Essays in Migration, Trade and the Labour Markets

Jiahong Han

A thesis submitted for the degree of

Doctor of Philosophy in Economics

Newcastle University Business School

September 2023

Abstract

Over the last thirty years, China's economic growth has been remarkable. Although much is still unknown about the process leading to this growth, a rapid structural transformation has also taken place over last three decades. Internal migration is central to this transformation. Meanwhile, China also opened its goods market to global trade and gradually became a key player in the world economy. Trade liberalisation has had a major impact on the Chinese economy. Workers' mobility and trade-led growth are connected and, indeed, trade and migration policies are often determined simultaneously. This thesis investigates how labour and goods markets frictions affect labour market outcomes in China. Specifically, it asks three research questions: (i) how do institutional restrictions to migration shape migration decisions in China?; (ii) does internal rural-urban migration affect urban natives' wages and how is this relationship affected by migration policies?; and (iii) are the effects of trade liberalisation at the sectoral level, sectoral wage premia in particular, related to migration restrictions?

To address the above questions, the thesis is organised as follows. Chapter 1 provides the background contextual information and motivation for the thesis. Chapter 2 examines the importance of migration restrictions in China by studying the role of the *Hukou* system in shaping migration decisions. Migration becomes costly in the context of the *Hukou* system as *hukou* prohibits migrants from accessing various social benefits in their actual destinations of residence. This chapter exploits a gravity equation modelling approach of migration stocks and a Bartik-type instrument and provides evidence that migration decisions are associated with economic factors such as wage and unemployment characteristics, with stronger effects in destinations where *hukou* restrictions are stronger. Also, migration restrictions can influence the relevance of other determinants of migration, such as public and amenity services in destinations. The results are robust across alternative estimation models. These findings suggest that internal migration is sensitive to *Hukou* policies, and migration restrictions are an important factor in understanding internal migration patterns in China.

Chapter 3 examines the impact of an increase in migration from rural to urban areas on native workers' wages in Chinese cities. Specifically, we exploit cross-city variations in local labour market conditions that arise from non-uniform *Hukou* reform

implementations. The identification strategy exploits variations in the relaxation of the internal migration restrictions across cities, and variations in the pre-reform migration flows across cities between 2000 and 2005. We find that, on average, an increase in rural-urban migration increases the wages of urban workers, with the effect being higher for high-skilled urban workers. Using a shift-share instrument to further examine the effect of rural-urban migrants on native workers' wages leads to similar evidence.

Chapter 4 looks at the relationship between trade openness and inter-industry wage differentials in China. Taking advantage of a rich household survey dataset, this chapter first estimates the wage premium for each industry conditional on individuals' characteristics. We then empirically assess the relationship between wage premia and trade openness. We find that trade openness has a positive effect on wage premia. Also, disaggregating sectors into tradable and non-tradable sectors shows that the positive effect of openness on the wages of the tradable sectors feeds into the wages of the non-tradable sectors. We show that the relationship between tradable and non-tradable sectors increases with the migration restrictions caused by household registration system in China.

Acknowledgements

First and foremost, I would like to thank my supervisors Giorgio Fazio and John Sessions for their support and advice, Pascal Mossay and Jong Kook Shin for their previous involvement in the supervision process. They have created a supportive and friendly environment for me to explore the topics I was truly interested in. Without their precious feedback and guidance, this thesis would have not been possible.

I am grateful for the guidance from Pascal Mossay and Jong Kook Shin who were part of the supervision team at the early stage of my PhD. In addition, I appreciate the advice on the PhD thesis from the Panel members, Nils Braakmann and Atanu Ghoshray.

I would also like to thank all colleagues and other faculty and professional support staff of the Business School who have provided me valuable support over the years, particularly those who have always stood by me and shared the ups and downs of the PhD process – together we supported each other and made the supposed lonely experience more pleasant. I gratefully acknowledge the three-year NUBS studentship and financial support from the Newcastle University.

Next, I must express my heartfelt gratitude to my parents, Changcui Zhou and Cunzhen Han, who give me priceless love and unconditional support. Without them, I would not have not been able to get through the difficult times. They are my everyday motivation and inspiration. I know I am biased, but they are the best parents in the world.

Finally, a special thanks to my friends Hui Sun, Suqiao Wang and Teng Zhang. You are always reliable and trustable, for now and forever.

Declaration

I declare that this thesis produces original work, under the guidance and supervision of Professor Giorgio Fazio and Professor John Sessions. They provided comments on drafts of all chapters of this thesis.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Research Background

China' economic growth has been impressive since 2000. A structural transformation of the Chinese economy and society has taken place over the last three decades, especially in terms of internal migration. Labour has moved from agricultural employment in rural areas to industry and services in urban areas (rural to urban migration), together with urban-to-urban migration between cities. In the meantime, China has gradually opened its goods market to global trade and joined the World Trade Organisation (WTO) in 2001. Trade liberalisation has had a major impact on the Chinese economy, contributing to increasing firm productivity (Yu, 2015; Brandt et al., 2017), product quality (Fan et al.,2015), increased skill premium (Chen et al.,2017; Li, 2018). The reduction of trade costs and the growth of the Chinese export opportunities increased the demand for labour. This should be followed by factor market adjustments and a flexible domestic labour market is needed. In 2014, there were 278 million individuals living in outside their hometowns for at least six months, which accounted for approximately 20% of the country's total population (Facchini et al., 2019). The world has never seen such "Great Migration" within such a short time. The large movement from low productivity sectors/regions to high productivity sectors/regions is the main driver of the unprecedented economic growth in China. Labour mobility and the export-led growth are connected. Indeed, at the country level, trade policies and migration policies are determined simultaneously. Low labour market and mobility frictions could enhance the benefits from international trade and lead to gains from both trade and migration.

However, the institutional impediments play an important role in the allocation of labour. Migrant workers face inequalities in China, mainly because migrants are denied the local status of dwellers as long as they do not hold a local household registration, or *hukou.* By granting differential access to social welfare between urban and rural residents and between local and non-local residents, the system can be a barrier to mobility. In addition, migrants often take jobs which are regarded as *3D* (dirty, dangerous, and demanding), and most of the local workers are unwilling to take these types of jobs. The labour markets for urban native workers and migrant workers are potentially segregated.

