-0.180***
	(0.059)	(0.069)	(0.010)	(0.045)	(0.032)	(0.026)
University	0.293***	0.267***	0.133***	0.135**	-0.613***	-0.217***
	(0.059)	(0.069)	(0.036)	(0.048)	(0.035)	(0.028)
Age	0.008**	-0.012***	0.011**	0.001	0.0003	-0.009**
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)
Age2	-0.0002***	0.0002***	-0.0001**	-0.0005	-0.0001	0.0001**
	(0.00004)	(0.00003)	(0.00004)	(0.0004)	(0.0004)	(0.00004)
Married	0.010	0.006	0.031**	0.009	-0.017	-0.039**
	(0.013)	(0.011)	(0.013)	(0.012)	(0.015)	(0.013)
Number of children	0.021**	0.007	0.004	-0.028***	-0.014*	-0.018**
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Male	0.018**	-0.088***	0.069***	-0.044***	0.091***	-0.045***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.009)	(0.008)

Note: (1) ***, ** and * indicate 1%, 5%, and 10% significant levels, respectively.

(2) Highest education level includes non-education attainment (base category), compulsory school, senior high school and university.

(3) Standard errors adjusted for individual-level clustering in parentheses.

(4) China Health and Nutrition Survey is utilized.

who graduate from senior high school, urban *hukou* workers are 17.0 percent more likely to be employed in the senior professional occupation, while the proportion of rural-urban migrants in this category is 3.8 percentage points. For the impact of other socio-demographic factors, with an additional year of age increasing, the probability of urban workers being engaged in senior professional and administrative/executive/manager increases but at a decreasing rate. The probability of rural-urban migrants being engaged in administrative/executive/manager also increases at a decreasing rate. Number of children has a significant and positive effect on the junior professional and administrative/executive/manager for rural-urban migrants, while it has a significant effect on senior professional, office staff, blue collar and service workers for urban workers. The individuals' family situation, such as marital status, seems to affect few occupational attainments for urban workers. Urban respondents who are married are more likely to be engaged in administrative and manager positions and less likely to be a service worker. The reason for these results might be that among the administrative and manager positions most have more work experience and accumulation of social wealth with their married status. In addition, workers in these occupations are older and married longer, in comparison to blue collar and service workers, therefore that may be why being married has a positive effect on the probability of being employed in administration or as a manager.

Turning to the effect of gender, a dummy variable for being male has similar effect on the two groups. Male workers are more likely to be employed in senior professional jobs (doctor, professions or lawyer), administration, executive and manager positions, and blue-collar occupations. However, they are less likely to be engaged in junior professional jobs (midwife, nurse, and teacher), office staff such as secretaries, and service workers. This issue should be attributed to the occupational segregation between genders, because certain occupations tend to be carried out by one group, such as jobs like nurse, midwife and secretary, which women are more likely to do, and occupations like craftsman or logger, which tend to be done by men.

The difference in occupational choice between urban people and rural-urban migrants indicates that urban people and rural-urban migrants might be treated differently in the urban labour market. In order to evaluate the extent of different

treatment for rural-urban migrants with regard to occupational choice, I also estimate the actual and predicted occupational choice of urban workers and rural-urban migrants in Table 2-8. The predicted occupational choice for rural migrants refers to the occupation distribution of rural workers who have the same occupation opportunities as urban *hukou* workers based on the urban *hukou* workers' occupational attainment model and using the rural estimated characteristics. The difference between actual and predicted occupational distributions for rural-urban migrants can be attributed to the different treatment of the migrants. The predicted occupational choice of urban *hukou* workers is based on the rural-urban migrants' occupational attainment model using urban estimated characteristics. The predicted occupation distributions for both urban workers and rural-urban migrants indicate what occupation attainment of urban workers and rural-urban migrants would have been if they had been treated as their counterparts in the labour market.

Table 2-8 shows the difference between actual and predicted occupational choices for both rural-urban migrants and urban *hukou* workers. According to Table 2-8, about 1.50% more rural-urban migrants would be hired in senior professional occupations if they were treated as urban *hukou* workers. Only 5.18% of rural-urban migrants were employed as the office staff, while 11.56% of rural-urban migrants should have been in the office staff category if they had been treated as urban *hukou* workers. In addition, only 49.58% would have been in blue-collar jobs if they were treated equally to urban workers. This situation is reversed for urban *hukou* workers. The actual percentage of urban *hukou* workers in senior professional jobs is 8.77%, while only 3.06% of urban *hukou* workers would have worked in this type of occupation if they were treated equally to rural workers. On the contrary, about 11.75% and 2.84% more urban *hukou* workers would have been in blue-collar and service jobs, respectively.

It seems that rural-urban migrants are treated differently from urban *hukou* people in terms of their job opportunities. If occupation segregation did not exist, more rural-urban migrants would have been engaged in white-collar occupations, while more urban *hukou* workers would have been in blue-collar occupations. Meng & Zhang (2001) also estimated the different occupation attainments between rural migrants and urban residents, and reached the same conclusion. They also indicated that occupation

segregation not only lowered rural-urban migrants' social status, but also led to a wage gap between urban *hukou* individuals and rural-urban migrants. Liu et al. (2004) also found similar results in their study of occupation segregation and wage differentials between Hong Kong natives and mainland immigrants. In their studies, they found that the actual percentage of people in high quality jobs (i.e. managers, professionals, and clerks) is lower than the predicted percentage, whilst the actual percentage of service workers is higher than predicted. If the mainland immigrants were treated equally to the Hong Kong natives, the percentage of mainland immigrants in the professions, such as clerks, managers, administrative, and plant operators, would increase. Démurger et al. (2009) analysed whether rural-urban migrants were second-class workers in the Chinese labour market and found that about 50% of urban *hukou* workers were professionals, technicians, or office workers, while more than 50% of rural-urban migrants did not have regular jobs and were self-employed workers.

Table 2- 8 Actual and Predicted Occupational Distribution for Urban People and Rural-Urban Migrants

Types of Occupation	Actual %	Predicted %	Difference %
Rural-urban migrants			
Senior professional	0.79	2.29	1.50
Junior professional	1.99	4.74	2.75
Administration/executive/manager	4.55	5.56	1.01
Office staff	5.18	11.56	6.38
Blue collar worker	61.56	49.58	-11.98
Service worker	25.96	26.27	0.31
Urban *hukou* workers			
Senior professional	8.77	3.06	-5.71
Junior professional	11.19	8.55	-2.64
Administration/executive/manager	12.07	11.81	-0.26
Office staff	18.09	12.09	-6.00
Blue collar worker	29.04	40.79	11.75
Service worker	20.84	23.68	2.84

Note: China Health and Nutrition Survey is utilized.

In summary, in light of the analysis above, it can be seen that urban people and rural-urban migrants have different characteristics; in addition, they face different treatments in occupational attainment (Table 2-8) and wages (Table 2-3). However, it is difficult to confirm the extent to which the different treatments of these two groups can

be attributed to individual characteristics or unexplained factors, such as discrimination. Most Chinese researchers have regarded occupation distribution as an exogenous variable to analyse differences in treatment by the Oaxaca & Ransom (1994) decomposition methodology. However, by using this methodology, they may ignore the potential discrimination in the different occupational attainments.

Unlike these previous Chinese studies, I take the occupational attainment as an endogenous variable. The wage differential from Brown et al. (1980) can be decomposed into within-occupation and inter-occupation gaps. I will also compare the decomposition results between Brown et al. (1980) and Oaxaca & Ransom (1994). Before analysing the wage differential by Oaxaca & Ransom (1994) approach, I consider the self-selection of migration for rural people and try to correct the selectivity bias. However, there is limited reference in terms of the theoretical and technical approach to show how to address the selection problem in Brown et al.'s (1980) decomposition; consequently, it is difficult to guarantee the feasibility of addressing selectivity. However, Neuman & Oaxaca (2004) present how to address the selection issue in Oaxaca's decomposition (Blinder-Oaxaca decomposition and Oaxaca-Ransom decomposition), and Jann (2008) gives a description of the technical calculations of selectivity in Oaxaca's decomposition. Due to the difficulties for addressing selectivity in Brown et al.'s (1980) decomposition, I intend to solve the self-selection for migration in Oaxaca & Ransom (1994) decomposition and compare the decomposition results with and without selectivity. These comparison results can provide evidence for whether the Oaxaca & Ransom (1994) decomposition results are sensitive to selectivity correction so that I can understand if the results of the Brown et al. (1980) decomposition will be influenced by ignoring selectivity. If there is no significant difference in the Oaxaca & Ransom (1994) decomposition results with and without selectivity, the estimation by the Brown et al. (1980) decomposition approach in this chapter would likely not be affected by ignoring selectivity and the estimations are more reliable. On the contrary, if there is a large difference when comparing the Oaxaca & Ransom (1994) decomposition with and without selectivity, the estimation results of the Brown et al. (1980) decomposition without selectivity in this chapter are likely biased. This caveat should be kept in mind when interpret the Brown et al. (1980) decomposition results.

Although the Oaxaca & Ransom (1994) decomposition is a benchmark for the Brown et al. (1980) decomposition, it is not the main content in this chapter. All of the information about the description of the theory, formulas for selectivity, selection model, exclusion restrictions, and probit estimations are shown in the Appendix. In the main results, I only demonstrate and compare the Oaxaca & Ransom (1994) decomposition results with and without selectivity. The estimation of the Brown et al. (1980) decomposition and the Oaxaca & Ransom (1994) decomposition with and without selectivity are compared in Table 2-9. Columns (A) and (B) show the results from the common decomposition approach, Oaxaca & Ransom (1994), without and with selectivity, respectively, which computes the decomposition by using the coefficients from a pooled model over both urban people and rural-urban migrants as the reference coefficients (Oaxaca & Ransom, 1994; Jann, 2008).

In column (A), according to the sample, the mean value of ln(wage) for urban people is 0.940 and is 0.676 rural migrants, which yields a raw wage differential of 0.264. The share of the explained portion is approximately 83.3%, whilst the unexplained part is around 16.7%. The estimation of the decomposition indicates that a majority part of the mean wage differential between urban people and rural migrants is ascribed to observed human capital or individual characteristic differentials. A small portion of the income gap is attributed to the unobserved factors, which are most likely to be discrimination against rural migrants. Among the explained portion, the most important contributing determinants are job characteristics (42.4%), education level (24.6%), and types of occupation (11.4%). Thus, the explained portion to total wage differential indicates that hourly wage differentials between urban people and rural migrants is mainly attributed to the difference in the different occupation distribution and their different education levels. The unexplained portion can be decomposed further into advantages of urban people (overpayment and favouritism) and disadvantages of rural migrants (underpayment and discrimination).

When the self-selection of migration for rural people is taken into account, the adjusted wage differential (shown in Eq. 2-21) between urban people and rural-urban migrants shown in column (B) increases to 0.370, with 0.218 explained (58.9%) and 0.152 unexplained (41.1%). In comparison with the decomposition without selectivity, the detailed decomposition results with selectivity for migration expresses that there is a

Table 2- 9 Wage Decomposition by Oaxaca-Ransom with and without Selectivity

	Oaxaca-Ransom (1994) without selection (A)		Oaxaca-Ransom (1994) with selection (B)	
	Coef.	Perc. (%)	Coef.	Perc. (%)
Wage differential	0.264	100%	0.370	100%
Total explained	0.220	83.3%	0.218	58.9%
Total unexplained	0.044	16.7%	0.152	41.1%
Advantage of urban	0.002	0.8%	-0.001	-0.3%
Disadvantage of rural	0.042	15.9%	0.153	41.4%
Total explained	**Coef.**	**Perc. (83.3%)**	**Coef.**	**Perc. (58.9%)**
Social demographic	-0.003	-1.1%	-0.003	-0.8%
Age (working experience)	0.015	5.6%	0.013	3.5%
Education level	0.065	24.6%	0.064	17.3%
Job characteristics	0.112	42.4%	0.111	30.0%
Occupations	0.030	11.4%	0.031	8.4%
Time dummies	0.001	0.4%	0.002	0.5%
Total unexplained	**Coef.**	**Perc. (16.7%)**	**Coef.**	**Perc. (41.1%)**
Social demographic	-0.081	-30.7%	-0.095	-25.7%
Age (working experience)	0.317	120.1%	0.372	100.5%
Education level	-0.041	-15.5%	-0.072	-19.5%
Job characteristic	-0.032	-12.1%	-0.032	-8.5%
Occupations	-0.068	-25.8%	-0.069	-18.6%
Time dummies	0.091	34.5%	0.095	25.6%
Constant	-0.142	-53.8%	-0.047	-12.7%

Note: (1) Social demographic factors contain gender and marital status. Education level include non-education, compulsory school, senior high middle school and university which non-educational attainment is regarded as reference group. Job characteristics include the dummy variables of working overtime and work unit.

(2) In the Oaxaca & Ransom (1994) decomposition, the occupation types are included as exogenous dummy variables. The occupation categories include six dummy variables that are senior professional, junior professional, administration/executive/manager, office staff, blue collar workers and service workers, respectively. In terms of the categorical variable e.g. occupation, the decomposition results reply on the choice of reference group. The Stata command "*categorical*" is used in this chapter to identify dummy variables for categorical variables and transform coefficient vectors. Consequently, deviations from the grand mean are presented and the coefficient for the base category is added (Yun, 2005; Jann, 2008). Such transformation of estimation can guarantee the decomposition results do not rely on different choice of omitted base category.

(3) The unexplained portion of Oaxaca & Ransom (1994) decomposition can be further decomposed into advantaged of urban people and disadvantaged of rural migrants.

(4) Decomposition approaches are usually utilized in cross-sectional dataset. There is very litter research conducting the panel analysis currently in decomposition method (including Oaxaca-Blinder, Oaxaca-Ransom, Brown decomposition) regardless of the structure of panel data. That is because it is difficult to obtain corresponding coefficient for those variables that do not change over time through the panel analysis to perform the further decomposition. In addition, while the CHNS is a panel dataset most individuals are only observed in one wave (see Table 2-12), which makes it difficult to control for fixed effects in this chapter. In China, new cities or new provinces

were added constantly into different waves of several panel dataset, due to Chinese specific situation. In addition, some rural areas or districts in one of the provinces did not continue to be investigated due to several special cases in local province. In order to consider wage gap across the whole country, one Chinese researcher Ge (2007) pooled different years together. In this chapter, I follow Ge (2007) to pool different years and decompose the wage differential between urban people and rural-urban migrants. The standard errors are clustered by ID in this chapter to control the individual effect.

(5) China Health and Nutrition Survey is utilized.

slightly different impact of the explained portion, whereas the unexplained portion varies to a large extent. The variation of the raw wage differential with selectivity is more likely to be ascribed to the unexplained portion. Similar to the Oaxaca & Ransom (1994) decomposition without selectivity, job characteristic (30.0%), education attainment (17.3%), and types of occupation (8.4%) also denote the major contributions to the explained portion. Therefore, these comparison estimations show that the extent of discrimination of decomposition results are likely to be downward biased when ignoring migration selection for rural people.

The reason why the coefficient effects (unexplained portion) in the Oaxaca & Ransom (1994) decomposition with selectivity increases a lot may be ascribed to that rural-urban migrants differing markedly from non-migrant rural *hukou* residents in terms of some unobservable factors. Given the similar observed human characteristics and unobservable productivity determinants, rural-urban migrants are more likely to have greater ambition, more motivation, and better capability of dealing with pressure and recovering, compared to non-migrant rural people (Lee, 2012). Rural-urban migrants, therefore, are likely to be a group who have these unobserved factors. In this case, those rural people who choose to migrate to urban cities may not be a random sample among all the rural *hukou* populations. In comparing the Oaxaca & Ransom (1994) decomposition with selectivity and without selectivity, I realize that the Brown et al. (1980) decomposition results in my analysis without correcting for selectivity shown in Table 2-10 may be downward biased and may underestimate the actual discrimination against rural -urban migrants. The failure to deal with sample selection in the Brown et al. (1980) decomposition is the limitation of this chapter.

As shown in Table 2-10, the estimated results from the Brown et al. (1980) decomposition change significantly, in comparison with results from the Oaxaca & Ransom (1994) decomposition without selectivity. The explained portion of the wage

gaps from the Brown et al. (1980) decomposition is smaller than the results achieved from the Oaxaca & Ransom (1994) decomposition, once occupational segregation is taken into account. Using the Brown et al. (1980) method, the total explained portion decreases to 69%, and the unexplained portion of the whole wage differential is 31%. This result is consistent with Brown's theory and other research. For example, Fearon & Wald (2011) analysed the earnings gap between black and white workers in Canada. They indicated that 41.2% was due to differences in the endowments of characteristics, while 58.8% was not attributed to the characteristics when using the Oaxaca & Ransom (1994) approach. However, only 21.1% of the wage differential was attributed to endowment differences, while 78.9% was obtained from an unexplained portion once the Brown et al. decomposition was used. Meng & Zhang (2001) obtained similar conclusions.

Table 2- 10 Decomposition Results of Wages Differentials between Urban *Hukou* Workers and Rural-Urban Migrants

	Brown et al. decomposition		Oaxaca-Ransom decomposition without selection	
	Log hourly wages	% of total	Log hourly wages	% of total
Total earnings differential	0.264	100	0.264	100
Within-occupation	0.101	38%		
Explained	0.037	14%		
Unexplained	0.064	24%		
Inter-occupation	0.163	62%		
Explained	0.146	55%		
Unexplained	0.017	7%		
Total explained	0.183	69%	0.220	83%
Total unexplained	0.081	31%	0.044	17%

Note: (1) China Health and Nutrition Survey is utilized.

(2) The reason why I compare the Brown et al. (1980) with Oaxaca & Ransom (1994) without selectivity correction is attributed to that the results of Brown et al. (1980) is this table is not dealt with selectivity.

The reason for this result is that, in the Oaxaca & Ransom (1994) decomposition, rural-urban migrants are compared to urban *hukou* workers, regardless of their occupational attainment. However, in the Brown et al. (1980) decomposition, the occupation distribution between urban individuals and rural-urban migrants is taken into account. Rural-urban migrants are compared with urban ones within each occupation.

As shown in Table 2-3 above, rural-urban migrants are more likely to be engaged in lower-status categories. Taking this into account, occupational distribution leads to the unexplained part of the whole wage differential rising. To sum up, the comparison between the two decompositions shows that the increase in the unexplained portion should be attributed to inter-occupational wage differentials between rural-urban migrants and urban *hukou* individuals.

Brown et al.'s (1980) method decomposes the wage differential into four categories, which are within occupational explained and unexplained, and inter-occupational explained and unexplained portions, all listed in Table 2-10. The average value of log hourly wage differentials between urban workers and rural-urban migrants is 0.264, 0.101 or 38% of the whole wage gap is attributed to within occupational wage differentials, and the rest (0.163 or 62%) should be attributed to inter-occupational wage gaps from differences in occupations. On one hand, of the 0.101 or 38% within occupational wage differentials, about 14% of the overall wage differential could be explained by different personal characteristics, such as age, a dummy of overtime, gender, marital status, and work units. Here, it means that the wage differential between urban people and rural-urban migrants is due to different human capital, work experience. The difference in individual characteristics between urban and rural-urban migrants is shown in Table 2-4. It can be seen that urban *hukou* workers are three years older than rural migrants, on average. In the Chinese labour market, especially in the public sector, wages are positively affected by length of service and ranking level; for example, the wages for civil servants would grow with their seniority. At the same ranking level, then, older individuals would be likely to earn more wages than younger persons. Urban *hukou* workers in this chapter might earn higher wages than their rural-urban migrants, as the average age for urban *hukou* workers is higher than that for rural ones. The variable "age" contributes to the explanation of the wage gaps. In addition, within each occupation, working overtime, marital status, gender, type of work unit, and education level also makes contributions to the wage gaps. The effect of these variables on wages was previously shown in Table 2-5. On the other hand, there is an unexplained portion, 0.064, or 24% of the entire within-occupation wage differential. These results show that 24% of the intra-occupational income differential is ascribed to the different treatment between urban people and rural-urban migrants, which are more likely to be attributed to discrimination in favour of urban workers and against

rural-urban migrants.

Kidd & Shannon (1996) indicated that some of the within-occupational wage differences might be inter-occupational gaps. According to Table 2-10, about 0.163 or 62% of the entire hourly wage is attributed to inter-occupational differences. Among the inter-occupational wage differentials, 7% is due to occupation segregation and 55% is attributed to human characteristics, such as marital status, number of children, gender, age, and education level. Tables 2-6 and 2-7 show the influence of these variables on occupational attainment. It can be seen that gender and education levels are major contributing factors to the explained portion of wage gaps. Further, 7% of occupational segregation indicates that rural-urban migrants are faced with different treatments while applying for jobs even if they have the same human characteristics as urban *hukou* workers. This portion may be attributed to discrimination in favour of urban *hukou* workers and against rural ones. Overall, 69% of the total income gap is attributed to personal endowments and 31% of the earnings differential can be ascribed to the unexplained part, which might be due to discrimination.

Comparing the within-occupational portion with the inter-occupational portion, the latter makes a major contribution to the total wage differential. It indicates that the occupation segregation shows a significant effect on wage gaps by oppressing rural-urban migrants in lower occupational categories. However, in Meng & Zhang (2001) and Fearon & Wald (2011), who also used the Brown et al. (1980) decomposition, the occupational segregation has a significant effect on wage differentials, although its measured impact on the wage gaps is quite small (Meng & Zhang, 2001). In their research, the within occupational portion is the major contributing factor to the entire wage differential, and according to Liu et al. (2004), the importance of intra-occupational wage differentials is almost the same as the inter-occupational differential. The effect of the within-occupational earnings gap is slightly higher than the inter-occupational gap. In their research, the intra-occupational portion is the major contributing factor to the entire wage differential. The reason for this result might be attributed to the different personal characteristics chosen in their research. If the variables in the wage equation were the major contributing factors to the total wage gaps, the influence of the intra-occupational differential would be higher than occupation segregation. For example, Meng & Zhang (2001) chose job experience,

current job tenure, years of educational schooling, a dummy variable for training, and other socio-demographic factors, including gender, marital status, and number of children in the wage model. They chose age, years of schooling, a dummy for training, and the same personal characteristics in the occupation distribution model. In their research, job experience and job tenure had a positive and significant effect on wages. The effects of job experience and job tenure might have made a significant contribution to the within-occupational wage gaps, which in turn probably influenced the entire wage differential. In this chapter, it is quite likely that occupational segregation makes a major contribution to the hourly wages because of the significant effects of higher educational levels on occupational attainment. Moreover, unlike Meng & Zhang (2001), this chapter analyses the wage differential between urban people and rural migrants all over the nine provinces, including the coastal city (Shanghai) and provinces and inland provinces such as Henan and Heilongjiang, which mainly focuses on heavy industry and manufacturing. As discussed in section 2.1, with economic transitions and the coverage of the survey expansion, an increasing number of industries and more job opportunities have appeared all over the China. Different treatments for applying for a job become more widespread than that only in Shanghai. Thus, this may be a reason why the inter-occupational wage differential is higher than the intra-occupation wage gap in this chapter. In addition, Messinis (2013) also explained that the educational level could lead to different opportunities in different occupations, which in turn influence the returns from education. The effects of the educational level on the wage gap mainly come from different occupational attainments, which may be why the effects of inter-occupational segregation outweigh that of intra-occupational wage gaps in this study.

Overall, my findings on the wage differential between urban people and rural-urban migrants are consistent with previous studies based on CHNS (i.e., Messinis, 2013 and Démurger et al., 2009), which is that a sizeable wage differential between urban and rural-urban migrants is the explained portion. However, Lee (2012) and Meng & Zhang (2001) have opposite findings, where they indicate that most of the difference cannot be explained by characteristic differences between the two groups, which implies that the rural-urban migrants are discriminated against. The difference in conclusions may be because of the different datasets and methodologies used in the different studies. For example, Lee (2012) used the China Urban Labour Survey, which covers five cities: Wuhan, Shenyang, Xi'an, Fuzhou, and Shanghai. His results suggest

that rural-urban migrants face significant discrimination, but this discrimination varies a lot between the different cities. In addition, although Meng & Zhang (2001) utilize the same Brown et al. (1980) decomposition method, they only analyse the different treatment of urban people and rural-urban migrants in the Shanghai province. In addition, the CHNS dataset in my research involves almost the whole of China, including 9 provinces (Guangxi, Guizhou, Henan，Hubei, Hunan, Jiangsu, Liaoning, Shandong and Heilongjiang). In addition, because of the limitations of my dataset from the China Health and Nutrition Survey (CHNS), the variables related to occupation and wages are limited. For example, there are no available variables related to job tenure, work experience, and so on. Therefore, choosing different human characteristics in the decomposition might lead to different results.

2.6 Conclusion

This chapter utilizes data from the China Health and Nutrition Survey (CHNS) to estimate the wage differential of urban workers and rural-urban migrants in the Chinese urban labour market. Although the existence of different treatment in the Chinese urban labour market is well-known, how the market is segregated between urban workers and rural-urban migrants has not been explored in detail. Specifically, only 23.16% of rural migrants have white collar jobs, while 51.63% of urban workers are in this category. However, most previous research regards the different occupation distribution as an exogenous variable. Unlike these previous studies, occupation segregation has been considered as an endogenous variable influencing the wage differential in this chapter. Utilizing the Brown et al. (1980) decomposition, the actual and predicted occupational distribution of urban residents and rural-urban migrants show that approximate 12% of rural-urban migrants who would have been working in white collar jobs if they are treated the same as urban workers are employed in blue collar jobs or service jobs. This occupation mismatch means that rural-urban migrants earn lower wages than what they would have earned, if discrimination did not exist.

Without considering the occupation segregation of urban residents and rural-urban migrants, the Oaxaca & Ransom (1994) decomposition without selectivity correction shows that about 83% of the wage differential is attributed to differences in human characteristics between urban residents and rural-urban migrants. 17% of raw wage gap

is the unexplained portion which is more likely to be ascribed to *hukou* discrimination. When taking occupational distribution into account, the estimated results show that the proportion of the explained part decreases to 69% and the unexplained portion increases to 31%. These results suggest that the results from the Oaxaca & Ransom (1994) decomposition underestimate the discrimination against rural-urban migrants.

My results show that a large proportion of the wage differential can be attributed to the inter-occupational wage gap (62%). The whole wage differential is mainly ascribed to occupation segregation. Specifically, I find that the main reason that rural migrants have lower wages is due to the fact that they are restricted from obtaining high quality jobs, not unequal pay within occupations. Among the inter-occupational wage differential, 55% of the wage gap is due to human characteristics, such as rural-urban migrants' lower education levels. Due to the positive association between education attainment and occupation distribution, the lower educational level of rural-urban migrants prohibits them from getting better jobs. In addition, about 7% of inter-occupation by wage gap is ascribed to an unobserved characteristic which is most likely to be due to discrimination (it may be the synthetic of prejudice and statistical discrimination). Although the causal factors of taste-based discrimination and statistical discrimination is slightly different, i.e., the former one is due to employer's animosity against rural people and statistical discrimination is attributed to employer's limited information on rural people which leads to the assessment of individual characteristics by this group's average situation, both of these two types of discrimination depend on employer's unobservd personal attitude. Due to that there is no difference in the appearance of urban *hukou* workers and rural-urban migrants and no available questions related to respondents' specific attitude to the minority group in CHNS, it is difficult to distinguish prejudice from statistical discrimination in China. In addition, Charles & Guryan (2008) and Doleac & Stein (2013) also claimed that animosity can be a product of statistical discrimination, and vice versa; the presence of one type did not preclude the presense of the other. Therefore, the discrimination found in this chapter may depend on the employers' prejudice against rural migrants. It may also be ascribed to that all rural-urban migrants are regarded according to their average situation, rather than their personal characteristics. Since, in general, rural-urban migrants' educational levels are lower than urban workers, those highly educated rural migrants may be treated unfairly, being prevented from obtaining better jobs corresponding to their

educational attainment.

In terms of the total explained and total unexplained part of the wage differential of urban workers and rural-urban migrants, 69% of the whole wage differential can be explained by individuals' characteristics and 31% cannot be explained by observable factors. However, it should be paid more attention to that the self-selection issue has not been dealt with in Brown et al. (1980) decomposition. By observing the Oaxaca & Ransom (1994) decomposition with selectivity and without selectivity, I found that the estimation will be downward biased while ignoring the selectivity correction. Therefore, the estimated results obtained from Brown et al. (1980) underestimate the actual discrimination against rural-urban migrants. This unsolved selectivity issue in Brown et al. (1980) decomposition is the limitation of this chapter. Regarding the policy implication, the results suggest that the policy should not only consider how to eliminate the difference in wages, but also should deal with the *hukou* discrimination when applying for jobs.

2.7 Appendix

China Health and Nutrition Survey (CHNS) is a longitudinal dataset with 9 waves, including 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009 and 2011. It contains multiple variables obtained over multiple time periods for the same individuals, recording information of individuals and households in a number of years.

2.7.1 Sample Changes

Table 2-11 states the integral information of the observations in each survey and how the process of observations' changes during these waves. In addition, family size and number of children in one family can be seen from Table 2-11. By comparing family size in each year, it can be seen how family size and number of children in one household changed during the waves. In row 3 of Table 2-11, "-" indicates the number of individuals leaving in that specified year, while "+" refers to the number of individuals added to the sample. According to the questionnaires, the reason for the added individuals can be attributed to new-born children, marriage and newly formed households, replacement households and all households in replacement communities;

meanwhile, subtracted individuals may be attributed to moving away or death. In Table 2-11, it is obvious that the average number of children in one household is slightly higher and the family size marginally larger than before 2004. This may be attributed to the fact that the One-Child policy was carried out in 1982 transforming the Law of the People's Republic of China on Population and Family Planning in 2001.Therefore, with the strict limitation of the law, an increasing number of persons give birth to one child, so that the average number of children after 2001 will decline gradually. In addition, in the early stages of the policy the restriction on rural residents was not rigorous. In rural areas, if the first child was a girl, a second child would be allowed. In addition, in light of the question "relationship with householders", 30.73% of respondents said they were the head of the household, and 27.20% that they were the spouse of the householder.

Table 2- 11 Information for Each Wave

Year	Observations	Changes in observations compared to the previous wave	Mean of family size in each wave	Mean of number of children under 18 in each wave
1989	14343	0	4.18	0.93
1991	13672	1989: -1842 1991: +1171	4.14	0.93
1993	13320	1991: -1757 1993: +1405	4.21	0.99
1997	13774	1993: -3920 1997: +4374	3.96	0.82
2000	14890	1997: -2430 2000: +3546	3.80	0.70
2004	11555	2000: -5489 2004: +2154	3.06	0.39
2006	11160	2004: -2825 2006: +2430	2.96	0.31
2009	11296	2006: -3171 2009: +3307	3.02	0.27
2011	14538	2009: -2731 2011: +5973	2.92	0.28

1989-2011(total): 118548 observations; 32877 individuals

Note: China Health and Nutrition Survey is utilized.

Among the rest of the respondents, more than 30% are kids of householders and their spouses, householders' parents occupying 2.44%, householders' grandsons/granddaughters taking over 3.72% and 3.99% being householders'

son-in-law/daughter-in-law. It is evident that there may be more than one generation in one family, which can explain family size.

Table 2-12 shows the frequency of individuals interviewed, meaning how often individuals were interviewed among the 9 waves. 28.47% (9360) of all the persons (32877) were interviewed once, which may be attributed to the new members in old households and all members in new households. In 2011, 5973 new members joined and 1842 persons left after 1989, which took up 23.8% of all the individuals, and these were interviewed just once. In addition, 7.49% of all individuals took part in all the nine surveys from 1989 to 2011. The average frequency of individuals interviewed is 3.606.

Table 2- 12 Frequency of Individuals Interviewed

Frequency of being interviewed	Observations	Individuals	Percentage of frequency being interviewed
1	9360	9360	28.47%
2	9824	4912	14.94%
3	13353	4451	13.54%
4	14496	3624	11.02%
5	15930	3186	9.69%
6	13278	2213	6.73%
7	8428	1204	3.66%
8	11712	1464	4.45%
9	22167	2463	7.49%
Mean=3.606			

Note: China Health and Nutrition Survey is utilized.

2.7.2 Questionnaire Information and Data Structure

2.7.2.1 Questionnaire Information

CHNS as a longitudinal data with nine waves, covering household, nutrition, physical examination, a never-married women survey and a community survey in 1989, 1991, 1993, 1997 and 2000, involving questions on detailed demographic information (age, marital status, family's relationship, education, employment and income) of head of households and other members, time allocation at home (e.g. elderly care, baby care, and other key home activities), economic activities (e.g. job, wage, expense occupation), health status, insurance, activities of daily living, nutrition (detailed diet) and so on. All

the surveys in each wave are not completely identical; however, the basic information of demographic and socioeconomic characteristics of household and individuals stayed the same. Table 2-13 shows an overview of the data CHNS domain for each year.

Table 2- 13 CHNS Domains

Questions domains	CHNS
Number of waves	9 (1989-2011)
Personal characteristics (age, gender)	All surveys
Marriage history	All surveys
Hukou status	1993-2011
Household composition (number of members)	1989-2006
Number of children	1989
Education	All surveys
Income/employment by sector of work	All surveys
Detailed health service use/insurance	All surveys
Time allocation/physical activity-inactivity	All surveys
Basic health examination (e.g. weight/height/ blood pressure)	All surveys

Note: China Health and Nutrition Survey is utilized.

Table 2- 14 *Hukou* Information

Waves	Obs.	*Hukou* distribution	Changing numbers	Percentage changeper wave
1993	13320	Urban:34.46% Rural:65.54%		
1997	13774	Urban: 37.44% Rural: 62.56%	1290	9.37%
2000	14890	Urban: 37.46% Rural: 62.54%	1801	12.10%
2004	11555	Urban: 41.33% Rural: 58.67%	1130	9.78%
2006	11160	Urban: 40.85% Rural: 59.15%	945	8.47%
2009	11296	Urban: 40.63% Rural: 59.37%	820	7.26%
2011	14538	Urban: 49.64% Rural: 50.36%	751	5.17%

Note: China Health and Nutrition Survey is utilized.

In this chapter, it is necessary to use *hukou* status to distinguish urban *hukou* workers and rural-urban migrants. However, before 1993, there were no questions related to this issue, so that I can just use 7 waves from 1993 to 2011 to analyse this issue. On the basis of the panel data recording dynamic change, Table 2-14 provides *hukou* distribution and *hukou* changes for each wave.

2.7.2.2 Data Structure

The data set CHNS has 20 independent documentations, based on the related information, which are agriculture, birth history, child care, children energy expense, education file, ever-married women, health care, household assets, ID file, income file, infant feeding files, marriage history, medical insurance, physical exam files, nutrition files, pregnancy files, relationship files, roster files, and time-use files.

The documentation on ID includes fundamental information about respondents' houses in nine waves, such as household number, member number of each individual in household, whether each family member lived in the household, interview date, province, etc. According to the ID file, it is evident that the household number (hhid) for any one family is constant during the nine waves, hence each individual and household can be tracked accurately by the ID number. In other documents, each individual is required to provide information such as demographic characteristics, education, and occupation under the corresponding household number, member number, and wave. Hence, household number and member number in each household and wave is the connection hub when merging different datasets.

The education file provides information on individuals' educational history and status (e.g., highest educational level and whether currently in school). The occupation documentation includes information on employment, such as whether presently working, whether retired, primary occupation, employment position, type of work unit, hours of work, regular wages, and secondary jobs. Questions were also asked about bonuses and subsidies (e.g., whether individuals received a bonus, the total bonus amount receiving for the whole year, whether subsidies were given, details of types of subsidies, and the total subsidy received for the whole year), as well as household and individual businesses, including type of business, average revenues, and average expenses.

The remaining documentation provides information about health status, child care, daily-living activities, assets, diet, and the agriculture activities in which rural residents are engaged. For example, child care involves the issue of who cares for the children in one household, whether the baby is cared for at home, whether child care takes place at a care centre or nursery school, how much household income goes to child care, and whether the family has a child subsidy. The agricultural documentation includes questions about whether individuals engaged in farming, fishing, livestock, or gardening. Total revenues and costs of these activities are also asked in this section. In the medical insurance documentation section, questions on the type of medical insurance, payments for health care, whether insurance is provided by employers, and so forth, were asked.

The variables I choose in this chapter can be divided into three sets, demographic information and educational and employment information. Thus, the datasets I used came from the four CHNS documents: ID, roster, education, and occupation files (including the occupation and wages datasets). In addition, the self-selection for migration is corrected for in the Oaxaca & Ransom (1994) decomposition so that the relevant information about health or disease history in physical examination files is taken into account as well. When I merge these six datasets (ID, roster, education, occupation, wages, and physical-examination), there are 41,186 valid observations (18,481 individuals) in this successfully merged sample. The study in this chapter mainly concentrates on the survey from 1993 to 2011, as the survey questionnaires have slight differences in the first two waves; for example, there are no available questions related to the *hukou* status and the given occupation categories are different with other waves. Therefore, 6,236 observations from 1989 or 1991 are deleted from the original sample. The socio-demographic factors, age, marital status, *hukou* status, and gender can be picked out from the two files (ID file and roster file). In addition, a majority of other independent variables that I bring forward in this chapter were selected from the occupation files (including the occupation and wages datasets), such as occupation type, employment position, work unit, working time, and regular wages.