The issues of China’s mobility barriers and of internal migration have long interested researchers and policy makers. Whether China has reached the “Lewis point” where rural labour supply is fully absorbed and unskilled wages begin to rise substantially has been hotly debated. A World Bank (2007) report dismisses the idea that China has reached the Lewis turning point. Meng (2012) argue that it is unlikely to be a result of labour shortages but rather institutional restrictions of migration. Ngai et al. (2019) argue that the barriers to mobility have led to lower levels of urbanisation and industrialisation because of the over-employment in the agricultural sector and under-employment in the non-agricultural sectors. Au and Henderson (2006) show that city sizes are smaller due to the existence of mobility barriers out of agriculture. Tombe and Zhu (2019) conclude that the observed more efficient labour allocation across prefectures in China account for 20% of TFP growth. Kinnan et al. (2018) find that improved access to migration smoothens the consumption volatility of households.

In addition to the already mentioned literature studying the effects of institutional mobility barriers, some papers study the interactions between trade liberalisation and migration policies, looking into issues such as how welfare gains from trade liberalisation are affected by migration frictions (Ma and Tang, 2020; Tombe and Zhu, 2019) or how the effects on migration depend of trade liberalisation (Facchini et al., 2019; Tian, 2022). Ma and Tang (2020) show that international trade liberalisation and migration are substitutes to each other as migration provides sufficient labour supply to export-led growth. Tian (2022) highlights the impact of trade liberalisation on changes in labour institutions that regulate internal migration.

This thesis investigates how labour and goods markets frictions affect labour market conditions in China by asking three research questions: (i) how do institutional restrictions to migration shape migration decisions in China?; (ii) does internal rural-urban migration affect urban natives’ wages and how is this relationship affected by migration policies?; and (iii) are the effects of trade liberalisation at the sectoral level, sectoral wage premia in particular, related to migration restrictions?

1.2 Institutional Background of the *Hukou* System

Despite the many structural changes in the Chinese economy, internal mobility has mostly remained restricted for a long period of time. While in most countries citizens are free to move internally, the unique household registration system (*hukou*) imposes substantial limits in internal mobility in China. Under the *hukou* system, each person is born with a *hukou* booklet that records a person's *hukou* registration place and *hukou* type. Each citizen is classified in an agricultural or non-agricultural *hukou* (sometimes referred to as rural or urban *hukou*). A *hukou* is primarily inherited from one's parents at the time of birth, and children born in urban areas to parents with rural *hukou* are similarly designated as rural *hukou* holders. People who live in their original *hukou* registration place are considered as local residents, while people living outside their registered place are considered as non-local residents. Based on this *hukou* system, the Chinese government is able to implement regulations on internal migration.

The *hukou* system also ties a person's access to various public services to her/his residential status such as medical services, public education services, etc. To be more specific, one may move to a new place but not have access to local social services, such as access to local schools, claim on local pension plan and so on. Also, those who have moved to new places may face difficulties to obtain the local citizenship (*hukou*). The stringency of the *hukou* system differs across cities/provinces. Each city sets its own residency requirements based on various criteria, such as what kind of jobs someone does and how long they have been involved in the local social-insurance system. Compared to big cities, smaller cities have fewer limits for migrants' access to the local *hukou* system. Specifically, as a government officer said, *"People who have legal dwellings, including rented homes, can apply for a small city hukou. If they want a Hukou for medium-sized cities with a population of 500,000 to one million, they must pay social insurance for at least one year. If they want to apply for household registration in big cities with a population of over one million, they have to pay social insurance for at least two years."*[1] Following is an example of the *hukou* registration criteria in 2014.

[1] This is from a piece of news report on the website http://www.china.org.cn.

Figure 1.1 An Example of the *Hukou* Registration Criteria (2014)

Who Gets a Hukou?

Migrants in China's cities have to meet a daunting array of requirements before they can apply for official resident status.

	County-level cities and other small towns (e.g. Dunhuang)	Population 500,000 to 1 million (e.g. Dandong)	Population 1 million to 5 million (e.g. Qingdao)	Population 5 million and above (e.g. Beijing)
Stable accommodation	✓	✓	✓	✓
Steady job		✓	✓	✓
Paid into local social security for minimum time		✓	✓	✓
Steady job for minimum time			✓	
Minimum continuous residency				✓
Other, unspecified requirements				✓

Sources: State Council of China; U.N. Population Division

The Wall Street Journal

Due to the economic reform and open policies in 1978, an increasing labour demand contributed to massive internal migration, which has been achieved also by the relaxation of the migration restrictions. Since 1980s, the government has taken steps to relax the *hukou* system to make it easier for people to move to a different place. Converting to a local *hukou* was relaxed for people migrating to small cities and towns. Without converting to a local *hukou*, migrants were denied accessing to most of social benefits, such as subsidised medical care, education for their children, pension, etc. These reforms, however, initially aimed to help rural people to settle down in urban areas of small cities and towns, while the relaxations in medium and big cities were limited. As a result, there was labour shortage in urban areas of medium- and large-sized cities, making the impact of *hukou* reform limited. The mismatch between labour demand and migration flows distorts labour allocation, investment and consumption.

For example, in the early years of industrialisation, the government encouraged the growth of urban private sectors, which employed mainly unskilled workers to produce labour-intensive goods.

However, inherited institutional impediments still play an important role in the allocation of labour. As a result of the above-mentioned migration discriminations, migrants do not see long-run future in cities. Typically, migrants leave their families behind hoping to earn as much as possible before returning to their hometowns. People may have to return to their original *hukou* registration place for different reasons. For example, women may have to return to hometowns when they have children, to take advantage of free social services; families may have return to their original places when they have children of school age and want to take advantage of free education service.

Starting from 2002, the reforms have been implemented to further relax the *hukou* system in medium- and large-sized cities. These were initially implemented by some local governments and then extended to the whole country over the following 12 years. In 2014, an easier *hukou* conversion policy has finally announced at the national level. The government published the "Opinions on Further Promoting the Reform of the *Hukou* System" to gradually abolish the distinction between urban *hukou* and non-urban *hukou*, to abolish the *hukou* restrictions in towns and small cities, to remove the *hukou* restrictions in middle-sized cities, to relax the restrictions in big cities, but to maintain the restrictions in the very large cities. The timing of reforms implementation varied across provinces and across cities. The main factors influencing the timing of reforms are the increase in urban labour demand, the desire by the local government to improve demand for local real estate, and political incentives.

There are local laws and regulations provide the details for the reform implementation. Several studies (Kinnan et al.,2018; Sun et al., 2011; Wang et al., 2021) have identified the information of local laws and regulation through several databases: Peking University's Chinalawinfo, Xihu Law Library (*www.law-lib.com*), Beijing Lawstar Tech Limited Company (*www.law-star.com*) and Zhengbao Online Education Company's database.

1.3 WTO Accession and Trade Openness

Exporting has been considered as an effective strategy to achieve industrialisation and economic growth in developing countries. Trade liberalisation in China has experienced two reforms. The first was the so called “open-door policy” in the early 1980s. The government began to reduce the barriers to foreign trade and investment. Before 1978, China had a centrally planned economy. Individuals and private corporations were not allowed to trade without approval from central government. Domestic commodity prices and international currency exchange were highly restricted. In 1980s, the government loosened the administrative restrictions on imports and exports with tariffs, quotas and licensing. Also, foreign direct investment was encouraged. In addition to loosening trade restrictions, special economic zones were established in Shenzhen, Zhuhai, Shantou and Xiamen in 1980, offering tax incentives to foreign investors, reducing administrative procedures for foreign investment projects, and favourable operating environment for importing and exporting. Beyond trade policies, industrial reforms were also implemented, allowing private enterprises to operate and state-owned enterprises to export excess output above the quota. The open-door policies boosted trade activities, with an increase in the trade share from 13% in the late 1980s to 20% in 2000. The second shock of trade liberalisation is China’s accession to the WTO. After fifteen years of negotiations, in 2001 China joined the WTO, which represent one of the most relevant economic changes in the new century. In the accession agreement, China and trade partners committed to reducing import tariffs, removing quotas and reducing other tariff barriers. For example, China reduced its tariffs significantly from 35% in 1994 to 17% in 1997, and the tariff rates were further reduced after China’s accession to the WTO. At the same time, China benefited from the lower tariff rates by its trading partners. Specifically, China started to enjoy the Most Favoured Nation (MFN) tariffs. Previously, regular Congressional approval was required to maintain MFN status and it was subject to yearly renewal.

This resulted in an uncertain trade environment for Chinese exporters. Without the MFN status, China would have faced a 35% tariff rate by the U.S. in 2000. By comparison, the MFN rate was only 4%. The uncertainty induced by China’s temporary MFN status was substantial and prevented further expansion on trade for Chinese exporters (Pierce and Schott, 2016; Handley and Limao, 2017). With China’s entry into the WTO, China received permanent MFN status, and the decline of export tariff was sizable. By receiving Most Favoured Nation (MFN) status after the accession to the WTO, tariffs on China's exports fell, which was followed export growth. For example, exports increased from 248 billion USD in 2000 to 1577 billion USD in 2011 (Li et al., 2019). During the last two decades, the average exports growth rate has been approximately 17%. In addition to the MFN tariff rates, there are bound tariff and effectively applied tariff rates. The bound tariff is the maximum MFN tariff level for a commodity line, which gives the WTO members flexibility to decrease or increase their tariffs as long as they do not raise them above their bound levels. The effectively applied tariff is defined as the lowest available tariff between the MFN and preferential tariffs. The preferential tariff is the tariff that falls under a preferential trade agreement. Normally, a country that joins a preferential trade agreement promises to give another country's products lower tariffs than their MFN rate.

After China's entry into the WTO, the tariff rate on Chinese goods in 1995 was 5.8 percentage points and declined to 4.1 percentage points in 2001 and declined further in 2007 to 3 percentage points (Tian, 2022). There effects of tariff reductions at the industry level have been substantial. Li et al. (2019) suggested that the higher the initial export tariff level, the larger the tariff reduction imposed on China's exporters after China's accession to the WTO. Tian (2022) estimates that a one percentage point reduction in export tariff rates faced by Chinese exporters induces a 13%-14% in export volumes. Importantly, the export tariff reductions are unlikely to be predicted in advance and unrelated with economic factors in China.

1.4 Outline of the Thesis

This thesis aims to understand how labour and goods market friction affect labour market conditions in China. The first part of the thesis focuses specifically on examining the effect of labour frictions (*hukou* system) on labour markets in China, especially whether the *hukou* reform is effective in booming the local economy by attracting migrants. The second part investigates the impact of trade openness on wages, which accounts for the effect of *hukou* restrictions on wage inequality. This thesis contains three core chapters around these central themes.

Chapter 2, titled "how does *Hukou* system affect internal migration in China" examines the importance of migration restrictions in China by studying the role of the *hukou* system in shaping migration decisions. Regional inequality in China has been a hotly debated topic. For example, the urban housing earnings were approximately 2.5 times rural household earnings in the 2000s (Sicular et al., 2007). Regarding regional inequality, the average earnings in coastal regions was 2215 RMB compared to 1652 RMB in interior regions (Benjamin et al.,2005). Although regional inequality has been considered as a key driver of migration, a very puzzling fact in China's context is the coexistence of large income disparity and migrant labour shortage. According to Combes and Zhu (2019), there were 49 million inter-provincial workers, which only accounted for 7.2 percent of total employment, in 2005. A critical question is why the workers do not take advantage of wage differentials to move across cities. One explanation is that of unique state policies towards migration management preventing workers from taking advantage of regional wage differentials.

The objective of chapter 2 is to investigate whether the *hukou* system is a barrier to internal migration and whether it enhances the responsiveness to economic factors (such as wages and employment probabilities). In this chapter, we first present a model across different destinations. This framework follows a random utility model where each migrant maximises their utility by choosing a location among potential destinations (Grogger and Hanson, 2011). The probability of migration depends on the differences between wages related to staying at origins or migrating to other places adjusted for the costs of migration, such as distance, and other impediments of migration. To facilitate empirical analysis, we estimate a regional-level panel data model and a multi-regional model that rests on the logic that migrants move to increase their wages. Specifically, we focus on the responsiveness of internal migration to *hukou* restrictions. *Hukou* policy is highly localised and there are large variations on *hukou* stringency across cities. By analysing *hukou* policy documents, Zhang et al. (2019) computed the *hukou* index for all policy categories in 120 cities using Projection Pursuit Model (PPM) in both 2000-2013 and 2014-2016. The *Hukou* index measures how difficult it is for migrants to obtain a local *hukou* in the place of their residence. The *hukou* index ranges from 0.2451 (Gansu Province) to 2.496 (Beijing City) in 2000-2013, and from 0.3507 (Shanxi Province) to 2.6284 (Beijing City) in 2014-2016, with a higher value indicating more stringent restrictions and thus greater difficulty for migrants to obtain local *hukou* registration. The main area of interest is the local labour market at the city-level. The data is collected from the 2005, 2010, 2015 China Population Census Survey and from various China City Statistics Yearbooks.