Firstly, since I investigate Chinese urban labour market discrimination, only the people within the legal working age range are selected. The minimum legal working age is 16 years old and retirement age for males and females is 60 and 55, respectively. Moreover, individuals within the legal working age who are currently studying at school

during the investigation stage are also not taken into account. The remaining sample that is within the legal working age and not studying at school includes 25,427 observations. Secondly, I remove those with missing values for demographic factors, education attainment, and health-related variables that cannot be replaced with relevant information from the previous or following wave. Thirdly, I delete the missing values of the remaining occupational related factors (e.g. work unit, working hours, occupation type) for those people who are employed in specific occupations and report their regular wages in the urban labour market. Moreover, the extremely high and extremely low value in terms of real income (top/bottom 1%) are also dropped from the sample in order to guarantee even income distribution. Lastly, as discussed in chapter 1, self-employed people do not suffer from employee discrimination. Due to the

Table 2- 15 Sample Clearance

The procedure for cleaning sample	Changes in the observations to the original sample	The remaining observations
The successfully merged sample	–	41186
Step1: Delete wave 1989 and wave 1991	-6236	34950
Step2: Delete those who are not in legal working age	-8359	26591
Step 3: Delete those within legal working age but still study at school during investigation period	-1164	25427
Step 4: Delete those irreplaceable missing values for demographic factors, educational attainment and health-related variables	-618	24809
Step 5: Delete the missing values for occupational related factors (working hours, work unit, occupation type) for those who are employed and report their regular wages in the urban labour market, and the extremely value in terms of real income	-1707	23102
Step 6: Delete all the self-employed individuals	-2175	20927

Note: China Health and Nutrition Survey is utilized.

difficulties in distinguishing rural migrants and urban people by appearance, self-employed individuals are less likely to suffer from customer discrimination in China; therefore, self-employed individuals are not considered when analysing urban labour market discrimination. Among the remaining observations, 2,175 individuals

reported a self-employment status and 90% of those are urban *hukou* individuals. After removing these self-employed observations, there are 20,927 observations remaining in the sample, where 9,678 observations have urban *hukou* status, 2,543 are rural-urban migrants, and 8,706 are rural non-migrants. The study in this chapter examines the *hukou* discrimination in the Chinese urban labour market, thus I mainly concentrate on urban people and rural migrants. There are 12,221 valid observations in this sample, which includes 9,678 urban observations and 2,543 rural-urban migrants. Table 2-15 also shows the process on how I obtain the sample for this chapter.

2.7.3 Selectivity in Oaxaca & Ransom decomposition

In terms of correcting the migration selection of the wage gap by the Oaxaca's decomposition method (including the Blinder-Oaxaca (1973) and Oaxaca & Ransom (1994) decomposition), the most straightforward approach is to deduct the selection effects from the overall wage differential and then apply the Oaxaca decomposition equation to this adjusted differential (Neuman & Oaxaca, 2004; Jann, 2008). This method, proposed by Heckman (1979), involves conducting a standard probit model that estimates the probability of a rural person migrating to the urban city to a set of determinants, and then applying the probit estimation to calculate the Mills ratio. The migration equation (Eq. 2-16) identifies the determinants that may influence a rural *hukou* individual's decision to migrate to an urban city. The wage function for rural-urban migrants is given by Equation (2-6), which has been introduced in the methodology section:

$$Migration_{i,r}^{*} = H_{i,r}'\gamma_r + \mu_{i,r} \tag{2-16}$$

$$Y_{it,r} = X_{it,r}'\beta_r + \varepsilon_{it,r} \tag{2-6}$$

where $Migration_{i,r}^{*}$ is a latent variable representing rural people's migration, $H_{i,r}'$ is a vector of variables affecting the migration. ε_{ij} and $\mu_{i,r}$ are identical and independently distributed (i.i.d.) error terms that follow a bivariate normal distribution (0, 0, $\sigma_{\varepsilon j}$, $\sigma_{\mu j}$, ρ_j). The probability of rural people's migration can be expressed as,

$$\mathrm{Pr}\,ob(Migration_{i,r}^{*} > 0) = \mathrm{Pr}\,ob(\mu_{i,r} > -H_{i,r}{}'\gamma_{r})$$

$$= \Phi(H_{i,r}{}'\gamma_{r}) \qquad (2\text{-}17)$$

where $\Phi(\cdot)$ is the standard normal cumulative density function (the variance of $\mu_{i,r}$ is normalized to 1). The dependent variable of the rural *hukou* individual's migration takes a value of 1 if rural people migrate to an urban city and 0 otherwise. Regular wages of rural migrants are observed for those where *Migration*$^{*}_{i, r}$>0, thus the expected earning equation shown in Eq. (2-6) of a rural migrant can be determined by,

$$E(Y_{i,r} \mid Migration_{i,r}^{*} > 0) = X_{i,r}{}'\beta_{r} + E(\varepsilon_{i,r} \mid \mu_{i,r} > -H_{i,r}{}'\gamma_{r})$$

$$= X_{i,r}{}'\beta_{r} + \theta_{r}\lambda_{i,r} \qquad (2\text{-}18)$$

where $\theta_{r} = \rho\sigma_{\varepsilon,r}$, $\lambda_{i,r} = \phi(H_{i,r}{}'\gamma_{r})/\Phi(H_{i,r}\gamma_{r})$, and $\phi(\cdot)$ is the standard normal density function. The estimating equation for rural migrants could be expressed as,

$$Y_{i,r} \mid Migration_{i,r}^{*} > 0 = X_{i,r}{}'\beta_{r} + \theta_{r}\lambda_{i,r} + error_{r} \qquad (2\text{-}19)$$

Equation (2-19) is estimated for rural migrants by the Heckman procedure. When regular income is decomposed in the presence of the selection issue, the wage differential based on the Oaxaca's decomposition is shown as,

$$\overline{Y}_{u} - \overline{Y}_{r} = (\overline{X}_{u}'\hat{\beta}_{u}) - (\overline{X}_{r}'\hat{\beta}_{r} + \hat{\theta}_{r}\hat{\lambda}_{r}) = (\overline{X}_{u}'\hat{\beta}_{u} - \overline{X}_{r}'\hat{\beta}_{r}) - \hat{\theta}_{r}\hat{\lambda}_{r} \qquad (2\text{-}20)$$

The most straightforward method mentioned by Neuman & Oaxaca (2004) and Jann (2008) to address the selection problem is to deduct the selection effects from the original income differential as shown in Eq. (2-21) below. The part on the left-hand side of this equation is the adjusted wage differential, and the part on the right-hand side can be decomposed by the same Oaxaca's decomposition formulas.

$$(\overline{Y}_{u} - \overline{Y}_{r}) - \underbrace{(-\hat{\theta}_{r}\hat{\lambda}_{r})}_{selectivity} = \overline{X}_{u}'\hat{\beta}_{u} - \overline{X}_{r}'\hat{\beta}_{r} \qquad (2\text{-}21)$$

In terms of choosing the suitable factors for selectivity correction in the Oaxaca decomposition, it must contain at least one variable that is in the probit selection model but not in the original earning function (Lee, 2012). As discussed previously,

rural-urban migrants are likely to be a group of people with larger ambition and better capability of dealing with pressure and recovering. In addition, according to the National Bureau of Statistics Survey, most of the rural-urban migrants are young adults. Regarding the factors influencing rural people's migration, Li (2003) conducts an analysis of push and pull factors for migration. The author finds that Chinese economic development, policy reform, family structure (i.e., number of children in the household), age, health status, frequency of contacting with neighbours, and unobserved motivation or ambition can influence rural people's aspirations for migration. The available information obtained from the CHNS includes the number of children and disease history. In the probit selection model for the Oaxaca & Ransom (1994) decomposition, I intend to take these two determinants into account because they satisfy the exclusion restrictions by not determining wages directly.

There is no available information in the CHNS of general self-reported health status (e.g., good health, standard health, or poor heath). However, the CHNS includes detailed health information and specific disease history; for example, "Has a doctor ever told you that you suffer from high blood pressure, diabetes, myocardial infarction?", "What treatment method did you use before?", "How much money did you spend on the medical service?", and so on. The dummy variable of chronic disease (blood high pressure, diabetes, and myocardial infarction) is taken into account in the selection model of this study. The dummy variable satisfies the exclusion restriction by not directly influencing earnings, as most people with chronic disease are elderly individuals who have already retired, and fewer young adults with chronic disease are influenced to participate in working. Even if young adults are diagnosed with chronic diseases, most of them are still working, as they have to earn money to support the family. Therefore, whether a person has a chronic disease has no direct impact on his regular wages in most cases. However, due to the fact that the available jobs for rural migrants are often limited, and most rural migrants are more likely to take part in the lower skilled, longer working time and physical exertion jobs, rural people without good health might be reluctant to migrate to and work in the urban city. Thus, I am concerned whether their decision for migrating to the urban city may be influenced by their status of health, which does not affect their earnings directly.

In addition, the number of children may positively influence migration as adults

bear the heavy economic burden of raising children and undertaking the educational expense (Li, 2003), which may push adults to migrate to the urban city to look for a job. In China, there is a very common situation where parents from both sides are more likely to help take care of child when the couple has a newborn baby. In addition, some of them also employ full-time nannies to take care of an infant aged below 12 months. Therefore, fewer young adults in China choose to stay at home and take care of the child. Most adults will return to work after their maternity leave and the child is more likely to be cared for by the grandparents during working hours. Consequently, earnings are less likely to be influenced by the number of children in China.

Table 2- 16 Probit Model for Rural *Hukou* Individuals' Migration

Dummy for Migration	Coef.	Robust SE.
Number of children	0.103***	0.021
Health status		
Dummy for chronic disease	0.142*	0.071
Socio-demographic factors		
Age	-0.047***	0.012
Age^2	0.0003*	0.0001
Male	0.210***	0.036
Married	-0.067	0.059
Education attainment (ref: no education)		
Compulsory school	0.252***	0.058
Senior high middle school	0.847***	0.068
University	1.714***	0.143
Year 1997	-0.295	0.197
Year 2000	-0.570**	0.073
Year 2004	-1.511***	0.065
Year 2006	-1.459***	0.065
Year 2009	-1.245***	0.062
Year 2011	-1.038***	0.064

Number of observations: 2543 rural migrants; 8706 rural *hukou* non-migrants

Note: (1) ***, **, * refer to 1%, 5% and 10% significant levels, respectively.

(2) China Health and Nutrition Survey is utilized.

Table 2-16 presents the results of the selectivity model for rural people's migration. The number of children in a household is found to have a positive correlation with the probability of migration for rural people, which is consistent with the suggestion from Li (2003). Specifically, the heavy economic burden to support a child pushes rural adults to work in an urban city to obtain greater economic benefits, which in turn leads to children left-behind and empty nesters in rural China being more widespread. It can

be seen from Table 2-16 that a rural person having been diagnosed with a chronic disease is positively related to the probability of migration at the 10% significance level, which is not line with Li (2003). The reason may be ascribed to the fact that these rural people who have a chronic disease consider that they can obtain better medical treatments in an urban city. Specifically, people who are diagnosed with a slight chronic disease (e.g., high blood pressure and diabetes) and take regular medication and treatment (e.g., hypotensive drugs, hypoglycemic agent, and control diet) are able to work normally. Moreover, few young adults in China with chronic diseases give up working, as they have to bear the financial burden of the whole family and medical expenses. Therefore, these rural people may consider that working in an urban city could bring more economic benefits so that they will be able to afford better medical treatment than before. In addition, the medical conditions in an urban city are also much better than that in the rural area; thus, they are likely to migrate for higher income in order to control their chronic disease more efficiently.

Regarding other control variables, results in Table 2-16 show that education attainment, gender, age, and time dummies have a significant relationship with the probability of migration. I find that education level positively influences rural individuals' aspiration for migration. It may be due to the fact that people with more educational attainment are more likely to have greater ambition. In terms of gender, males seem to be more likely to migrate to an urban city, which is in line with the statistics from National Bureau of the People's Republic of China. It describes that men account for the majority of all rural migrants, but the percentage of female migrants has continued to increase in recent years. In addition, rural-urban migrants are mainly energetic young people under the age of 30[9]. With age increasing, the willingness of rural people to migrate to an urban city declines gradually. Consequently, the results in Table 2-16 show the negative association between age and the probability of migration.

There is a negative correlation between time dummies and the probability of migration, which may be due to that the population structure in the rural area is gradually changing. Specifically, the number of children in households is decreasing due to the one child policy, and the aging problem in China is getting more serious. Consequently, the phenomena of empty nesters and children left-behind in the rural area

---

[9] National Bureau of Statistics of the People's Republic of China.

have become more prevalent during the recent decade. The available labour supply of young rural adults is declining so that the Chinese statistics surveys show that the growth rate of rural people migrating to an urban city is decreasing gradually. Based on the current rural population structure, the probability of migration for those remaining rural residents decreases inevitably, and the time dummies indicate the falling trend in the migrations.

<center>**Chapter Three**</center>

<center>**The Impact of *Hukou* Discrimination on Income Aspirations of**</center>

<center>**Rural-Urban Migrants**</center>

3.1Background and Introduction

As explained in the previous two chapters, all Chinese people are divided into urban *hukou* and rural *hukou* individuals, based on the household registration system established by The Public Security Organ Household Management Institutions. Chapter 2 has already investigated the effects of *hukou* discrimination in terms of income inequality and occupation segregation between urban *hukou* workers and rural-urban migrants in Chinese urban labour market. Most previous research has also analysed the impact of discrimination in these areas. However, during recent years, an increasing number of studies (e.g. Bigler et al., 2003; Alagaraja et al., 2016), have pointed out the psychological impact of social discrimination on the devalued group. Following these studies and Chapter 2, I intend to analyse whether *hukou* discrimination has other impacts on rural-urban migrants, and on society as a whole.

Alagaraja et al. (2016) have analysed the negative impact of racial discrimination from a psychological perspective, in particular the influence on the devalued group's aspirations. They state that the negative effect on individual aspirations is significant to social development because discrimination in the long term creates a sense of passivity and apathy. This passive acceptance of discrimination can create a cycle of silence around issues of power and privilege (Sims, 2010), which in turn lowers the devalued group's ambitions and aspirations regarding career and income. Many in the devalued group lose passion and motivation to work, which exacerbates the gap between the preference group and the devalued group. In addition, several individuals in the devalued group underestimate their own capabilities regarding career and income, although they have more educational attainment than many people in the preferred group. Therefore, Alagaraja et al. (2016) point out that there is an outflow of talent from

<center>73</center>

society due to discrimination, as many people do not make full use of their talents due to the context in which they live and work.

Rural-urban migrants in China are still discriminated against in the urban labour market, in spite of the increasing number of job opportunities for rural *hukou* workers. Many are engaged in low-skilled jobs ("dangerous, dirty and demeaning") with low incomes, such as manufacturing, construction, and service jobs[10]. Urban *hukou* workers still receive priority for the professional jobs, with higher incomes and better working conditions (Cai, 2011). Following Alagaraja et al. (2016), it is important to consider the aspirations of the people in the disadvantaged group, as their lower aspirations and ambitions will negatively influence their social productivity. In addition, the large number of rural-urban migrants in the urban labour market (about 20 million in 2010[11]) intensifies the necessity to care about rural-urban migrants' aspirations. If the aspiration or ambition of rural-urban migrants is influenced by *hukou* discrimination, their lower enthusiasm for work will negatively influence social development (Alagaraja et al. 2016).

Previous studies (Easterlin, 2001; Stutzer, 2004) have investigated the determinants of the absolute aspirations among American people and residents of Switzerland. Individuals with higher incomes and a higher social status tend to have larger ambitions and higher aspirations in terms of their income, career and social status. Thus, following these previous studies, urban *hukou* individuals with higher incomes and superior occupational conditions should have higher aspirations compared to rural *hukou* individuals in China. However, whether *hukou* discrimination between urban *hukou* individuals and rural-urban migrants contributes to the difference in aspirations is unknown. In this chapter, I intend to compare the income aspirations of urban *hukou* people and rural-urban migrants, and investigate the effect of *hukou* discrimination on their aspirations. In the next section, I will introduce the causal mechanism in this chapter, i.e. the causal relationship between aspirations and discrimination.

3.2 Causal Mechanism between Aspirations and Discrimination

---

[10] National Bureau of Statistics of the People's Republic of China.
[11] National Bureau of Statistics of the People's Republic of China.

The idea of this chapter was inspired by previous studies that investigate the impact of racial discrimination on career aspirations following social cognitive career theory. Social cognitive career theory (SCCT), which is derived from Bandura's (1986) general social cognitive theory, aims at understanding three interrelated aspects of career development: (1) how the career-relevant interests develop, (2) how occupational and educational selection are made, (3) how educational and occupational achievement can be obtained (Lent, Brown, & Hackett, 1994). It involves a series of factors in earlier career theories (e.g., abilities, values and environmental factors), and has also been found to have close relationship with individual's career development. When analysing the influence mechanisms of SCCT on the career development process, the main focus is on three social cognitive elements: self-efficacy belief, outcome expectations and goals.

Self-efficacy belief refers to individuals' personal self-evaluation of their abilities to perform activities and particular behaviour (Bandura, 1986). This judgement can be derived from personal previous performance achievement, previous experiences and emotional states. Self-efficacy belief not only affects one's choice of activities, it also influences one' effort expenditure, persistence, patterns of thinking, and attitude in the face of barriers (Lent, Brown, & Hackett, 1994, 2000). In addition, self-efficacy has a dynamic trait and is specific to particular performance domains. An individual's judgement of their self-efficacy will vary in different environment, academic and career domains (Lent, Brown, & Hackett, 1994). For example, one person may have confidence in his/her capabilities in making social connections but feel less confident about their capabilities in maths. Thus, people are more likely to be interested and have better performance in the domains which they report higher self-efficacy belief, in general.

Outcome expectation indicates one's beliefs about the potential consequences or outcomes of particular activities or behaviours (e.g., if I do this, what will happen?) (Lent, Brown, & Hackett, 1994). One's choice behaviour (e.g., the choice of activities he/she prefers to participate in, the effort expenditure and persistence in the activities) depends on both individual's judgement of self-efficacy and the probable outcome of participating in the activity. To be specific, a person is more likely to participate in the activities which lead to a positive outcome or have greater value utility (e.g., reward,

attractive welfare benefits). However, Bandura (1986) pointed out that career behaviour is determined by both self-efficacy beliefs and outcome expectations, with self-efficacy playing a more important role in influencing the behaviour. Because there are many examples indicating that many people would like to avoid such activities which they are not confident to do even if they can obtain more valuable outcomes.

Personal goals refer to one's determination to participate in a particular activity or to achieve a certain level of performance. Personal goals can help people to guide and regulate the behaviour and persist in their target when suffering from setbacks. SCCT points out that these three variables (self-efficacy, outcome expectation and personal goals) are highly interrelated with each other. To be specific, people are more likely to set practical goals which are consistent with the assessment of own abilities (self-efficacy) and expected outcomes from particular activities. On the contrary, whether the goals can be achieved or not will inversely affect and alter one's judgement of capabilities and outcome expectations.

SCCT highlights that various background factors such as social demographic factors, social economic status and racial background may directly influence one's career development (Lent, Brown, & Hackett, 1994). Such contextual factors influence career development and performance by affecting the motivational role of self-efficacy, outcome expectations and goal. To be specific, for two individuals with similar capabilities, the one having higher self-efficacy belief is more likely to expect the positive outcomes for particular activities, which in turn establish higher performance goals. Therefore, controlling for individuals' capability, these people with higher self-efficacy are more likely to organize their skills effectively, and insist on the goals even if they are confronted with obstacles. As a result, they can obtain more success than people with lower self-efficacy and outcome expectations (Bandura, 1986, 1989). On the contrary, a person may be restricted by his/her background factor such as gender or racial-ethnic background and underestimates their talents and capabilities. Thereby, they are more likely to set lower performance goals and less expectations of positive outcomes, compared to the people who have similar capabilities and are more confident. In the face of setbacks, these people with lower self-efficacy may give up more quickly and attain less success consequently. Such contextual factors may lower a person's assessment of their capability for work and restrict their aspirations. Therefore, the

SCCT theory is increasingly applied to the study of career behaviour in a number of countries and cultural contexts. Specifically, there are a number of studies (e.g., Bigler et al., 2003; Brown & Segrist, 2016) exploring how racial-ethnic background influences people's career aspirations, based on social cognitive career theory.

As discussed in Chapter 2 urban *hukou* individuals usually obtain preferential treatment and rural-urban migrants are discriminated against in urban cities. Following the principle of SCCT that highlights the effect of racial background on one's career development and aspiration by affecting self-efficacy beliefs, outcome expectation and goals (Lent, Brown, & Hackett, 1994), this chapter investigates whether rural-urban migrants' judgement of their capabilities (self-efficacy belief) is influenced by the impact of *hukou* discrimination as well. If so, rural-urban migrants' expectation of outcomes and their career aspirations will also be negatively affected. Based on the close relationship between careers and income, income aspirations are utilized as an available indicator to quantify career aspirations in this chapter. As mentioned before, SCCT focuses on how self-efficacy, outcome expectations and goal mechanism interrelate with other personal characteristics (e.g., gender, education, social economic status, racial background) to shape career assessment, development and expectations. In the next section, I will review previous studies based on socio-demographic factors, social economic status and discrimination, respectively.

3.3 Literature Review

Based on the discussion in section 3.2, I will review previous studies investigating the determinants of income aspirations from these three aspects (socio-demographic factors, social economic status and discrimination). It aims to understand which factors contribute to income aspirations more generally and to assess whether the differential aspirations are due to differences in these determinants, especially the discrimination.

3.3.1 Personal Socio-demographic Factors

In terms of personal factors, an individuals' present income is a common factor to be analysed (Fuchs & Landsberger, 1973; Centers & Cantril, 1946; Portes et al., 1978; Stutzer, 2004; McBride, 2010; Dalton et al., 2016). Income aspirations are generally

found to have a positive relationship to individuals' current income (Fuchs & Landsberger, 1973 and Stutzer, 2004). Fuchs & Landsberger (1973) examined the income aspirations of the elected presidents of 123 blue-collar local unions in Chile. Their results showed that current income was the most important determinant of income aspirations and aspiration income increases with increases in present income. This result is supported by the findings of other studies, e.g. Centers & Cantril (1946) and McBride (2010). Previous Chinese research also points out the positive relationship between personal current income (actual income) and income aspirations. Knight & Gunatilaka (2012b) compared subjective well-being between rural-urban migrants and rural residents, and found that rural-urban migrants are not happier than rural residents, although the rural-urban migrants have a higher income. They ascribe this to the difference in income aspirations between rural residents and rural-urban migrants. When empirically examining the determinants of absolute income aspirations, it was found that people's income aspirations depend positively on their actual income, which is consistent with previous research. Therefore, although rural-urban migrants earn higher wages than rural residents, the migrants experience higher income aspirations, which lower rural-urban migrants' happiness.

In addition to present income, some other socio-demographic factors also play an important role in determining aspirations. Centers & Cantril (1946) utilized a survey by the Office of Public Opinion Research in Princeton to investigate income satisfaction and income aspirations. All respondents were given a card indicating different weekly income classes. Based on the card, respondents answered two questions related to their average weekly wages and their income aspirations. Apart from present income, they also point out that occupation, age, gender and educational attainment have impacts on income aspirations. The impact of these factors on aspirations may vary, depending on background and the period in which the survey was carried out. Therefore, due to its age, I will not review this study in detail.

Stutzer (2004) investigated the effect of income aspirations on individuals' happiness in Switzerland, and also studied how income aspirations are formed. He found that absolute aspirations increase with people's income levels, which is consistent with the findings of previous research. He also investigated the effect of other socio-demographic factors, such as financial situation or social status, the size and

composition of the household, age, education and gender on absolute income aspirations. Financial situation and social class have a positive effect on individuals' income aspirations. In terms of the size and composition of the household, the number of children in the household has a positive relation with adults' income aspirations. Income aspirations are inversely U-shaped with age and increase with education level. No sizable differences in income aspirations levels exist between men and women.

Portes et al. (1978) analysed the determinants of occupational and income aspirations among Mexican and Cuban immigrants arriving in the United States in 1973. During 1970, net immigration accounted for nearly one-fifth of total US population growth. Numerically and socially, immigrants form an important component of American life (Portes et al., 1978), and one of the major characteristics that immigrants bring is aspirations. The authors investigate aspirations from two perspectives, namely rationality and motivational needs. From the rationality perspective, they suggest that aspirations can be regarded as the rational products of individuals' previous experiences and their assessment of their own abilities. Based on individuals' education levels, previous occupational status and past income attainment, they set goals in accordance with their reasonable aspirations. The higher the educational, occupational and income attainment is, the higher the aspirations will be. On the other hand, aspirations are formed based on attitudes and more basic personality needs. Individuals do not set aspirational goals only based on their previous achievements, but also based on their inherent preference of what they want to do in the future. Regarding their personal attitudes, Portes et al. (1978) argue that respondents' family background, such as their father's educational attainment, can positively influence respondents' attitudes towards career aspirations, which in turn may positively affect income aspirations. Therefore, the more modern and achievement-oriented immigrants will have higher aspiration levels.

Overall, the existing research reviewed in this section demonstrates how personal factors determine individual aspirations. These previous studies demonstrate the basic determinants of individual income aspirations, which include present income and social demographic factors (educational attainment, occupational status, family size, family background, financial situation and age). Regarding these factors, urban residents are more likely to be engaged in professional jobs with higher incomes and superior

working conditions. Rural *hukou* individuals are more likely to have lower educational attainment, as found in Chapter 1. Therefore, these factors that have been proposed by previous research are likely to play an important role in differential aspirations, as this can result from the significant difference in present income, occupational status, financial status and educational attainment between urban and rural *hukou* people.

3.3.2 Social Comparisons

In addition to the personal factors reviewed in the previous sub-section, it is necessary to consider whether income aspirations are influenced by impersonal factors such as different living environments. Urban and rural *hukou* people have unequal incomes and segregated occupations, thus having differences in working conditions and living environments as well. In terms of the effect of environment on aspirations, most previous research has focused on human factors. In other words, they primarily emphasize how the relationship with neighbours, friends and colleagues influence individual aspirations (Stutzer, 2004; Ray, 2006; McBride, 2010 and Dalton et al., 2016).

In addition to individual characteristics, Stutzer (2004) also studied the effect of social comparisons on people's absolute aspirations. The average income in the community where respondents live was considered to be a potential determinant of respondents' income aspirations. A higher income of fellow residents has a positive influence on individuals' income aspirations. This may be because individuals in a community with higher income levels have higher price levels for individual services, which in turn leads to higher income aspirations. Furthermore, Stutzer (2004) also found that the effect of average community income on individual income aspirations is larger for people who interact with other members in the community frequently.

Similar to Stutzer (2004), McBride (2010) also investigates the relationship between money, happiness and aspirations. In addition to exploring the effect of aspiration on happiness, he examines how absolute aspirations form and adapt over time. Firstly, the author finds that individual aspiration levels rely positively on past income, which is consistent with most previous research. Secondly, he examines how social comparisons influence individuals' aspiration formation, finding that respondents prefer

to compare themselves to those who are most similar to themselves when making social comparisons, and their aspirations are positively influenced by the outcomes of others in their comparison group. To be specific, individuals prefer to perform better relative to others in their comparison group. Moreover, aspiration levels will adapt to the environment and comparison group variations, therefore an income increase in the comparison group stimulates increasing aspirations.

Following Stutzer (2004), Knight & Gunatilaka (2010a) have also analysed how income aspirations affect subjective well-being and the determinants of absolute income aspirations in rural China. Similar to Stutzer (2004), they not only explore the effect of current income and social demographic factors (age, education, gender, marital status, and composition of household) on income aspirations, but also take social comparisons into account. Firstly, they show the effect of comparator income on respondents' income aspirations. More than 65% of respondents reported that their main reference group was within the village. The comparator income of the reference group has a positive impact on the respondents' income aspirations. A higher income of the main reference group raises income aspirations substantially (Knight & Gunatilaka, 2010a). Secondly, respondents assess their own income and answer whether their income is much above, above, at, below or much below the average income of their village, which is regarded as the reference group. They found that, in comparison to individuals at the average income level, those respondents whose incomes were far below the village average had significantly higher aspirations. However, this result was only found for those with materialistic attitudes who firmly believed money was very important. For those without such attitudes, this finding does not hold.

Knight & Gunatilaka (2010b, 2012b) divided their rural sample into rural households and rural-urban migrants in China. When utilizing income aspirations of rural-urban migrants to explain the reason for rural-urban migrants with higher income in urban areas having lower happiness levels than rural residents, Knight & Gunatilaka (2010b, 2012b) also analyse the determinants of income aspirations. They ascribe the reason for rural-urban migrants' failure to foresee how they are likely to have to adapt their aspirations in their new situation. Before migration, rural-urban migrants and rural households may hold similar aspirations regarding income and living standards, but aspirations are likely to rise after making the decision to migrate. Although rural-urban

migrants earn higher wages and can access superior individual services, their comparison group has changed as well. Migrants tend to compare their own situations with others living in their new surroundings, and the higher standard of those living in their new environment will lead them to develop higher income aspirations.

In this section, I have reviewed previous research that investigates the effects of different environments or comparison groups on aspirations. The next section focuses on the extent to which discrimination restricts the aspirations of the disadvantaged group.

3.3.3 Discrimination Impacts

In this sub-section, I intend to review studies which analyse the effect of discrimination on aspirations. Several previous studies have analysed the career aspirations of African American individuals who had encountered racism in the United States, and have examined whether racial discrimination or oppression can restrict this disadvantaged groups' career aspirations. In general, the disadvantaged group is discriminated against in terms of both career and earnings, which is similar to how rural-urban migrants encounter occupation segregation and income inequality in China. In the United States, a complex history of occupational inequality has existed between European Americans and African Americans. African Americans are overrepresented in lower occupational status and lower paid jobs. Although the representation of African Americans in some professional fields has improved, African Americans are still discriminated against and remain under-represented in careers related to high-skilled occupations (Alagaraja et al., 2016).

In order to understand the career development of African Americans, some researchers have focused on career aspirations. They have discussed whether being discriminated against or lack of representation in higher status occupations in the long term can exacerbate African American's generally low aspirations (Evans & Herr, 1994; Jackson & Nutini, 2002; Bigler et al., 2003; Tovar & Murray et al., 2012; Brown & Segrist, 2016 and Alagaraja et al., 2016). Evans & Herr (1994), Jackson & Nutini (2002), Bigler et al. (2003) and Tovar & Murray et al. (2012) focused on a sample of African American children and college students to investigate whether racial

discrimination and oppression constrained their career aspirations. Evans & Herr (1994) explored the possible impacts of perceptions of discrimination on the traditional career aspirations of 111 African American college student volunteers and proposed that discrimination can predict career aspirations. Their results do not support the hypothesis, which they ascribe to the fact that African Americans have learned to adapt to the disadvantaged status resulting from long term racial discrimination.

Although Jackson & Nutini (2002), Bigler et al. (2003) and Tovar & Murray et al. (2012) focus on a sample of children and students, their results are not consistent with Evans & Herr (1994), suggesting that racial discrimination can influence career choices and lower aspirations. Jackson & Nutini (2002) conducted an analysis with students in a low-income, culturally diverse and inner-city public school and examined the possible influence of discrimination on career learning. They posit that individuals who are sensitive to discrimination encounter potentially unobserved barriers in their career assessment. In terms of career understanding, the awareness of systemic career barriers may be different for culturally diverse individuals. Moreover, this awareness can partly determine their direct and indirect learning experiences. Behaviours and feelings that racially discriminated individuals learn from the inequality and discrimination they experience can restrict the utilization of existing resources and constrain their exploration of areas that might be beneficial to their career development (Jackson & Nutini, 2002). In the long term, the disadvantaged group cannot conduct accurate career assessments, which lowers their career aspirations.

Both Bigler et al. (2003) and Tovar & Murray et al. (2012) have examined whether racism-related stress influences judgments on occupational planning. It was found that most African American people are more likely to select jobs in which African Americans are traditionally engaged. On one hand, the lower occupational aspirations have been ascribed to their assessment of potential limitations afforded by their socio-economic status. On the other hand, they have also been attributed to the fact that African Americans have an awareness that they are more likely to face prejudicial attitudes in traditional occupational settings, which is consistent with Jackson & Nutini (2002). The results of both Bigler et al. (2003) and Tovar & Murray et al. (2012) show that the coefficient for racism stress was negative on African American career goals, suggesting an inverse relationship. Racism plays an important role in restricting African

American respondents' occupational goals. Their career aspirations decrease as their racism-related stress increases.

In addition, both Bigler et al. (2003) and Tovar & Murray et al. (2012) investigated the factors that moderate African Americans' occupational judgments and alleviate the career barriers they encounter. Tovar & Murray et al. (2012) report racial or ethnic identity as a psychological buffer that protects African Americans from racism regarding career aspirations, moderating the negative effects of racial discrimination on respondents' self-assessments, which suggest that African Americans will report higher levels of career aspirations when they strongly identify with their racial identity. On the other hand, Bigler et al. (2003) suggested that children's socioeconomic status can moderate their occupational judgments, especially during adolescence. African-American children with a higher socioeconomic status have more interest in the jobs that European Americans, or both African American and European Americans, undertake than they do in jobs viewed as involving only African Americans (Bigler et al., 2003). The relationship between socioeconomic status and career aspirations suggests that economic support is significant in African American children's career decisions. To be specific, lower socioeconomic status children have the awareness that they lack sufficient financial resources to attain a high educational level, which in turn restricts their pursuit of high-skilled jobs.

In the United States, previous research has focused mostly on the career aspirations and development of children and adolescents. Among the sample of African-American adults from 18 to 62 years of age, Brown & Segrist (2016) have investigated whether African-American career aspirations are influenced by internalized racism, based on Social Cognitive Career Theory (SCCT) (Lent et al., 1994). Brown & Segrist (2016) argue that racial discrimination or racial oppression creates many barriers to African American's career development by lowering respondents' self-evaluation for capabilities. Therefore, it is necessary to take the experience of racial discrimination into account when exploring the career aspirations of this disadvantaged group. An inverse relationship between internalized racial discrimination and career aspirations was found. The most serious consequence of racial oppression and discrimination may be the internalisation of the idea of inferiority. African Americans are more likely to feel that they are not as worthy as others; therefore, racial oppression can constrain

African-American individuals' motivations and aspirations.

Although Dalton et al. (2016) do not indicate whether discrimination exists between the poor and the rich, the difference in access to relevant information is apparent. They construct a theoretical model and analyse the relationship between poverty and "aspiration failure", defined as the failure to aspire to one's own potential. They divide all individuals into rich and poor persons and assume that both the rich and the poor share the same preferences and behaviours in setting career aspirations. They found that poverty can cause behavioural bias or internal constraints, such as a lack of foresight, a lack of willpower and a lack of aspirations. In addition, poverty can also impose external limitations on the poor, which intensify the negative impacts of the behavioural bias in setting career aspirations. The external limitations are not only financial resource barriers, but also include other constraints that make it more difficult for the poor to achieve a given outcome, such as less influential contacts or less access to relevant information (Dalton et al., 2016). Such limitations make the poor more vulnerable to aspiration failure; thus, these individuals are more likely to have a relatively lower level of aspirations. Furthermore, empirical evidence suggests that individual aspirations and efforts are jointly determined. The lower level of aspirations can lead to a low level of effort relative to the optimal level, which in turn strengthens the lower aspirations. In urban China, rural-urban migrants may also suffer from differential access to relevant opportunities and beneficial information, due to *hukou* discrimination. Therefore, following Dalton et al. (2016), it is reasonable to suspect that the different treatments deriving from discrimination can restrict rural-urban migrants' aspirations.

To sum up, most previous studies have examined how current income, socio-demographic factors and social comparisons influence individuals' income aspirations. Following previous research, a study of the Chinese context by Knight & Gunatilaka (2012b) also investigates the determinants of income aspirations for individuals in rural China. They examined the influence of present income, socio-demographic factors (age, education, gender, marital status, household size) on the formation of income aspirations, and tested how the comparator income of the reference group influenced respondents' income aspirations. In addition, Knight & Gunatilaka (2010a, 2012b) compared the income aspirations between rural-urban

migrants and rural residents and explained the difference in income aspirations between these two groups. They also reported that aspirations depend on income compared with the relevant reference group. Rural-urban migrants will adapt to their new environments and compare their own situations with the new members in their communities. Therefore, rural-urban migrants have higher income aspirations in comparison to rural residents. Overall, although previous studies of the Chinese context have identified the determinants of income aspirations, there are limitations and gaps to be filled, as explained below.

Firstly, there is a limited amount of Chinese research investigating the determinants of income aspirations. In particular, previous research by Knight & Gunatilaka (2010a, 2012b) only analyses the determinants of income aspirations for rural *hukou* individuals: no research compares the income aspirations between urban *hukou* individuals and rural-urban migrants. As discussed in Chapter 2, income inequality and occupational segregation between urban people and rural-urban migrants are not only attributed to educational attainment, working experience and family background, but also to *hukou* discrimination. Following previous research reviewed in Section 3.2.3, African-American people, who are also overrepresented in lower occupational status with lower incomes, suffer from barriers to setting their career aspirations. Their aspirations are restricted by racial discrimination and/or racial oppression in the United States. However, no previous Chinese research has examined the impact of *hukou* discrimination on rural-urban migrants' aspirations in China. Thus, following these studies which analyse the impacts of discrimination on aspirations, I intend to utilize the China General Social Survey (CGSS 2010) to obtain specific information on respondents' income aspirations and compare income aspirations between urban *hukou* individuals and rural-urban migrants. The purpose of this chapter is to investigate whether *hukou* discrimination constrains rural-urban migrants' income aspirations and whether any differences in income aspirations between urban *hukou* individuals and rural-urban migrants is attributable to *hukou* discrimination.

3.4 Methodology

As discussed in Chapter 2, rural-urban migrants suffer from discrimination in urban labour market. In order to investigate whether the *hukou* discrimination affects

rural-urban migrants' income aspirations, all individuals are divided into two groups as Chapter 2: urban *hukou* individuals and rural-urban migrants. Propensity score matching (PSM) is used to estimate the effect of *hukou* discrimination on aspirations, by matching urban *hukou* individuals and rural-urban migrants with similar characteristics and comparing their reported income aspirations. This section discusses the evaluation framework, conditions and procedures of PSM implementation.

Propensity score matching (PSM) is a technique to estimate causal treatment effects (Rosenbaum & Rubin, 1983). It aims to estimate the effect of a treatment or policy program by accounting for the covariates that predict the reception of the treatment. PSM applies to situations where there are two groups, a group of treated individuals (the treatment group) and a group of untreated individuals (the control group). An estimate of the treatment effect is obtained from the difference between potential outcomes of individuals under treatment conditions (treatment outcome) and under control conditions (potential outcomes). Assuming a case of binary treatment, $D_i=1$ refers to individual *i* receiving the treatment, and $D_i=0$ otherwise. Potential outcomes for each individual *I* are denoted $Y_i (D_i)$, where $i=1, 2,…, N$. Therefore, the effect of treatment (δ_i) can be defined as the difference between the potential outcomes $Y_i (1)$ and $Y_i (0)$.

$$\delta_i = Y_i(1) - Y_i(0) \qquad (3\text{-}1)$$

However, the problem arises that it is impossible to observe both outcomes for the same individuals at the same time. The unobserved outcome is called the counterfactual outcome (Rosenbaum & Rubin, 1983). For example, if I intend to analyse the impact of having a Master's degree on a person's income, it would be impossible for a Master's student to observe their income without a Master's degree. Therefore, the income without a Master's degree for those Master's students is the counterfactual outcome. As a result, researchers have to concentrate on the average treatment effect (ATE) (Caliendo & Kopeinig, 2008):

$$\delta_{ATE}=E(\delta_i) = E[Y_i(1) - Y_i(0)] \qquad (3\text{-}2)$$

The average treatment effect (ATE) is defined as the difference between the expected outcome with and without treatment. Heckman (2008) has pointed out that the

estimation of the average treatment effect (ATE) includes the effect on individuals who will never be treated. Therefore, an alternative parameter ATT (the average treatment effect on the treated) can be utilized instead. It focuses directly on individuals who actually participate in the treatment, and is defined as the difference between the expected outcomes of those individuals with and without treatment (Caliendo & Kopeinig, 2008).

$$\delta_{ATT} = E(\delta_i \mid D = 1) = E[Y_i(1) \mid D = 1] - E[Y_i(0) \mid D = 1] \tag{3-3}$$

Since the mean value of the untreated outcome for those being treated, $E[Y_i(0) \mid D=1]$, is not observed, researchers have to utilize the mean outcomes for the control individuals, $E[Y_i \mid D=0]$, as a substitute to estimating ATT. However, using the mean outcome of control individuals in non-experimental studies is not a good idea, as participants and non-participants differ in terms of their characteristics, which in turn may cause selection bias (Rosenbaum & Rubin, 1983; Caliendo & Kopeinig, 2008).

$$
\begin{aligned}
& E[Y_i(1) \mid D = 1] - E[Y_i(0) \mid D = 0] \\
&= \{E[Y_i(1) \mid D = 1] - E[Y_i(0) \mid D = 1]\} + \{E[Y_i(0) \mid D = 1] - E[Y_i(0) \mid D = 0]\} \\
&= \delta_{ATT} + \{E[Y_i(0) \mid D = 1] - E[Y_i(0) \mid D = 0]\} \\
&= \delta_{ATT} + selectionBias
\end{aligned}
\tag{3-4}
$$

The matching method is used to solve selection bias in non-experimental studies. The basic idea of matching is to find individuals in the non-participant group who are similar to their participant counterparts in terms of their observable characteristics. In non-experimental studies, some conditions (unconfoundedness and common support) need to be met to solve the selection issue in equation (3-4) (Blundell et al., 2005). The conditions are:

Condition 1. Unconfoundedness: $Y_i(0), Y_i(1) \perp D \mid X$

where " $\perp$ "refers to independence, indicates that potential outcomes do not rely on the treatment assignment, given a set of observable covariates, *X*. In other words, there is a need for all variables influencing treatment participation and potential outcomes to be observed.

Condition2. Common support: $0 < P(D = 1 \mid X) < 1$

Common support indicates that all individuals have a positive probability of being participants and non-participants (Heckman et al., 1997). This is also called the overlap condition, which can be observed visually by the density distribution of the propensity scores in both groups. The propensity score refers to the probability of participating in treatment for each individual. There is a potential match in the control group for every individual in the treated group (Bryson et al., 2002).

In terms of unconfoundedness, if *X* simply includes one observable characteristic (*X=x*), the average treatment for the treated for a particular value of *X* can be written as follows:

$$
\begin{aligned}
\delta_{ATT}(x) &= E[Y(1) - Y(0) \mid D = 1, X = x] \\
&= E[Y(1) \mid D = 1, X = x] - E[Y(0) \mid D = 1, X = x] \\
&= E[Y(1) \mid D = 1, X = x] - E[Y(0) \mid D = 0, X = x]
\end{aligned}
\tag{3-5}
$$

Here, the outcome of control group individuals, *E[Y(0)| D=0, X=x]*, can be utilized to replace the unobservable outcome of treated people, *E[Y(0)| D=1, X=x]*, with the same characteristics. If *X* contains multiple covariates, it is difficult to make a judgment regarding how close the control group is to the treated group. Therefore, in order to cope with the dimensionality problem, Rosenbaum & Rubin (1983) proposed a balancing score, *b(X)*, that is a function of relevant observed covariates, *X*. They proved that matching based on the balancing score will be valid when based on the relevant covariates. Moreover, if the outcomes are independent of treatment, conditional on covariates *X*, the outcomes are independent, conditional on the balancing score, *b(X)* (Caliendo & Kopeinig, 2008). One possible balancing score is the propensity score that is the probability for individuals to participate in the treatment, given the observed covariates, *X*. The basic idea of propensity score matching (PSM) is to find in a large control group of non-participating individuals who are similar to the participants from the treatment group in terms of the probability of being treated. When researchers find participants and non-participants with similar propensity scores, the differential outcomes between these two groups can be estimated, in order to explain the effect of the treatment on the treated group. The propensity score can be written thus:

$$P(D = 1 \mid X) = P(X) \tag{3-6}$$

The propensity score the matching (PSM) estimator for the average treatment for the treated (ATT), can be written as:

$$\delta_{ATT}^{PSM} = E_{P(X) \mid D=1}\{E[Y(1) \mid D = 1, P(X)] - E[Y(0) \mid D = 0, P(X)]\} \tag{3-7}$$

This gives the difference of expected outcomes, weighted by the propensity score distribution of participants (Caliendo & Kopeinig, 2008). A key step to implementing PSM is to estimate the propensity score and find applicable nonparticipants in the control group with similar propensity scores to the participants. Below, I discuss the procedures of implementation of propensity score matching (PSM).

Step 1. Estimating the propensity score

In general, for a binary treatment, researchers utilize a probit or logit model to estimate the probability of participants receiving the treatment. The model can be written as

$$y_i = \begin{cases} 1 & if \quad y_i^* > 0 \\ 0 & otherwise \end{cases} \tag{3-8}$$

The latent variable y^* is specified as

$$y_i^* = X_i \beta + \varepsilon_i \tag{3-9}$$

where X contains factors, which influence the probability of a person' *hukou* status. The dummy variable *hukou* status (y_i) divides all the observations into the treatment group (rural-urban migrants) and control group (urban *hukou* individuals).

$$y_i = \begin{cases} 0 & urban \quad hukou \\ 1 & rural-urban \quad migrants \end{cases} \tag{3-10}$$

Assuming that the error term, ε_i, follows a logistic distribution, the propensity score, which is the probability for an individual to participate in the treatment given his observed covariates, X, is given by

$$prob(y = 1) = \frac{e^{\beta'X}}{1 + e^{\beta'X}} = \Lambda(\beta'X) \tag{3-11}$$

All the individuals will obtain a specific propensity score, which is constrained to lie between 0 and 1. Based on these participants' propensity scores, researchers can choose different matching methods to find the control group individuals that are most similar to the treated individuals in terms of their propensity score.

Step 2. Choosing a matching algorithm

The matching algorithms search for potential matches in the control group based on the propensity score, which has been computed in the first step. The traditional and most popular matching estimator is nearest-neighbour matching (Rosenbaum & Rubin, 1983; Dehejia & Wahba, 1999; Caliendo & Kopeinig, 2008). The basic idea of nearest-neighbour matching is to find the individual in the control group who is closest in terms of their propensity score for each participant in the treatment group. Nearest-neighbour matching can be done either "with replacement" or "without replacement". In nearest-neighbour (NN) matching with replacement, individuals in the control group can be utilized as a potential match more than once, whilst a control group individual can be used only once in the case of NN without replacement. Rosenbaum & Rubin (1983) and Caliendo & Kopeinig (2008) suggest that the quality of matching with replacement will be higher and the bias will decrease (Rosenbaum & Rubin, 1983; Caliendo & Kopeinig, 2008). In addition, matching without replacement will reduce successful matched samples, as each potential partner comparison group can be used only once. The second problem of matching without replacement is that it relies on the order in which observations are matched (Smith & Todd, 2005). Following previous research, this study therefore uses nearest-neighbour matching with replacement.

Step 3. Checking common support or overlap

Successful implementation of matching requires that the condition of common support can be satisfied, because Heckman et al. (1997) point out a bias will be caused by the violation of this condition. In other words, comparing the incomparable untreated persons with the treatment group will lead to bias in the estimates of ATT. Dehejia & Wahba (1999) also point out that only individuals from the comparison group who are matched and comparable to the people in the treatment group should be utilized in the

analysis, in order to calculate the average treatment of the effect on the treated (ATT). Therefore, checking the overlap between the treatment and control group is an important step before estimating the ATT.

Different methods have been proposed to check the region of common support between the treatment and the control group[12]. Lechner (2008), Bryson et al. (2002) and Caliendo & Kopeinig (2008) suggest that the most straightforward way is a visual inspection of the density distribution of the propensity score for the treatment and comparison group. Therefore, in this study, it is essential to check the overlap between urban *hukou* individuals and rural-urban migrants before estimating ATT. In other words, there is a need to check that a subset of the urban *hukou* individuals in the control group is comparable to the rural-urban migrants in the treatment group.

Step 4. Estimating average treatment for the treated (ATT)

Once the treated individuals have been matched to control group individuals based on propensity score, the difference between the actual and the counterfactual income aspirations[13] for the rural-urban migrants can be estimated by

$$\delta_{ATT}^{PSM} = E_{P(X)|rural=1}\{E[(Y(1)\,|\,rural=1,P(X)] - [E(Y(0)\,|\,rural=0,P(X)]\} \qquad (3\text{-}12)$$

Step 5. Sensitivity analysis

Finally, it is necessary to check the robustness of the estimated results after obtaining an estimate of the average treatment for the treated (ATT). Dehejia & Wahba (1999), Lechner (2008), Caliendo & Kopeinig (2008) point out that failure of the common support condition can be a significant source of estimation bias. This is because the individuals outside the region of common support are disregarded from the

---

[12] See Rosenbaum & Rubin (1983), Caliendo & Kopeinig (2008), Bryson *et al.* (2002) for more details of theoretical discussion.

[13] For those rural-urban migrants working and living in urban cities, fewer of them have opportunities to obtain urban *hukou* status unless they participate in further professional or technical training or receive further educational attainment. In this case, such rural-urban migrants have a higher probability to obtain the opportunities for promotion and better jobs so that they are likely to get urban *hukou* status. Thus, for rural migrants who still hold rural *hukou* status at the present stage, the counterfactual income aspirations refer to their income aspiration if they had urban *hukou* status. The counterfactual income aspirations for these rural migrants is unobserved, thus I utilize urban *hukou* individuals with similar endowments to replace these counterfactual values.

sample and the estimation only focuses on the sub-sample within common support. Nevertheless, disregarding those individuals outside the common support also disregards useful information, especially if treatment effects are heterogeneous (Caliendo & Kopeinig, 2008). The most straightforward and simple approach to check the robustness of the estimated average treatment effect is to change the range of common support in order to test whether the results are sensitive to the variation of common support (Lechner, 2008). In addition, some researchers use different matching algorithms to test the robustness of ATT because different matching algorithms look for potential partners based on different principles (Zhao, 2004; Heckman, 2008 and Johar, 2009). To be specific, similar to one-to-one nearest-neighbour matching in this chapter, one-to-five matching means that five individuals from the control group are picked as matching partners for one treated individual with the closest propensity score. Radius matching means choosing all the individuals from the control group as matching partners for one treated individual whose propensity scores lie within a pre-specified range. Kernel matching and local linear matching are non-parametric matching estimators which utilize weighted averages of approximately all control individuals, depending on different kernel functions and bandwidth parameters (Caliendo & Kopeinig, 2008). All samples in both treated and control group are matched, depending on the estimated outcome. In this chapter, I will compare the results from these different matching approaches. For the multiple nearest-neighbour matching, one-to-five nearest-neighbour matching is utilized. The radius of radius matching is 0.01. The bandwidth of kernel and local linear matching is set to the default value (0.06)[14].

3.5 Data and Choice of Variables

The data utilized in this chapter comes from the China General Social Survey (CGSS) 2010, implemented by the Department of Sociology of Renmin University of China and the Survey Research Centre of Hong Kong University of Science and Technology. The CGSS, which started in 2003, is the first continuous national social survey project in China. Respondents in households and communities are randomly selected. By the end of 2008, the CGSS team had carried out five data collection rounds and finished the first section of the entire project. The second stage started in 2010 and

---

[14]See Aakvik (2001), Rosenbaum (2002), Imbens (2004), Caliendo et al. (2008) for a discussion of sensitivity analysis in detail.

the plan is to establish an investigation on a two-year basis, with 5 rounds of data collection to be carried out until 2019. The CGSS 2010, which is utilized in this chapter, is the latest available version, carried out in 2010 and published in 2014.

The advantage of CGSS is that I can obtain information of individuals' income aspirations directly. Regarding the measurement of income aspirations, Stutzer (2004) utilized income evaluation measures from two aspects that were individuals' adaption provided by their consumption habits and comparison with relevant and similar others. Through the comparison with respect to consumption and income status, respondents were required to evaluate their income based on their job and family. Following Stutzer (2004), the information I can obtain from CGSS 2010 is related to social comparison. Income aspirations can be captured from two questions that are asked one after the other. Firstly, all respondents are asked "Are you satisfied with your current yearly income, based on your educational attainment, job, working experience, and other social demographic factors?". Secondly, they are asked "If not satisfied, what do you think you would have earned based on your educational attainment, job, working experience, and other social demographic factors?". I delete those respondents who do not reply to the both questions (4%). Among these respondents, about 2% have never worked before. The other 2% respondents refuse to answer questions. For other respondents, I obtain a measure of income aspirations mainly from the second question. About 74% respondents give specific answers to the second question which is used to measure their income aspirations. The remaining 22% respondents who do not answer the second question, are decomposed to that approximately 8% respondents reply that they are satisfied or relatively satisfied with their current income and 14% respondents report that they feel unhappy with their current income and do not provide an answer to the second question. In terms of the 8% respondents, I replace the missing value of income aspirations with their actual income. However, the rest 14% respondents are dropped from the sample and that how their basic social demographic factors differ from the remaining sample is shown in Appendix Table 3-11. As shown in Table 3-11, there is no huge difference in these social demographic factors between the remaining and dropped samples, so that the reduction of sample does not have enormous impact on the analysis. In addition, they are also required to give their *hukou* status. Information regarding the survey location is also available in CGSS 2010, which can distinguish rural-urban migrants from the rural *hukou* individuals.

When carrying out propensity score matching, it is important to carefully choose the covariates in the analysis. Since propensity score matching (PSM) is based on the unconfoundedness condition, all variables influencing treatment participation should be controlled for to avoid estimation bias. The choice of variables can rely on economic theory, previous research and institutional settings (Heckman et al., 1997; Dehejia & Wahba, 1999 and Smith & Todd, 2005). Regarding the choice of variables, Heckman et al. (1997) suggest that only those influencing participation and outcome simultaneously should be included in the analysis. Augurzky & Schmidt (2001) divide the potential variables into three groups. The first contains those that affect the treatment participation but are unrelated to outcomes; the second includes variables that influence outcomes but are not related to treatment participation; the third group contains variables influencing both. They point out that the model with the third group, which includes all variables influencing treatment participation and potential outcomes simultaneously, has better estimated results of average treatment effect for the treated group (ATT).

However, Bryson et al. (2002) indicate that more information in the dataset can advance the quality of matching. Including non-significant variables in the matching specification will not lead to bias in the estimation of ATT, but it will increase the variance. Therefore, they suggest choosing all available variables in the analysis which are relevant to the outcome. Rubin & Thomas (1996) argue that a variable can be excluded from the analysis in two cases. Firstly, based on previous research and economic theory, this variable is not relevant to the outcome. Secondly, the variable you chose is not a relevant covariate for the treatment participation and potential outcome (Caliendo & Kopeinig, 2008). In this chapter, I intend to choose all the available covariates potentially affecting income aspirations from CGSS (2010) firstly, and then test which of the variables influence treatment participation.

Based on previous research on the determinants of income aspirations, the effect of current income, socio-demographic factors (e.g. age, gender, education attainment and marital status), occupation status, size of household, family background, health status and self-evaluated SES on income aspirations have been discussed. First of all, for the effect of current income on income aspirations (Fuchs & Landsberger, 1973; Centers & Cantril, 1946; Portes et al., 1978; Stutzer, 2004; McBride, 2010; Dalton et al., 2016),

Table 3- 1 Description of Variables

Variables	Description
Dependent variable	
Income aspirations	Individuals' aspirations regarding total income
Treatment dummy	
Rural	1-Respondents with rural *hukou* status
	0-respondents with urban *hukou* status
Independent variables	
Current income	Actual total income per year
Social demographic factors	
Male	1- Male
	0- Female
Age	Age of respondent in years. Male (16 to 60); female (16 to 55)
Married	1- Married
	0- Unmarried
Children	Number of children in the family
Respondents' highest education level	
Non-education	1- No education
	0- Otherwise
Compulsory level	1- Highest education level is compulsory school
	0- Otherwise
Senior high school	1- Highest education level is senior level (professional high school, senior high school and technical school)
	0- Otherwise
University	1- Highest education level is university
	0- Otherwise
Type of Occupation	
Private company	1- Engaging in private organizations
	0- Otherwise (engaging in public organizations)
Father's highest education level	
Non-education for father's highest education	1- No education
	0- Otherwise
Compulsory level for father's highest education	1- Highest education level is compulsory school
	0- Otherwise
Senior high school for father's highest education	1- Highest education level is senior level (professional high school, senior high school and technical school)
	0- Otherwise
University	1- Highest education level is university
	0- Otherwise
Self-evaluated SES	Individuals' self-evaluated socio-economic status on a 10-point scale
	1- Significantly below average level
	10- Significantly above average level
Health status	
Good health	1- Good health
	0-Otherwise
Standard	1-Standard health status
	0-Otherwise
Poor health	1-Poor health
	0-Otherwise

Note: China General Social Survey 2010 is utilized.

total income and occupational incomes per year are provided in the dataset. Since the income aspirations variable refers to individuals' total income aspirations, I decided to use the total income variable as the measure of current income. In addition, fewer

people reported extremely high and extremely low value in terms of their income aspirations and total income. However, there is no available information to explain these extreme issues. Therefore, I delete the top and bottom 1% of total income aspirations and total income from the sample to guarantee even income distribution.

Regarding the socio-demographic variables, information about age, gender, educational attainment and marital status can be found in the CGSS 2010. For the educational attainment, I use the highest education level, which is a categorical variable. Following the "Nine-year Compulsory Education Law", decreed in 1986, all school-age children and teenagers must receive nine years' compulsory education, which includes 6 years of primary school and 3 years of lower middle school. The government exempts all tuition fees for compulsory education. It is the obligation of the compulsory schools, parents and the community to let school-age children receive the appropriate education. After the nine years of compulsory education, some students choose to study further in higher middle school or technical school, while some students may give up studying and devote their time to working. After graduating from higher middle school, some individuals choose to study further at university. Therefore, this categorical variable "highest education level" is divided into 4 dummy variables: no educational attainment, compulsory education, senior higher school and university.

Marital status is a categorical variable as well, with 6 categories, including unmarried; married; living together but not married; divorced; widowed; and separated. In order to transfer this categorical variable into a dummy variable, people living together are regarded as married, whilst divorced, those separated but not divorced and widowed individuals are treated as unmarried. The variable health status is an ordinal variable with 5-point scale: 1-completely unhealthy, 2-unhealthy, 3-standard, 4-healthy, 5-completely healthy. I restructured this ordinal variable into three dummy variables, which are good health (healthy and completely healthy), standard health and poor health status (unhealthy and completely unhealthy).

For occupation status, specific information on detailed occupation types is not available but I can get information related to the broad type of occupation, including farmers, self-employed, employed in national organizations, such as government or public organizations, and employed in private enterprises. As the purpose of this study

is to discuss how *hukou* discrimination influences rural-urban migrants' income aspirations, I only focus on employed individuals in the urban labour market. As a consequence, people who work in farming and the self-employed are deleted from the analysis. The occupation can be transferred into a dummy variable: 0 refers to being employed in public organizations and 1 refers to being employed in a private company. In terms of family background, I follow Portes et al. (1978) in investigating whether a father's educational background has a positive association with respondents' income aspirations. As for the respondents' educational attainment variable, the father's education attainment is also divided into four dummy variables: no education, compulsory education, senior higher school and university. Regarding household size, I follow Stutzer (2004), choosing the number of children as a measure of the size and composition of the family.

When estimating the effect of social comparisons on income aspirations, Stutzer (2004), McBride (2010), Knight & Gunatilaka (2012a) and Knight & Gunatilaka (2012b) suggest that comparisons with neighbours, close friends, colleagues, other members in the community and the average standard of the whole society play an important role in individual aspirations. Specific information about respondents' neighbours, friends and colleagues and the average standard of the whole society is unfortunately not available in the CGSS 2010. However, I can obtain information about individuals' self-evaluated social status instead from the question "Can you make a self-evaluation of your social status into a given 10-point scale from bottom (1-significantly below average level) to the top (10-significantly above average level), based on your own income, education level, and other endowment factors (e.g., the extent of being respected)?". Following these objective criteria such as income and education level, all respondents evaluate their own socio-economic status differently depending on their relative positions. They are more likely to assess their own status by comparing with proximal persons around them, such as their neighbours, relatives, fellows, close friends or the people in the same community. Even if a person with lower income and is not highly educated, his self-evaluated social status may be similar to a person who has higher education level and higher income. That is because these two persons are in different relative position and environment. Table 3-1 shows a description of the outcome variable, participation variable and a set of covariates which potentially have an impact on income aspirations. After choosing the covariates influencing potential outcomes, I tested which factors

among these covariates also influence treatment participation. Following the introduction in the Chapter 1, I can obtain detailed information of the implementation, reform, development and how *hukou* status has changed. In terms of variations in *hukou* status, most variables among the covariates influencing income aspirations have an impact on *hukou* status as well. To be specific, the educational attainment of the head of household, age, educational attainment, marital status, present income and occupation can determine an individuals' *hukou* status.

For other factors, such as the number of children and SES, no existing literature discusses whether *hukou* status can be influenced by these two factors. However, both Rubin & Thomas (1996) and Bryson et al. (2002) indicate that all covariates potentially influencing potential outcomes should be included in the propensity score matching. Following previous research, it is likely that family size and social comparisons play a significant role in determining income aspirations. Therefore, I keep these two variables (number of children and self-evaluated SES) when conducting the propensity score matching.

Table 3-2 summarizes the characteristics of all the covariates for both urban *hukou* and rural-urban migrants. The ratio of the sample sizes for urban *hukou* and rural-urban migrants is in line with Bryson et al. (2002) and Caliendo & Kopeinig (2008), who suggest that the sample size of the treated group should be smaller than the sample size of the control group. The treated group (rural-urban migrants) has 633 observations, in comparison to the control group (urban *hukou*) with 2242 observations. It can be seen from Table 3-2 that urban *hukou* individuals report higher aspirations for income than rural-urban migrants on average, and also have higher incomes. In addition, urban *hukou* individuals are more likely to be educated and their father's highest education level is higher than that for rural-urban migrants. To be specific, about 75% of urban *hukou* individuals had graduated from senior higher school or university compared to only 33% rural-urban migrants. Regarding occupation, the percentage of urban *hukou* individuals (34%) working for public organizations was about twice as high as for rural-urban migrants. In terms of self-evaluated SES, it can be seen that rural-urban migrants in different positions have different attitudes to their socio-economic status with urban people, by the objective criteria such as their own income and education attainment. Although rural-urban migrants have lower income and less education

attainment, their reference comparison populations are mainly individuals with the similar income level and education attainment as themselves. Therefore, they report relatively similar socio-economic status to urban *hukou* workers. Regarding

Table 3- 2 Descriptive Statistics

Variables	Mean	Std. Dev	Mean	Std. Dev
	Urban *hukou* (2242 observations)		Rural-urban migrants (633 observations)	
Dependent variable				
Income aspirations	45420.44	35157.91	31941.07	25937.76
Independent variable				
Current income	26496.20	23377.69	15073.20	14143.49
Social demographic				
Age	42.01	10.96	35.17	10.80
Male	0.50	0.50	0.50	0.50
Married	0.82	0.39	0.79	0.40
Children	1.01	0.65	1.18	0.89
Respondents' highest education level				
Non-education	0.01	0.11	0.02	0.15
Compulsory school	0.24	0.43	0.65	0.48
Senior high school	0.33	0.47	0.24	0.42
University	0.42	0.49	0.09	0.29
Type of Occupation				
Public organization	0.34	0.47	0.13	0.33
Private company	0.66	0.47	0.87	0.34
Father's highest education level				
Non-education	0.25	0.44	0.30	0.46
Compulsory school	0.48	0.50	0.56	0.50
Senior high school	0.18	0.38	0.13	0.33
University	0.09	0.29	0.01	0.11
Self-evaluated SES	4.33	1.66	4.06	1.58
Health status				
Good health	0.67	0.47	0.76	0.41
Standard health	0.24	0.42	0.18	0.36
Poor health	0.09	0.28	0.06	0.24

Note: (1) China General Social Survey 2010 is utilized.

(2) Self-evaluated SES utilized in this chapter is continuous variable. Due to the propensity score matching being utilized, the exact propensity score is required to match the urban *hukou* individuals and rural-urban migrants. If I combine several different levels together into three dummy variables, the PS score obtained from this approach may not be precise enough for matching. Therefore, I think it is much better to regard the SES as a continuous variable to obtain the propensity score in this Chapter. However, I have also tried to re-run analysis using the dummy variables of self-evaluated SES and show the results in Table 3-14 of Appendix. When comparing the estimation results by using continuous SES (Table 3-3 and Table 3-4) with the estimation by using dummy variable of SES (Table 3-14 in Appendix), I find that both the results from OLS and propensity score matching have no differences in the significance and direction of all the variables.

socio-demographic factors, urban *hukou* and rural-urban migrants had similar characteristics, in terms of gender, marital status and the number of children. However, the average age of urban *hukou* respondents was higher than that of rural-urban migrants.

In line with the previous research on the determinants of income aspirations, it is clear from Table 3-2 that rural-urban migrants with lower income, less educational attainment and lower self-evaluated SES are expected to report lower absolute income aspirations. However, it is not clear whether their *hukou* status plays a role in decreasing their income aspirations. The propensity score matching (PSM) can disentangle this issue. By controlling for all the determinants of income aspirations, PSM finds urban *hukou* individuals with similar characteristics to match with the rural-urban migrants. The difference between the actual and counterfactual outcome for the rural-urban migrants can be used to explain whether *hukou* status has an impact on income aspirations.

3.6 Results and Discussion

3.6.1 Ordinary Least Squares Estimation Results

Before describing the results of the propensity score matching, the results of a basic ordinary least squares (OLS) regression of income aspirations are shown in Table 3-3. The covariates are those described in Table 3-2. I also include the dummy variable for *hukou* status in order to test whether this dummy treatment has an effect on income aspirations.

As shown in Table 3-3, most covariates have a significant relationship with income aspirations, which is consistent with the previous literature. As expected, current income is positively and significantly associated with individuals' income aspirations, in line with Centers & Cantril (1946), Fuchs & Landsberger (1973), Stutzer (2004), and McBride (2010). Income aspirations increase by 0.083% given a 1% increase in current income, holding other factors constant.

In terms of socio-demographic factors, age, gender, marital status, and the number

of children is significantly related to individuals' income aspirations. As shown in Table 3-3, the positive coefficient on age and the negative coefficient of age squared on income aspirations show that income aspirations will increase by 2.9%, at a decreasing rate, which is in line with Stutzer (2004) who suggests an inversely U-shaped relationship. As in Centers & Cantril (1946), men have higher aspirations and ambition for income, probably because they usually undertake the major economic burden in most families in China. Table 3-3 shows that the income aspiration for men is 24.0% higher than that for women. Moreover, respondents who are married aspire to a 14.4% higher income than those who have never been married or divorced. The number of children is positive related to income aspirations, which is consistent with Stutzer (2004). Specifically, respondents' income aspirations will increase by 6.5% for every additional child in the household.

For respondents' education, it is clear that individuals' income aspirations have positive association with their educational attainment, in accordance with Centers & Cantril (1946) and Stutzer (2004). In addition, the more education they receive, the higher their income aspirations. In comparison with the reference group (no education), those respondents who have finished compulsory education have 28.4% higher income aspirations, while the aspirations for those people graduating from senior higher school is 44.2% higher than those who have not completed compulsory education. Those who graduated from university have the highest income aspirations compared with people who have never received an education. The father's educational attainment has a similar relationship with income aspirations: the more education the respondents' father has, the higher the income aspirations of the respondents. Compared to individuals whose father has not completed compulsory education, income aspirations for respondents whose fathers graduate from university increases by 27.8%, while the income aspirations for those whose fathers have only finished compulsory school increases by 16.2%. From 1966 to 1976, China experienced the Cultural Revolution. The college entrance examination was abolished during this period and was not resumed until October 1977. Therefore, individuals who graduated from senior high school had no choice but to work in factories and government in this stage. When the college entrance examination resumed, some individuals abandoned their jobs to take the college entrance examination. Such individuals who have courage to quit job and take part in the examination were more likely to be those who have excellent academic

Table 3- 3 Ordinary Least Square Estimation

Dependent variable	Coefficient	S.E.
Log of income aspiration		
Log of current income	0.080***	(0.007)
Rural	-0.042	(0.036)
Socio-demographic factors		
Age	0.029***	(0.010)
Age2	-0.0003***	(0.0001)
Male	0.240***	(0.025)
Married	0.144***	(0.038)
Number of children	0.065***	(0.022)
Respondents' highest educational level		
(reference category is no education)		
Compulsory school	0.284***	(0.102)
Senior high school	0.442***	(0.104)
University	0.871***	(0.107)
Type of occupation		
Private company	0.083***	(0.030)
Father's highest educational level		
Compulsory school	0.162***	(0.033)
Senior high school	0.192***	(0.044)
University	0.278***	(0.055)
Self-evaluated SES	0.024***	(0.008)
Health status		
Good health	0.074	(0.048)
Standard health	0.110**	(0.052)
Observations	2875	
R-squared	0.303	

Note: (1) Sig: * p<0.1, ** p<0.05, *** p<0.01 for two-tailed t-tests; Standard errors in parentheses.

(2) The OLS estimation results are not sensitive to the assumption that set few people's missing income aspirations equal to their actual income when they are satisfied or relative satisfied with income but give no answers to their income aspirations. There is no significant difference in the results when taking these people out of my sample. The estimated results when taking these people out of the sample is shown in the Table 3-12 in Appendix.

(3) China General Social Survey 2010 is utilized.

performance, larger ambitions, and higher aspirations for their future career development (Liu & Liu, 2017). When these highly educated individuals educate their children, such characteristics could influence the children's personality positively and unconsciously. In addition, highly educated people can provide their children with better educational resources and more opportunities to broaden their horizons. Consequently, individuals with a highly educated father are more likely to have greater ambitions and higher aspirations.

When looking at the effect of occupation on income aspirations, the income aspirations for respondents engaged by private companies are 8.3% higher than for the reference group (respondents in public organizations). This may be attributed to the fact that income is more flexible in private companies, according to Cooke (2004). He has stated that public sector funding in setting the basic income structure is under the tight control of the state. However, employees can get a rise in salary or obtain an extra reward in a private company in light of their outstanding performance. As income in private companies is more flexible, individuals in these companies might hope to receive more money or reward because they believe they can obtain what they expect through outstanding achievement.

The estimated coefficient for self-evaluated SES shows that aspiration income increases by 2.4% for every unit increase in SES (10 classes totally), when holding other determinants constant. This is in line with Stutzer (2004), Ray (2006) and Knight & Gunatilaka (2010b, 2012a, 2012b), who indicate that a higher social status or financial situation, based on comparisons with neighbours, colleagues, friends or the average standard of the whole society, has a positive associations with income aspirations.

In terms of the key hypothesis regarding the effect of *hukou* status on income aspirations, the OLS estimation results in Table 3-3 show that the dummy of rural-urban migrants' variable has no significant relation with income aspirations, which indicates that the lower income aspirations for rural-urban migrants is not attributed to their rural *hukou* status. In order to test the robustness of this result, propensity score matching (PSM) was utilized to match urban *hukou* individuals to rural-urban migrants, based on the same determinants. The propensity score matching (PSM) results are presented in the next section.

3.6.2 Propensity Score Matching-Balancing Test

Before matching the urban *hukou* individuals and rural-urban migrants, a logit model of the treatment is utilized to calculate the propensity score, as explained in the methodology section. The estimation results of the logit model are shown in Appendix Table 3-8, as the determinants of the rural *hukou* dummy are not the primary concern in

this chapter. As shown in Table 3-8, most determinants of income aspirations also have relationship with the treatment dummy. The estimated propensity score is used to match the urban *hukou* individuals in the control group with the rural-urban migrants in the treated group using nearest neighbour one-to-one matching. The next stage after preparing the matching is to carry out a balancing test in order to check the overlap and the quality of matching. Appendix Table 3-9 reports the mean value of the covariates for the matched and unmatched samples. Following Rosenbaum & Rubin (1983), the successful matching should guarantee that the differences between treated and untreated individuals for every factor in the matched sample should be smaller than that in the unmatched sample.

Figure 3-1 Propensity Score Density Distribution of One-to-One Nearest-Neighbor Matching
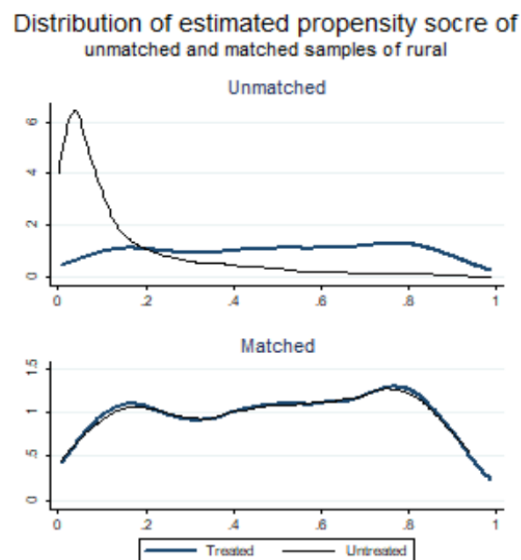
As shown in Table 3-9, the difference in the mean value of covariates between the treated and control groups after matching is smaller than that before matching in general. For example, the log of current income difference in the matched sample is 0.15 (|T-C|=|8.58-8.73|), in comparison to 0.96 (|T-C|=|8.58-9.54|) in the unmatched sample. For respondent's education attainment, the difference of finishing compulsory education between rural-urban migrants and urban individuals in the matched sample, 0.01 (|T-C|=|0.65-0.64|), is lower than that in the unmatched sample, at 0.41 (|T-C|=|0.65-0.24|). The estimated results of the joint test indicate no significant difference between rural-urban migrants and urban *hukou* individuals in the matched

sample (p=0.872). Thus, the results shown in Table 3-9 confirm that the matching procedure is successful and substantially reduces the bias in terms of most covariates. Figure 3-1 depicts the overlapping of PS densities distribution between rural-urban migrants and urban *hukou* individuals in the matched sample, in comparison to unmatched sample. It is clear that the overlapping distribution between the treated and untreated groups is much better in the matched sample. These results suggest that the personal characteristics of the rural-urban migrants and urban *hukou* individuals are similar after matching.

3.6.3 Propensity Score Matching Results

This section presents the results from the last two steps of the propensity score matching: estimation of ATT and sensitivity analysis. Table 3-4 shows the average estimated treatment effect for the treated group. 2784 observations have been matched successfully, where 633 observations come from the treated group and 2151 from the control group. All members in the treated group (rural-urban migrants) have been matched to individuals in the control group (urban *hukou* individuals). The estimated ATT shows that the mean difference in log income aspirations between rural-urban migrants and urban *hukou* people is -0.06, when holding another factors constant. This estimated ATT is not significant at the usual levels. This verifies the estimated results in the ordinary least square (OLS) analysis in Table 3-3, that no difference in income aspirations exist between rural-urban migrants and urban *hukou* individuals after controlling for a wide range of personal characteristics. In other words, the PS matching confirms that neither *hukou* status nor *hukou* discrimination has any impact on individuals' income aspirations.

Following Rosenbaum & Rubin (1983) and Heckman (2008), it is necessary to conduct a sensitivity analysis to test the robustness of the estimation results. I follow Liu & Ma (2009) and test the sensitivity by using different methods, including multiple nearest-neighbour matching (one-to-five NN matching), radius matching, kernel matching and local linear matching [15]. Following the procedure of one-to-one nearest-neighbour matching, it is necessary to conduct a balancing test for the other algorithms before estimating the ATT. Figure 3-3, Figure 3-4, Figure 3-5 and Figure 3-6

---

[15] See Rosenbaum and Rubin (1983); Caliendo & Kopeinig (2008) for more details.

in the Appendix show the propensity score distributions of the treated and untreated groups in the matched samples compared to the unmatched samples for each of the additional matching algorithms. In general, the results in these four figures suggest an overlap and the balancing test is successful in supporting a further estimation of the average treatment effect. In addition, these results suggest that the PS distribution of one-to-five nearest-neighbour and local linear matching are better than for radius and kernel matching, which is consistent with Liu & Ma (2009).

Table 3- 4 Average Treatment Effect for Rural Group on Income Aspirations

Variable	Sample	Treated	Controls	Difference	S.E.	T-statistic
Ln(Income aspirations)	Unmatched	10.09	10.45	-0.36	0.03	-10.46
	ATT	10.09	10.15	-0.06	0.06	-0.87

Treatment assignment	Common support		Total
	Off support	On support	
Untreated	91	2151	2242
Treated	0	633	633
Total	91	2784	2875

Notes: (1) T-statistics are based on a two-tail test.