Regional wage differentials are a main driver of interregional migration in China, with higher wages in one region pulling stronger immigration. However, the positive impact may reflect causation in the opposite direction. The inflow of migrants may positively affect wages in the city of destination. In particular, migrants are mostly low-skilled workers in China and their skills complement those of native workers (Combe et al., 2015). The complementarities between natives and migrants increase the wages of natives when the labour supply of migrants increases. In addition, destinations income may be influenced by other factors which are potentially related to the migration costs of the inflow of migrants, such as institution quality and first-nature geography. Endogeneity issue arises when estimating a gravity model. To reduce endogeneity concerns, following Tombe and Zhu (2019), we construct a Bartik-style instrumental variable. That is, wage is instrumented by each city's expected wage which is constructed by interacting national average earnings weighted by sector composition with the distribution of employment shares across sectors in each city.

The results demonstrate that, all else equal, migrants tend to favour destinations with higher wage differentials and favourable probability of finding employment. Importantly, *hukou* stringency plays an essential role in internal migration. Increased migration restrictions increase the effect of migration cost and enhance the importance of economic factors such as wage differentials and employment probabilities. Chapter 2 is also the first to quantitatively estimate the effect of migration regulations on migration flows in the Chinese context using specific measures of *hukou* stringency. This study can serve as guidance for policy makers to improve migration management, to achieve rapid urbanisation and industrialisation targets.

It is costly for people to move from urban-to-urban areas and from rural to urban areas. There was surplus labour in rural areas and labour shortage in urban areas (Meng, 2012). Since the 2000s, China has launched some *hukou* reforms to eliminate the distinction between rural and urban *hukou* types and to attract rural workers to urban areas. Such a policy change could be seen as a positive labour supply shock to the urban labour market by allowing more rural workers to move to the urban areas. Much efforts have been made to estimate the effect of rural-to-urban migration on the wages of urban natives in the local labour market in China and they have contradictory findings. In chapter 3, we contribute to this literature by examining the impact that an increase in migration from rural to urban areas has on native workers' wages in Chinese cities. To do this, chapter 3, titled "Internal Migration and Urban Wage: Evidence from *Hukou* Reform in China", exploits variations in the relaxation of the internal migration restrictions across cities. The aim of chapter 3 is to investigate whether internal migration affects urban natives' wages and whether this relationship is influenced by *hukou* policies and *hukou* reforms using the 2005 China Population Census Survey. Relaxing the *hukou* restrictions reduces rural-urban income disparity, increases the urbanization rate (Au and Henderson,2006; Ngai et al.,2019; Whalley and Zhang, 2007), increases productivity and reduces welfare distortions (Hsu and Ma, 2019; Tombe and Zhu, 2019). However, most papers focus on aggregate analysis. Chapter 3 provides evidence from the micro-level perspective on how the *hukou* reforms affect the average wage of local labour markets through shaping migration patterns.

In the first instance, chapter 3 employs an Oaxaca-Blinder decomposition to examine which factors influence China's unequal wage structure between rural migrants and urban natives. The results show that the main sources of wage differential between rural migrants and urban natives are productivity-linked characteristics and differences in employment distributions across sectors, ownership and occupation. This evidence indicates job-market segregation between rural-urban migrants and urban natives, as a result of which, rural migrants are unlikely to substitute urban natives. After the above decomposition, two stage least square estimation is employed in order to estimate the causal relationship between migrant share and urban native wages. The identification strategy follows Nunn and Qian (2014) and Sequeira et al. (2020) to construct a DID style instrument that exploits two facts: variations in the relaxation of the internal migration restrictions across cities, and variations in the pre-reform migration flows across cities between 2000 and 2005. A Shift-share instrument is also constructed as a robustness check. The interaction term between *hukou* reforms ("the shift") and historical migration patterns between prefectures ("the share") is employed to construct exogeneous predictor of migrant shares.

The results of chapter 3 suggest rural-urban migrants have positive effect on the wages of urban native wages, with a larger effect on high-skilled urban natives. The positive migrant impact can be attributed to gains from complementarity with natives in the production function. The results support the wage gains that can be expected from further migration and urbanisation in China. The findings in chapter 3 provide evidence of rural migrants' contribution to China's urbanisation and economic growth. These findings also offer important policy implications. Though there are fewer mobility restrictions, the Chinese labour market is still far from being a flexible labour market. The positive impact of rural-urban migrants on urban labour markets imply that further labour mobility relaxation may be necessary to increase productivity in urban labour markets.

At the country level, trade policies and migration policies are often determined simultaneously. Countries with faster economic growth might choose both a more open trade policy and a more flexible migration policy. Chapter 4, titled "Trade Openness and Wage Premium: Evidence from China", moves away from examining the direct effect of migration policies on labour markets and moves towards investigating the effect of trade policies on local labour market combined with that of internal migration policies. This chapter contributes to the impact of trade openness in China. Recent studies investigate the local labour market adjustment in terms of employment, wages, migration as a result of China's WTO accession (Cheng and Potlogea, 2015; Fachini et al., 2019; Li, 2018; Zi, 2021). Chapter 4 aims to contribute to the literature which examines the relationship between trade openness and wage inequality and addresses three main questions. First, whether China's trade openness affects the industry wage premium. Second, whether the effects of trade openness vary across sectors. Third, whether tradable and non-tradable sectors respond differently to trade openness and how one influences the other.

To this end, Chapter 4 draws upon trade-induced industry productivity difference and trade-related Balassa-Samuelson effects. The influential work of Melitz (2003) has been accepted as a central model of inter-industry reallocation due to trade liberalisation. Melitz (2003) theoretically demonstrates that exposure to trade will improve aggregate industry productivity through market effects (i.e., the exit of low productive firms from an industry) or resource reallocation towards more productive firms. Industrial productivity increases more strongly in liberalised sectors than less liberalised sectors, and trade openness could potentially affect industry wage premia through productivity growth. One important channel is that the trade-induced greater productivity in tradable sectors translate into a rise in the wages of non-tradable sectors, which is normally considered as trade-induced effect.