(2) Similar to OLS estimation, the PSM results are not sensitive to the assumption that set few people's missing income aspirations equal to their actual income when they are satisfied or relative satisfied with income but give no answers to their income aspirations. There is also no significant difference in the ATT between rural-urban migrants and urban *hukou* individuals, when taking these people out of my sample. This estimation result is shown in Table 3-12 in Appendix.

(3) China General Social Survey 2010 is utilized.

Table 3-5 demonstrates that the estimated results are not very sensitive to the choice of matching algorithm. While the estimated ATT decreases for a one-to-five nearest-neighbour, radius, kernel and local linear matching, the estimates are not significant, in line with the estimation results of one-to-one nearest-neighbour matching. The relatively smaller standard error of the other algorithms compared to one-to-one nearest-neighbour matching may be attributed to the fact that, in the later method, only one member in the control group is chosen as a matching partner, whilst in the former multiple matching partners are picked out from the untreated sample. More information may help to reduce the standard error (Liu & Ma, 2009). To sum up, the estimated result from one-to-one nearest-neighbour matching appears to be robust and not sensitive to different matching algorithms, which confirms that *hukou* status does not contribute to

individuals' income aspirations, conditional on other factors.

Table 3- 5 Sensitivity Analysis on using Different Algorithms

Propensity Arithmetic	Difference of ATT	S.E.	T-value
One-to-one nearest-neighbour matching	-0.06	0.06	-0.87
One-to-five nearest-neighbour matching	-0.02	0.05	-0.45
Radius matching	-0.04	0.05	-0.67
Kernel matching	-0.03	0.05	-0.54
Local linear matching	-0.03	0.05	-0.52

Note: (1) T-statistic is based on two-tail test; Radius is 0.01.

(2) China General Social Survey 2010 is utilized.

3.6.4 Further Analysis of Occupational Impacts

The insignificant effect of *hukou* status on income aspirations may be attributed to two reasons. Although rural-urban migrants are discriminated against in the urban labour market, their aspiration income is not restricted to the *hukou* discrimination. The insignificant effect of *hukou* discrimination on income aspirations may be attributed to the fact that most rural-urban migrants face discrimination when they are applying for a job. Therefore, the occupation aspirations of rural-urban migrants may be restricted by *hukou* discrimination. As discussed in the literature review, many studies suggest that racial discrimination influences African Americans' career assessment and can restrict career aspirations. Similarly, in China, rural-urban migrants not only suffer from income discrimination but also from occupational segregation. As discussed in the second chapter, most rural *hukou* individuals are blue collar and service workers. Even when rural-urban migrants and urban *hukou* individuals have similar educational attainment, the probability of being engaged in senior professional, junior professional, blue collar and service jobs is different between the two groups. Also, a large proportion of the effect of *hukou* discrimination on rural-urban migrants is attributed to discrimination when applying for jobs (occupation segregation). As a consequence, *hukou* discrimination may influence aspiration for occupations instead of income.

In order to analyse whether the occupational segregation contributes to the effect of *hukou* status on income aspirations, the matching analysis was run again, after removing the occupation variable from the propensity score model. In addition, as current income and SES are positively influenced by occupation types, occupational segregation can

also lead to inequitable wages and SES. Thus, current income and SES were taken out of the analysis as well. Other covariates, including social demographic factors and father's highest education level, were retained. If a large proportion of the discrimination in the urban labour market derives from occupation segregation, the estimated difference in income aspirations should be higher after removing occupation, current income and SES from the model. Table 3-6 shows the results from an OLS regression without occupation, current income and SES as explanatory variables. Compared to Table 3-3, the effect of dummy variable for rural-urban migrants on income aspirations has changed in terms of the magnitude, to a small extent. The estimate indicates that having rural *hukou* status decreases income aspirations by 5.5%, which is not significant.

Table 3- 6 Estimation without Occupation, Income and Social Status

Dependent variable	Coefficient	S.E.
Log of income aspirations		
Rural	-0.054	(0.037)
Socio-demographic factors		
Age	0.039***	(0.010)
Age2	-0.0001***	(0.00002)
Male	0.302***	(0.025)
Married	0.137***	(0.039)
Number of children	-0.089***	(0.023)
Respondents' highest educational level		
Compulsory school	0.369***	(0.105)
Senior high school	0.539***	(0.106)
University	1.035***	(0.108)
Father's highest educational level		
Compulsory school	0.165***	(0.033)
Senior high school	0.204***	(0.045)
University	0.301***	(0.057)
Health status		
Good health	0.152***	(0.049)
Standard health	0.186***	(0.053)
Observations	2875	
R-squared	0.26	

Note: (1) * p<0.1, ** p<0.05, *** p<0.01 for two-tailed t-tests;

    (2) Standard errors in parentheses.

    (3) China General Social Survey 2010 is utilized.

Similar to the previous process of propensity score matching, it is necessary to

conduct the balancing test before calculating the average treatment for the treated (ATT). As shown in Appendix Table 3-10, the difference in the mean value of most covariates between the treated and control groups after matching is smaller than that before matching, in line with the balancing test for the sample with all covariates. Consequently, the balancing test confirms the success of the matching strategy.

Figure 3-2 depicts the distribution of the propensity score for rural-urban migrants and urban *hukou* individuals for the unmatched and matched samples. It can be seen that the overlap between urban *hukou* individuals and rural-urban migrants is much better in the matched sample compared to the unmatched sample. Therefore, all the results suggest that matching is successful and substantially diminishes the bias in terms of personal characteristics and family background.

Table 3-7 shows the estimates of ATT, based on all the algorithms. The results of one-to-one nearest-neighbour matching indicates that the mean difference in income aspirations between rural-urban migrants and urban *hukou* people has decreased, after removing occupation, income and SES from the PS model. Specifically, the average effect of *hukou* status on income aspirations is -0.03, in comparison to -0.06 when matching with all covariates. However, the estimate ATT is still not significant, which is consistent with the OLS results. When conducting sensitivity analysis, the estimated ATT among one-to-five nearest-neighbour, radius, kernel and local linear matching are also not significant on income aspirations, in line with the results of one-to-one nearest-neighbour matching. Consequently, the estimation from one-to-one nearest-neighbour matching is robust and not sensitive to different matching algorithms.

Following the estimation results in the Chapter 2 and previous studies, I expect the significance and magnitude of the effect of *hukou* status on income aspirations to vary when I remove occupations and the related covariates. However, these variations are not obvious in either OLS estimation or propensity score matching. Thus, the estimations do not accord with my expectation, which may be attributed to two possible explanations. The first possible explanation is that *hukou* discrimination has no impact on restricting rural-urban migrants' income aspirations, even though rural-urban migrants are faced with income discrimination and occupation segregation in the urban labour market. The second possible explanation is that *hukou* discrimination influences other factors, which

in turn determine income aspirations. Specifically, since rural-urban migrants are discriminated against in the long term, they may consider that it is difficult for them to obtain high quality jobs in the urban labour market, even if they achieve higher educational attainment.

Figure 3- 2 Propensity Score Density Distribution of One-to-One Nearest-Neighbor Matching for the New Sample

Distribution of Estimated Propensity Socre of
unmatched and matched samples of rural
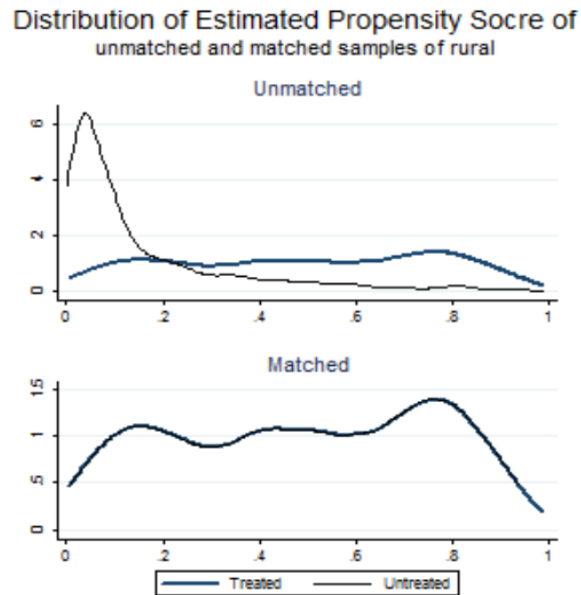
Therefore, rural *hukou* individuals may not devote much time to human capital investment, which in turn reduces their income aspirations. In addition, the *hukou* system leads to area isolation. In general, the economic status of urban households is better than for rural households in China. Some rural households cannot support the economic burden involved with sending their children to university. Therefore, their children give up studying further and devote themselves to work after graduating from high middle school. Furthermore, *hukou* discrimination leads to an uneven distribution of information. To be specific, high quality education resources are concentrated in urban cities so that urban students have a higher probability of passing the university entrance examination. Hence, rural *hukou* students who are not admitted to a university seek work in the labour market. To sum up, rural-urban migrants' educational attainment is limited by *hukou* discrimination, to some extent. In effect, due to the positive relationship between educational level and respondents' income aspirations, rural-urban migrants' lower educational levels may be indirectly influenced by

discrimination.

Table 3- 7 ATT using Different Algorithms

Propensity Arithmetic	Difference of ATT	S.E.	T-value
One-to-one nearest-neighbour matching	-0.03	0.07	-0.42
One-to-five nearest-neighbour matching	-0.02	0.05	-0.40
Radius matching	-0.03	0.05	-0.64
Kernel matching	-0.03	0.05	-0.66
Local linear matching	-0.03	0.07	-0.51

Note: China General Social Survey 2010 is utilized.

3.7 Conclusion

Following Chapter 2, which suggests that rural-urban migrants suffer from wage inequality and occupation segregation in the urban labour market, this chapter has investigated whether *hukou* discrimination has other negative influences on rural-urban migrants. According to Alagaraja et al. (2016), a devalued group's aspiration and ambition are restricted by discrimination. Individuals in the disadvantaged group are more likely to lose the motivation to work, which exacerbates the gap between the preference group and disadvantaged groups, and weakens social stability. Motivated by Alagaraja et al. (2016), Chapter 3 compares income aspirations between rural-urban migrants and urban residents, and also explores whether *hukou* discrimination constrains rural-urban migrants' income aspirations, as an increasing number of rural-urban migrants flow into the urban labour market as a result of the recent reform of the *hukou* system.

Using data from the China General Social Survey 2010, this chapter contributes to previous Chinese literature and analyses empirically the effect of *hukou* discrimination on income aspirations between rural-urban migrants and urban residents. It is found that rural-urban migrants have lower income aspirations than urban residents. The OLS estimate shows that respondents' education attainment and their family background (father's education attainment) have a positive association with income aspirations. In addition, being married, male and having a lower number of children in one's family also leads to higher income aspirations. However, the dummy variable of rural *hukou* status has no significant effect on income aspirations. The OLS results suggest that, although rural-urban migrants are faced with discrimination, such as income inequality and occupation segregation, their income aspirations are not directly influenced by

hukou discrimination.

Propensity score matching is also utilized to estimate the effect of *hukou* discrimination on income aspirations, by matching urban workers and rural-urban migrants with similar characteristics and comparing their reported income aspirations. The propensity score matching results show that there is no difference between actual income aspirations and counterfactual income aspirations. In other words, after matching rural-urban migrants and urban *hukou* workers with similar characteristics, there is no significant difference in the income aspirations between these two groups, which is consistent with the OLS estimates. The explanation of the insignificant results may be due to the following possible explanations: firstly, *hukou* discrimination may have no impact on restricting rural-urban migrants' income aspirations, even though they suffer from different treatment in the urban labour market; secondly, *hukou* discrimination may influence income aspirations indirectly. *Hukou* discrimination may lead to rural-urban migrants considering that they cannot get high quality jobs anyway, in spite of having more educational attainment. In addition, discrimination leads to uneven information distribution between rural and urban areas, so that rural *hukou* individuals are less likely to receive a good quality education. Moreover, university tuition fees are a huge economic burden for most rural families, compared to urban households. Therefore, due to these potential influences, rural *hukou* individuals are more likely to start working rather than undertake further study, and their lower education attainment leads to their lower income aspirations. Therefore, the lower income aspirations of rural-urban migrants may be ascribed to that their highest educational levels being negatively influenced by *hukou* discrimination, which in turn restricts their aspirations.

However, there is still one limitation in this chapter, and that is the unsolved endogeneity issue. When I conduct propensity score matching, the logit model analysing the correlates for rural *hukou* status is applied to estimate the PS score. While discussing the determinants of *hukou* status changing in the introduction chapter, I explained that age, marital status, occupation, income, and so forth are all factors that may influence the *hukou* status. Consequently, I take these factors into account while conducting the logit model estimation in Table 3-8. However, *hukou* status has an impact on respondents' income as well. As shown in Chapter 2, rural-urban migrants'

lower income is negatively influenced by their *hukou* status and restricted by *hukou* discrimination, to some extent. Therefore, the reverse causality can cause endogenous issues in this model, which in turn leads to biased estimation results. Following most previous studies that consider the endogeneity for income (Chau et al., 2007; Knight & Gunatilaka, 2010; Li et al., 2011; Knight & Gunatilaka, 2012), father's education level is the most common variable that has been utilized as an instrumental variable to solve the endogeneity issue for income. However, as discussed in Chapter 1, the father's *hukou* status and education level can directly determine the new-born's *hukou* status. Father's education level is utilized as a determinant for rural *hukou* status in the logit model, and the results in Table 3-8 present that it has negative correlation with the rural *hukou* status; therefore, father's education level is not suitable for use as an effective instrumental variable to deal with the endogeneity issue. Due to the limitation of dataset, it is difficult to obtain another suitable instrumental variable. In addition, as the CGSS 2010 is a cross-sectional dataset, I cannot apply the PSM in combination with DID to deal with the endogeneity, so the unsolved endogeneity issue is the limitation of this study. If there is a new and better dataset in the future, I will try to deal with the endogeneity issue and improve this study further.

3.8 Appendix

Table 3- 8 Logit Estimation Results of Rural Migrants

Dependent variable: Rural Dummy	Coefficient	S.E.
Ln(Current income)	-0.040	(0.026)
Social demographic factors		
Age	-0.134***	(0.007)
Male	0.220*	(0.117)
Married	0.076	(0.160)
Number of children	0.801***	(0.093)
Respondents' highest educational level		
Compulsory school	-0.150	(0.361)
Senior high school	-1.608***	(0.373)
University	-2.990***	(0.398)
Property of occupation		
Private company	0.449***	(0.153)
Father's highest educational level		
Compulsory school (father's)	-0.605***	(0.146)
Senior high school (father's)	-0.895***	(0.204)
University (father's)	-1.713***	(0.413)
Self-evaluated SES	0.021	(0.035)
Observations	2875	
Pseudo R-squared	0.322	

Note: (1) * $p<0.1$, ** $p<0.05$, *** $p<0.01$ for two-tailed t-tests;

(2) Standard errors in parentheses.

(3) China General Social Survey 2010 is utilized.

Table 3- 9 Balancing Test for Matching Process

Variable	Sample	Mean		% Bias	% Bias reduction
		Treated	Control		
Ln(Current income)	Unmatched	8.58	9.54***	-43.6	
	Matched	8.58	8.73	-6.5	85.1
Social demographic factors					
Age	Unmatched	35.17	42.01***	-62.9	
	Matched	35.17	35.65	-4.4	93.0
Male	Unmatched	0.50	0.50	0.9	
	Matched	0.50	0.50	0.0	100
Married	Unmatched	0.79	0.82	-6.0	
	Matched	0.79	0.80	0.0	100.0
Number of children	Unmatched	1.18	1.01***	21.0	
	Marched	1.18	1.19*	-1.8	91.3
Respondents' highest educational level					
Compulsory school	Unmatched	0.65	0.24***	90.4	
	Matched	0.65	0.64	2.8	96.9
Senior high school	Unmatched	0.23	0.33***	-21.5	
	Matched	0.23	0.23	0.4	98.4
University	Unmatched	0.09	0.41***	-80.5	
	Matched	0.09	0.09	0.0	100.0
Property of occupation					
Private company	Unmatched	0.87	0.66***	52.0	100.0
	Matched	0.87	0.87	0.0	
Father's highest educational level					
Compulsory school (father's)	Unmatched	0.56	0.48**	15.5	
	Matched	0.56	0.55	1.9	87.7
Senior high school (father's)	Unmatched	0.13	0.17**	-13.3	
	Matched	0.13	0.10	6.6	50.2
University (father's)	Unmatched	0.01	0.09***	-35.4	
	Matched	0.01	0.01*	2.9	91.7
Self-evaluated SES	Unmatched	4.06	4.33***	-16.6	
	Matched	4.06	4.02	3.6	78.2
Good health	Unmatched	0.78	0.68	22.6	
	Matched	0.78	0.79	-1.1	95.2
Standard health	Unmatched	0.16	0.23	-19.6	
	Matched	0.16	0.14	3.6	81.6
Propensity score	Unmatched	0.49	0.14***	155.3	
	Matched	0.49	0.49	0.1	99.9

Joint Hypothesis Test

Sample	Ps R²	LR Chi2	p>chi2	Mean bias	Med bias	B	R	% Var
Unmatched	0.323	978.36	0.000	41.1	22.1	161.7*	1.03	69
Matched	0.006	9.90	0.872	2.7	2.3	17.7	1.03	13

Notes: (1) propensity score are estimated at the baseline using 1-to-1 nearest matching (two-tail t-test);

(2) Significance: *** p<0.01; ** p<0.05; * p<0.1.

(3) China General Social Survey 2010 is utilized.

Figure 3- 3 Propensity Score Density Distribution of One-to-Five Nearest-Neighbor Matching

Distribution of estimated propensity socre of
unmatched and matched samples of rural
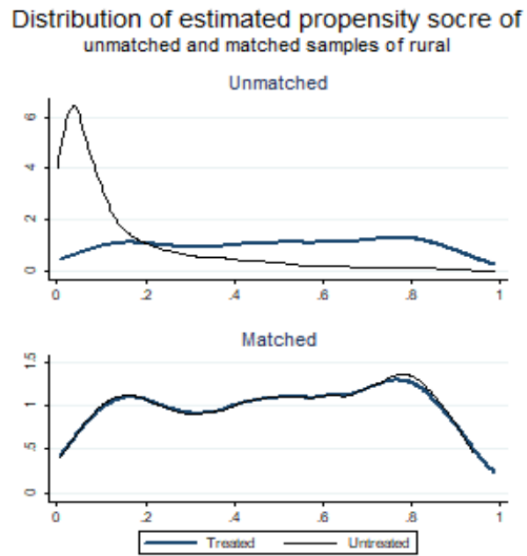
Figure 3- 4 Propensity Score Density Distribution of Radius Matching

(Radius is 0.01)

Distribution of estimated propensity socre of
unmatched and matched samples of rural
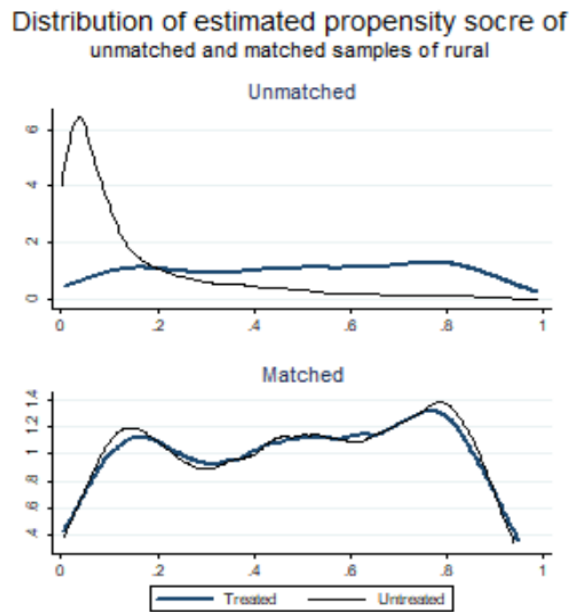
Figure 3- 5 Propensity Score Density Distribution of Kernel Matching

Distribution of estimated propensity socre of
unmatched and matched samples of rural
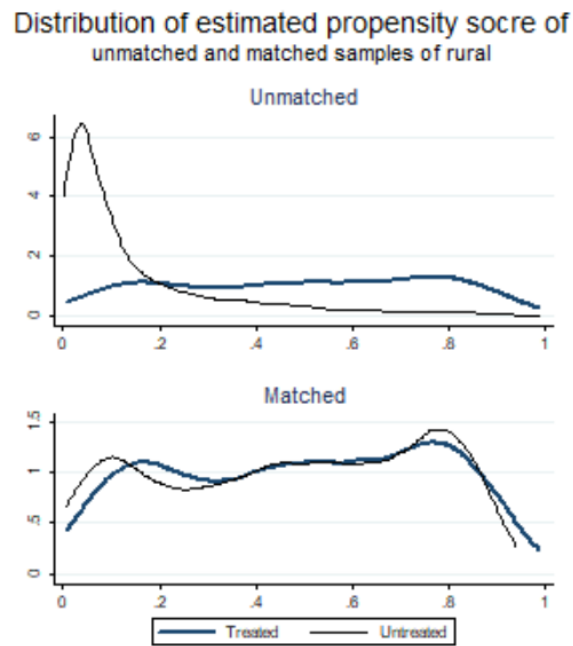
Figure 3- 6 Propensity Score Density Distribution of Local Linear Matching

Distribution of estimated propensity socre of
unmatched and matched samples of rural
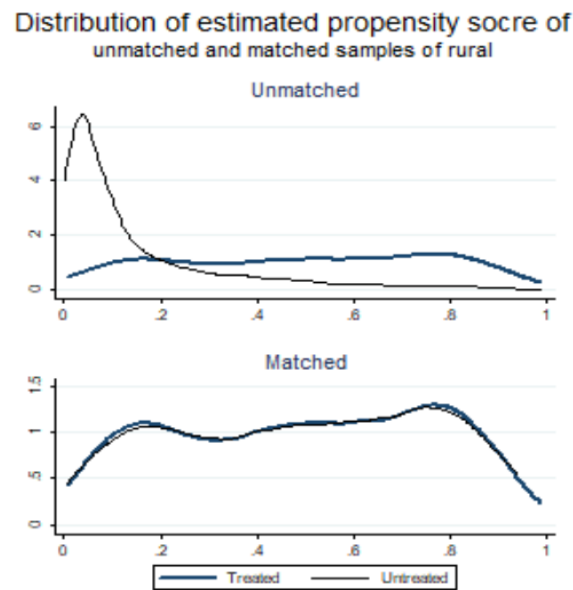
Table 3- 10 Balancing Test for Sample without Occupation, Income and Social Status

Variable	Sample	Mean		% Bias	% Bias reduction
		Treated	Control		
Social demographic factors					
Age	Unmatched	35.17	42.01***	-62.9	
	Matched	35.17	35.79	-5.7	90.9
Male	Unmatched	0.50	0.50	0.9	
	Matched	0.50	0.51	-1.6	-81.2
Married	Unmatched	0.79	0.82	-6.0	
	Matched	0.79	0.77	7.2	-19.3
Number of children	Unmatched	1.18	1.01***	21.0	
	Marched	1.18	1.19	-1.8	91.3
Respondents' highest educational level					
Compulsory school	Unmatched	0.65	0.24***	90.4	
	Matched	0.65	0.67	-3.8	95.8
Senior high school	Unmatched	0.23	0.33***	-21.5	
	Matched	0.23	0.22	-2.5	88.5
University	Unmatched	0.09	0.41***	-80.5	
	Matched	0.09	0.08	-2.0	97.6
Father's highest educational level					
Compulsory school (father's)	Unmatched	0.56	0.48**	15.5	
	Matched	0.56	0.51	10.5	32.4
Senior high school (father's)	Unmatched	0.13	0.17**	-13.3	
	Matched	0.13	0.15	-5.8	56.8
University (father's)	Unmatched	0.01	0.09***	-35.4	
	Matched	0.01	0.01	1.5	95.9
Good health	Unmatched	0.78	0.68	22.6	
	Matched	0.78	0.82	-9.0	60.4
Standard health	Unmatched	0.16	0.23	-19.6	
	Matched	0.16	0.16	12.8	34.6
Propensity score	Unmatched	0.49	0.14***	154.8	
	Matched	0.49	0.49	0.0	99.9

Joint Hypothesis Test

Sample	Ps R^2	LR Chi2	p>chi2	Mean bias	Med bias	B	R	% Var
Unmatched	0.318	965.50	0.000	41.8	21.5	160.2*	1.05	69
Matched	0.009	16.50	0.124	5.7	3.8	29.9	1.13	23

Notes: (1) propensity score are estimated at the baseline using 1-to-1 nearest matching (two-tail t-test);

(2) Significance: *** p<0.01; ** p<0.05; * p<0.1.

(3) China General Social Survey 2010 is utilized.

Table 3- 11 Social Demographic Variables Statistics of Individuals Who are Dropped

Variables	Mean	Std. Dev	Mean	Std. Dev
	Urban (610 observations)		**Rural (556 observations)**	
Age	42.02	12.04	39.73	8.80
Age^2	1910.16	983.16	1657.23	708.32
Number of children	1.10***	0.89	1.77***	1.16
Male	0.47**	0.50	0.49**	0.50
Married	0.84***	0.44	0.88***	0.37
Educational attainment				
Non-education	0.05***	0.23	0.07***	0.38
Compulsory education	0.29***	0.48	0.70***	0.47
Senior school	0.30***	0.45	0.20***	0.35
University level	0.36***	0.46	0.03***	0.17
Health status				
Good health	0.62**	0.49	0.69**	0.49
Standard health	0.26**	0.44	0.22**	0.43
Poor health	0.12*	0.32	0.09*	0.37

Note: (1) * refers to whether there is significantly difference between variables

(2) *, **, *** indicate 10%, 5%, 1% significant levels.

(3) China General Social Survey 2010 is utilized.

Table 3- 12 The Sensitive Analysis for Sample

Dependent variable	Coefficient	S.E.
Log of income aspiration		
Log of current income	0.076***	(0.007)
Rural	-0.049	(0.036)
Socio-demographic factors		
Age	0.028**	(0.010)
Age2	-0.0003***	(0.0001)
Male	0.239***	(0.026)
Married	0.141***	(0.038)
Number of children	0.072***	(0.023)
Respondents' highest educational level		
(reference category is no education)		
Compulsory school	0.298***	(0.103)
Senior high school	0.454***	(0.105)
University	0.879***	(0.108)
Type of occupation		
Private company	0.091***	(0.030)
Father's highest educational level		
Compulsory school	0.162***	(0.033)
Senior high school	0.183***	(0.044)
University	0.305***	(0.057)
Self-evaluated SES	0.025***	(0.008)
Health status		
Good health	0.094*	(0.048)
Standard health	0.125**	(0.052)

Average Treatment Effect for Rural Group on Income Aspirations (ATT)

Variable	Sample	Treated	Controls	Difference	S.E.	T-statistic
Ln (Income aspirations)	Unmatched	10.10	10.47	-0.37	0.04	-10.42
	ATT	10.10	10.15	-0.05	0.07	-0.68

Note: (1) * refers to whether there is significantly difference between variables

(2) *, **, *** indicate 10%, 5%, 1% significant levels.

(3) China General Social Survey 2010 is utilized.

Table 3- 13 Estimation by Dummy Variables for Economic Status
Ordinary Least Square

Dependent variable	Coefficient	S.E.
Log of income aspiration		
Log of current income	0.081***	(0.007)
Rural	-0.041	(0.255)
Socio-demographic factors		
Age	0.029***	(0.010)
Age2	-0.0003***	(0.0001)
Male	0.238***	(0.025)
Married	0.146***	(0.038)
Number of children	0.064***	(0.022)
Respondents' highest educational level		
(reference category is no education)		
Compulsory school	0.284***	(0.102)
Senior high school	0.444***	(0.104)
University	0.876***	(0.107)
Type of occupation		
Private company	0.083***	(0.030)
Father's highest educational level		
Compulsory school	0.164***	(0.033)
Senior high school	0.195***	(0.044)
University	0.282***	(0.055)
Self-evaluated SES (reference category: low economic status)		
Average status	0.038	(0.021)
High economic status	0.108**	(0.049)
Health status		
Good health	0.083	(0.048)
Standard health	0.114**	(0.052)

Average Treatment Effect for Rural Group on Income Aspirations (ATT)

Variable	Sample	Treated	Controls	Difference	S.E.	T-statistic
Ln(Income aspirations)	Unmatched	10.09	10.45	-0.36	0.03	-10.46
	ATT	10.09	10.09	-0.02	0.06	-0.03

Note: (1) * refers to whether there is significantly difference between variables

(2) *, **, *** indicate 10%, 5%, 1% significant levels.

(3) China General Social Survey 2010 is utilized.

Chapter Four

How Income Aspirations Influence Subjective Well-Being between

Urban Residents and Rural-Urban Migrants

4.1 Introduction

As discussed in Chapter 2, I find that rural-urban migrants suffer from a disparity in average income and occupation choice, even though the *hukou* system has been reformed. Moreover, due to the unequal information distribution, rural *hukou* individuals also suffer from different access to social service and welfare. With the transition of Chinese economic characteristics, China is faced with a labour supply shortage, which stimulates an increasing number of rural-urban migrants to enter the urban labour market. A large number of rural-urban migrants are, therefore, an indispensable factor to urban city construction, economic growth, and Chinese development. Jiang et al. (2012) indicated that the inequality between advantaged and disadvantaged groups could increase the gap of their subjective well-being[16] (SWB), which in turn leads to lower economic growth and a higher crime rate. Although the urban residents are the advantaged group by their *hukou* status, they also feel unhappy because of the social unrest that is caused by identity-related inequality[17]. In this dualistic society, the difference in the SWB between the preferred group and devalued group plays an important role in determining social stability (Knight et al., 2009). In Chapter 3, I found that rural-urban migrants report lower income aspirations than urban *hukou* individuals, but this difference in aspirations is not due to discrimination. It seems to be driven by different characteristics of rural-urban migrants including fewer educational opportunities and attainment. Motivated by the

---

[16] The subjective well-being (SWB) contains two distinct spheres: happiness and satisfaction. Although the definitions between happiness and satisfaction seem similar and are usually utilized interchangeably in previous studies, there is slight difference between these two variables. As discussed in Pavot & Diener (2008), subjective well-being is conceptualized as multifaceted with both affective and cognitive components. Life satisfaction as a measure of satisfaction with life scale represents a cognitive evaluation of the quality of one's life. Happiness is more likely to be a measurement of affect. An assessment of SWB includes separate measure of both life satisfaction and the affective components of SWB (happiness). In this chapter, SWB refers to the respondents' happiness, and the detailed information of happiness will be presented in the data section.

[17] It is a between-group inequality that is caused by the *hukou* system.

findings of the previous two chapters, Chapter 4 intends to assess whether rural-urban migrants also report lower SWB, and how income aspirations shape their SWB, in comparison to urban *hukou* individuals.

Most previous studies on SWB have been concentrated in developed countries, while few investigations on SWB are found in developing countries. Among these few Chinese studies on SWB, Knight et al. (2009) and Knight & Gunatilaka (2010b) have only examined the SWB of rural households and rural *hukou* individuals. Although Knight & Gunatilaka (2010a) and Knight & Gunatilaka (2012b) consider the determinants of expectation or aspiration income, their studies mainly focus on the influence of the rural-urban migrants' aspirations on their SWB, in comparison to other rural *hukou* individuals living in the rural area. The authors also discuss how rural-urban migrants' aspirations change when they settle in cities and how these changes influence their SWB. Unlike these previous Chinese studies, I contribute to the existing literature by examining the role of income aspirations on SWB across rural-urban migrants and urban *hukou* individuals.

As shown in Chapter 3, rural-urban migrants have lower income aspirations in comparison to urban *hukou* individuals, while Chapter 2 shows that rural-urban migrants also suffer from inequality in terms of income and occupational opportunities. Economic theory suggests that income positively influences SWB and would increase the gap of SWB between rural-urban migrants and urban *hukou* individuals. However, there are studies analysing the impact of income aspirations on SWB, which suggest that the absolute income aspirations reduce SWB and offset the positive impact of absolute income on SWB (Stutzer, 2004). Although rural-urban migrants have lower income, they also report lower absolute income aspirations. The suggestion is that the differential income aspirations between rural-urban migrants and urban *hukou* individuals would reduce the gap in SWB. There is no previous Chinese research investigating how the difference in income aspirations shapes the SWB between rural-urban migrants and urban *hukou* individuals. I, therefore, am able to fill the gap of previous literature and examine the influence of income aspirations on SWB for these two groups.

This chapter is structured as follows: Section 2 provides a brief review of the literature on the development of SWB, and the literature mainly focuses on the

relationship between actual income, aspirations and relative income after respondent comparison, which includes income and social and economic circumstances. Section 3 outlines a discussion of the data, which is followed by methodology in Section 4, Section 5 displays and discusses the results from the analysis, and the paper concludes with a discussion of the implications of the findings.

4.2 Literature Review

SWB has been a topic of interest in psychology and sociology for a long time[18]. During recent decades, an increasing number of economists have begun to focus on SWB. Among all the economic determinants of SWB, income is a dominating factor, where many studies have examined (Clark et al., 2008). This section mainly reviews those studies which explore the impact of income on SWB. Regarding the research on the impact of income, most previous economists analyse the impact of income from two aspects: absolute income and relative income. In this section, I review those studies related to absolute income and relative income separately. Afterwards, studies examining the effect of socio-demographic factors are also discussed to identify suitable control variables.

4.2.1 Absolute Income

Absolute income refers to actual income during the period of investigation. Some researchers, such as Veenhoven (1991) and Diener et al. (1993), found that richer persons are more likely to be happier than poorer ones. In addition, persons in affluent countries are much happier than those in poorer countries. Veenhoven (1991) analyses determinants of SWB for poor countries and suggests that life satisfaction is directly determined by how innate bio-psychological needs are gratified. To be specific, whether individuals have adequate and ample food and clothing or suffer from war largely determines the level of their happiness. Individuals cannot feel happy in an awful environment, where there is hunger, poverty, danger, war, contagion and isolation, even if their neighbours are worse off.

Diener et al. (1993) utilizes the data based on two studies to test the relationship

---

[18] See more details in Diener (1984).

between SWB and income within and across countries. Regarding the two studies, one is based on 18032 college students in 39 countries, and the other relies on 4942 American adults. They are longitudinal datasets with two waves over a 9-year interval. The results show that income correlates with SWB, both within-country and cross-country. The correlation is larger in the cross-country sample than in the within-country sample. In addition, they also explore whether relative status determines happiness. However, the results are unexpected, show that relative standard of income, including social comparison and adaption, has no influence on SWB, which is consistent with Veenhoven (1991). They ascribe the reason to the fact that people are more likely to compare themselves with to proximal people, which comes from the theory of social comparison (Festinger, 1954).

Following the German reunification, Frijters et al. (2004) examine in how the life satisfaction of East Germans changed. Using the data from German Socio-Economic Panel (GSOEP), they mainly examine how real household income improves the life satisfaction as a result of German reunification. They find that real household income plays a significant role in life satisfaction variation, which is consistent with Veenhoven (1991) and Diener et al. (1993). They find the increase in real household income contributes to 40 percent to the increase in life satisfaction.

In addition, Stevenson & Wolfers (2013) test the relationship between SWB and income, and analyse whether the happiness has a satiation point. In particular, they examine whether income only matter for the SWB among those whose basic needs have not been gratified, but has no influence on happiness beyond specific income threshold. They propose two hypotheses that are: firstly, there is no relationship between income and SWB when basic needs are satisfied; secondly, the relationship between income and happiness among the poor is different with the correlation among the rich. Their analyses confirm the significance role of income in SWB for both the poor and the rich. They find that proportional increases in income have same increase in SWB among the wealthy people, in comparison to the poor, which is not consistent with their hypotheses.

4.2.2. Relative Income

Several researchers put forward different viewpoints and suggest that absolute

income is not a decisive factor for SWB (Easterlin, 1974; Brickman & Campbell, 1971; Campbell et al., 1976; Brickman et al., 1978; Diener, 2000). Some previous researchers such as Brickman & Campbell (1971) and Brickman et al. (1978) indicate that the adaptation level theory suggests that a sudden increase of income can affect individuals' happiness growth during a short period. Winning a lottery can bring a huge pleasure and improve current living standards, but people will adapt to the pleasure and new status; therefore, the peak pleasures will reduce eventually, and the income growth cannot make people happy for a long time. In addition, some researchers demonstrate that individual's happiness is not only attributed to his absolute income, but also his income relative to other people in society, such as his neighbours and friends. An individual's happiness is determined by whether his income is higher or lower than the income of the comparison reference group (Duesenberry, 1952). Specifically, the happiness will increase when his income is higher than the income of other members in the reference group, but will decrease while his income is below other members' income.

In terms of the impact of income on subjective well-being, Easterlin (1974) observes a paradox in that within a given country, individuals with higher income are happier than those people with lower income, but happiness levels are relatively flat across time despite rapid rises in incomes, and moreover between countries there appears to be very little association between GDP and happiness. Easterlin (1974) gives the two interpretations. The first possible interpretation is based on relative income. Individuals' utility is positively influenced by their own income and negatively influenced by the average income of the whole society. If income of everyone increased, an increase in the income of any individuals would remain unchanged as the whole average income increases and income of the comparison reference group increases as well. Other possible interpretations may be due to those external non-economic factors such as health, social capital, pollution, and so forth. If these external non-economic factors have a negative association with individual's happiness, the positive relation between the economic factors with happiness can be offset by these non-economic determinants, to some extent. Therefore, with GDP and income increasing, individual's happiness remains unchanged.

In general, the most common measurement of relative income relies on social comparison with different reference groups, as seen in Easterlin (1995), Graham & Felton (2005), Vendrik & Woltjer (2007), Luttmer (2005) and McBride (2001).