Using the data of China General Social Survey (CGSS), in chapter 4 we first estimate industry-level wage premia conditional on worker, firm, or job characteristics. When estimating the relationship between trade openness and industrial wage premia, we use two measures of trade openness. The first is the trade shares in gross output. The second is the tariffs faced by China exporters. The results suggest that trade openness has positive effect on industrial wage premia. In the case of tariffs, as a lower tariff is associated with higher trade openness, a negative sign of the coefficients indicates that the wages increase with higher trade openness (lower tariffs). Subsequently, in chapter 4 we examine whether the Balassa-Samuelson theory can be reconciled with numerical evidence in terms of Chinese cases. Imperfect mobility of labour across regions is considered when the transmission from tradable sectors to non-tradable sectors is examined. Taking into consideration imperfect mobility of labour across regions, the results suggest the impact of tradable sectors on non-tradable sectors is most pronounced in high restrictive regions. Finally, the estimates provide evidence on consumer preference, which can be another important channel to explain the impact of tradables on non-tradables.

Chapter 4 departs from existing literature by disaggregating industries into tradable and non-tradable sectors. Such disaggregation is necessary because by doing so we gain more insights into direct and indirect influences of trade openness on industry wage premium. As China’s development process has entered a new stage in which service sectors play an important role, it is important to study the causes and consequences of the rise of traditional non-tradable industries. In addition, chapter 4 is among the first, to the best of our knowledge, to empirically test the role of labour mobility in determining the effect of trade on wages.

Finally, chapter 5 concludes the thesis by summarising the main results, the contribution and policy implications of each chapter. The limitations and possible research avenues are also discussed.

Chapter 2 How does *Hukou* System Affect Internal Migration in China?

2.1 Introduction

In China, rising spatial wage inequality is a big concern. At the urban-city level, there was a dramatic rise in inequality in urban wages between 1988 to 2008. Appleton et al. (2014) document that the ratio of the 90^{th} and 10^{th} percentile of wage rises from 2.82 in 1988 to 6.43 in 2008. By comparison, Martins and Pereira (2004) present that the ratio of wages at the 90^{th} and 10^{th} percentile are generally between 2 and 3 for most of the developed countries in 1990s, for example, Austria (2.28), Denmark (2.39), France (2.73). The exceptions are UK (3.33), US (3.45), Portugal (4.58) and Ireland (4.74). In the early 2000s, wage inequality ranged from 2.2 in Sweden to 5.8 in South Korea across OECD countries (Broecke, 2016). According to the OECD Income Distribution Database, income inequality has risen in most advanced countries as well as developing countries over the past three decades. Income inequality is also typically higher in developing countries than in advanced countries. The largest increases among the developing countries have happened in China, India and South Africa.

Increasing inequality has been a hotly debated topic since China's economic reform in 1979. There are especially large disparities between coastal regions and interior regions. In general, the mean earnings in coastal provinces were 2,215 yuan compared with 1,652 yuan in the interior provinces (Benjamin et al.,2005). In the 2000s, urban household earnings were approximately 2.5 times rural household earnings. That said, large rural-urban earnings gap exists in addition to interregional disparities (Sicular et al., 2007). According to Combes and Zhu (2019), there were 49 million inter-provincial workers, which only accounted for 7.2 percent of total employment, in 2005.

A critical question is why workers do not take advantage of wage differentials to move across cities. One explanation for the persistence of the wage gap is that internal mobility restrictions, such as the *hukou* registration system, prevent workers from taking advantage of regional wage differentials.[2] In this chapter, we focus on how migration decisions respond to migration costs and economic factors. Specifically, this

[2] *Hukou* system (household registration system) is used as a tool to control population mobility. Under this system, each citizen must register with a local government as either an agricultural or non-agricultural hukou. The following section provides more details of this registration system.

chapter aims to investigate whether the *hukou* system is a barrier to internal migration and whether it enhances the responsiveness of migration decisions to economic factors (such as wages and employment probabilities). The *hukou* (household registration) system is profoundly connected to social welfare provision and depends on the policies of the local authority. Most workers will be discouraged from migrating by the associated loss of social provision. Although *hukou* restrictions are gradually being relaxed, this system is still a determinant of internal migration.

There is a vast theoretical and empirical literature on the research effects of mobility restrictions in China. Much of this literature examines the economic consequences of a particular policy barrier to internal labour mobility, the *hukou* system of household registration (Bosker et al., 2012; Kinnan et al., 2018; Ngai et al., 2019; Tombe and Zhu, 2019; Whalley and Zhang, 2007). Specifically, past papers either concentrate on the impact of the *hukou* system on income inequality, or alternatively on employment allocation efficiency and productivity. Less attention has been devoted to exploring the determinants of internal migration with respect to economic and demographic factors, especially with respect to the effect of wage differentials and migration obstacles. The main goal of this chapter is to provide a quantitative analysis of the trade-off between income and migration policy restrictions.

Using data from the 2005 China Population Census Survey, we provide evidence regarding migration restrictions and their impact in shaping migration decisions. We use a newly developed *Hukou* Registration Index (Zhang et al., 2019) of Chinese cities to measure the stringency of *hukou* regulations across Chinese regions. Although the role of *hukou* registration constraints on internal mobility is clearly important, few studies specifically quantify the impact of the stringency of *hukou* controls. One paper in this direction is Bao et al. (2011), who use proxy variables, such as the ratio of registered population to total population at year's end, to measure the stringency of *hukou* registration at the provincial level. Unlike Bao et al. (2011), we use the *Hukou* Registration Index[3] for 120 Chinese cities to measure the stringency of *hukou* controls

[3] Zhang et al. (2019) construct city-level *hukou* index for the period of 2000-2013 and 2014-2016 for 120 cities using the Projection Pursuit Model (PPM).

rather than rely on the realised migration volumes as a measure of policy restriction differences.

To facilitate the empirical analysis, we estimate a regional-level panel data model and a multi-regional model that rest on the logic that migrants move to increase their wages.[4] However, the literature on immigration emphases the inflow of migrants may positively or negatively affect natives' wages in the city of destination, endogeneity can represent a major concern. Specifically, when the skills of migrants complement those of natives, the complementarities between natives and migrants bring positive effect on natives' wages. While migrants have negative effects on natives' wages when they substitute to natives. To address this issue, following Tombe and Zhu (2019), we construct a Bartik-style instrument by calculating the expected earnings based on each city's distribution of employment shares across sectors, which is uncorrelated with unobserved labour supply shocks. We find that the regional wage differential has a positive effect on the migration of origin-destination pairs. The IV estimation suggests that when the regional wage differential increases by 1%, the propensity of migrating increases by 0.23%. In particular, the wage differential has a larger effect on migration if a destination has higher *hukou* stringency index, i.e., migrants tend to favour jobs with higher wages if these have stricter *hukou* policies, likely to mitigate the higher costs of migrating due to the restrictions. These results are robust to the use of the alternative estimators: PPML, Eaton-Kortum Tobit model and Heckman selection model.