Individuals may choose reference groups according to their socio-demographic factors, such as age, gender, and region. For example, a 45-year old female may like to compare herself to other women whose age is between 40-50. Individuals in Shanghai like to compare themselves to other persons in the same city. In addition, they define relative income as a socio-economic standard (SES), based on individuals' own judgments of their income, wealth, savings, occupation and education attainment, in comparison to their colleagues, neighbours, friends or the general standard of the whole society. Thirdly, the reference group is the respondents themselves. Specifically, individuals can compare their current situation to their previous one, such as between current and previous income.

Easterlin (1995) acknowledges that individuals with higher income are happier. However, he also proposes that the impact of actual income cannot be maintained for a long time and emphasizes the importance of individuals' comparisons to other people. Income increases do not increase happiness because individuals' judgment of SWB is based on a comparison of their own status to others in society as a whole. If one person's income increases substantially while others' income stays at the same, he would be happier than before. On the contrary, if his income remains identical while others increase substantially, he would feel less happy. Therefore, SWB is positively influenced by respondents' own income, while being negatively influenced by the comparison group. If the income of all individuals increases, individual's SWB does not increase, since the positive effect of his own income on SWB is offset by the negative effect of the general growth in income.

Luttmer (2005) utilizes a comparison of income between respondents and their neighbours to analyse social comparisons and SWB. He suggests that individuals feel worse off when their neighbours earn more. He utilizes the U.S. National Survey of Families and Households data, which contains life satisfaction measures, and explores how the economic status of neighbours influences respondents' SWB. He finds that it is negatively influenced by neighbours' higher earnings, when controlling for respondents' own income. Furthermore, he also investigates the reason why neighbours' higher income lowers respondents' SWB. He finds that the negative impact of neighbours' income is most likely to be ascribed to relative consumption patterns. If one individual were surrounded by neighbours' with higher incomes, he would have more disagreements

about goods, since his consumption behaviour is influenced by neighbours' consumption patterns and he may generate a larger consumption desire that exceeds his actual income.

Graham & Felton (2005) rely on a dataset from Latin America to study the effects of inequality, which is related to race, income, and the interviewer's judgment of socio-economic status (SES) on SWB. The authors regard income as a proxy for wealth and decompose income into average income and relative income. Average income is calculated for each city within each country. They define the relative income as the discrepancy between an individual's own income and the average income. The difference in relative income matters for SWB. When holding individual incomes constant, the average wealth (average income) of one country has no significant correlation with SWB, whilst the discrepancy of the individuals' own income and average income shows great contribution to SWB. To be specific, the relative income (discrepancy between own income and average income) negatively influences SWB for those individuals with below average income levels. However, for the wealthy, the relative income has a positive impact on their SWB. They are happier with the higher difference between their own income and average income.

Similar to Graham & Felton (2005), Vendrik & Woltjer (2007) also analyse SWB, relying on the comparison between individual income and average income in the whole society. They propose a hypothesis that individuals' life satisfaction is negatively related to household income in the reference group, holding household size, composition and other relevant variables constant. They hypothesise that life satisfaction operates primarily through consumption and is determined by household income rather than individual income. Household income is, therefore, a more adequate determinant of satisfaction. Another relevant hypothesis indicates that life satisfaction relies positively on the difference between respondents' household income and the average income of a reference group (relative income), controlling for household income (absolute income). The findings reveal that SWB is influenced by both household income and relative income but largely influenced by household income.

Unlike the previously reviewed research, McBride (2001) utilizes a different method to define relative income. He asks respondents to compare their current status with their parents' situation at the same age and defines the difference between the

parents' standard and individuals' own status as relative income. Using 1994 General Social Survey (GSS) data, which is an annual US household survey conducted by the National Opinion Research Centre, McBride (2001) investigates the presence of relative-income effects on life satisfaction and draws a similar conclusion, namely that relative income contributes greatly to SWB. The finding shows that an increase in an individuals' own income influences SWB positively for a short period, while the difference between the parents' standard and individuals' current status reduces the levels of SWB. The size of the difference can lower individuals' well-being. In addition, he also finds that the strength of the effects on SWB change with an increase in respondents' income over time. The effects of relative income are stronger at higher income levels, while the effect of relative income on SWB is smaller at lower income levels.

By reviewing previous literature, I find that most researchers define relative income by social comparison with others, including neighbours, friends, colleagues, their parents, and the average income in society. A few researchers utilize a different method to define relative income as aspiration income: Easterlin (2001) and Stutzer (2004). Easterlin (2001) finds that material aspirations are fairly similar across individuals at the start of the lifecycle; therefore, a higher income can bring greater happiness in the beginning. However, with increasing income, individuals' material aspirations grow correspondingly, which offset the favourable effect of income growth on SWB. SWB increases with income growth, while it decreases with increasing aspirations. Although the income growth guarantees individuals' ability to enjoy more products, the favourable effect of these goods on SWB can be offset by the desire to consume more over the lifecycle.

Similar to Easterlin (2001), Stutzer (2004) also considered the impacts of income aspirations on SWB. Firstly, he investigates the respective effect of actual income and aspiration income on SWB. The findings show that respondents' SWB is negatively influenced by their income aspirations, and the positive effect of income growth on SWB will be offset by the negative effect of aspiration income, which is similar to Easterlin (2001). In addition, Stutzer (2004) finds the positive and negative effect of income and aspirations on SWB are similar in absolute size and asks whether the equal increase of income and aspiration can counterbalance the respective impacts on life satisfaction. He establishes the discrepancy between aspiration income and actual income

to estimate the effect of people's aspirations relative to their actual income on life satisfaction further. The findings show that SWB is unrelated to income per se but depends on the discrepancy between aspiration income and actual income. Specifically, the larger this gap, the lower SWB. Furthermore, he also analyses the factors that would determine the absolute aspiration income, and states that absolute aspiration income is systematically influenced by the average income within the community. Richer residents around respondents can stimulate higher aspirations, and the effect of aspiration income has a greater influence for those individuals who often interact within the community.

4.2.3 Socio-demographic Factors

Apart from income, SWB is also influenced by other socio-demographic factors. In this section, I also review studies on the impact of socio-demographic factors, in order to help choose suitable control variables in this chapter. Oswald (1997), Blanchflower & Oswald (2004) find a U-shaped relationship between age and happiness. To be specific, happiness is lowest around the age range of 45 to 50 and higher for younger and older individuals. The reason for this may be that individuals whose age is around 40 have to give up their leisure time to earn more money in order to support their family. On the other hand, the reason why the elderly is happier is ascribed to the fact that they have lower job and wage aspirations and more experience to cope with negative life events. For young people, most have good health. In addition, the young, after graduating from university, are less likely to have a child, therefore they have fewer financial burdens. Furthermore, young people have larger career ambitions, so that they are enthusiastic about work and life. Therefore, the young and the elderly feel happier than the middle-aged. Dolan et al. (2008) and Frey & Stutzer (2010) also suggest the U-shaped curve between age and SWB.

The relationship between marital status and SWB (Helliwell, 2003; Dolan et al., 2008; Frey & Stutzer, 2010) indicates that married individuals are happier than those who have never married, are divorced or have been widowed. Marital status has a positive effect on SWB because couples can be affected positively by each other. Moreover, marriage can provide benefits, including sharing financial burdens and psychological pressures. Stutzer & Frey (2006) found a similar positive correlation

between marriage and SWB . Their results reveal that married individuals are happier or feel more satisfied with their lives. The single individuals were divided into two groups, where the first group involved individuals who stayed single throughout the research period and the other group comprised persons who were married during the research period. Stutzer & Frey (2006) also found that singles married during the survey period reported higher levels of satisfaction than persons who stayed single for a long time. In addition, marriage can provide psychological support and enable couples to share stress with each other, in line with Frey & Stutzer (2010).

Regarding the effect of children, Dolan et al. (2008) point out that the impact on SWB is different across countries and studies. Some studies suggest the positive relationship between children and SWB in general (Lelkes, 2006). However, Alesina et al. (2004) indicate that children may be an additional financial burden for the poor family. Thus, having children seem to affect SWB negatively for these poor family. Some studies find that educational attainment positively influence individuals' SWB (e.g. Inglehart, 1990; Blanchflower & Oswald, 2004). They suggest that highly educated individuals report higher SWB, since they can get highly skilled jobs with better conditions. Nevertheless, some studies have a contrary finding that well educated people have higher aspiration levels; therefore the highly educated individuals may suffer more pressure on work and feel more frustrated than their less educated counterparts when they do not obtain their expected achievement.

Regarding the impact on health, Dolan et al. (2008) and Frey & Stutzer (2010) find a high correlation between health status and SWB, which is probably due to the fact that people with poor health spend more money treating their illness and improving their health status. Some people with poor health can only work in lower work load jobs with lower wages, as their health status cannot support them in bearing a large work load. Shields & Price (2005) also state that several health problems such as heart attacks and some chronic disease lower individuals' SWB, in particular the elderly person.

4.2.4 Subjective Well-being (SWB) in China

There is limited research on the determinants of SWB in China. In this section, I will review existing Chinese studies and discuss their limitations before delineating my

own contribution to the literature. Using China Household Income Project, Jiang et al. (2012) investigate how the between-group inequality (BI) that refers to the social inequality and is measured by the income gap between rural-urban migrants and urban residents, influence SWB. In addition, they divide sample into three groups: urban residents with urban *hukou*, rural-urban migrants without urban *hukou*, and people who were born as rural residents and acquired urban *hukou* at certain point of time after migrating to urban city. They examine the attitude to between-group inequality of all individuals among these three groups. Firstly, they find that urban residents report higher levels of happiness than rural-urban migrants without urban *hukou* status, to a small extent. Secondly, the inequality related to *hukou* status lowers SWB, regardless of the residents having local urban *hukou* (urban residents) or not (rural migrants) (Jiang et al., 2012). Lastly, all the people in their sample feel unhappy with the between-group *hukou* identity-related inequality [19]. Rural-urban migrants feel more dissatisfied regarding *hukou* identity-related inequality, since they belong to the disadvantaged group. Those urban residents with local urban *hukou* who are primarily born in rural areas and acquire urban *hukou* subsequently have similar attitudes towards the *hukou* identity-related inequality with rural migrants without local urban *hukou*, since they may be discriminated against in the labour market before they obtain urban *hukou* identity.

Knight et al. (2009) analyse the economic determinants of happiness in rural China, based on a household survey organized by the Institute of Economics, Chinese Academy of Social Sciences on income distribution in China. The results suggest that relative income between villages and relative income over time are significant for current satisfaction, while current income has a smaller impact. In addition, current mood, attitudes, and social comparisons[20] reveal significant and positive associations with SWB. Regarding the social comparison, all respondents evaluate their household income and living standard by comparing themselves with the average standard for their village. The answers are classified to much below average, below average, average, above average and much above average level. Although rural residents have lower

---

[19] Inequality is attributed to *hukou* system.
[20] Social comparison in Knight et al. (2009) refers to that individuals evaluate their socio-economic status by the relative positions, depending on the objective criteria such as individual income, household income, wealth, occupation and education attainment. In general, respondents are more likely to compare their income, education level, wealth and positions with their close friends, neighbours, fellows, relatives and others in the same community to make self-assessment.

wages in comparison to their urban counterparts, it is uncertain that these disadvantaged rural people feel less happy. Even if most rural residents have lower income, lower socio-economic status and experience worse living conditions, they are found to confine their reference group to others in the village, since they have limited information regarding other situations (Knight et al., 2009).

Akay et al. (2012) examine the impact of relative income on rural-urban migrants' SWB. Using the Rural-to-Urban Migration in China and Indonesia (RUMICI) dataset, they focus on the role of different reference groups' relative income on rural-urban migrants' SWB. They find that rural-urban migrants' happiness is negatively influenced by the income of other rural-urban migrants, and also by the income of rural workers in rural areas. The latter result is likely to be attributed that rural *hukou* individuals migrate to cities to improve their financial status. Secondly, the higher relative income of other migrants may evoke the feelings of envy and competition. On the contrary, Akay et al. (2012) find positive relative income effect when they regard urban workers as reference group. The first reason for the positive relation is attributed to that most rural-urban migrants regard the relative income of urban workers as indicator of the benefits and higher standard public services of living in wealthy areas and with rich neighbours. In other words, the relative income of urban workers is taken as an indicator of better conditions. In addition, rural-urban migrants regard the relative income of urban *hukou* individuals as an informative signal of their future income, which in turn increases their SWB.

In addition, there is existing but limited research investigating the influence of respondents' aspiration levels on SWB in China (Knight & Gunatilaka, 2010a; Knight and Guantilaka, 2012b). Both Knight & Gunatilaka (2010a) and Knight & Gunatilaka (2012b) investigated the reason why rural-urban migrants reported lower average levels of happiness despite their higher mean income, in comparison with rural residents. They argue that rural-urban migrants are likely to overestimate the quality of urban life, which causes their expectations of living conditions and working environment to be biased upwards. In addition, migrants fail to foresee how their aspirations will change as their living or working environment changes. One of the reasons why unhappy rural-urban migrants do not return to their rural origins is that a majority of migrants expect their income will rise over the subsequent five years. Therefore, rural-urban migrants

experience lower levels of happiness, although they have better living and working conditions, in comparison with rural residents.

These existing Chinese studies, like Knight et al. (2009) and Knight & Gunatilaka (2010b), focus on the determinants of rural *hukou* individuals' SWB from factors such as individual income, household income and SES based on comparison. In this chapter, I compare the SWB between rural-urban migrants and urban *hukou* residents, and also consider how income aspirations shape the SWB between these two groups. As discussed in Chapter 3, rural-urban migrants have lower income on average report lower absolute income aspirations. Although urban workers have high quality jobs with higher wages, they are more likely to have higher income aspirations as well. Therefore, is not clear that urban residents are happier than rural-urban migrants because of their higher income aspirations. In Chapter 4, I intend to follow Stutzer (2004) to explore how aspiration income influences SWB and test whether there is a difference in the impact of income aspirations on happiness between urban residents and rural-urban migrants.

4.3 Data and Hypotheses

4.3.1 Data Description and Summary Statistics

The data used in this chapter is the China General Social Survey (CGSS) 2010, which have been utilized in Chapter 3 as well. Similar to Chapter 3, the dummy variable of *hukou* status and information of survey location can be used to categorize the respondents into urban people and rural-urban migrants. In terms of dependent variable SWB, all the respondents are asked "In general, do you think you are happy?". The ordinal variable of happiness is measured on a 5-point scale: 1-not at all happy; 2-not happy; 3-neutral; 4-happy; 5-very happy. About three-quarters of respondents expressed being happy and very happy. The estimation sample has 3019 observations, consisting of 2076 urban *hukou* people and 943 rural-urban migrants.

Similar to Chapter 3, aspiration income also comes from those two questions "Are you satisfied with your current yearly income, based on your educational background and other factors?" and "If not satisfied, what do you think you should have earned based on your educational attainment, working experience and other factors?". The

detailed measurement and definition of income aspirations have been explained in Chapter 3. Income is a leading factor in most economic research. Regarding the definition of income, there are two relevant variables: total income and occupational income per year. Since aspiration income in this survey refers to individuals' total income, I chose the total income as the measure of income, in order to keep aspiration income and absolute income consistent. Similar to Chapter 3, I also dropped the top 1% and bottom 1% of total income and income aspirations from the sample to guarantee even income and aspirations distribution. The mean value of total income and aspiration income after removing the top 1% and bottom 1% from the sample was 22,749 yuan and 42,946 yuan. The distributions of total income and aspiration income are shown in Figure 4-1 and Figure 4-2.

Figure 4- 1 Distribution of Total Income
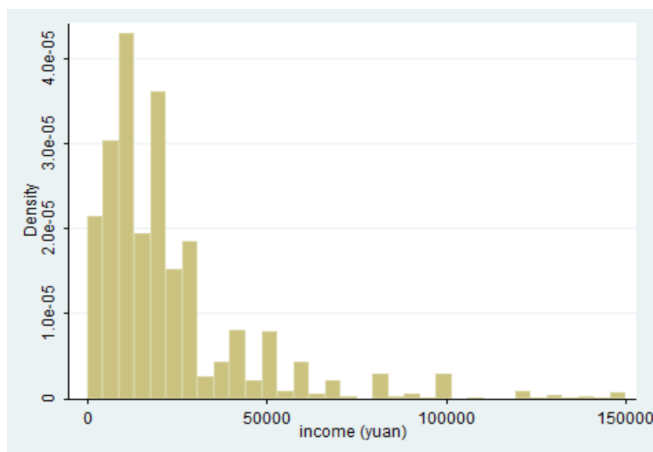
Figure 4- 2 Distribution of Aspiration Income
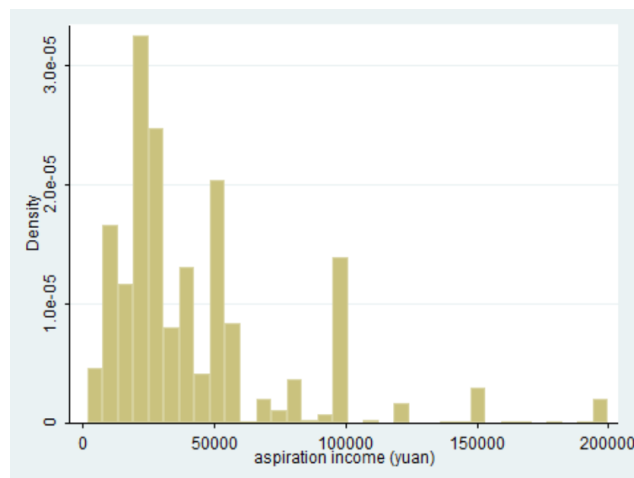
Since this chapter investigates the well-being of people in the urban labour market, I

also took individuals' working situations into account and dropped 735 individuals who had never worked before. After making this adjustment, the sample only includes individuals who were currently employed or those currently unemployed but who had had an occupation previously. The reason for including this latter group is that these individuals also provided information in terms of their income from their previous job and aspiration income for their new jobs.

Regarding the socio-demographic factors, age, number of children, marital status, gender, educational attainment and health are taken into account (Frey & Stutzer, 2010). Given the focus on income aspirations, the sample only chooses respondents aged (16 to 60 for male and 16 to 55 for female). The variable of number of children comes from the question, "How many sons and daughters do you have, respectively?". The number of sons and daughters separately are added as the total number of children. Two thirds of individuals had just one child, which is likely a result of the Chinese one child policy. The marital status is a categorical variable with 6 categories (unmarried, married, living together but not married, divorced, widowed, and separated) and I combined some categories in order to transfer this into a dummy variable. Similar to Chapter 3, people living together are regarded as married, whilst divorced, those separated but not divorced and widowed individuals are treated as unmarried. The variable "highest education level" is a categorical variable which is divided into multiple dummy variables. As with Chapter 3, this categorical variable of "highest education level" is divided into four dummy variables: no educational attainment, compulsory education, senior high school and university. The same to previous chapter, health status is an ordinal variable with 5-point scale: 1-completely unhealthy, 2-unhealthy, 3-standard, 4-healthy, 5-completely healthy. I restructured this ordinal variable into three dummy variables, which are good health (healthy and completely healthy), standard health and bad health status (unhealthy and completely unhealthy). In the sample, more than 50% of individuals reported that they had good health in general (healthy and completely healthy). Approximately 8.6% of individuals reported that they had bad health (unhealthy and completely unhealthy), and 23% considered their health situation to be 'standard'. Individual health status is included in the analysis as it can influence individual happiness directly. Individuals with poor health are less likely to feel happy, especially the elderly (Frey & Stutzer, 2010).

As shown in Table 4-2, the mean value of happiness of rural-urban migrants is similar to urban workers. Approximately 77.02% of urban individuals reported being very happy and happy, while 77.09% of rural-urban migrants reported the same. In terms of respondents' actual income, urban individuals have higher actual income (24721.63 yuan) than rural-urban migrants (18405.12 yuan).

Table 4- 1 Description of Variables

Happiness	1-	Not at all happy;
	2-	Not happy;
	3-	Neutral (between happy and unhappy);
	4-	Happy;
	5-	Very happy
Rural	1-	Rural-urban migrants
	0-	Urban *hukou* workers
Actual income		Real total income per year
Aspiration income		Individuals' aspirations on income
Partner income		Income per year of respondents' partner
Age		Individuals within the range of working age (16-60 for male and 16-55 for female respondents)
Children		Number of children for each individual
Male	1-	Male
	0-	Female
Married	1-	Married
	0-	Unmarried
Non-educational attainment	0-	No education
	0-	otherwise
Compulsory school	1-	Highest education level is compulsory school (primary or junior middle)
	0-	Otherwise
Senior high school	1-	Highest education level is senior level (professional high school, senior high school and technical school)
	0-	Otherwise
University	1-	Highest education level is university
	0-	Otherwise
Good Health	1-	Good health
	0-	Otherwise
Standard Health	1-	Standard health status
	0-	Otherwise
Poor Health	1-	Poor healthy
	0-	Otherwise

Note: China General Social Survey 2010 is utilized.

However, urban *hukou* respondents have correspondingly higher absolute aspirations (46215.31 yuan) than their rural-urban counterparts (35749.73 yuan) as well. Considering the impact of socio-demographic factors, the average value of age, marital status, gender distribution, was similar between urban *hukou* workers and rural-urban migrants. However, rural-urban migrants are healthier than their urban counterparts, in general. Regarding educational attainment, only 3% of rural-urban migrants graduated

from university, in comparison to 34% of urban people. Moreover, 33% of urban *hukou* workers only received compulsory education attainment while more than 70% of rural-urban migrants gave up further study when they finished the compulsory education.

Table 4- 2 Summary Statistics for Urban People and Rural Migrants

Variables	Mean	Std. Dev	Mean	Std. Dev
	Urban (2076 observations)		Rural (943 observations)	
Happiness				
Not at all happy	0.01	0.11	0.01	0.10
Not happy	0.05	0.23	0.06	0.23
Neutral	0.16	0.37	0.16	0.37
Happy	0.60	0.49	0.59	0.49
Very happy	0.17	0.38	0.18	0.39
Independent variables				
Income categories				
Actual income	24721.63	24021.63	18405.12***	17663.21
Aspiration income	46215.31	36433.38	35749.73***	29073.31
Partner income	23259.42	33146.65	12347.33***	17078.31
Social demographic factors				
Age	43.45	9.65	39.75***	8.79
Age^2	1980.65	834.37	1657.34***	708.34
Number of children	1.20	0.61	1.61***	0.82
Male	0.50	0.50	0.62***	0.49
Married	0.97	0.05	0.96	0.07
Health status				
Good health	0.65	0.47	0.74***	0.44
Standard health	0.25	0.48	0.19***	0.40
Poor health	0.10	0.29	0.07**	0.38
Educational attainment				
Non-education	0.01	0.14	0.04***	0.21
Compulsory education	0.33	0.47	0.66***	0.44
Senior school	0.32	0.46	0.27***	0.44
University level	0.34	0.47	0.03***	0.18

Note: (1) * in the column of rural-urban migrants refers to whether the mean value of each variable is significantly

different with mean value of urban *hukou* individuals. *, **. *** indicate 10%, 5%, 1% significant levels.

    (2) The sample of this chapter is slightly different with the sample of Chapter 3 shown in Table 3-2. Although these two chapters involve the same variable (income aspiration), as mutually independent chapters they have different research topics. Chapter 4 analyses subjective well-being, which is not considered in Chapter 3. Thus, the dependent variable is different in each case. Moreover, some Chinese individuals are reluctant to answer questions related to personal income, personal or family property, family background and personal attitudes. The subjective well-being and income aspirations depend on respondents' self-evaluation and their personal attitude. Therefore, some interviewees are not willing to answer questions related to happiness in CGSS 2010. Many reasons may be adduced for respondents being reluctant to respond to questions on subjective well-being, such as economic reasons, special experiences, physical health, and so on. In addition, as these two chapters have different research topics, the corresponding independent variables are slightly different, as explained above. In the preceding chapter, which investigates the determinants on income aspirations, some respondents were unwilling to report any information related to family background, e.g., father's education, family property, and so on. This may be attributed to parental divorce, paternal death, and/or disharmonious family relationships from childhood. Thus, due to the different purposes of and variables at work in the two chapters, there is a slight difference between the two samples.

    (3) China General Social Survey 2010 is utilized.

4.3.2 Hypotheses and Corresponding Specifications

    Considering the similar SWB between urban individuals and rural migrants shown in Table 4-2, this chapter aims at explaining this issue from the impact of income aspirations on SWB. As is well-known, SWB not only depends on individuals' own income, but also replies on income aspirations or income of the reference group that may be the average income of all individuals in the same group or the income of respondents' friends, neighbour, and colleagues. Moreover, income aspirations and the income of the reference group are more likely to have a negative impact on SWB. Some studies suspect whether the positive and negative coefficient for absolute income and aspiration income (or income of reference group) have a similar magnitude, and if the increase of actual income and aspiration income may counterbalance the respective impact on SWB. Therefore, in addition to analysing the respective effect of actual income and income aspirations (or income of the reference group) itself on happiness, these studies also put forward the discrepancy between income aspirations (or income of reference group) and actual income as relative income to estimate its impact on SWB further.

    In general, when these studies investigate the impact of aspirations or other people's income on respondents' SWB, the usual estimation process includes studying the effect of absolute aspirations or income of a reference group (e.g., the average

income of all individuals in the same group or income of respondents' friends, neighbour, relatives and colleagues) on SWB in the first step. Secondly, when controlling the respondents' actual income, the discrepancy between income aspirations (or income of the reference group) and actual income is established to further estimate the impact of relative income on SWB (Stutzer, 2004; Ferrer-i-Carbonell, 2005; Graham & Felton, 2005; and Vendrik & Woltjer, 2007). For example, when Stutzer (2004) investigates the role of income aspiration on individuals' happiness, the author also estimates the actual income and absolute income aspirations on SWB first. Following aspiration level theory, an individual's subjective well-being is determined by the difference between aspiration and achievement, as well (Michalos, 1991; Inglehart, 1990); thus Stutzer (2004) studies the income aspirations relative to the actual income (discrepancy between aspiration income to actual income) on respondents' happiness further. Consequently, following these previous studies, the empirical analysis in this chapter will be based on two different specifications to test for various hypotheses regarding the effect of income and aspirations on SWB.

The first simple specification is one that includes absolute actual income and aspiration income as a determinant of SWB. A common assumption in economics is that individuals' actual income has a positive association with SWB (Veenhoven, 1991; Diener et al., 1993). However, as shown in Table 4-2, the similar SWB is found between urban people and rural migrant although urban workers report higher actual income (24721.63 yuan) in comparison to rural-urban migrants (18405.12 yuan). That is inconsistent with the correlation between SWB and income suggested by previous studies.

Firstly, a possible reason may be ascribed to differences in the effect of actual income on SWB between urban workers and rural-urban migrants, holding other variables constant. If actual income matters more for rural-urban migrants' SWB, an additional actual income would contribute more to rural-urban migrants' happiness. In contrast, a 1% increase of actual income contributes less to urban individuals' SWB. Consequently, higher actual income may present a positive effect on urban respondents' happiness, to the same extent as lower income affects rural-urban migrants' SWB. The direction of the coefficients for an interaction term between the dummy variable of rural status and actual income can be checked in order to explore whether the impacts of

additional actual income for rural-urban migrants are greater than that for urban ones. I propose the first hypothesis.

Hypothesis 1: There is a differential effect of actual income on SWB for rural-urban migrants with a given increase in income generating larger gains in SWB for rural migrants, in comparison with urban *hukou* individuals. Thus, an interaction term between actual income and a dummy variable for rural-urban migrants should be positive.

In terms of partner income, Kahneman & Deaton (2010) and Brown & Gray (2014) find a positive relationship between household income and SWB. When holding individuals' actual income constant, partner income can positively influence SWB. As shown in Table 4.2, the partner's income for urban individuals (23259.42 yuan) is twofold in comparison to rural-urban migrants (12347.33 yuan); therefore, urban workers should report higher levels of SWB. However, the happiness of urban workers is similar to rural-urban migrants. Following Hypothesis 1, the reason for the similar happiness levels may be attributed to the fact that partner income matters more for rural-urban migrants' happiness levels. An additional partner income can contribute more to rural-urban migrants' SWB; hence, an interaction term between the dummy variable of rural *hukou* status and partner income is also generated to explore the difference in the impact of partner income between urban workers and rural-urban migrants.

Hypothesis 2: There is a difference in the impact of partner income on SWB for rural-urban migrants with a given rise in partner income generating larger gains in SWB for rural migrants compared with urban residents. Thus, an interaction term between partner income and a dummy variable for rural migrants should be positive.

Another reason why urban respondents with higher incomes have similar happiness to rural-urban migrants may be due to some factors offsetting the positive effect of actual income or partner income on SWB. Stutzer (2004) suggests that individuals with higher incomes are more likely to have higher aspirations regarding their current status, and he also reveals that aspiration income has a negative effect on SWB. The negative correlation between aspiration and SWB can offset the increasing

happiness that is brought by higher actual income or partner income. As can be seen from Table 4-2, even though urban workers have a higher income and higher partner income, they also report higher absolute aspiration income (46215.31 yuan) than rural-urban migrants (35749.73 yuan). I, therefore, propose a third hypothesis to explore whether similar happiness can be attributed to aspiration income offsetting the positive effect of actual income or partner income on SWB.

Hypothesis 3: Aspiration income has a negative impact on SWB.

Following hypothesis 2 and hypothesis 3, I generate an interaction term between income aspirations and a dummy variable of rural-urban migrants, and expect that rural-urban migrants are less affected by income aspirations. Knight & Gunatilaka (2010b) have suggested that urban *hukou* individuals are more materialistic and have more requirements, in comparison to rural-urban migrants. Thus, urban *hukou* individuals are more likely to have higher aspirations and care more about whether their aspirations can be satisfied. If the aspiration income has a negative effect on SWB and matters less for rural-urban migrants, the difference in SWB between urban *hukou* individuals and rural-urban migrants can be reduced, and the similar SWB in this chapter can be explained. I propose the fourth hypothesis.

Hypothesis 4: There is a differential effect of aspiration income on SWB for rural-urban migrants with a given increase in aspiration income leading to a smaller loss in SWB for rural-urban migrants, in comparison with urban residents. Thus, an interaction term between aspiration income and a dummy variable for rural migrants should be positive.

The description above shows the first specification to test the impact of actual income and aspiration income itself on SWB, respectively. In addition, as discussed by Michalos (1991) and Ingelhart (1990), subjective well-being is also determined by the gap between aspiration and achievement. Therefore, I follow Stutzer (2004) in considering the effect of aspiration income relative to individuals' actual income. The second specification estimates how aspiration income influences SWB in the form of relative status, i.e., how the discrepancy between individual aspiration income and actual income determines SWB. This can be conducted by taking the difference between the logarithm of each respondent's income aspirations and the logarithm of their actual

income, i.e. ln (aspiration income)-ln (actual income). I found that, although urban individuals report higher absolute income aspirations, the gap between their aspirations and actual income is slightly lower than that for rural-urban migrants. The reason for this may be attributed to that actual income of rural-urban migrants is too low to reach the normal standard they deserve. The actual income of rural-urban migrants is negatively restricted by *hukou* discrimination, whilst the findings of Chapter 3 show that their income aspirations are not negatively influenced by *hukou* discrimination. Therefore, the discrepancy between rural-urban migrants' income aspirations and actual income (ln (aspiration income)-ln (actual income)) is larger. As suggested by Stutzer (2004), the greater discrepancy between aspirations and income, the lower subjective well-being. Thus, following Stutzer (2004), I assume that the discrepancy between aspiration income and actual income is negatively associated to people's SWB.

However, as rural-urban migrants who experience larger discrepancy between aspirations and actual income report the similar happiness with their urban counterparts, I expect that the negative impact of this gap on rural migrants is far less than that for urban individuals. As only a 1% increase in the discrepancy has a greater negative impact on urban people, the similarity in happiness levels between rural migrants and urban people can be explained. Therefore, following hypothesis 4, I generated an interaction term between the dummy variable of rural-urban migrants and the gap between income aspirations and actual income, expecting that rural-urban migrants care less about this distance between their aspirations and what they actually earned. Based on description above, hypothesis 5 and hypothesis 6 are established as follows.

Hypothesis 5: The gap between aspiration income and actual income is negatively associated to people's SWB.

Hypothesis 6: There is a differential effect of this gap on SWB for rural-urban migrants with a given increase in the discrepancy between individuals' aspiration income and actual income, generating a quite smaller loss in SWB for rural-urban migrants, in comparison to urban *hukou* individuals. Thus, an interaction term between the dummy variable for rural-urban migrants and the discrepancy between income aspirations and actual income should be positive.

4.4. Methodology

The ordered probit model will be used to estimate the effect of actual income and aspiration income on happiness. Following Long & Freese (2006), the ordered probit model can be derived from a model in which a latent variable y^* ranging from $-\infty$ to ∞ is mapped to an observed variable y. The dependent variable (SWB) is thought of as providing incomplete information about an underlying y^* according to the following equation:

$$y_i = happiness = m \quad if \quad \tau_{m-1} \le y_i^* < \tau_m \quad \text{for m=1 to 5} \tag{4-1}$$

The $\tau's$ are thresholds and the extreme categories 1 and 5 are defined with $\tau_0 = -\infty$ and $\tau_5 = \infty$. The dependent variable 'happiness' is measured on a 5-point scale: 1-not at all happy; 2-not happy; 3-netural; 4-happy; 5-very happy. Then the observed y (happiness) is related to y^* according to the following equation:

$$y_i = happiness_i = \begin{cases} 1 & if & \tau_0 = -\infty \le y_i^* < \tau_1 \\ 2 & if & \tau_1 \le y_i^* < \tau_2 \\ 3 & if & \tau_2 \le y_i^* < \tau_3 \\ 4 & if & \tau_3 \le y_i^* < \tau_4 \\ 5 & if & \tau_4 \le y_i^* < \tau_5 = \infty \end{cases} \tag{4-2}$$

The latent variable y^* is specified as:

$$y_i^* = X_i \beta + \varepsilon_i \tag{4-3}$$

where the matrix **X** contains actual income, partner income, aspiration income and a set of control variables including age, number of children, male, married, rural *hukou* status, medical and pension insurance, debt and highest educational level. In addition, I also take the interaction terms between rural *hukou* dummy variable and actual income, and between partner income and aspiration income into account when estimating the different effect on SWB between urban people and rural-urban migrants. The reason why I pool urban and rural migrants together and generate an interaction term between rural dummy variable with income aspiration, actual income, partner income is attributed to that I cannot only see the difference effect of actual income or aspiration income on happiness between urban people and rural migrants, but also can clearly

observe if the differential impact is significant to SWB. This advantage of generating interaction terms is that regression for urban people and rural-urban migrants separately cannot be achieved[21]. $\boldsymbol{\beta}$ is a column vector of coefficients and ε_i is the error term which is a standard normally distributed constant (Long & Freese, 2006). For the first specification, the latent variable can be expressed specifically as:

$$y_i^* = happiness_i^* = \beta_{0i} + \beta_{1i} \ln W^{inc} + \beta_{2i} \ln W^{part} + \beta_{3i} \ln W^{aspir} + \beta_{4i}R + \beta_{5i}R \ln W^{inc}$$
$$+ \beta_{6i}R \ln W^{part} + \beta_{7i}R \ln W^{aspir} + \beta_{8i}Z + \varepsilon_i$$

$$(4\text{-}4)$$

where R is the dummy variable referring to rural *hukou* status. W^{inc}, W^{part} and W^{aspir} are actual income, partner income and aspiration income, respectively. In order to express the equation (4-4) briefly, I used Z to represent all the remaining control variables.

The probability that a random variable is between two values is the difference between the cumulative distribution function at these values. The probability of any observed outcome $y=m$, given **X,** is:

$$pr(y_i = m \mid X_i) = pr(\tau_{m-1} \le y_i^* < \tau_m \mid X_i)$$
$$= pr(\tau_{m-1} - X_i\boldsymbol{\beta} \le \varepsilon_i < \tau_m - X_i\boldsymbol{\beta} \mid X_i)$$
$$= pr(\varepsilon_i < \tau_m - X_i\boldsymbol{\beta} \mid X_i) - pr(\varepsilon_i \le \tau_{m-1} - X_i\boldsymbol{\beta} \mid X_i)$$
$$= F(\tau_m - X_i\boldsymbol{\beta}) - F(\tau_{m-1} - X_i\boldsymbol{\beta})$$

$$(4\text{-}5)$$

I use F to represent the cumulative distribution function and f to represent the probability density function of ε_i (Wooldridge, 2009).

Interpretation of the modelling results is carried out by computing the partial change (marginal effect) or discrete change (incremental effect) for the probabilities resulting from marginal/discrete changes in one of the independent variables. The partial derivative of any observed outcome, $y=m$, with respect to X_k is given by:

---

[21] I have tried the regression for urban people and rural-urban migrants separately in the Table 4-13 of Appendix.

$$\frac{\partial pr(y_i = m \mid \boldsymbol{X}_i)}{\partial x_k} = \frac{\partial F(\tau_m - \boldsymbol{X}_i\boldsymbol{\beta})}{\partial x_k} - \frac{\partial F(\tau_{m-1} - \boldsymbol{X}_i\boldsymbol{\beta})}{\partial x_k}$$

$$= \beta_k f(\tau_{m-1} - \boldsymbol{X}_i\boldsymbol{\beta}) - \beta_k f(\tau_m - \boldsymbol{X}_i\boldsymbol{\beta}) \qquad (4\text{-}6)$$

$$= \beta_k \left[f(\tau_{m-1} - \boldsymbol{X}_i\boldsymbol{\beta}) - f(\tau_m - \boldsymbol{X}_i\boldsymbol{\beta}) \right]$$

The marginal effect is the slope of the curve relating X_k to $Pr(y=m/k)$, holding all other variables constant. However, the interpretation of a dummy variable is very different. In this case, a measure of discrete change or incremental effect is much more informative. The discrete change for a dummy variable, x_k, is the change in the probability resulting from a change in x_k from 0 to 1.

$$\Delta pr(y_i = m \mid \boldsymbol{X}_i) = pr(y_i = m \mid \boldsymbol{X}_i, x_k = 1) - pr(y_i = m \mid \boldsymbol{X}_i, x_k = 0) \qquad (4\text{-}7)$$

The predicted probability of outcome, m, changes by $\Delta pr(y_i = m \mid \boldsymbol{X}_i)$, holding all other variables constant, while the independent variable, x_k, changes from 0 to 1.

As explained earlier, I also include interaction terms between rural *hukou* status and actual income, and between partner income and aspiration income in the model. According to Abdel-Aty (2003), Ai & Norton (2003), Norton et al. (2004), Drichoutis & Nayga Jr (2011), and Karaca & Mandic et al. (2012), the marginal effect of variables that are interacted is different. The marginal effect of the interaction term on the probability of outcome, m, can be calculated in two steps. Firstly, the correct marginal effect for the log of actual income on the probability of outcome *m*, is:

$$\frac{\partial pr(y_i = m \mid \boldsymbol{X}_i)}{\partial \ln W^{inc}} = (\beta_{1i} + \beta_{5i}R)\left[f(\tau_{m-1} - \boldsymbol{X}_i\boldsymbol{\beta}) - f(\tau_m - \boldsymbol{X}_i\boldsymbol{\beta}) \right] \qquad (4\text{-}8)$$

The correct marginal effect for the log of partner income on the probability of outcome, m, is:

$$\frac{\partial pr(y_i = m \mid \boldsymbol{X}_i)}{\partial \ln W^{part}} = (\beta_{2i} + \beta_{6i}R)\left[f(\tau_{m-1} - \boldsymbol{X}_i\boldsymbol{\beta}) - f(\tau_m - \boldsymbol{X}_i\boldsymbol{\beta}) \right] \qquad (4\text{-}9)$$

In addition, the correct marginal effect of log of aspiration income on the probability of outcome *m*, is:

$$\frac{\partial pr(y_i = m \mid \boldsymbol{X}_i)}{\partial \ln W^{aspir}} = (\beta_{3i} + \beta_{7i}R)\left[f(\tau_{m-1} - \boldsymbol{X}_i\boldsymbol{\beta}) - f(\tau_m - \boldsymbol{X}_i\boldsymbol{\beta}) \right] \qquad (4\text{-}10)$$

The correctly estimated discrete change for the dummy variable, R, (rural *hukou*) is:

$$\frac{\Delta pr(y_i = m \mid X_i)}{\Delta R}$$

$$= pr(y_i = m \mid X_i, R = 1) - pr(y_i = m \mid X_i, R = 0)$$

$$= F(\tau_m - \beta_{0i} - \beta_{1i} \ln W^{inc} - \beta_{2i} \ln W^{part} - \beta_{3i} \ln W^{aspir} - \beta_{4i} - \beta_{5i} \ln W^{inc} - \beta_{6i} \ln W^{part} - \beta_{7i} \ln W^{aspir} - \beta_{8i} Z)$$

$$\quad - F(\tau_{m-1} - \beta_{0i} - \beta_{1i} \ln W^{inc} - \beta_{2i} \ln W^{part} - \beta_{3i} \ln W^{aspir} - \beta_{4i} - \beta_{5i} \ln W^{inc} - \beta_{6i} \ln W^{part} - \beta_{7i} \ln W^{aspir} - \beta_{8i} Z)$$

$$\quad - F(\tau_m - \beta_{0i} - \beta_{1i} \ln W^{inc} - \beta_{2i} \ln W^{part} - \beta_{3i} \ln W^{aspir} - \beta_{8i} Z)$$

$$\quad + F(\tau_{m-1} - \beta_{0i} - \beta_{1i} \ln W^{inc} - \beta_{2i} \ln W^{part} - \beta_{3i} \ln W^{aspir} - \beta_{8i} Z)$$

$$\tag{4-11}$$

The correct marginal effect of the interaction term involves a mixture of equation (4-8) and (4-11), a mixture of equations (4-9) and (4-11), and the mixture of equations (4-10) and (4-11). That is, I first calculate the derivative with respect to actual income, partner income and aspiration income, and then calculated the difference in probability for a change in rural *hukou* from 0 to 1. The marginal effect of outcome *y=m* with respect to the interaction between the log of actual income and rural *hukou* is given by:

$$\frac{\Delta[\dfrac{\partial pr(y_i = m \mid X_i)}{\partial \ln W^{inc}}]}{\Delta R}$$

$$= \frac{\partial pr(y_i = m \mid X_i)}{\partial \ln W^{inc}}\Big|_{R=1} - \frac{\partial pr(y_i = m \mid X_i)}{\partial \ln W^{inc}}\Big|_{R=0}$$

$$= (\beta_{1i} + \beta_{5i})[f(\tau_{m-1} - \beta_0 - \beta_1 \ln W^{inc} - \beta_2 \ln W^{part} - \beta_3 \ln W^{aspir} - \beta_4 - \beta_5 \ln W^{inc} - \beta_6 \ln W^{part} - \beta_7 \ln W^{aspir} - \beta_8 Z)$$

$$\quad - f(\tau_m - \beta_0 - \beta_1 \ln W^{inc} - \beta_2 \ln W^{part} - \beta_3 \ln W^{aspir} - \beta_4 - \beta_5 \ln W^{inc} - \beta_6 \ln W^{part} - \beta_7 \ln W^{aspir} - \beta_8 Z)]$$

$$\quad - \beta_{1i}[f(\tau_{m-1} - \beta_0 - \beta_1 \ln W^{inc} - \beta_2 \ln W^{part} - \beta_3 \ln W^{aspir} - \beta_8 Z)$$

$$\quad - f(\tau_m - \beta_0 - \beta_1 \ln W^{inc} - \beta_2 \ln W^{part} - \beta_3 \ln W^{aspir} - \beta_8 Z)]$$

$$\tag{4-12}$$

The marginal effect of outcome *y=m* with respect to the interaction between log of partner income and rural *hukou* is:

$$\frac{\Delta[\dfrac{\partial pr(y_i = m \mid X_i)}{\partial \ln W^{part}}]}{\Delta R}$$

$$= \frac{\partial pr(y_i = m \mid X_i)}{\partial \ln W^{part}}\Big|_{R=1} - \frac{\partial pr(y_i = m \mid X_i)}{\partial \ln W^{part}}\Big|_{R=0}$$

$$= (\beta_{2i} + \beta_{6i})[f(\tau_{m-1} - \beta_0 - \beta_1 \ln W^{inc} - \beta_2 \ln W^{part} - \beta_3 \ln W^{aspir} - \beta_4 - \beta_5 \ln W^{inc} - \beta_6 \ln W^{part} - \beta_7 \ln W^{aspir} - \beta_8 Z)$$

$$\quad - f(\tau_m - \beta_0 - \beta_1 \ln W^{inc} - \beta_2 \ln W^{part} - \beta_3 \ln W^{aspir} - \beta_4 - \beta_5 \ln W^{inc} - \beta_6 \ln W^{part} - \beta_7 \ln W^{aspir} - \beta_8 Z)]$$

$$\quad - \beta_{2i}[f(\tau_{m-1} - \beta_0 - \beta_1 \ln W^{inc} - \beta_2 \ln W^{part} - \beta_3 \ln W^{aspir} - \beta_8 Z)$$

$$\quad - f(\tau_m - \beta_0 - \beta_1 \ln W^{inc} - \beta_2 \ln W^{part} - \beta_3 \ln W^{aspir} - \beta_8 Z)]$$

$$\tag{4-13}$$

The marginal effect of outcome *y=m* with respect of to the interaction between log of

aspiration income and rural *hukou* is:

$$\frac{\Delta[\frac{\partial pr(y_i = m \mid X_i)}{\partial \ln W^{aspir}}]}{\Delta R}$$

$$= \frac{\partial pr(y_i = m \mid X_i)}{\partial \ln W^{aspir}}\Big|_{R=1} - \frac{\partial pr(y_i = m \mid X_i)}{\partial \ln W^{aspir}}\Big|_{R=0}$$

$$= (\beta_{3i} + \beta_{7i})[f(\tau_{m-1} - \beta_0 - \beta_1 \ln W^{inc} - \beta_2 \ln W^{part} - \beta_3 \ln W^{aspir} - \beta_4 - \beta_5 \ln W^{inc} - \beta_6 \ln W^{part} - \beta_7 \ln W^{aspir} - \beta_8 Z)$$

$$- f(\tau_m - \beta_0 - \beta_1 \ln W^{inc} - \beta_2 \ln W^{part} - \beta_3 \ln W^{aspir} - \beta_4 - \beta_5 \ln W^{inc} - \beta_6 \ln W^{part} - \beta_7 \ln W^{aspir} - \beta_8 Z)]$$

$$- \beta_{3i}[f(\tau_{m-1} - \beta_0 - \beta_1 \ln W^{inc} - \beta_2 \ln W^{part} - \beta_3 \ln W^{aspir} - \beta_8 Z)$$

$$- f(\tau_m - \beta_0 - \beta_1 \ln W^{inc} - \beta_2 \ln W^{part} - \beta_3 \ln W^{aspir} - \beta_8 Z)]$$

$$(4\text{-}14)$$

As discussed in section 4.3.2., in addition to estimating the respective effect of actual income and absolute aspiration income itself on SWB, I also follow Stutzer (2004) to consider the impact of individual aspirations relative to actual income on SWB in the following analysis. Following the approach by Stutze (2004), Ferrer-i-Carbonell (2005), Graham & Felton (2005) and Vendrik & Woltjer (2007), this can be implemented by replacing aspirations with the discrepancy between aspiration income and actual income, i.e. ln (aspiration income)- ln (actual income).

$$Relative\ aspiration = \ln W^{aspir} - \ln W^{inc} \qquad (4\text{-}15)$$

According to the Eq. (4-4), the latent variable based on the aspiration income in the form of relative status can be transformed as:

$$y_i^* = happiness_i^* = \beta_{0i} + (\beta_{1i} + \beta_{3i}) \ln W^{inc} + \beta_{2i} \ln W^{part} + \beta_{3i}(\ln W^{aspir} - \ln W^{inc}) + \beta_{4i} R +$$

$$(\beta_{5i} + \beta_{7i}) R \ln W^{inc} + \beta_{6i} R \ln W^{part} + \beta_{7i} R(\ln W^{aspir} - \ln W^{inc}) + \beta_{8i} Z + \varepsilon_i$$

$$(4\text{-}16)$$

Eq. (4-16) separates the effect that is attributed to the difference between log aspiration income and log actual income and the pure effect of log of actual income. In comparison to the first specification shown in Eq. (4-4), the coefficient of discrepancy between aspiration income and actual income is equal to the coefficient of pure aspiration income. Moreover, the coefficient of actual income in Eq. (4-16) is equal to the sum of the respective coefficient of actual income and aspiration income in Eq. (4-4).

For all the marginal effects, I present the average of marginal effects.

4.5 Results and Discussion

4.5.1 Main Results

Regarding the impact of the respective effect of actual income and aspiration income on SWB in first specification, Table 4-3 presents the results of the ordered probit model and the marginal effects of reporting being 'very happy'. Columns A and B present the results of the basic model, which includes absolute actual income, income aspirations, and other basic control variables.

Table 4-3 shows that actual income has a significant and positive effect on SWB. The marginal effect of actual income in column B shows that a 1% increase in actual income leads to a 0.93 percent increase in the probability of "reporting being very happy", which is consistent with Veenhoven (1991) and Diener et al. (1993). In addition, the coefficient on partner income also shows a positive correlation with SWB. Based on the marginal effect of partner income in column B, a 1% higher partner income leads to a 1.23 percent greater probability of reporting the top category, in line with Kahneman & Deaton (2010) and Brown & Gray (2014). Consequently, the higher the income, the better basic needs can be satisfied, which in turn leads to higher SWB.

Regarding the differential effect of actual income between rural-urban migrants and urban workers, the marginal effect of the interaction term[22] shows positive and significant impacts on SWB. The positive relationship between the interaction term and happiness reveals that actual income matters more in terms of rural-urban migrants' happiness. As can be seen from Column B of Table 4-3, the marginal effect of interaction term between actual income and dummy variable of rural-urban migrants presents a positive effect on SWB. To be specific, with 1 % actual income increasing, rural-urban migrants experience a 0.38[23] higher percentage points increase in the probability of reporting "very happy", in comparison with their urban counterparts. As a consequence, an additional actual income increase contributes more to happiness of rural-urban migrants, in comparison with their urban counterparts, which supports the first hypothesis that an interaction term between the dummy variable of rural migrants and actual income has a positive effect on SWB. This result may be attributed to the

---

[22] In general, the marginal effect of the interaction term is calculated by the difference between rural=1 and rural=0 across the different continuous variables (e.g., actual income and income aspirations).

[23] In general, it is calculated by taking the difference in the marginal effect between rural=1 and rural=0 across actual income. The marginal effect of actual income for rural=1 is 0.01189 and 0.00809 for rural=0.

different marginal benefit for rural-urban migrants and urban residents. To be specific, rural-urban migrants have lower income and therefore are less able to meet their basic needs than urban *hukou* residents. Therefore, with 1% increase in actual income, rural-urban migrants' basic needs can be better achieved, in comparison to their previous status. In addition, for those rural migrants who left the elderly, children and even their wives at home, they feel happier as they can send money back home to improve the living standards for the left-behind family members which has been discussed in Chapter 1 while describing the nature of migration.

Meanwhile, the marginal effect of the interaction between the dummy of rural migrants and partner income shows a negative effect on SWB, which suggests that an increase in partner income contributes less to the well-being of rural-urban migrants. In Column B of Table 4-3, the negative marginal effect of the interaction term between partner income and the dummy variable of rural migrants shows that rural-urban migrants have a 0.29[24] lower percent increase in terms of the probability of reporting "very happy" than urban workers, with a 1% partner income increase. Additional partner income contributes less to rural-urban migrants' SWB, relative to urban individuals. This estimation result is not consistent with the second hypothesis that partner income contributes more to rural migrants' happiness; therefore, the similar SWB between rural-urban migrants and urban workers cannot be explained by the differential effect of partner income.

In terms of the previous two hypotheses, I can conclude that the interaction term of partner income presents an opposite effect to the interaction term of actual income on SWB. Actual income contributes more to rural-urban migrants' happiness while partner income has more contribution to urban workers' SWB. Whether the similar SWB between urban people and rural-urban migrants can be explained by the first hypothesis should be discussed further. It is not certain whether the effect of interaction terms between the dummy of rural migrants and partner income can offset the positive impacts of interaction variables between the dummy of rural-urban migrants and actual income. By considering the marginal effect of these two interaction terms, the magnitude of the marginal effects can be compared to investigate the differential effect

---

[24] The marginal effect of partner income for rural=1 is 0.01015 and 0.013091 for rural=0. Thus, the marginal effect of interaction term between rural and partner income is negative.

Table 4- 3 Ordered Probit Results for Different Specifications

	Basic model		Extended model	
	Coefficient (A)	Marginal effect of very happy (B)	Coefficient (C)	Marginal effect of very happy (D)
Income categories				
Actual income	0.0335**	0.0093**	0.0287**	0.0079**
	(0.0134)	(0.0032)	(0.0134)	(0.0032)
Aspiration income	0.0375	0.0041	0.0297	0.0030
	(0.0378)	(0.0077)	(0.0378)	(0.0077)
Partner income	0.0542***	0.0123***	0.0508***	0.0112***
	(0.0105)	(0.0021)	(0.0106)	(0.0021)
***Hukou* status**				
Rural	0.7475	0.0138	0.6263	0.0134
	(0.6234)	(0.0140)	(0.6249)	(0.0144)
Interaction between *hukou* status and income categories				
Rural× actual income	0.0131	0.0038**	0.0107	0.0031**
	(0.0317)	(0.0016)	(0.0318)	(0.0014)
Rural×aspiration income	-0.0658	-0.0162	-0.0551	-0.0136
	(0.0638)	(0.0159)	(0.0639)	(0.0159)
Rural×partner income	-0.0144	-0.0029**	-0.0110	-0.0021**
	(0.0162)	(0.0005)	(0.0163)	(0.0009)
Social demographic factors				
Age	-0.0527**	-0.0006	-0.0548**	-0.0006
	(0.0183)	(0.0006)	(0.0184)	(0.0007)
Age^2	0.0006**		0.0006**	
	(0.0002)		(0.0002)	
Number of children	0.0698**	0.0173**	0.0821**	0.0202**
	(0.0221)	(0.0079)	(0.0326)	(0.0080)
Male	-0.0088	-0.0022	-0.0016	-0.0004
	(0.0450)	(0.0111)	(0.0451)	(0.0111)
Married	0.8757**	0.1380***	0.8322**	0.1341***
	(0.3226)	(0.0264)	(0.3228)	(0.0284)
Educational level				
Compulsory education	0.4675***	0.1165***	0.4565**	0.1133***
	(0.1255)	(0.0313)	(0.1257)	(0.0313)
Senior school	0.4965***	0.1336**	0.4795**	0.1280**
	(0.1303)	(0.0374)	(0.1308)	(0.0372)
University level	0.6082***	0.1694***	0.5786***	0.1596***
	(0.1387)	(0.0421)	(0.1394)	(0.0418)
Health status				
Good health	0.6842***	0.1490***	0.6606***	0.1439***
	(0.0705)	(0.0146)	(0.0754)	(0.0147)
Standard Health	0.2323**	0.0605**	0.2175**	0.0562**
	(0.0806)	(0.0220)	(0.0808)	(0.0219)
Financial status				
Debt			-0.1749**	-0.0411***
			(0.0508)	(0.0114)
Social security				
Medical insurance			0.2588***	0.0583***
			(0.0624)	(0.0128)
Pension insurance			-0.0006	-0.0002
			(0.9901)	(0.0120)

Note: (1) *, **, *** indicate 10%, 5%, 1% significant levels, respectively. Figures in parentheses are standard errors.

(2) ME presented in this table is average marginal effect.

(3) China General Social Survey 2010 is utilized.

on SWB between the two terms. From Column B of Table 4-3, we can see that the magnitude of the marginal effect of the interaction term between actual income and the dummy of rural-urban migrants is larger than that of the interaction term between partner income and the dummy of rural migrants. With an increase in actual income, both the probability of replying "very happy" for rural-urban migrants and urban people is increased, but the magnitude of the probability increasing for rural-urban migrants is 0.38 percentage points larger than that of urban individuals. However, with an increase in partner income, the magnitude of increase in the probability of choosing "very happy" for a rural-urban migrant is 0.29 percentage points smaller than that of urban workers. In other words, the magnitude of difference in the effect of actual income between rural-urban migrants and urban people is higher than the difference of partner income between rural-urban migrants and urban residents. The positive differential effect of actual income between rural-urban migrants and urban workers is not completely offset by the negative difference in the effect of partner income between rural-urban migrants and urban individuals. Consequently, the similar SWB between urban people and rural-urban migrants can be partly explained by the first hypothesis that actual income has more contribution to rural-urban migrants' SWB.

Following Stutzer (2004), the relationship between income aspirations and SWB are also investigated. Hypothesis 3 is proposed to investigate whether the impact of income aspirations has a negative association with SWB and offsets the positive effect of income on SWB. Table 4-3 shows that aspiration income itself has no statistically significant effect on latent SWB, which is not consistent with Easterlin (2001) and Stutzer (2004), who both find aspiration income has a negative relationship with SWB. In addition, the marginal effect of aspiration income is not significant either, which does not support Hypothesis 3. The positive effect of actual income and partner income on SWB cannot be offset by the effect of higher aspiration income. Although urban *hukou* workers report higher income aspirations, their similar SWB with rural-urban migrants is not ascribed to their higher aspirations. Moreover, the ordered probit results for the marginal effect of "very happy" illustrates that the interaction term between the dummy of rural-urban migrants and aspiration income has no statistically significant effect on SWB, which is not consistent with Hypothesis 4. To sum up, the results from testing Hypothesis 3 and Hypothesis 4 reveal that the similar SWB between urban workers with higher income and rural-urban migrants cannot be explained by the fact that urban

workers report higher income aspirations.

For the effect of those basic socio-demographic factors other than income on SWB, the dummy variable for gender does not have a significant effect on SWB. As shown in Table 4-3, the coefficients of age and age squared in Column B show that SWB has a "U-shaped" age pattern, in line with Inglehart (1990), Oswald (1997), Blanchflower & Oswald (2004), Frey & Stutzer (2010), Hansen et al. (2008) and Plagnol (2011).

SWB is positively influenced by education. The marginal effect in Column B shows that, in comparison with people who have never received an education, respondents attaining compulsory education experience 11.65 percentage points increase in the probability of replying "very happy". In addition, respondents with more educational attainment have a greater chance of being in this top category. Respondents having attained a senior education had a 13.36 percentage point increase in the probability of being in the top category whilst people graduating from university reported a 16.94 percentage points increase in the probability of the same category, relative to the reference group: non-educated people. The estimation results in this chapter are consistent with Witter et al. (1984) and Inglehart (1990), who found that education is positively related to adults' SWB. There have been mixed findings on the impacts of educational attainment on SWB in previous research. The results of education impact contradict the findings of Frey & Stutzer (2010). Since highly educated individuals have higher ambitions, the pressure may lead to an inverse impact on SWB (Frey & Stutzer, 2010). My finding may be attributed to the fact that highly educated individuals are more likely to be engaged in better jobs. Holding income constant, highly educated people can enjoy many fringe benefits from better jobs, such as better welfare, higher standard working conditions, and a better environment. Therefore, the positive relationship between education attainment and SWB found in this chapter may be due to better educated people being employed in higher skilled jobs with better welfare and conditions.

Being married is positively related to SWB. According to the marginal effect in Column B of Table 4-3, married people report a 13.80 percent higher probability of choosing "very happy", compared to individuals who have never married, been divorced, or are widowed. This is in line with Frey & Stutzer (2010) who suggest that

married individuals report higher SWB since marriage can provide additional sources through sharing financial burdens and provide psychological support to each other. With an increasing of number of children, the probability of being happiness will increase as well.

For the effect of health status on SWB in Column B of Table 4-3, it can be seen that both dummy variables, 'good health status' and 'standard health status', have positive impacts on SWB. In comparison with unhealthy respondents, people who reply "good health" or "standard health status" report higher levels of happiness than people who self-describe as "unhealthy", in line with Okun et al. (1984), Gerdtham & Johannesson (2001) and Frey & Stutzer (2010). In addition, the better the health status is, the higher the SWB. Individuals in good health have a 14.90 percent higher probability of replying "very happy", whilst people who report having "standard health status" have only a 6.05 percent higher probability of reporting the top category, compared to those who describe themselves as "unhealthy".

On the other hand, few researchers have proposed that there may be some other factors influencing the impact of these traditional economic variables on SWB, such as the effect of income on SWB (Roberts et al., 1998; Brown et al., 2005; Esser & Palme, 2010; Liao et al., 2012 and Keng & Wu, 2014). Other studies focus on aspects of security, specifically financial security and social security. Roberts et al. (1998) and Brown et al. (2005) explore the effect of debt on respondents' happiness. They find that respondents with debt are more likely to suffer from psychological distress, since they are faced with financial restrictions. The negative effect of debt can be offset by the positive effect generated by a rise in monthly income or annual savings, otherwise normal expenditure has to be sacrificed to repay debt. Sometimes, individuals may need to borrow money to repay their debt, which induces a vicious circle of financial debt (Brown et al., 2005). Therefore, even if a person has higher income, his happiness will be restricted if he suffers from the pressure of debt.

In terms of social security[25], I intend to follow latest studies (Esser & Palme, 2010;

---

[25] Not only the public sector but also private companies will provide social security including pension insurance and medical insurance to the permanent staff with a regular labour contract. As long as employees sign the legal labour contract with a normal private company, they are provided with social security, which is similar to those in the public

Liao et al., 2012 and Keng & Wu, 2014) to examine its impact on SWB from two aspects, medical insurance and pension insurance, due to the incomplete medical care system and pension system in China. The catastrophic health expenditure can compel a middle class family to be poor due to health shocks (Zhang et al., 2010; Keng & Wu, 2014). Although one person without medical insurance has higher wages, he may be worried about the huge expenditure for an unexpected health shock in the future. He is thus more likely to sacrifice several expenses to build a deposit against future medical expenditures. The same applies to the pension system. One individual without pension insurance should consider saving more money to finance daily expenses after retirement. Medical insurance and pension insurance can alleviate the financial loss for those unexpected health shocks and retirement. A deficiency of medical and pension insurance may lower individuals' happiness even if they earn higher wages. Keng & Wu (2014) also find that the implementation of National Health Insurance (NHI) has increased the levels of happiness of elderly persons, especially elderly women, to a large extent. In CGSS 2010, information on debt related to financial security, and medical insurance and pension insurance related to social security, was available. Consequently, I follow these studies and estimate the impact of financial security and social security on SWB by controlling for other economic factors. The basic model is extended by taking a dummy variable of debt, and two dummy variables of pension and medical insurance into account in order to check the robustness of the main results. Results of the extended model are shown in Columns C and D of Table 4-3.

The results of all independent variables in the extended model are consistent with the basic model. Results in Table 4-3 show that having debt leads to lower levels of SWB. An individual with debt has a 4.11 percent lower probability of reporting being "very happy" compared to people with no debt, in line with Roberts et al. (1998) and Brown et al. (2005). In addition, the positive relationship shown in Column C of Table 4-3 between medical insurance and SWB reveals that people having medical insurance are found to report higher SWB than those without medical insurance, as seen in Liao et al. (2012) and Keng & Wu (2014). People with medical insurance are 5.83 percent more likely to report being "very happy", as individuals with medical insurance do not have to worry about incurring large medical expenses when in need of medical treatment.

sector. However, only temporary staff or informal employees in the private organizations are not able to enjoy benefits from social security.

156

However, results in Table 4-3 show that pension insurance does not influence SWB, which is not consistent with Esser & Palme (2010) and Bender (2012).

Overall, according to the relationship between these social demographic factors and SWB, urban individuals should report higher SWB. When comparing these characteristics between urban residents and rural-urban migrants, it can be seen from Table 4-2 that 66% of urban people had graduated from senior school or university, whilst only 30% of rural-urban migrants reported the same. Due to the positive relationship between educational attainment and happiness, urban respondents with more educational attainment should have been happier than rural-urban migrants. In addition, when taking financial security and social security into account, I found that 70% of urban individuals had medical insurance, compared to 29% of rural-urban migrants. Thus, urban people should have reported higher SWB as well, due to the negative effect of debt and positive impacts of medical insurance on happiness. However, urban residents and rural-urban migrants showed similar levels of SWB.

In the first specification above, I know that individuals' aspiration income has no direct impact on happiness, which is not in line with my expectations. Individuals' aspiration income does not offset the positive impact of actual income on SWB. This may be due to the fact that SWB is more likely to be affected by the difference between the aspiration and actual income. When people realize that there is a huge gap between their aspirations and actual status or their own real status has already achieved the aspiration levels, the happiness may be correspondingly affected. Similarly, other people's income may not have a profound effect on the happiness; however, when a person compares his own income with others and finds that his income is much lower, he is likely to be unhappy. As discussed in hypotheses section, the common approach for analysing the relative income on SWB is to substitute the reference income itself (e.g., aspiration income or other individuals' income) with the distance between the reference income and actual income in the second step, on controlling the individuals' actual income. Michalos (1991) and Stutzer (2004) also indicate that individuals' well-being is determined by the difference between aspirations and achievement. Following previous studies, the second specification of this chapter is to examine the impact of the discrepancy between aspirations and actual income on SWB in order to explain the similar happiness between urban people and rural-urban migrants.

Table 4-4 presents the estimation results for this second specification in which the aspiration income is substituted by the discrepancy between aspiration income and actual income. As shown in Eq. (4-16), the estimation of other variables (e.g., education attainment, social-demographic factors, partner income, health status, financial and social security) remain consistent in terms of their sign, significance and magnitude. The coefficient of actual income in Table 4-4 is equal to the sum of the respective coefficient of actual income and aspiration income in Table 4-3, which is in line with Eq. (4-16). Similarly, the coefficient of the interaction term between rural *hukou* status and actual income in Table 4-4 is also the sum of respective coefficient of interaction for actual income and aspiration income in Table 4-3, which presents negative impact on SWB but it is not significant. Regarding the estimation of the discrepancy between aspiration income and actual income, it is consistent with estimation of aspiration income alone in Table 4-3. Specifically, the discrepancy between aspiration income and actual income has a sizable positive impact on the reported happiness, but the impact is also statistically insignificant. Furthermore, the ordered probit results for "very happy" in Table 4-4 also show that the interaction term between the dummy of rural *hukou* status and the discrepancy between aspiration income and actual income is statistically insignificant to SWB as well. Similar to the results of hypotheses 3 and 4, the estimated results in Table 4-4 do not support hypotheses 5 and 6. Therefore, based on all the estimations above, both income aspirations and the discrepancy between individuals' aspirations and actual income have no effects on SWB. The similar happiness between rural-urban migrants and urban *hukou* individuals cannot be explained by respondents' aspiration level, no matter which measurement of income aspiration (aspiration itself or aspiration in the form of relative status) is utilized in the model. Overall, self-evaluated aspiration has no contribution to balance the happiness between rural-urban migrants and urban people.

The reason why aspirations have no significant impact on SWB may be attributed to the fact that Chinese people are not really concerned about their own aspiration levels. In Chinese cultural context, there is a proverb that people's desire is always imagined to be canyon that can never be filled up. Chinese individuals may realize that even if their current aspirations are satisfied, they will have new aspiration further and the desire will never be satisfied. As a result, most Chinese people may be more concerned with their current situation such as the current income and the current status, in comparison with

others. Secondly, the insignificant effect may be because social comparisons matter more for SWB in the Chinese cultural context. For example, the measurement of the respondents' aspiration income depends on the respondent's personal evaluation, but the comparison is determined by an evaluation of the whole families' financial status. For those respondents with lower actual income and lower aspiration income, their higher SWB may not be influenced by their own lower aspirations but may be influenced by their whole household's social and economic status (SES). As previous studies suggest, social comparisons play a critical role in determining individuals' SWB (Graham & Felton, 2005; Vendrik & Woltjer, 2007; Wu & Tam, 2015 and Lam & Liu, 2014). I, therefore, investigate the effect of SES on SWB to establish whether it is in fact SES, rather than income aspirations per se, that matter in the Chinese background. This will be discussed in section 4.5.2.

To sum up, the analysis in the first specification shows that the similar SWB between urban residents and rural-urban migrants can be partly explained by hypothesis 1, that actual income contributes more to rural-urban migrants' levels of happiness. Although the results are inconsistent with hypothesis 2, that partner income matters more for rural-urban migrants, the differential effect of actual income is larger and outweighs the differential effect of partner income. Consequently, the reason for similar SWB is partly ascribed to higher contributions of actual income to rural-urban migrants' happiness, based on the first specification. Hypotheses 3 and 4, that aspiration income has a negative effect on SWB, were proposed in order to analyse whether the negative impact of aspiration income can offset the positive effect of actual income, in turn leading to similar SWB. However, aspiration income is shown to make no contribution to SWB. In addition, the relative measurement of aspirations in the second specification, the discrepancy between aspiration income and actual income, shows no contribution to SWB, as well. The estimation results of this discrepancy between aspirations and actual income also do not support hypotheses 5 and 6. Therefore, I can conclude that income aspiration income has no contribution to the similar SWB between urban workers and rural-urban migrants in China. Following most previous studies that suggest that income influence happiness relatively (Graham & Felton, 2005; Vendrik & Woltjer, 2007; Knight et al., 2009; and Lam & Liu, 2014), the most likely economic factor for happiness is social comparison. Further analysis and discussion about the impact of social comparison on SWB are given in the next subsection.

Table 4- 4 Ordered Probit Results for the Discrepancy between Income Aspirations and Actual Income

	Basic model		Extended model	
	Coefficient (A)	Marginal effect of very happy (B)	Coefficient (C)	Marginal effect of very happy (D)
Income categories				
Actual income	0.0710**	0.0134**	0.0584**	0.0109**
	(0.0337)	(0.0054)	(0.0225)	(0.0034)
Discrepancy between aspirations	0.0375	0.0041	0.0297	0.0030
and income (= aspiration level (ln)-	(0.0378)	(0.0077)	(0.0378)	(0.0077)
actual income (ln))				
Partner income	0.0542***	0.0123***	0.0508***	0.0112***
	(0.0105)	(0.0021)	(0.0106)	(0.0021)
***Hukou* status**				
Rural	0.7475	0.0138	0.6263	0.0134
	(0.6234)	(0.0140)	(0.6249)	(0.0144)
Interaction between *hukou* status and income categories				
Rural× actual income	-0.0527	-0.0124	-0.0444	-0.0105
	(0.0601)	(0.0131)	(0.0602)	(0.0101)
Rural× discrepancy between	-0.0658	-0.0162	-0.0551	-0.0136
aspirations and income	(0.0638)	(0.0159)	(0.0639)	(0.0159)
Rural× partner income	-0.0144	-0.0029**	-0.0110	-0.0021**
	(0.0162)	(0.0005)	(0.0163)	(0.0009)
Social demographic factors				
Age	-0.0527**	-0.0006	-0.0548**	-0.0006
	(0.0183)	(0.0006)	(0.0184)	(0.0007)
Age^2	0.0006**		0.0006**	
	(0.0002)		(0.0002)	
Number of children	0.0698**	0.0173**	0.0821**	0.0202**
	(0.0221)	(0.0079)	(0.0326)	(0.0080)
Male	-0.0088	-0.0022	-0.0016	-0.0004
	(0.0450)	(0.0111)	(0.0451)	(0.0111)
Married	0.8757**	0.1380***	0.8322**	0.1341***
	(0.3226)	(0.0264)	(0.3228)	(0.0284)
Educational level				
Compulsory education	0.4675***	0.1165***	0.4565**	0.1133***
	(0.1255)	(0.0313)	(0.1257)	(0.0313)
Senior school	0.4965***	0.1336**	0.4795**	0.1280**
	(0.1303)	(0.0374)	(0.1308)	(0.0372)
University level	0.6082***	0.1694***	0.5786***	0.1596***
	(0.1387)	(0.0421)	(0.1394)	(0.0418)
Health status				
Good health	0.6842***	0.1490***	0.6606***	0.1439***
	(0.0705)	(0.0146)	(0.0754)	(0.0147)
Standard Health	0.2323**	0.0605**	0.2175**	0.0562**
	(0.0806)	(0.0220)	(0.0808)	(0.0219)
Financial status				
Debt			-0.1749**	-0.0411***
			(0.0508)	(0.0114)
Social security				
Medical insurance			0.2588***	0.0583***
			(0.0624)	(0.0128)
Pension insurance			-0.0006	-0.0002
			(0.9901)	(0.0120)

4.5.2 Further Analysis

Graham & Felton (2005), Vendrik & Woltjer (2007), Knight et al. (2009) and Lam & Liu (2014) suggest that social comparison and social-economic status (SES) plays a critical role in determining individuals' SWB. Knight et al. (2009) define the relative income by classifying respondents' household income and living standard to much below average, below average, average, above average, or much average in their rural village. Therefore, following the measurement of socioeconomic status in Knight et al. (2009), further analysis in this chapter focuses on the effect of social comparison to test whether the similar happiness can be explained by SES[26]. All individuals were asked to evaluate their own economic standard relative to others in their community, based on their own income, their household income, savings and other assets. The answers are given on a 10-point scale from the bottom 1-significantly below average level to the top significantly above average level. In this chapter, levels 1, 2 and 3 are combined in the dummy variable "low economic status". Levels 4, 5 and 6 are combined in a dummy variable "average economic status". Levels 7, 8, 9 and 10 are combined in a dummy variable "high economic status".

It can be seen from Table 4-5 that, although urban workers have higher income and partner income, they report similar economic status to rural-urban migrants. Specifically, about 61.65% of urban people consider they are located at the average or high economic level, compared to 62.8% of rural migrants. The reason for the similar economic status

---

[26] The measurement of SES in this chapter is different with that in the preceding Chapter 3. In Chapter 3, propensity score matching is utilized so that I need calculate the exact propensity score to match the urban *hukou* individuals and rural-urban migrants. If I used same method in Chapter 4 and combined several different levels together into three dummy variables, the PS score obtained from this approach is not precise enough for matching. For example, if levels 1, 2 and 3 were combined in the "low economic status", respondents reporting any levels among this category (lower economic status) would report the same PS score. Therefore, I think it is much better to regard the SES as a continuous variable to obtain the propensity score in Chapter 3. However, in most studies which analyze the effect of social status on subjective well-being, the variable of social status is utilized to classify all the individuals into different groups. People with relatively close answers can be classified as one group. Therefore, following previous studies, I transform the 10-point scale ordinal variable into few simple and straightforward dummy variables including lower socio-economic status, average status and higher socio-economic status.

is ascribed to the variable's definition that respondents compare their economic status to other people in their own communities. Following the theory of social comparison (Festinger, 1954), individuals compare themselves to proximal others. Therefore, respondents compare their own income or household income to their friends, neighbours, colleagues and other people around them.

Table 4- 5 Summary Statistics of Economic Status of Urban Worker and Rural Migrants

	Urban (2076 observations)	Rural (943 observations)
Variables	**Mean**	**Mean**
Low economic status	0.3834	0.3720
Average status	0.5178	0.5536
High economic status	0.0987	0.0744

Note: China General Social Survey 2010 is utilized.

Although rural-urban migrants report lower income and partner income, they may not have a lower evaluation of their economic status because they are found to confine or narrow their reference groups (Knight et al., 2009). Most people around them, such as their friends, neighbours and colleagues, experience a similar situation. In addition, as discussed in Chapter 1 (introduction chapter), rural adults who work in urban cities left their parents, children or wife in countryside, and transfer income back to support their families in rural area. Thus, these rural-urban migrants can improve the living standards of families in rural area, in comparison to those families in which there is no one working in an urban city. Therefore, these migrants consider that the economic status of the entire family is in good condition. Due to the positive relationship between SWB and socio-economic status (Graham & Felton, 2005; Vendrik & Woltjer, 2007; and Lam & Liu, 2014), Hypothesis 7 is proposed.