In terms of the other determinants of migration, the results show that the effects of unemployment and geographical distance are larger when the *hukou* system of the destination city is more restrictive. The effect of public services on migrants gradually decreases with more stringent *hukou* regulations. The rationale for this result might be that public services are only available to the residents who are entitled to local *hukou* while the primary objective here is to investigate the migration decisions of those with non-local *hukou* who are not entitled to receive local public services. Thus, the accessibility of public service is less important for migrants. Finally, we explore the

[4] This chapter focuses on gross migration rather than net migration since the interest of this chapter is to investigate the individual migration decision determinants.

possible different impact of *hukou* restrictions on high-skill and low-skill migrants. The results indicate the effect is larger for high-skilled migrants.

2.2 Literature Review

This chapter contributes to the literature that analyses the interplay between migration policy and economic incentives and their migration effects.

First, our work is related to the literature that examines the response of migrants to economic incentives. Many studies directly or indirectly answer the question of which factors determine international and internal migration. Sprenger (2013) concludes that the key determinants of international migration are factors such as the income and unemployment differentials, age structure, education attainment and social network effects. Although many of the determinants discussed in the literature focus on international migration, some of the factors may apply to internal migration as well.

Concerning the Chinese case, Wu et al. (2018) examine the spatiotemporal patterns and the determinants of inter-provincial migration in the period 1995-2015. Utilising inter-provincial migration data, Wu et al. (2018) find that inter-provincial migration is mainly driven by the influence of regional disparities and industrial upgrading. Poncet (2006) analyses the workers' motion law in China by the logic that workers move to obtain higher incomes. Moreover, a number of empirical papers on the determinants of migration have studied the effects of wage and unemployment differentials on migration flows using a gravity equation between origins and destinations (e.g., Grogger and Hanson, 2011; Ortega and Peri, 2009; Mayda, 2010; Ortega and Peri, 2013). This literature finds that differences in the level of wages and unemployment between destination and origin city have, respectively, positive and negative effects on bilateral migration flows (and stocks), which our results are in line with.

Second, some studies have looked at the effect of migration regulations on migration flows. Bertocchi and Strozzi (2008) evaluate the impact of migration institutions on attracting international migrants. Similarly, Chassamboulli and Peri (2020) analyse the role of immigration policies together with incentives on determining the inflows of immigrants. For the Chinese case, Sun et al. (2011) test the effect of *hukou* reforms, modelled by dummy variables, on rural-urban migration flows. One paper that quantifies the effects of *hukou* policies is Bao et al. (2011). They use proxy variables –

such as the ratio of registered population to total population at year’s end, to measure the stringency of *hukou* registration at the provincial level. In this chapter, we quantitatively examine the effects of economic and demographic factors on internal migration, especially the influence of *hukou* restrictions on city-level migration. Unlike Bao et al. (2011), we use the *hukou* index for 120 Chinese cities constructed by Zhang and Lu (2018) to measure the stringency of *hukou* control. By focusing on city-level migration patterns, we can account for the substantial variations in economic and demographic conditions compared with the variations at provincial level investigation. This chapter assesses the mobility barrier effect on internal migration taking into consideration interactions with economic factors. The work presented is motivated by the literature on the determinants of migration where the role of migration policy has remained mainly discussed in a theoretical setting without specific empirical assessment of the role of stringency in local *hukou* policies.

Thirdly, this chapter studies a mechanism to explain spatial inequality and underdevelopment in many developing economies. Academics have investigated the persistence of spatial wage inequality and asked why migration cannot mitigate the inequality. If it is costly for workers to relocate in places where it is more productive, labour may be unable to move out of low-income places. This will negatively affect the efficiency of aggregate productivity and long-run economic growth. Morten and Oliveira (2016) construct and estimate a spatial equilibrium model to identify the bilateral costs of moving between two regions. They try to understand migration decisions focusing on the cost of migrating and show the implications of migration costs on labour mobility for regional economic development and the allocation of resources. This chapter tries to shed light on the effect that policies easing labour frictions may have on the efficient allocation of labour across space.

The remainder of this chapter is organised as follows: section 2.3 introduces the institutional background; section 2.4 presents the theoretical model which motivates the empirical specification and section 2.5 describes the data, followed the estimation results reported in section 2.6. Section 2.7 analyses the heterogeneous impacts in terms of different sizes of cities and different skills for migrants. Section 2.8 concludes.

2.3 Institutional Background and Descriptive Evidence

2.3.1 Hukou System

The *hukou* system came into effect in 1958 with the implementation of the People's Republic of China *Hukou* Registration Regulation to control population mobility. Each citizen was classified in an agricultural or non-agricultural *hukou* (sometimes referred to as rural or urban *hukou*). A *hukou* is primarily inherited from one's parents at the time of birth; children born in urban areas to parents with rural *hukou* are seemingly designated as rural *hukou* holders. Social services, such as the eligibility to purchase a local house, access local public school, make claims on a local pension plan, are *hukou*-based. That is, citizens who register their *hukou* at different places will receive different social benefits which vary according to different administrative units. On top of that, temporary migrant workers without local *hukou* have restricted access to local social provision and face higher costs for medical care and children's education.

From 1950 to 1980, those residents who held non-agricultural (urban) *hukou* status received benefits which were not available to rural counterparts, such as providing food/ housing/ free medicines. People need to get approval from the local government if they want to change their *hukou* status. Before 1980s, it was difficult to get approval and change one's *hukou* from rural to urban areas, or from small cities to large cities. Since 1980, the government has taken steps to relax the *hukou* system, but, still, people who work outside their *hukou* location had to apply for a temporary residence permit which is also difficult to obtain. Since 2003 some provinces have abolished the requirements of temporary residence permit for migrant workers. Even without residence permit requirements, however, the costs of being a migrant worker are still high in terms of opportunity costs, sunk costs and lack of access to public services.