Hypothesis 7: Self-reported economic status has a positive impact on SWB.

In addition, as with hypothesis 1 and hypothesis 2, the differential effect of economic status on SWB between urban people and rural-urban migrants should be taken into account. If economic status contributed more to urban *hukou* people, their happiness will increase more with the same extent of increase in economic status. Consequently, if this assumption is accepted, the similar happiness cannot be explained by economic status even though urban *hukou* respondents reported similar economic

status to their rural migrants. Therefore, I generated interaction terms between the economic status and a dummy variable of rural status, in order to investigate whether SWB is positively impacted on by the interaction terms, and propose Hypothesis 8.

Hypothesis 8: There is a differential impact of economic status on SWB for rural-urban migrants with an increase in their economic status generating larger gains in SWB for rural migrants. Therefore, an interaction term between economic status and a dummy for rural migrants should be positive, in comparison with urban *hukou* workers.

Table 4- 6 Ordered Probit Results for Different Specifications including Economic Status

	Basic model		Extended model	
	Coefficient (A)	Marginal effect of very happy (B)	Coefficient (C)	Marginal effect of very happy (D)
Income categories				
Actual income	0.0255** (0.0128)	0.0052* (0.0031)	0.0215* (0.0129)	0.0043 (0.0025)
Partner income	0.0425*** (0.0106)	0.0088*** (0.0021)	0.0402*** (0.0107)	0.0085*** (0.0021)
***Hukou* status**				
Rural	0.2890 (0.3082)	0.0026 (0.0136)	0.2556 (0.3100)	0.0021 (0.0141)
Interaction between *hukou* status and income categories				
Rural× actual income	0.0121 (0.0302)	0.0028** (0.0010)	0.0113 (0.0303)	0.0027* (0.0012)
Rural× partner income	-0.0195 (0.0166)	-0.0046*** (0.0008)	-0.0161 (0.0166)	-0.0071** (0.0022)
Social demographic factors				
Age	-0.0425** (0.0184)	-0.0003 (0.0006)	-0.0448** (0.0185)	-0.0006 (0.0006)
Age^2	0.0005*** (0.0002)		0.0005** (0.0002)	
Number of children	0.0043 (0.0323)	0.0104 (0.0078)	0.0534 (0.0328)	0.0129 (0.0089)
Male	-0.0154 (0.0446)	-0.0037 (0.0108)	-0.0100 (0.0447)	-0.0024 (0.0108)

Table 4-6 Ordered Probit Results for Different Specifications including Economic Status (Continued)

	Basic model		Extended model	
	Coefficient (A)	Marginal effect of very happy (B)	Coefficient (C)	Marginal effect of very happy (D)
Married	0.9315**	0.1411***	0.8957**	0.1382***
	(0.3233)	(0.0245)	(0.3235)	(0.0259)
Education level				
Compulsory education	0.4501***	0.1099***	0.4426***	0.1078***
	(0.1258)	(0.0307)	(0.1260)	(0.0307)
Senior school	0.4516**	0.1181**	0.4407**	0.1147**
	(0.1304)	(0.0363)	(0.1309)	(0.0362)
University level	0.5045***	0.1351**	0.4845***	0.1288**
	(0.1379)	(0.0399)	(0.1386)	(0.0398)
Health status				
Good Health	0.6034***	0.1309***	0.5888***	0.1278***
	(0.0756)	(0.0148)	(0.0759)	(0.0149)
Standard Health	0.2055**	0.0521**	0.1952**	0.0493**
	(0.0808)	(0.0214)	(0.0810)	(0.0223)
Economic Status				
Average status	0.4271***	0.1030***	0.4091***	0.0980***
	(0.0546)	(0.0107)	(0.0548)	(0.0107)
High economic status	0.6345***	0.1825***	0.6090***	0.1720***
	(0.0942)	(0.0262)	(0.0945)	(0.0260)
Interaction between *hukou* status and financial status				
Rural× average status	0.0261	0.0080***	0.0184	0.0058***
	(0.0963)	(0.0030)	(0.0965)	(0.0003)
Rural× high status	0.0253	0.0061***	0.0393	0.0110***
	(0.1668)	(0.0021)	(0.1669)	(0.0023)
Financial security				
Debt			-0.1363***	-0.0318**
			(0.0511)	(0.0115)
Social security				
Medical insurance			0.2120**	0.0477***
			(0.0628)	(0.0131)
Pension insurance			-0.0051	-0.0012
			(0.0487)	(0.0118)

Note: (1) *, **, *** indicate 10%, 5%, 1% significant levels, respectively.

(2) Figures in parentheses are standard errors.

(3) ME presented in this table is average marginal effect.

(4) China General Social Survey 2010 is utilized.

Table 4-6 shows the results of the ordered probit model, replacing income aspirations with economic status. As with Table 4-3, Columns A and B report the ordered probit coefficients and marginal effect of the basic model. Columns C and D report the ordered probit coefficients and marginal effects of the extended model. In general, the results for all the variables which are included in Table 4-3 remain the same.

For example, the interaction term between actual income and dummy of rural-urban migrants presents a positive effect on SWB. While the interaction between the dummy of rural migrants and partner income is negative as found in Table 4-3, the magnitude of the marginal effect of the interaction term between the dummy variable of rural *hukou* and actual income is smaller than that of the interaction between dummy of rural-urban migrants and partner income. In terms of economic status, individuals with higher economic status are happier than those reporting a lower economic status, supporting Hypothesis 7, which suggests that economic status has a positive association with SWB. To be specific, it can be seen from basic model of Table 4-6 that individuals reporting average economic status have a 10.30 percentage points higher probability of reporting being "very happy", compared to people reporting low economic status. In addition, people with high economic status are 18.25 percentage points more likely to report being in the top category.

Considering the differential effect of economic status on SWB between urban people and rural-urban migrants, the marginal effect of the interaction terms between dummy of rural-urban migrants and economic status shows a significant and positive effect on SWB. This reveals that economic status matters more for rural-urban migrants' happiness. As can be seen from Column B of Table 4-6, the difference in the effect of average economic status on SWB between rural-urban migrants and urban people is 0.80 percentage points. In addition, rural migrants, having higher economic status has a 0.61 percentage points greater positive effect on SWB for rural-urban migrants, compared to their urban counterparts. Economic status contributes more to rural-urban migrants' happiness compared to their urban counterparts, which supports Hypothesis 7. Consequently, a similar economic status between urban residents and rural-urban migrants is a potential reason to explain why the reported SWB of rural-urban migrants and urban people are similar. Overall, the differential effect of economic status helps to explain the similar SWB across rural-urban migrants and urban individuals, while income aspirations appear to have very little impact on SWB for both groups. Respondents' economic status can balance happiness between urban individuals and rural-urban migrants, in line with many previous economic studies which explore the impact of income on happiness relatively.

4.6 Conclusion

According to Knight et al. (2009), the difference in SWB between the preferred group and disadvantaged group is important to social stability. As discussed in Chapter 2, rural-urban migrants are discriminated against and suffer from inequality in terms of income and occupation. Therefore, the difference in happiness levels between urban residents and rural-urban migrants plays an important role in determining Chinese social stability. However, few studies have considered the difference in SWB between these two groups. Therefore, this chapter contributes to the previous Chinese literature by comparing SWB between urban workers and rural-urban migrants, and explores the determinants of their SWB. In addition, motivated by Chapter 3, this chapter also contributes to the previous literature by estimating the impact on SWB from the aspect of income aspirations.

Using the data CGSS 2010, which is the same as with Chapter 3, it is found that, rural-urban migrants report similar happiness to urban *hukou* individuals. Although rural-urban migrants have lower income and lower partner income, they also report lower aspiration income. However, the results show that aspiration income has no contribution to SWB. Although urban *hukou* individuals have higher income aspirations, the positive effect of their higher income cannot be offset by the aspiration income. In addition, further analysis is conducted to explain the similar SWB between rural-urban migrants and urban *hukou* individuals, in terms of social comparison. It has found that economic status has positive impact on SWB and matters more for rural-urban migrants' SWB, compared to urban people. Thus, social comparison rather than income aspirations explains the similar SWB between rural-urban migrants and urban *hukou* individuals.

Since the similar happiness level between urban residents and rural-urban migrants cannot be adequately explained by aspiration income, I follow previous studies which suggest that income influences happiness relatively, in order to estimate whether the social comparison can explain the similar happiness level between urban people and rural-urban migrants. Although rural-urban migrants have lower income, they evaluate themselves as having similar economic status to urban people. The results of the social comparison on happiness show that economic status positively influences SWB. People with higher economic status are happier than those with lower economic status. In addition, economic status contributes more to rural-urban migrants' happiness, in

comparison to their urban counterparts. Hence, it is social comparison rather than income aspirations that can help to explain why rural-urban migrants have similar levels of happiness.

As with most studies, there is limitation to this chapter's model, as an endogeneity issue may exist. Firstly, there might be reverse causation as discussed in Stutzer (2004), unhappy people are likely to report higher income aspirations because these unhappy people consider that the levels of their happiness can be improved by higher income. In addition, there may be some unobserved and omitted factors, such as traits or characteristics that may not only influence the dependent variable SWB, but also have a relationship with the explanatory variables. This can cause the explanatory variables to be related to the disturbance term, which in turn leads to endogeneity problems. For example, an extroverted and positive person is more likely to feel happy. These people are also active in taking part in many social activities and communicating with their neighbours, friends, relatives and fellows during leisure time. These individuals may also be more likely to be influenced by others around them, which in turn influences their assessment of their aspirations. In this circumstance, the estimation of the impact of aspiration income on SWB may be upward-biased.

Similarly, due to reverse causation and omitted factors, there may be endogeneity of actual income as well. Lyubomirsky, King, & Diener (2005) state that positive affect can foster skills, positive behaviour, sociability and activity, and effective conflict resolution capability. Consequently, happier people may be better at generating income or benefits. To be specific, an outgoing and optimistic person may be more positive and communicative in daily life and work, which helps them to improve themselves by communication with others. Therefore, these positive individuals are more likely to perform better and earn higher incomes. In addition, if there are also some unobserved and omitted factor such as personal traits of characteristics positively influenced respondents' income as their impact on aspirations, the estimated coefficient of actual income on SWB would be upward-biased as well.

An appropriate empirical approach to deal with endogeneity is an instrumental variable, which ideally will influence the endogenous explanatory variables only with no impact on the SWB. Following previous studies like Chau et al. (2007), Knight &

Gunatilaka (2010), Li et al. (2011), Knight & Gunatilaka (2012), potential instrumental variables for actual income are language ability, siblings' income, and father's education attainment. In CGSS 2010, only the father's education level is available. When I tried to utilized father's education attainment to deal with the potential endogenous problem for income, I found that father's education is not relevant with the actual income unfortunately so that it is not suitable as instrument variable (IV) for income, based on my dataset.[27]

In considering potential instrumental variables for aspiration income, Stutzer (2004) indicates that it is difficult to find a suitable instrument, as both aspiration and SWB are attitudes. Stutzer (2004) applies the proportion of rich people in the same community and its interaction term with the frequency of contact with neighbours as an instrument to overcome the endogeneity. However, due to the limitation of the dataset (CGSS 2010), there is no available information for such instrumental variables for aspiration income, unfortunately. Lagged variables of endogenous explanatory variables can be also utilized as instrumental variables; however, this approach of seeking can only be applied for time series or panel data. Unfortunately, CGSS 2010 is a cross-sectional dataset and is not suitable for this approach. Overall, considering the potential endogenous problems in this chapter, there is limited information to obtain available and valid instrumental variables. Consequently, it is difficult for me address the endogenous problem at the moment, a limitation of this chapter's study.

---

[27] Both the estimation from the conditional mixed process (CMP) in ordered probit model and the two-stage least square (2SLS) in linear regression model show that father's education attainment is not relevant with actual income unfortunately. The estimation results are shown in Table 4-14 in the appendix.

4.7 Appendix

Table 4- 7 Marginal Effect of Basic Model

	Not at all happy	Not happy	Neutral	Happy	Very happy
Income categories					
Actual income	-0.0010**	-0.0035**	-0.0063**	0.0014**	0.0093***
	(0.0004)	(0.0012)	(0.0022)	(0.0006)	(0.0032)
Aspiration income	-0.0005	-0.0016	-0.0028	-0.0007	0.0041
	(0.0008)	(0.0029)	(0.0052)	(0.0013)	(0.0077)
Partner income	-0.0014***	-0.0046***	-0.0083***	0.0019***	0.0123***
	(0.0003)	(0.0008)	(0.0014)	(0.0005)	(0.0021)
***Hukou* status**					
Rural	-0.0021	-0.0065	-0.0105	0.0053	0.0138
	(0.0015)	(0.0049)	(0.0091)	(0.0039)	(0.0140)
Interaction between *hukou* status and income categories					
Rural× actual income	-0.0002	-0.0010**	-0.0022**	0.0005**	0.0038**
	(0.0003)	(0.0004)	(0.0010)	(0.0002)	(0.0016)
Rural× aspiration income	0.0018	0.0061	0.0109	-0.0025	-0.0162
	(0.0013)	(0.0044)	(0.0089)	(0.0022)	(0.0159)
Rural× partner income	0.0006**	0.0017**	0.0024**	-0.0017**	-0.0029**
	(0.0001)	(0.0007)	(0.0011)	(0.0008)	(0.0005)
Social demographic factors					
Age	-0.0006	-0.0001	-0.0001	0.0006*	-0.0006
	(0.0006)	(0.0002)	(0.0003)	(0.0003)	(0.0006)
Number of children	-0.0019**	-0.0064**	-0.0117**	0.0027**	0.0173**
	(0.0009)	(0.0030)	(0.0054)	(0.0011)	(0.0079)
Male	0.0002	0.0008	0.0015	-0.0003	-0.0022
	(0.0012)	(0.0041)	(0.0075)	(0.0017)	(0.0111)
Married	-0.0592	-0.1206*	-0.1260***	0.1678**	0.1380***
	(0.0409)	(0.0554)	(0.0271)	(0.0760)	(0.0264)
Educational attainment					
Compulsory education	-0.0144**	-0.0430***	-0.0723***	0.0132**	0.1165***
	(0.0051)	(0.0119)	(0.0172)	(0.0043)	(0.0313)
Senior school	-0.0124**	-0.0406***	-0.0759***	-0.0047	0.1336**
	(0.0038)	(0.0099)	(0.0174)	(0.0084)	(0.0374)
University	-0.0125***	-0.0465***	-0.0950***	0.0155	0.1694***
	(0.0032)	(0.0094)	(0.0190)	(0.0130)	(0.0421)
Health status					

Good Health	-0.0230***	-0.0736***	-0.1184***	0.0661***	0.1490***
	(0.0047)	(0.0103)	(0.0127)	(0.0120)	(0.0146)
Standard health	-0.0060**	-0.0200**	-0.0368**	0.0023	0.0605**
	(0.0022)	(0.0066)	(0.0120)	(0.0025)	(0.0220)

Note: (1) *, **, *** indicate 10%, 5%, 1% significant levels, respectively.

(2) Figures in parentheses are standard errors.

(3) ME presented in this table is average marginal effect.

(4) China General Social Survey 2010 is utilized.

Table 4- 8 Marginal Effect of Extended Model

	Not at all happy	Not happy	Neutral	Happy	Very happy
Income categories					
Actual income	-0.0008*	-0.0029**	-0.0054**	0.0012**	0.0079**
	(0.0004)	(0.0012)	(0.0022)	(0.0006)	(0.0032)
Aspiration income	-0.0004	-0.0012	-0.0020	0.0005	0.0030
	(0.0008)	(0.0029)	(0.0052)	(0.0012)	(0.0077)
Partner income	-0.0013***	-0.0043***	-0.0079***	0.0018***	0.0112***
	(0.0003)	(0.0008)	(0.0014)	(0.0005)	(0.0021)
***Hukou* status**					
Rural	-0.0019	-0.0060	-0.0100	0.0045	0.0134
	(0.0015)	(0.0050)	(0.0093)	(0.0039)	(0.0144)
Interaction between *hukou* status and income categories					
Rural× actual income	-0.0002*	-0.0007**	-0.0018**	-0.0004**	0.0031**
	(0.0001)	(0.0003)	(0.0008)	(0.0001)	(0.0014)
Rural× aspiration income	0.0014	0.0049	0.0091	-0.0020	-0.0136
	(0.0013)	(0.0046)	(0.0087)	(0.0018)	(0.0159)
Rural× partner income	0.0005**	0.0013**	0.0020**	-0.0015**	-0.0021**
	(0.0002)	(0.0006)	(0.0009)	(0.0007)	(0.0009)
Social demographic factors					
Age	-0.00002	-0.00004	0.0003	0.0003**	-0.0006
	(0.00006)	(0.0002)	(0.0004)	(0.0001)	(0.0007)
Number of children	-0.0022**	-0.0075**	-0.0137**	0.0031**	0.0202**
	(0.0009)	(0.0030)	(0.0054)	(0.0014)	(0.0080)
Male	0.0001	0.0001	0.0003	-0.0001	-0.0004
	(0.0012)	(0.0041)	(0.0075)	(0.0017)	(0.0111)
Married	-0.0526	-0.1123**	-0.1225***	0.1533*	0.1314***
	(0.0375)	(0.0548)	(0.0306)	(0.0736)	(0.0284)
Educational attainment					
Compulsory education	-0.0137**	-0.0416**	-0.0706**	0.0126**	0.1133***
	(0.0049)	(0.0118)	(0.0173)	(0.0041)	(0.0313)
Senior school	-0.0117**	-0.0391***	-0.0736***	-0.0037	0.1280**
	(0.0036)	(0.0099)	(0.0177)	(0.0079)	(0.0372)
University	-0.0117***	-0.0443***	-0.0909***	-0.0128	0.1596***
	(0.0030)	(0.0095)	(0.0194)	(0.0121)	(0.0418)
Health status					
Good Health	-0.0214***	-0.0701***	-0.1146***	0.0621***	0.1439***
	(0.0045)	(0.0101)	(0.0129)	(0.0118)	(0.0147)
Standard status	-0.0056**	-0.0187**	-0.0345**	0.0025	0.0562**
	(0.0022)	(0.0066)	(0.0121)	(0.0022)	(0.0219)

Financial security

Debt	0.0051**	0.0170***	0.0296**	-0.0105**	-0.0411***
	(0.0018)	(0.0053)	(0.0087)	(0.0044)	(0.0114)

Social security

Medical insurance	-0.0083**	-0.0264***	-0.0439***	0.0203**	0.0583***
	(0.0026)	(0.0072)	(0.0108)	(0.0077)	(0.0128)
Pension insurance	0.0001	0.0001	-0.0001	-0.00002	-0.0002
	(0.0013)	(0.0044)	(0.0008)	(0.001)	(0.0120)

Note: (1) *, **, *** indicate 10%, 5%, 1% significant levels, respectively.

(2) Figures in parentheses are standard errors.

(3) ME presented in this table is average marginal effect.

(4) China General Social Survey 2010 is utilized.

Table 4- 9 Marginal Effects of the Ordered Probit Results for Discrepancy between Income Aspirations and Actual Income (Basic Model)

	Not at all happy	Not happy	Neutral	Happy	Very happy
Income categories					
Actual income	-0.0015**	-0.0051**	-0.0091*	0.0007**	0.0134**
	(0.0007)	(0.0020)	(0.0051)	(0.0002)	(0.0054)
Discrepancy between aspirations and income (aspiration income(ln)-actual income (ln))	-0.0005	-0.0016	-0.0028	-0.0007	0.0041
	(0.0008)	(0.0029)	(0.0052)	(0.0013)	(0.0077)
Partner income	-0.0014***	-0.0046***	-0.0083***	0.0019***	0.0123***
	(0.0003)	(0.0008)	(0.0014)	(0.0005)	(0.0021)
***Hukou* status**					
Rural	-0.0021	-0.0065	-0.0105	0.0053	0.0138
	(0.0015)	(0.0049)	(0.0091)	(0.0039)	(0.0140)
Interaction between *hukou* status and income categories					
Rural× actual income	0.0016	0.0051	-0.0087	-0.0020	-0.0124
	(0.0012)	(0.0039)	(0.0059)	(0.0015)	(0.0131)
Rural× discrepancy between aspiration and actual income	0.0018	0.0061	0.0109	-0.0025	-0.0162
	(0.0013)	(0.0044)	(0.0089)	(0.0022)	(0.0159)
Rural× partner income	0.0006**	0.0017**	0.0024**	-0.0017***	-0.0029**
	(0.0001)	(0.0007)	(0.0011)	(0.0008)	(0.0005)
Social demographic factors					
Age	-0.0006	-0.0001	-0.0001	0.0006*	-0.0006
	(0.0006)	(0.0002)	(0.0003)	(0.0003)	(0.0006)
Number of children	-0.0019**	-0.0064**	-0.0117**	0.0027**	0.0173**
	(0.0009)	(0.0030)	(0.0054)	(0.0011)	(0.0079)
Male	0.0002	0.0008	0.0015	-0.0003	-0.0022
	(0.0012)	(0.0041)	(0.0075)	(0.0017)	(0.0111)
Married	-0.0592	-0.1206*	-0.1260***	0.1678**	0.1380***
	(0.0409)	(0.0554)	(0.0271)	(0.0760)	(0.0264)
Educational attainment					
Compulsory education	-0.0144**	-0.0430***	-0.0723***	0.0132**	0.1165***
	(0.0051)	(0.0119)	(0.0172)	(0.0043)	(0.0313)
Senior school	-0.0124**	-0.0406***	-0.0759***	-0.0047	0.1336**
	(0.0038)	(0.0099)	(0.0174)	(0.0084)	(0.0374)
University	-0.0125***	-0.0465***	-0.0950***	0.0155	0.1694***
	(0.0032)	(0.0094)	(0.0190)	(0.0130)	(0.0421)
Health status					
Good Health	-0.0230***	-0.0736***	-0.1184***	0.0661***	0.1490***
	(0.0047)	(0.0103)	(0.0127)	(0.0120)	(0.0146)

Standard health	-0.0060**	-0.0200**	-0.0368**	0.0023	0.0605**
	(0.0022)	(0.0066)	(0.0120)	(0.0025)	(0.0220)

Note: (1) *, **, *** indicate 10%, 5%, 1% significant levels, respectively.

(2) Figures in parentheses are standard errors.

(3) ME presented in this table is average marginal effect.

(4) China General Social Survey 2010 is utilized.

Table 4- 10 Marginal Effects of the Ordered Probit Results for Discrepancy between Income Aspirations and Actual Income (Extended Model)

	Not at all happy	Not happy	Neutral	Happy	Very happy
Income categories					
Actual income	-0.0012**	-0.0041**	-0.0074*	0.0017	0.0109**
	(0.0007)	(0.0019)	(0.0039)	(0.0010)	(0.0034)
Discrepancy between aspiration and income(aspiration(ln)-actual income(ln))	-0.0004	-0.0012	-0.0020	-0.0005	0.0030
	(0.0008)	(0.0029)	(0.0052)	(0.0012)	(0.0077)
Partner income	-0.0013***	-0.0043***	-0.0079***	0.0018***	0.0112***
	(0.0003)	(0.0008)	(0.0014)	(0.0005)	(0.0021)
***Hukou* status**					
Rural	-0.0019	-0.0060	-0.0100	0.0045	0.0134
	(0.0015)	(0.0050)	(0.0093)	(0.0039)	(0.0144)
Interaction between *hukou* status and income categories					
Rural× actual income	0.0012	0.0042	0.0073	-0.0024*	-0.0105
	(0.0011)	(0.0033)	(0.0055)	(0.0012)	(0.0101)
Rural× discrepancy between aspiration and actual income	0.0014	0.0049	0.0091	-0.0020	-0.0136
	(0.0013)	(0.0046)	(0.0087)	(0.0018)	(0.0159)
Rural× partner income	0.0005**	0.0013**	0.0020**	-0.0015**	-0.0021**
	(0.0002)	(0.0006)	(0.0009)	(0.0007)	(0.0009)
Social demographic factors					
Age	-0.00002	-0.00004	0.0003	0.0003**	-0.0006
	(0.00006)	(0.0002)	(0.0004)	(0.0001)	(0.0007)
Number of children	-0.0022**	-0.0075**	-0.0137**	0.0031**	0.0202**
	(0.0009)	(0.0030)	(0.0054)	(0.0014)	(0.0080)
Male	0.0001	0.0001	0.0003	-0.0001	-0.0004
	(0.0012)	(0.0041)	(0.0075)	(0.0017)	(0.0111)
Married	-0.0526	-0.1123**	-0.1225***	0.1533*	0.1314***
	(0.0375)	(0.0548)	(0.0306)	(0.0736)	(0.0284)
Educational attainment					
Compulsory education	-0.0137**	-0.0416**	-0.0706**	0.0126**	0.1133***
	(0.0049)	(0.0118)	(0.0173)	(0.0041)	(0.0313)
Senior school	-0.0117**	-0.0391***	-0.0736***	-0.0037	0.1280**
	(0.0036)	(0.0099)	(0.0177)	(0.0079)	(0.0372)
University	-0.0117***	-0.0443***	-0.0909***	-0.0128	0.1596***
	(0.0030)	(0.0095)	(0.0194)	(0.0121)	(0.0418)
Health status					
Good Health	-0.0214***	-0.0701***	-0.1146***	0.0621***	0.1439***
	(0.0045)	(0.0101)	(0.0129)	(0.0118)	(0.0147)
Standard status	-0.0056**	-0.0187**	-0.0345**	0.0025	0.0562**
	(0.0022)	(0.0066)	(0.0121)	(0.0022)	(0.0219)
Financial security					

Debt	0.0051**	0.0170***	0.0296**	-0.0105**	-0.0411***
	(0.0018)	(0.0053)	(0.0087)	(0.0044)	(0.0114)
Social security					
Medical insurance	-0.0083**	-0.0264***	-0.0439***	0.0203**	0.0583***
	(0.0026)	(0.0072)	(0.0108)	(0.0077)	(0.0128)
Pension insurance	0.0001	0.0001	-0.0001	-0.00002	-0.0002
	(0.0013)	(0.0044)	(0.0008)	(0.001)	(0.0120)

Note: (1) *, **, *** indicate 10%, 5%, 1% significant levels, respectively.

(2) Figures in parentheses are standard errors.

(3) ME presented in this table is average marginal effect.

(4) China General Social Survey 2010 is utilized.

Table 4- 11 Marginal Effect of Basic Model including Economic Status

	Not at all happy	Not happy	Neutral	Happy	Very happy
Income categories					
Actual income	-0.0006*	-0.0019*	-0.0035*	0.0008*	0.0052*
	(0.0003)	(0.0010)	(0.0021)	(0.0005)	(0.0031)
Partner income	-0.0010**	-0.0033***	-0.0059***	0.0014**	0.0088***
	(0.0003)	(0.0008)	(0.0014)	(0.0004)	(0.0021)
***Huko*u status**					
Rural	-0.0010	-0.0025	-0.0031	0.0039	0.0026
	(0.0016)	(0.0050)	(0.0089)	(0.0061)	(0.0136)
Interaction between *hukou* status and income categories					
Rural× actual income	-0.0004	-0.0011**	-0.0020*	0.0006*	0.0028**
	(0.0005)	(0.0004)	(0.0010)	(0.0004)	(0.0010)
Rural× partner income	0.0006*	0.0018*	0.0032*	-0.0009**	-0.0046***
	(0.0003)	(0.0010)	(0.0017)	(0.0004)	(0.0008)
Social demographic factors					
Age	-0.00004	-0.00004	0.00008	0.0003**	-0.0003
	(0.00006)	(0.0002)	(0.0004)	(0.0001)	(0.0006)
Number of children	-0.0011	-0.0039	-0.0070	0.0016	0.0104
	(0.0009)	(0.0029)	(0.0053)	(0.0012)	(0.0078)
Male	0.0004	0.0014	0.0025	-0.0006	-0.0037
	(0.0012)	(0.0040)	(0.0073)	(0.0016)	(0.0108)
Married	-0.0629	-0.1260**	-0.1296***	0.1774*	0.1411***
	(0.0415)	(0.0545)	(0.0247)	(0.0950)	(0.0245)
Educational attainment					
Compulsory education	-0.0132**	-0.0402**	-0.0688***	0.0124**	0.1099**
	(0.0047)	(0.0116)	(0.0172)	(0.0041)	(0.0307)
Senior school	-0.0108**	-0.0363**	-0.0687**	-0.0023	0.1181**
	(0.0035)	(0.0097)	(0.0178)	(0.0041)	(0.0363)
University	-0.0101***	-0.0387***	-0.0794**	-0.0069	0.1351**
	(0.0028)	(0.0094)	(0.0199)	(0.0097)	(0.0399)
Health status					
Good Health	-0.0183***	-0.0618***	-0.1038***	0.0531**	0.1309***
	(0.0039)	(0.0095)	(0.0131)	(0.0110)	(0.0148)
Standard status	-0.0052**	-0.0174**	-0.0322**	0.0026*	0.0521**
	(0.0021)	(0.0065)	(0.0120)	(0.0019)	(0.0214)
Financial status					
Average economic status	-0.0108***	-0.0399***	-0.0739***	0.0217***	0.1030***
	(0.0020)	(0.0049)	(0.0082)	(0.0047)	(0.0107)

High economic status	-0.0096***	-0.0406***	-0.0935***	-0.0388**	0.1825***
	(0.0016)	(0.0043)	(0.0104)	(0.0132)	(0.0262)

Interaction between *hukou* status and financial status

Rural× average status	-0.0001***	-0.0017***	-0.0051***	0.0015***	0.0080***
	(0.00001)	(0.0005)	(0.0019)	(0.0006)	(0.0030)
Rural× high status	-0.0009***	-0.0023***	-0.0034***	-0.0004	0.0061***
	(0.00003)	(0.0002)	(0.0009)	(0.0010)	(0.0021)

Note: (1) *, **, *** indicate 10%, 5%, 1% significant levels, respectively.

(2) Figures in parentheses are standard errors.

(3) ME presented in this table is average marginal effect.

(4) China General Social Survey 2010 is utilized.

Table 4- 12 Marginal Effect of Extended Model including Economic Status

	Not at all happy	Not happy	Neutral	Happy	Very happy
Income categories					
Actual income	-0.0005	-0.0016	-0.0029	0.0007	0.0043
	(0.0003)	(0.0010)	(0.0022)	(0.0005)	(0.0025)
Partner income	-0.0009**	-0.0031***	-0.0057***	0.0013**	0.0085***
	(0.0003)	(0.0008)	(0.0014)	(0.0004)	(0.0021)
Hukou **status**					
Rural	-0.0009	-0.0023	-0.0028	0.0038	0.0021
	(0.0016)	(0.0051)	(0.0092)	(0.0061)	(0.0141)
Interaction between *hukou* status and income categories					
Rural× actual income	-0.0004**	-0.0010*	-0.0018	0.0005**	0.0027*
	(0.0001)	(0.0005)	(0.0025)	(0.0001)	(0.0012)
Rural× partner income	0.0005**	0.0015*	0.0026	-0.0008**	-0.0071**
	(0.0002)	(0.0009)	(0.0017)	(0.0003)	(0.0022)
Social demographic factors					
Age	-0.00009	0.00007	0.0003	0.0003**	-0.0006
	(0.00006)	(0.0002)	(0.0004)	(0.0001)	(0.0006)
Number of children	-0.0014	-0.0048	-0.0087	0.0020	0.0129
	(0.0009)	(0.0029)	(0.0054)	(0.0013)	(0.0089)
Male	0.0003	0.0009	0.0016	-0.0004	-0.0024
	(0.0011)	(0.0040)	(0.0073)	(0.0016)	(0.0108)
Married	-0.0574	-0.1195**	-0.1271***	0.1658*	0.1382***
	(0.0389)	(0.0542)	(0.0275)	(0.0935)	(0.0259)
Educational attainment					
Compulsory education	-0.0127**	-0.0393**	-0.0677***	0.0120**	0.1078**
	(0.0046)	(0.0115)	(0.0173)	(0.0040)	(0.0307)
Senior school	-0.0104**	-0.0353**	-0.0677**	-0.0018	0.1147**
	(-0.0034)	(0.0098)	(0.0180)	(0.0068)	(0.0362)
University	-0.0096***	-0.0372**	-0.0764**	-0.0056	0.1288**
	(0.0027)	(0.0095)	(0.0203)	(0.0091)	(0.0398)
Health status					
Good Health	-0.0174***	-0.0598***	-0.1013***	0.0508***	0.1278***
	(0.0038)	(0.0094)	(0.0132)	(0.0108)	(0.0149)
Standard status	-0.0048**	-0.0165**	-0.0306**	0.0026	0.0493**
	(0.0020)	(0.0065)	(0.0121)	(0.0018)	(0.0223)
Financial status					
Average economic status	-0.0101***	-0.0377***	-0.0704***	0.0202***	0.0980***
	(0.0019)	(0.0049)	(0.0083)	(0.0045)	(0.0107)

High economic status	-0.0092***	-0.0391***	-0.0896***	-0.0342**	0.1720***
	(0.0016)	(0.0044)	(0.0106)	(0.0125)	(0.0260)

Interaction between *hukou* status and income categories

Rural× average status	-0.0002***	-0.0008***	-0.0038***	-0.0015***	0.0058***
	(0.00004)	(0.00009)	(0.0005)	(0.0004)	(0.0003)
Rural× high status	-0.0009***	-0.0028***	-0.0051***	0.0021**	0.0110***
	(0.00006)	(0.0005)	(0.0012)	(0.0005)	(0.0023)

Financial security

Debt	0.0037**	0.0127**	0.0226**	-0.0072**	-0.0318**
	(0.0016)	(0.0050)	(0.0086)	(0.0030)	(0.0115)

Social security

Medical insurance	-0.0063**	-0.0207**	-0.0353**	0.0145**	0.0477***
	(0.0023)	(0.0050)	(0.0106)	(0.0065)	(0.0131)
Pension insurance	-0.0001	0.0005	0.0008	-0.0002	-0.0012
	(0.0013)	(0.0043)	(0.0079)	(0.0018)	(0.0118)

Note: (1) *, **, *** indicate 10%, 5%, 1% significant levels, respectively.

(2) Figures in parentheses are standard errors.

(3) ME presented in this table is average marginal effect.

(4) China General Social Survey 2010 is utilized.

Table 4- 13 Order Probit Model for Rural Migrants and Urban People Separately

	Rural-urban migrants		Urban people	
	Coefficient (A)	Marginal effect of very happy (B)	Coefficient (C)	Marginal effect of very happy (D)
Income categories				
Actual income	0.0339**	0.0110**	0.0289**	0.0078**
	(0.0138)	(0.0035)	(0.0136)	(0.0033)
Aspirations income	-0.0248	-0.0062	0.0340	0.0077
	(0.0547)	(0.0138)	(0.0383)	(0.0093)
Partner income	0.0453**	0.0111***	0.0523***	0.0134***
	(0.0140)	(0.0035)	(0.0108)	(0.0026)
Social demographic factors				
Age	-0.0605**	-0.0010	-0.0504**	-0.0006
	(0.0183)	(0.0012)	(0.0235)	(0.0008)
Age^2	0.0011**		0.0006**	
	(0.0004)		(0.0003)	
Number of children	0.0798**	0.0201**	0.9002***	0.0291**
	(0.0221)	(0.0084)	(0.0306)	(0.0108)
Male	-0.0490	-0.0123	-0.0013	-0.0002
	(0.0891)	(0.0222)	(0.0527)	(0.0128)
Married	0.7875**	0.1713***	0.8439**	0.1124***
	(0.2915)	(0.0191)	(0.2351)	(0.0330)
Educational level				
Compulsory education	0.4890***	0.1268***	0.3973**	0.1011**
	(0.1771)	(0.0333)	(0.1806)	(0.0483)
Senior school	0.5192***	0.1555**	0.4457**	0.1113**
	(0.1938)	(0.0619)	(0.1842)	(0.0494)
University level	0.6078**	0.1861**	0.5097**	0.1393***
	(0.2734)	(0.0925)	(0.1914)	(0.0522)
Health status				
Good health	0.6039***	0.1148***	0.6942***	0.1504***
	(0.1498)	(0.0296)	(0.0879)	(0.0173)
Standard health	0.1985**	0.0528	0.2755**	0.0634**
	(0.0705)	(0.0414)	(0.0935)	(0.0253)
Financial status				
Debt	-0.2084**	-0.0505**	-0.1545**	-0.0397**
	(0.0815)	(0.0190)	(0.0654)	(0.0144)
Social security				
Medical insurance	0.2696**	0.0619**	0.2907***	0.0578***
	(0.1139)	(0.0237)	(0.0759)	(0.0150)
Pension insurance	-0.1082	-0.0278	-0.0461	-0.0113
	(0.0831)	(0.0218)	(0.0607)	(0.0150)

Note: (1) It can be seen that there is no difference in the significance and direction for the variables except for income, aspiration income and partner income between urban people and rural-urban migrants. Although there is a slight difference in the magnitude, this chapter mainly focuses on the contribution of aspiration income and actual income to SWB so that I only set up interaction terms for these key variables in my analysis.

(2) China General Social Survey 2010 is utilized.

Table 4- 14 Relationship between Income and Father's Education (Potential Instrument Variable)

	Conditional Mixed Process (CMP)		Two Stage Least Square (2SLS)	
	Coef.	S.E.	Coef.	S.E.
Father education level				
Compulsory education	-0.0911	0.0725	-0.1292	0.0898
Senior school	-0.0085	0.1114	-0.0166	0.0919
University level	0.0365	0.1554	0.0032	0.1408
Other exogenous variables				
Rural	3.6286***	0.9168	1.3242**	0.6120
Aspiration income	0.8726***	0.0529	0.9019***	0.0672
Partner income	0.0304*	0.0158	0.0218	0.0200
Age	0.0258	0.0271	-0.0050	0.0256
Age square	-0.0001	0.0003	0.0002	0.0003
Number of kid	-0.0458	0.0475	-0.0639	0.0451
Male	0.5659***	0.0659	0.4077***	0.0581
Married	0.2088	0.4973	0.4109	0.5848
Compulsory education	0.1247	0.0895	0.1622**	0.0776
Senior school	0.2262**	0.1170	0.2791**	0.1098
University level	0.9120***	0.2089	0.9247***	0.2162
Good health	0.3822**	0.1120	0.3420**	0.1312
Standard health	0.4149**	0.1215	0.4004**	0.1358

Note: China General Social Survey 2010 is utilized.

Chapter Five

Conclusion

5.1 Summary of Findings

The *hukou* system (household registration system) divides the Chinese labour market into a dual structure which is composed of urban *hukou* and rural *hukou* individuals. This thesis focuses on exploring whether the dual characteristics of Chinese society has a series of negative impacts on the labour market in three aspects, including discrimination, economic development and social stability. Chapter 2 describes the phenomenon of inequality in the Chinese urban labour market, in terms of income and occupation distribution, and analyses whether the inequality in the Chinese urban labour market is attributed to *hukou* discrimination. Chapter 3 investigates whether discrimination in the urban labour market restricts rural-urban migrants' aspirations, motivation or ambitions, which in turn negatively influence the economic development of China. Chapter 4 compares the subjective well-being (SWB) between urban people and rural-urban migrants in urban China, and also estimates the impact of aspiration income on SWB. The main findings of each chapter are discussed below. The policy implications, limitations and suggestions for future work are also discussed at the end of this section.

With the reform of the *hukou* system, an increasing number of people in rural areas have migrated to urban cities and looked for suitable jobs. However, most of them suffer from unequal treatment in urban China, in comparison to their urban counterparts. Motivated by the inequality in the Chinese urban labour market, Chapter 2 investigates whether the different treatment of rural-urban migrants and urban people are ascribed to *hukou* discrimination, using data from the China Health and Nutrition Survey (CHNS). Unlike most previous Chinese studies, I take occupation segregation into account and utilize the Brown et al. (1980) decomposition in Chapter 2. I find that 11.64% more rural-urban migrants would have been employed in white collar jobs (senior professional, junior professional, administrative/executive/manager, office staff) if they

were treated the same as their urban counterparts. In addition, about 12% less rural-urban migrants would have been employed in blue collar jobs if they were treated the same as urban people. In comparison with the results from using the Oaxaca & Ransom (1994) decomposition without selectivity correction, the estimated results taking occupation segregation into account show that the total explained part from the decomposition results decreases to 69% and the unexplained part increases to 31%, which suggests that Oaxaca & Ransom (1994) decomposition underestimates the unexplained portion (discrimination). In addition, a large proportion of the wage differential is attributed to occupation segregation. The lower wages for rural-urban migrants is mainly due to the fact that they are restricted to obtain a higher standard job. Overall, the findings from the analysis suggest that occupational segregation faced by rural-urban migrants is more important in urban China, and this situation is mostly ascribed to the rural migrants' human capital investments, i.e. their lower education level.

Motivated by the wage discrimination and occupation segregation between rural-urban migrants and urban people investigated in Chapter 2, Chapter 3 contributes to previous Chinese studies by examining empirically whether *hukou* discrimination restricts rural-urban migrants' aspirations. Using data from the China General Social Survey 2010 (CGSS 2010), it is found that rural-urban migrants report lower income aspirations in comparison to urban people. Both the estimated results from OLS regressions and propensity score matching show that although rural-urban migrants suffer from discrimination in urban labour market, their income aspirations are not directly influenced by *hukou* discrimination. The explanation may be due to the following reasons: firstly, *hukou* discrimination may have no impact on restricting rural-urban migrants' income aspirations, even though they suffer from different treatment in the urban labour market. Rural-urban migrants may have already become accustomed to long-term discrimination. Secondly, *hukou* discrimination may determine rural-urban migrants' income aspiration indirectly. To be specific, people in rural areas are faced with different information and resources, and have more limited access to a good quality education. Moreover, unlike urban families, most rural families cannot afford the expensive tuition fees to attend university. Consequently, most of them are more likely to start working rather than pursuing further study. Their lower educational levels, indirectly caused by *hukou* discrimination, can restrict their income aspirations.

Following the previous two chapters, rural-urban migrants are faced with *hukou* discrimination in terms of wage and occupation opportunities, but their lower income aspirations seem not to be driven by *hukou* discrimination. Motivated by the previous two chapters, Chapter 4 investigates whether rural-urban migrants also have lower subjective well-being, and how their income aspirations shape their SWB in comparison to urban *hukou* individuals. Using the same dataset (CGSS 2010) as in Chapter 3, it is found that rural-urban migrants report similar SWB as their urban counterparts, although they have lower income. However, it is also found that the income aspirations have no significant impacts on individuals' SWB. The higher income aspirations of urban people do not offset the positive effect of their higher income on SWB. Moreover, there is no difference in the effect of aspirations on SWB between rural-urban migrants and urban people as well. Therefore, the similar SWB between urban people and rural-urban migrants cannot be explained by aspirations. In order to explain the similar SWB between rural-urban migrants and urban individuals, I follow previous studies and consider whether social comparisons shape their SWB. It is found that economic status positively influences SWB, and that it has a larger impact on rural-urban migrants' happiness compared to urban individuals. Therefore, the findings suggest that it is social comparison rather than aspiration income that explains the similar SWB between rural-urban migrants and urban people.

In summary, the conclusion of the whole thesis is that the *hukou* system still influences Chinese society and its urban labour market. Rural-urban migrants suffer from occupation segregation, which in turn exacerbates wage differentials. Moreover, rural-urban migrants have less access to social welfare. Although rural-urban migrants are still discriminated against in the urban labour market, their ambition or aspirations are not found to be directly influenced by *hukou* discrimination. Their lower income aspirations may be attributed to other factors, such as their educational background. In addition, although rural-urban migrants suffer from different treatment in the urban labour market, they still have similar levels of SWB to urban *hukou* residents.

5.2 Policy Implications

Although this thesis cannot be regarded as evidence for policy formulation, it can provide useful information for policy makers. Based on the empirical analysis in this

thesis, the distribution of resources between rural and urban areas and the reform of the *hukou* system should be highlighted. Firstly, the resource allocation should be balanced between the urban and rural areas. The educational resources in urban areas are superior to those in rural areas; thus the probability of urban students being admitted to university is much higher than for rural students. As a result, these highly educated urban people are more likely to obtain a good job with higher wages and better welfare. Additionally, rural people are aware of the poor educational resources in rural areas, and, as a result, they are reluctant to spend much money and time on educational investments. Individuals with such awareness will devote themselves to work instead. These outcomes can negatively affect long-term economic growth in China. Under the situation where the *hukou* system still exists, it is necessary to guarantee the educational resources as equal as possible in urban and rural areas, to ensure that rural students have the same probability as equally talented urban students of receiving a higher education or being admitted to universities.

In addition to the equalization of the resource distribution, policy formulation should also concentrate on the *hukou* system itself. For example, policy makers should focus on how rural *hukou* individuals can obtain urban *hukou* status or how rural people can enjoy the same rights as urban people, as it is currently difficult for rural people to change their *hukou* status. First of all, the reform of the *hukou* system should gradually relax the restrictions on population movement and variations of *hukou* status. Secondly, the two-tier society in China should be transformed into an integrated one, specifically the integration of the social security systems between urban and rural individuals. It should actively play the role of the government in embodying the principle of fairness in the social security system. In 2016, the Interim Regulations of Residence Permit was decreed in order to decrease the differences in social security treatments and basic public services between urban people and rural-urban migrants. Following the Interim Regulations of Residence Permit, if one person migrates to an urban city with a year legal residency and job, he/she can obtain a residence permit. In addition, he/she can enjoy the same basic public services as other local urban residents. However, there are still some limitations to the resident permit (i.e., the resident permit cannot replace *hukou* registration). Although rural-urban migrants having resident permits can enjoy most basic public services, there are still some rights that migrants cannot enjoy. For example, rural migrants have no access to buying a house or a car in the urban areas that

they migrate to, nor can they gain legal assistance. Moreover, in the absence of local urban *hukou*, there is no access to schools for rural-urban migrants' children in local urban area. These public services should be the responsibility of the local government, but these problems continue to result from citizens' *hukou* status.

In addition, basic public services do not include social welfare and social security, such as pensions insurance, unemployment benefits, allowances, resident medical insurance, and housing subsidies. These are implemented based on *hukou* status. If a rural migrant wishes to enjoy the same standard of social welfare and security, they must have a local urban *hukou*. The extent of public services that migrants can enjoy varies across cities. For some higher standard items, such as health care, there is still no access for rural migrants with a resident's permit. These examples show that equality between citizens with rural *hukou* and urban *hukou* has not been achieved by the Interim Regulations of Residence Permit. Recently, several coastal cities have started to carry out a "credit settling" policy. Migrants can obtain corresponding credit based on their work achievements, length of residence, educational background, and so on. Migrants with higher educational attainments and better jobs have a higher probability of obtaining a local urban *hukou*. At the moment, "credit settling" is a step in the right direction to obtain equality between rural-urban migrants and urban residents in urban cities, but it still maintains the duality of the *hukou* system and is only available in certain areas. The national government should implement a form of "credit settling" for the whole country in order to help capable rural-urban migrants obtaining local urban *hukou*.

5.3 Research Limitations and Further Study

There are some limitations to the analysis in the thesis. Firstly, due to very limited references in terms of the theoretical and technical approach on addressing the selection issue in the Brown et al. (1980) decomposition, it is difficult for me to obtain appropriate information to address the selection issue in Chapter 2 using the Brown et al. (1980) decomposition. When I compare the estimation of the Oaxaca & Ransom decomposition (1994) with and without selectivity correction, it can be seen that the sample selection issue has an impact on the estimated results. The unexplained portion in the Oaxaca & Ransom decomposition (1994) without selectivity correction is

downward biased; therefore, my estimation results based on the Brown et al. (1980) decomposition underestimate the actual discrimination faced by rural-urban migrants in the Chinese urban labour market. In the future, if a more systematic and specific approach to address the selectivity in the Brown et al. (1980) decomposition is introduced, or some other more advanced approach is put forward, I will consider the wage differential and occupation segregation in Chinese urban labour market in more depth.

In addition, the second limitation of this thesis should be ascribed to the dataset China General Social Survey 2010 (CGSS 2010) used in Chapter 3 and Chapter 4. Due to the problems of reverse causation and omitted variables, there are potential endogeneity problems in both Chapter 3 and Chapter 4. The limitations of the China General Social Survey (CGSS 2010) also make it difficult for me to find appropriate and corresponding instrumental variables for those endogenous explanatory variables (e.g. actual income and income aspirations). As the endogeneity issue cannot be addressed effectively by the CGSS 2010 dataset, the estimated results of these chapters may be biased. Nevertheless, the information of income aspirations can only be obtained from the CGSS 2010 at the moment. If there is a new survey dataset that includes information of aspirations appears in the future, it is worthwhile to investigate the questions in Chapters 3 and Chapter 4 further and try to overcome the existing endogeneity problems.

References

Aakvik, A. (2001). Bounding a Matching Estimator: The Case of a Norwegian Training Program. *Oxford Bulletin of Economics and Statistics*, 63(1), 115–143.

Abdel-Aty, M. (2003). Analysis of Driver Injury Severity Levels at Multiple Locations Using Ordered Probit Models. *Journal of Safety Research*, 34(5), 597-603.

Afridi, F., Li, S. X., & Ren, Y. (2015). Social Identity and Inequality: The Impact of China's Hukou System. *Journal of Public Economics*, *123*, 17-29.

Ai, C. & Norton, E. C. (2003). Interaction Terms in Logit and Probit Models. *Economics Letters*, 80(1), 123-129.

Akay, A., Bargain, O. & Zimmermann, K. F. (2012). Relative Concerns of Rural-to-Urban Migrants in China. *Journal of Economic Behavior & Organization*, 81(2), 421–441.

Alagaraja, M., McLean, G., Sims, C. & Hirudayaraj, M. (2016). The Impact of Colorism on the Career Aspirations and Career Opportunities of Women in India. *Advances in Developing Human Resources*, 18(1), 38-53.

Alesina, A., Di Tella, R. & MacCulloch, R. (2004). Inequality and Happiness: Are Europeans and Americans Different? *Journal of Public Economics*, 88(9), 2009–2042.

Altonji, J. G. & Blank, R. M. (1999). Race and Gender in the Labor Market. *Handbook of Labor Economics*, 3, 3143-3259.

Andrews, F. M. & Withey, S. B. (2012). *Social Indicators of Well-Being: Americans' Perceptions of Life Quality*. Springer Science & Business Media.

Arrow, K. (1973). The Theory of Discrimination. *Discrimination in Labour Markets*, 3 (10), 3-33.

Augurzky, Boris and Schmidt, Christoph M., The Propensity Score: A Means to an End (March 2001). *IZA Discussion Paper No. 271*. Available at SSRN: https://ssrn.com/abstract=270919.

Bandura, A. (1986). Social Foundations of Thought and Action: A Social Cognitive Theory. *Englewood Cliffs, NJ: Prentice-Hall.*

Bandura, A. (1989). Human Agency in Social Cognitive Theory. *American Psychologist,* 44 (9), 1175.

Becker Gary, S. (1957). *The Theory of Discrimination*. Chicago: University of Chicago Press.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, 70(5, Part 2), 9-49.

Becker, G. (1964). Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. *Published By National Bureau of Economic Research, New York*.

Bender, K. A. (2012). An Analysis of Well-Being in Retirement: The Role of Pensions, Health, and 'Voluntariness' of Retirement. *The Journal of Socio-Economics*, 41(4), 424-433.

De Beyer, J. & Knight, J. B. (1989). The Role of Occupation in the Determination of Wages. *Oxford Economic Papers*, 41(3), 595-618.

Bigler, R. S., Averhart, C. J. & Liben, L. S. (2003). Race and the Workforce: Occupational Status, Aspirations, and Stereotyping among African American Children. *Developmental Psychology*, 39(3), 572.

Blanchflower, D. G. & Oswald, A. J. (2004). Well-Being over Time in Britain and the USA. *Journal of Public Economics*, 88(7), 1359-1386.

Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *Journal of Human Resources*, 436-455.

Blundell, R., Dearden, L. & Sianesi, B. (2005). Evaluating the Effect of Education on Earnings: Models, Methods and Results from the National Child Development Survey. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(3), 473-512.

Boskin, M. J. (1974). A Conditional Logit Model of Occupational Choice. *Journal of Political Economy*, 82(2, Part 1), 389–398.

Brickman, P. & Campbell, D. T. (1971). Hedonic Relativism and Planning the Good Society. *Adaptation-Level Theory*, 287-305.

Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery Winners and Accident Victims: Is Happiness Relative? *Journal of Personality and Social Psychology, 36*(8), 917-927.

Brown, D. L. & Segrist, D. (2016). African American Career Aspirations: Examining the Relative Influence of Internalized Racism. *Journal of Career Development*, 43(2), 177-189.

Brown, R. S., Moon, M. & Zoloth, B. S. (1980). Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials. *Journal of Human Resources*, 3-28.

Brown, S. & Gray, D. (2014). Household Finances and Well-Being: An Empirical

Analysis of Comparison Effects. *IZA Discussion Paper No. 8530*. Available at SSRN: https://ssrn.com/abstract=2514703.

Brown, S., Taylor, K. & Price, S. W. (2005). Debt and Distress: Evaluating the Psychological Cost of Credit. *Journal of Economic Psychology*, 26(5), 642-663.

Bryson, A., Dorsett, R. & Purdon, S. (2002). The Use of Propensity Score Matching in the Evaluation of Active Labour Market Policies. *Working Paper No. 4, Department for Work and Pensions.*

Cai, F. (2011). Hukou System Reform and Unification of Rural–urban Social Welfare. *China & World Economy*, 19(3), 33-48.

Caliendo, M. & Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, 22(1), 31–72.

Caliendo, M., Hujer, R. & Thomsen, S. L. (2008). The Employment Effects of Job-Creation Schemes in Germany: A Microeconometric Evaluation. *In Modelling and Evaluating Treatment Effects in Econometrics*, (pp. 381–428). Emerald Group Publishing Limited.

Campbell, A., Converse, P. E. & Rodgers, W. L. (1976). *The Quality of American Life: Perceptions, Evaluations, and Satisfactions*. Russell Sage Foundation.

Card, D. & Krueger, A. B. (1991). *School Quality and Black-White Relative Earnings: A Direct Assessment.* National Bureau of Economic Research.

Centers, R. & Cantril, H. (1946). Income Satisfaction and Income Aspiration. *The Journal of Abnormal and Social Psychology*, 41(1), 64.

Chan, K. W. (2010). The Household Registration System and Migrant Labor in China: Notes on a Debate. *Population and Development Review*, 36(2), 357–364.

Charles, K.K. & Guryan, J. (2008). Prejudice and Wages: An Empirical Assessment of Becker's the Economics of Discrimination. *Journal of Political Economy*, vol.116(5), pp.773-809.

Chatterji, M., Mumford, K., & Smith, P. N. (2011). The Public–Private Sector Gender Wage Differential in Britain: Evidence from Matched Employee-Workplace Data. *Applied Economics*, *43*(26), 3819–3833.

Chau, T. W., Hongbin, L., Liu, P. W. & Zhang, J. (2007). Testing the Collective Model of Household Labor Supply: Evidence from China. *China Economic Review*, 18(4), 389-402.

Cheng, T. & Selden, M. (1994). The Origins and Social Consequences of China's

Hukou System. *The China Quarterly*, 139, 644-668.

Chi, W., Li, B. & Yu, Q. (2011). Decomposition of the Increase in Earnings Inequality in Urban China: A Distributional Approach. *China Economic Review*, 22(3), 299-312.

Clark, A. E. (1999). Are Wages Habit-Forming? Evidence from Micro Data. *Journal of Economic Behavior & Organization*, 39(2), 179-200.

Clark, A. E., Frijters, P. & Shields, M. A. (2008). Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles. *Journal of Economic Literature*, 46(1), 95-144.

Clark, A. E. & Oswald, A. J. (1994). Unhappiness and Unemployment. *The Economic Journal*, 104(424), 648-659.

Clark, A. E. & Oswald, A. J. (1996). Satisfaction and Comparison Income. *Journal of Public Economics*, 61(3), 359-381.

Cooke, F. L. (2004). Public-Sector Pay in China: 1949–2001. *The International Journal of Human Resource Management*, 15(4-5), 895-916.

Costa, P. T. & McCrae, R. R. (1980). Influence of Extraversion and Neuroticism on Subjective Well-Being: Happy and Unhappy People. *Journal of Personality and Social Psychology*, 38(4), 668.

Dalton, P. S., Ghosal, S. & Mani, A. (2016). Poverty and Aspirations Failure. *The Economic Journal*, 126(590), 165-188.

Dehejia, R. H. & Wahba, S. (1999). Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs. *Journal of the American Statistical Association*, 94(448), 1053-1062.

Démurger, S., Gurgand, M., Li, S. & Yue, X. (2009). Migrants as Second-Class Workers in Urban China? A Decomposition Analysis. *Journal of Comparative Economics*, 37(4), 610-628.

Diener, E. (1984). Subjective Well-Being. *Psychological Bulletin*, 95(3), 542–575.

Diener, E. (2000). Subjective Wellbeing: The Science of Happiness and a Proposal for a National Index. *American Psychological Association*, 55: 34-43.

Diener, E. & Larsen, R. J. (1984). Temporal Stability and Cross-Situational Consistency of Affective, Behavioral, and Cognitive Responses. *Journal of Personality and Social Psychology*, 47(4), 871.

Diener, E., Sandvik, E., Seidlitz, L. & Diener, M. (1993). The Relationship between Income and Subjective Well-Being: Relative or Absolute? *Social Indicators Research*,

28(3), 195-223.

Dolan, P., Peasgood, T. & White, M. (2008). Do We Really Know What Makes Us Happy? A Review of the Economic Literature on the Factors Associated with Subjective Well-Being. *Journal of Economic Psychology*, 29(1), 94–122.

Doleac, J.K. & Stein, C.D. (2013). The Visible Hand: Race and Online Market Outcomes. *The Economic Journal*, 123(November), F469-492.

Drichoutis, A. C. & Nayga Jr, R. M. (2011). Marginal Changes in Random Parameters Ordered Response Models with Interaction Terms. *Econometric Reviews*, 30(5), 565-576.

Drolet, M., & Mumford, K. (2012). The Gender Pay Gap for Private-Sector Employees in Canada and Britain. *British Journal of Industrial Relations*, *50*(3), 529–553.

Duesenberry, J.S. (1952). *Income, Saving and the Theory of Consumer Behavior.* Cambrige, Massachusetts: Harvard University Press.

Easterlin, R. A. (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence. *Nations and Households in Economic Growth*, 89, 89-125.

Easterlin, R. A. (1995). Will Raising the Incomes of All Increase the Happiness of All? *Journal of Economic Behavior & Organization*, 27(1), 35-47.

Easterlin, R. A. (2001). Income and Happiness: Towards a Unified Theory. *The Economic Journal*, 111(473), 465-484.

Easterlin, R. A., McVey, L. A., Switek, M., Sawangfa, O. & Zweig, J. S. (2010). The Happiness–Income Paradox Revisited. *Proceedings of the National Academy of Sciences*, 107(52), 22463-22468.

Esser, I. & Palme, J. (2010). Do Public Pensions Matter for Health and Wellbeing among Retired Persons? Basic and Income Security Pensions across 13 Western European Countries. *International Journal of Social Welfare*, 19(s1), S103-S120.

Evans, K. M. & Herr, E. L. (1994). The Influence of Racial Identity and the Perception of Discrimination on the Career Aspirations of African American Men and Women. *Journal of Vocational Behavior*, 44(2), 173-184.

Fairlie, R. W. (1999). The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment. *Journal of Labor Economics*, 17(1), 80-108.

Fearon, G. & Wald, S. (2011). The Earnings Gap between Black and White Workers in Canada: Evidence from the 2006 Census. *Relations Industrielles/Industrial Relations*,

324-348.

Ferrer-i-Carbonell, A. (2005). Income and Well-Being: An Empirical Analysis of the Comparison Income Effect. *Journal of Public Economics*, 89(5), 997–1019.

Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, 7(2), 117-140.

Frey, B. S. & Stutzer, A. (2010). *Happiness and Economics: How the Economy and Institutions Affect Human Well-Being*. Princeton University Press.

Frijters, P., Haisken-DeNew, J. P. & Shields, M. A. (2004). Money Does Matter! Evidence from Increasing Real Income and Life Satisfaction in East Germany Following Reunification. *The American Economic Review*, 94(3), 730–740.

Fuchs, C. J. & Landsberger, H. A. (1973). 'Revolution of Rising Expectations' or 'Traditional Life Ways'? A Study of Income Aspirations in a Developing Country. *Economic Development and Cultural Change*, 21(2), 212-226.

Gabriel, P. E. & Schmitz, S. (1989). The Effect of Occupational Segregation on the Relative Earnings of Black Males. *Economics Letters*, 30(1), 77-83.

Ge, Y. (2007). Sectoral Choice and the Gender Wage Gap: 1988-2001. *China Economic Quarterly-Beijing-*, *6*(2), 607. (In Chinese)

Gerdtham, U.-G. & Johannesson, M. (2001). The Relationship between Happiness, Health, and Socio-Economic Factors: Results Based on Swedish Microdata. *The Journal of Socio-Economics*, 30(6), 553-557.

Graham, C. & Felton, A. (2005). Does Inequality Matter to Individual Welfare? An Initial Exploration Based on Happiness Surveys from Latin America. *Center on Social and Economic Dynamics Working Paper Number 38*. Available at SSRN: https://ssrn.com/abstract=1028066 or http://dx.doi.org/10.2139/ssrn.1028066.

Greene, W. H. & Hensher, D. A. (2010). *Modeling Ordered Choices: A Primer*. Cambridge University Press.

Hamilton, G. G. (1990). Patriarchy, Patrimonialism, and Filial Piety: A Comparison of China and Western Europe. *British Journal of Sociology*, 77-104.

Han, C. (2015). Explaining the Subjective Well-Being of Urban and Rural Chinese: Income, Personal Concerns, and Societal Evaluations. *Social Science Research*, 49, 179-190.

Hansen, T., Slagsvold, B. & Moum, T. (2008). Financial Satisfaction in Old Age: A Satisfaction Paradox or a Result of Accumulated Wealth? *Social Indicators Research*,

89(2), 323-347.

Heckman, J. J. (2008). Econometric Causality. *International Statistical Review*, 76(1), 1-27.

Heckman, J. J., Ichimura, H. & Todd, P. E. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies*, 64(4), 605-654.

Helliwell, J. F. (2003). How's Life? Combining Individual and National Variables to Explain Subjective Well-Being. *Economic Modelling*, 20(2), 331–360.

Horwitz, A. V. (1984). The Economy and Social Pathology. *Annual Review of Sociology*, 10(1), 95-119.

Inglehart, R. (1990). *Culture Shift in Advanced Industrial Society*. Princeton University Press.

Imbens, G. W. (2004). Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. *Review of Economics and Statistics*, 86(1), 4–29.

Jackson, M. A. & Nutini, C. D. (2002). Hidden Resources and Barriers in Career Learning Assessment with Adolescents Vulnerable to Discrimination. *The Career Development Quarterly*, 51(1), 56-77.

Jann, B. (2008). The Blinder-Oaxaca Decomposition for Linear Regression Models. *The Stata Journal*, 8(4), 453-479.

Jiang, S., Lu, M. & Sato, H. (2012). Identity, Inequality, and Happiness: Evidence from Urban China. *World Development*, 40(6), 1190-1200.

Johar, M. (2009). The Impact of the Indonesian Health Card Program: A Matching Estimator Approach. *Journal of Health Economics*, 28(1), 35-53.

Kahneman, D. & Deaton, A. (2010). High Income Improves Evaluation of Life but Not Emotional Well-Being. *Proceedings of the National Academy of Sciences*, 107(38), 16489-16493.

Karaca-Mandic, P., Norton, E. C. & Dowd, B. (2012). Interaction Terms in Nonlinear Models. *Health Services Research*, 47(1pt1), 255–274.

Keng, S.-H. & Wu, S.-Y. (2014). Living Happily Ever after? The Effect of Taiwan's National Health Insurance on the Happiness of the Elderly. *Journal of Happiness Studies*, 15(4), 783-808.

Kidd, M. P. & Shannon, M. (1996). Does the Level of Occupational Aggregation Affect Estimates of the Gender Wage Gap? *Industrial & Labor Relations Review*, 49(2),

317-329.

Knight, J. (2008). Reform, Growth, and Inequality in China. *Asian Economic Policy Review*, 3(1), 140-158.

Knight, J. & Gunatilaka, R. (2010a). Great Expectations? The Subjective Well-Being of Rural–Urban Migrants in China. *World Development*, 38(1), 113-124.

Knight, J. & Gunatilaka, R. (2010b). The Rural–Urban Divide in China: Income but Not Happiness? *The Journal of Development Studies*, 46(3), 506-534.

Knight, J. & Gunatilaka, R. (2012a). Income, Aspirations and the Hedonic Treadmill in a Poor Society. *Journal of Economic Behavior & Organization*, 82(1), 67-81.

Knight, J. & Gunatilaka, R. (2012b). Aspirations, Adaptation and Subjective Well-Being of Rural-Urban Migrants in China. *Adaptation, Poverty and Development: The Dynamics of Subjective Well-Being*, 91-110.

Knight, J., Lina, S. & Gunatilaka, R. (2009). Subjective Well-Being and Its Determinants in Rural China. *China Economic Review*, 20(4), 635-649.

Knowles, J., Persico, N. & Todd, P. (2001). Racial Bias in Motor Vehicle Searches: Theory and Evidence. *Journal of Political Economy*, 109(1), 203-229.

Konings, J., & Vanormelingen, S. (2015). The Impact of Training on Productivity and Wages: Firm-Level Evidence. *Review of Economics and Statistics*, 97(2), 485-497.

Lam, K.-C. J. & Liu, P.-W. (2014). Socio-Economic Inequalities in Happiness in China and US. *Social Indicators Research*, 116(2), 509-533.

Lang, K. (2011). *Poverty and Discrimination.* Princeton University Press.

Lang, K. & Manove, M. (2011). Education and Labor Market Discrimination. *The American Economic Review*, 101(4), 1467-1496.

Lechner, M. (2008). A Note on the Common Support Problem in Applied Evaluation Studies. *Annales d'Économie et de Statistique*, 217-235.

Lee, L. (2012). Decomposing Wage Differentials between Migrant Workers and Urban Workers in Urban China's Labor Markets. *China Economic Review*, 23(2), 461-470.

Lelkes, O. (2006). Knowing What Is Good for You: Empirical Analysis of Personal Preferences and the 'Objective Good'. *The Journal of Socio-Economics*, 35(2), 285–307.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a Unified Social Cognitive Theory of Career/Academic Interest, Choice, and Performance. *Journal of Vocational Behavior*, 45, 79–122.

Lent, R. W., Brown, S. D., & Hackett, G. (2000). Contextual Supports and Barriers to Career Choice: A Social Cognitive Analysis. *Journal of Counseling Psychology,* 47, 36–49.

Liao, P.-A., Chang, H.-H. & Sun, L.-C. (2012). National Health Insurance Program and Life Satisfaction of the Elderly. *Aging & Mental Health*, 16(8), 983-992.

Liu, F. & Ma, H. (2009). Sensitivity Analysis on the Propensity Score Matching Method. *Statistics & Information Forum*, vol.24 (10) (In Chinese).

Liu, H. & Liu, L. (2017). Forty Years' Development Since the Recovery of Gaokao and Its Changes. *Journal of Higher Education,* vol.38 (10). (In Chinese).

Liu, P.-W., Zhang, J. & Chong, S.-C. (2004). Occupational Segregation and Wage Differentials between Natives and Immigrants: Evidence from Hong Kong. *Journal of Development Economics*, 73(1), 395-413.

Long, J. S. & Freese, J. (2006). *Regression Models for Categorical Dependent Variables Using Stata*. Stata Press.

Luttmer, E. F. (2005). Neighbors as Negatives: Relative Earnings and Well-Being. *The Quarterly Journal of Economics*, 120(3), 963-1002.

Lykken, D. & Tellegen, A. (1996). Happiness Is a Stochastic Phenomenon. *Psychological Science*, 7(3), 186-189.

Lyubomirsky, S., King, L. & Diener, E. (2005). The Benefits of Frequent Positive Affect: Does Happiness Lead to Success? *Psychological Bulletin*, 131(6), 803.

Mallick, D. (2008). Marginal and Interaction Effects in Ordered Response Models. *Munich Personal Repec Archive Paper No. 9617*.

Maxwell, N. L. (1994). The Effect on Black-White Wage Differences of Differences in the Quantity and Quality of Education. *Industrial & Labor Relations Review*, 47(2), 249-264.

McBride, M. (2001). Relative-Income Effects on Subjective Well-Being in the Cross-Section. *Journal of Economic Behavior & Organization*, 45(3), 251-278.

McBride, M. (2010). Money, Happiness, and Aspirations: An Experimental Study. *Journal of Economic Behavior & Organization*, 74(3), 262-276.

Meiyan, W. & Fang, C. (2010). Future Prospects of Household Registration System Reform. *The China Population and Labor Yearbook*, 2, 173-186.

Meng, X. & 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, Jacob (1958). "Investment in Human Capital and Personal Income Distribution". *Journal of Political Economy*. 66 (4): 281–302.

Mincer, J. (1974). Schooling, Experience, and Earnings. *Human Behavior & Social Institutions* No. 2.

Messinis, G. (2013). Returns to Education and Urban-Migrant Wage Differentials in China: IV Quantile Treatment Effects. *China Economic Review*, 26, 39-55.

Mumford, K., & Smith, P. N. (2007). The Gender Earnings Gap in Britain: Including the Workplace. *The Manchester School*, 75(6), 653–672.

Norton, E. C., Wang, H. & Ai, C. (2004). Computing Interaction Effects and Standard Errors in Logit and Probit Models. *Stata Journal*, 4, 154-167.

O'Neill, J. (1990). The Role of Human Capital in Earnings Differences between Black and White Men. *The Journal of Economic Perspectives*, 4(4), 25-45.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 693-709.

Oaxaca, R. L., & Ransom, M. R. (1994). On Discrimination and the Decomposition of Wage Differentials. *Journal of Econometrics*, 61(1), 5-21.

Okun, M. A., Stock, W. A., Haring, M. J. & Witter, R. A. (1984). Health and Subjective Well-Being: A Meta-Analyis. *The International Journal of Aging and Human Development*, 19(2), 111-132.

Oswald, A. J. (1997). Happiness and Economic Performance. *The Economic Journal*, 107(445), 1815-1831.

Pavot, W., & Diener, E. (2008). The Satisfaction with Life Scale and the Emerging Construct of Life Satisfaction. *The Journal of Positive Psychology*, 3(2), 137–152.

Peiro, A. (2006). Happiness, Satisfaction and Socio-Economic Conditions: Some International Evidence. *The Journal of Socio-Economics*, 35(2), 348–365.

Plagnol, A. C. (2011). Financial Satisfaction over the Life Course: The Influence of Assets and Liabilities. *Journal of Economic Psychology*, 32(1), 45-64.

Portes, A., McLeod Jr, S. A. & Parker, R. N. (1978). Immigrant Aspirations. *Sociology of Education*, 241-260.

Qiang, L. (2003). An Analysis of Push and Pull Factors in the Migration of Rural Workers in China [J]. *Social Sciences in China*, 1, 125-136. (In Chinese)

Ray, D. (2006). Aspirations, Poverty, and Economic Change. *Understanding Poverty*,

409-421.

Ritter, J. A. & Taylor, L. J. (2011). Racial Disparity in Unemployment. *The Review of Economics and Statistics*, 93(1), 30–42.

Roberts, R., Golding, J. & Towell, T. (1998). *Student Finance and Mental Health*. The Psychologist.

Rosenbaum, P. R. (2002). Observational Studies. *In Observational Studies* (pp. 1-17). Springer, New York, NY.

Rosenbaum, P. R. & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 41-55.

Rosenbaum, P. R. & Rubin, D. B. (1984). Reducing Bias in Observational Studies Using Subclassification on The Propensity Score. *Journal of the American Statistical Association*, 79(387), 516-524.

Rosenbaum, P. R. & Rubin, D. B. (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *The American Statistician*, 39(1), 33-38.

Rubin, D. B. & Thomas, N. (1996). Matching Using Estimated Propensity Scores: Relating Theory to Practice. *Biometrics*, 249-264.

Schmid, P. & Strauss, R. P. (1975). The Prediction of Occupation Using Multiple Logit Model. *International Economic Review*, 47–86.

Scott Long, J. (1997). Regression Models for Categorical and Limited Dependent Variables. *Advanced Quantitative Techniques in the Social Sciences*, 7.

Shields, M. A. & Price, S. W. (2005). Exploring the Economic and Social Determinants of Psychological Well-Being and Perceived Social Support in England. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(3), 513–537.

Sims, C. (2010). The Impact of African American Skin Tone Bias in the Workplace: Implications for Critical Human Resource Development. *Online Journal for Workforce Education and Development*, 3(4), 5.

Smith, J. A. & Todd, P. E. (2005). Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators? *Journal of Econometrics*, 125(1), 305-353.

Stevenson, B. & Wolfers, J. (2013). Subjective Well-Being and Income: Is There Any Evidence of Satiation? *The American Economic Review*, 103(3), 598–604.

Stutzer, A. (2004). The Role of Income Aspirations in Individual Happiness. *Journal of Economic Behavior & Organization*, 54(1), 89-109.

Stutzer, A. & Frey, B. S. (2006). Does Marriage Make People Happy, or Do Happy People Get Married? *The Journal of Socio-Economics*, 35(2), 326–347.

Tang, Z. (2014). They Are Richer but Are They Happier? Subjective Well-Being of Chinese Citizens Across the Reform Era. *Social Indicators Research*, 117(1), 145-164.

Tovar-Murray, D., Jenifer, E. S., Andrusyk, J., D'Angelo, R. & King, T. (2012). Racism-Related Stress and Ethnic Identity as Determinants of African American College Students' Career Aspirations. *The Career Development Quarterly*, 60(3), 254-262.

Veenhoven, R. (1991). Is Happiness Relative? *Social Indicators Research*, 24(1), 1-34.

Vendrik, M. C. M. & Woltjer, G. B. (2007). Happiness and Loss Aversion: Is Utility Concave or Convex in Relative Income? *Journal of Public Economics*, 91(7), 1423-1448.

Witter, R. A., Okun, M. A., Stock, W. A. & Haring, M. J. (1984). Education and Subjective Well-Being: A Meta-Analysis. *Educational Evaluation and Policy Analysis*, 6(2), 165-173.

Wooldridge, J. (2009). Introductory Econometrics: A Modern Approach (Mason, OH: South-Western). *R. Desbordes and V. Verardi*, 181.

Xu, X. (2015). The Trend of Intergenerational Transmission of Income Gap of Chinese Urban and Rural Residents: 2002-2012. *China Industrial Economics*, 324 (3), 5-17. (In Chinese)

Yakovlev, P. & Leguizamon, S. (2012). Ignorance Is Not Bliss: On the Role of Education in Subjective Well-Being. *The Journal of Socio-Economics*, 41(6), 806-815.

Yang, Z. (2017). An Analysis of the Development of China's Rural Migrant Workers with Chinese Characteristics. *Chinese Rural Economy*, (10), 38-48. (In Chinese)

Zhang, L., Cheng, X., Tolhurst, R., Tang, S. & Liu, X. (2010). How Effectively Can the New Cooperative Medical Scheme Reduce Catastrophic Health Expenditure for the Poor and Non-Poor in Rural China? *Tropical Medicine & International Health*, 15(4), 468–475.

Zhang, X., Yang, J. & Wang, S. (2011). China Has Reached the Lewis Turning Point. *China Economic Review*, 22(4), 542-554.

Zhao, Z. (2004). Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence. *Review of Economics and Statistics*, 86(1), 91-107.

Zhou, G., Tan, H., and Li. L. (2017). Does Migration Experience Promote Entrepreneurship in Rural China? *China Economic Quarterly*, 16(2), 793-814. (In Chinese).