Migrants need to satisfy several criteria in order to obtain a local *hukou*, such as stable jobs, paying social insurance for several consecutive years in local residence places, investing in local cities, etc. The local governments in China have been granted greater autonomy in deciding *hukou* policies and the eligibility criteria for issuing local *hukou* to migrants. For instance, local governments can grant local *hukou* to people whose parents or children already possess local *hukou*; or permit local *hukou* to those who meet the criterion of *hukou* qualification through investment and tax payment channel;

or to migrants who have been recruited in state-owned enterprises for several years. In big cities, the local government aims for strict control and a decrease in the non-*hukou* migrants by limiting the *hukou* registration. According to a report based on a monitoring survey conducted by the National Bureau of Statistics in 2017, only 38% of rural migrants feels localised in their current residential city without holding local *hukou*. The report also shows that the larger the city, the weaker the sense of social belonging for rural migrants. Residential segregation, poor housing conditions and unequal schooling are the main problems. More than half of the migrants report facing some difficulties in enrolling their children in local public schools (NBS, 2017). Also, rural migrants and poor urban migrants tend to live in “urban villages” (or “villages within cities”), construction sites or factory dormitories, which are cheap, highly-densely populated and lacking basic infrastructure.

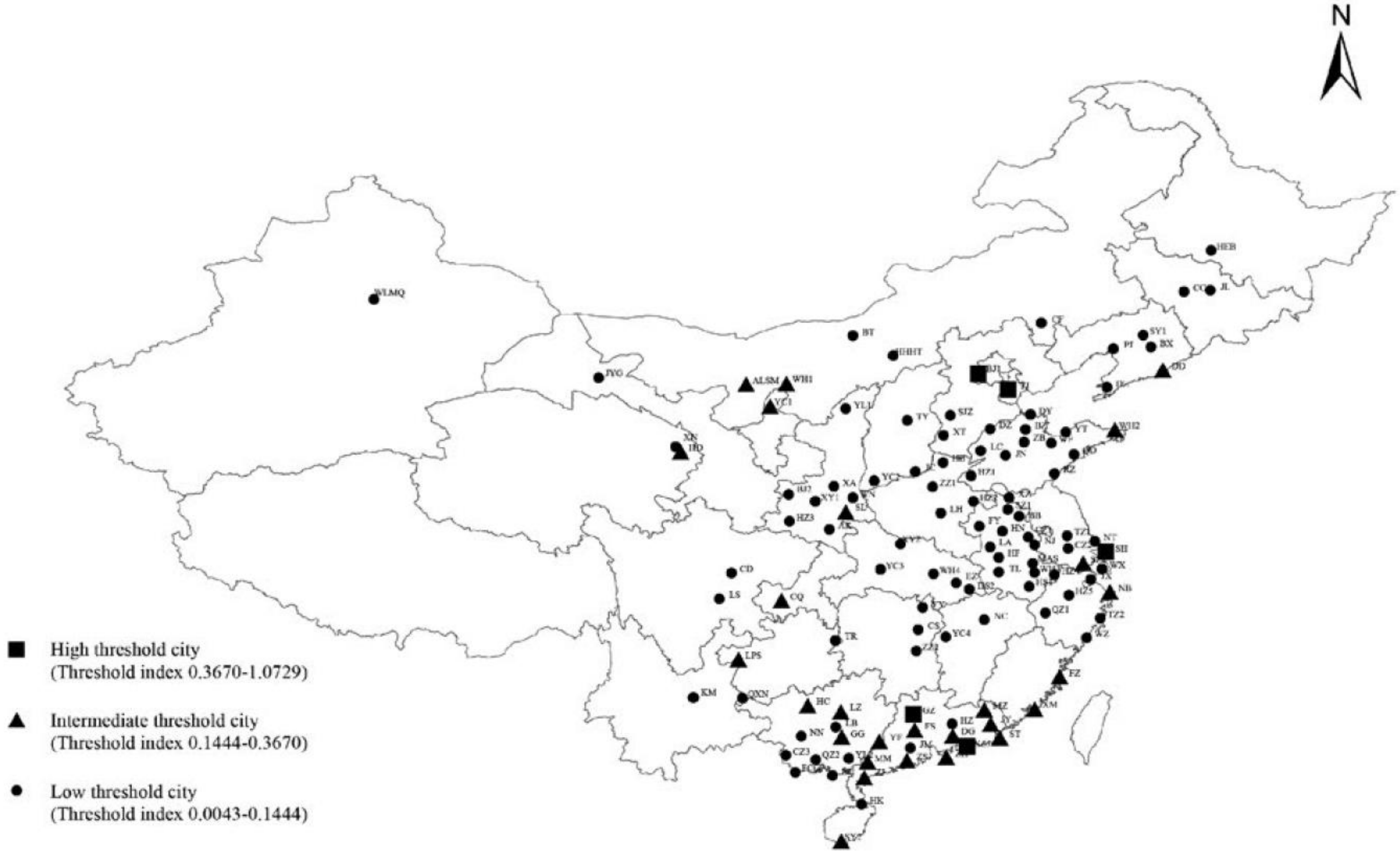
Since the 1980s, the government has taken steps to reform the *hukou* system by implementing some reform policies. Especially in 2014, the China government published the first national urbanisation blueprint entitled `National New-type Urbanisation Plan, 2014-2020'. The plan specified a target of granting 100 million new urban *hukou* in 2015-2020 to enable migrants to settle in their current residential locations. The granting priority would be given to skilled migrants, the college-educated, and the long-term migrants. However, big cities still restrict migrants. Some big cities took this opportunity to clear out the migrants, especially the low-income rural migrants. Additionally, big cities further restrict migrant children's access to education. These strict measures make the life of migrants more complicated.

The *hukou* system is still very restrictive for migrants relocating to the cities, which segments the Chinese population into local and non-local, rural and urban. It also brings social discrimination and segregation. Overall, as one of the most important institutional regulations, the *hukou* system prevents the free flow of labour and undermines workers’ mobility as well as discourages population from benefiting higher earning potentials.

2.3.2 Hukou Index

The China *Hukou* Registration Index (Zhang et al.,2019) measures the stringency of *hukou* restrictions in Chinese cities from 2000 to 2016. The higher the index is, the more stringent the *hukou* restriction. Generally speaking, there are four main channels for migrants to obtain local *hukou*: investment, house purchase, talent programme and employment. Zhang et al. classify the policies of migrants to apply for resident status into different categories and compute the *hukou* index for all policy categories in 120 cities using Projection Pursuit Model (PPM) in two periods of 2000-2013 and 2014-2016. Since 1978, the evolution of *hukou* system has undergone. The major changes of *hukou* system began in early 2000s. As stated above in section 1.2, a *hukou* reform has announced at national level which is considered as a new stage of the relaxation of migration restrictions. Because of the significant changes after 2014, Zhang et al. (2019) separate the *hukou* reforms into two stages of 2000-2013 and 2014-2016. They computed the *hukou* stringency index of each city in two periods of 2000-2013 and 2014-2016. That is, the data of *hukou* registration index provides the composite indexes for 120 cities before and after 2014. The index is highest for Beijing, followed by Shanghai, Guangzhou and other cities. Figure 2.1 presents the spatial distribution of the city-level *hukou* registration index in 2000-2013. Cities with relatively high *hukou* registration index are mainly located in coastal regions in eastern China and the cities in middle China have relatively low requirement for *hukou* registration. This figure presents the spatial variations of *hukou* registration index and the analysis in this chapter relies on these spatial variations.

Figure 2.1 Spatial Distribution of the City-level *Hukou* Registration Index in 2000-2013



Source: The China *Hukou* Registration Index (Zhang et al.,2019)

2.3.3 Migration and Income: Some Stylised Facts

In this section, we first compare relative incomes across Chinese provinces. The data is collected from the China Statistical Yearbook, reporting average wages in each province. To calculate relative incomes, we first compute real incomes by deflating the average wages using the consumer price index in 2000 price levels. Relative incomes are then calculated by deflating real incomes by average real income across provinces. In figure 2.2, we can see that there is large dispersion of income across regions in China.

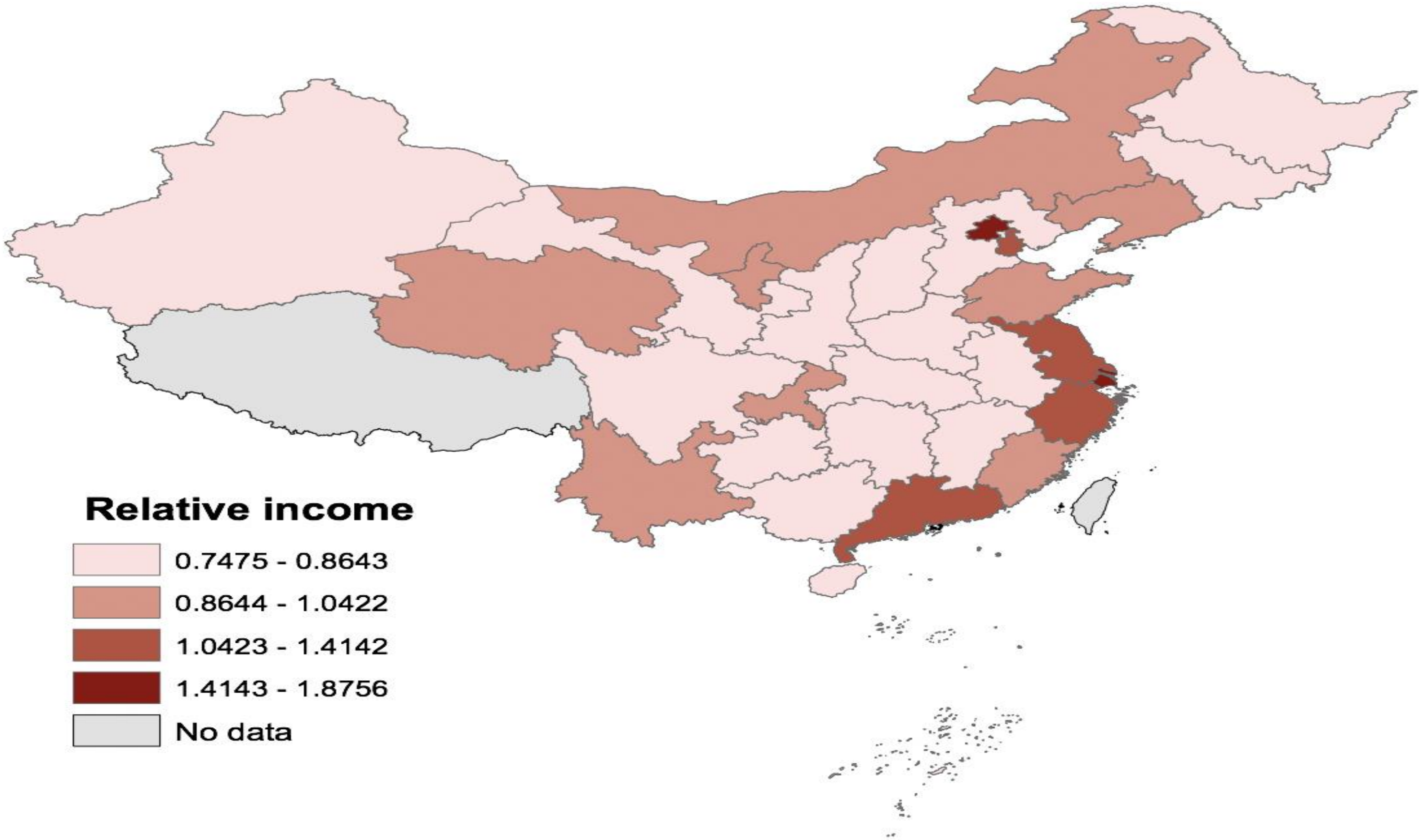
The data of migrant shares is constructed using the 2005 population census data[5]. Any worker who is working in a province other than the province where their *hukou* is registered is defined as a migrant. This is the definition of migrants in this thesis. In reality, there are two types of migrants: non-local *hukou* registered migrants and local *hukou* registered migrants. Local *hukou* registered migrants refer to those who are residing in a place other than the place where their original *hukou* is registered but now already obtained the local *hukou* at where they are currently residing. This thesis is mainly interested in non-local *hukou* registered migrants thus we use migrants to refer to non-local *hukou* registered migrants.

Figure 2.2 panel B shows the dispersion of the calculated migrant share across provinces in China. Figures 2.2 panel A and panel B show a similar pattern and there seems to be positive correlation between income and migration share of employment. There is spatial divide of relative income and migration shares. The relative income is higher in coastal regions compared to middle regions. In Midwest regions also have relatively higher income. There are similar spatial patterns for migration shares.

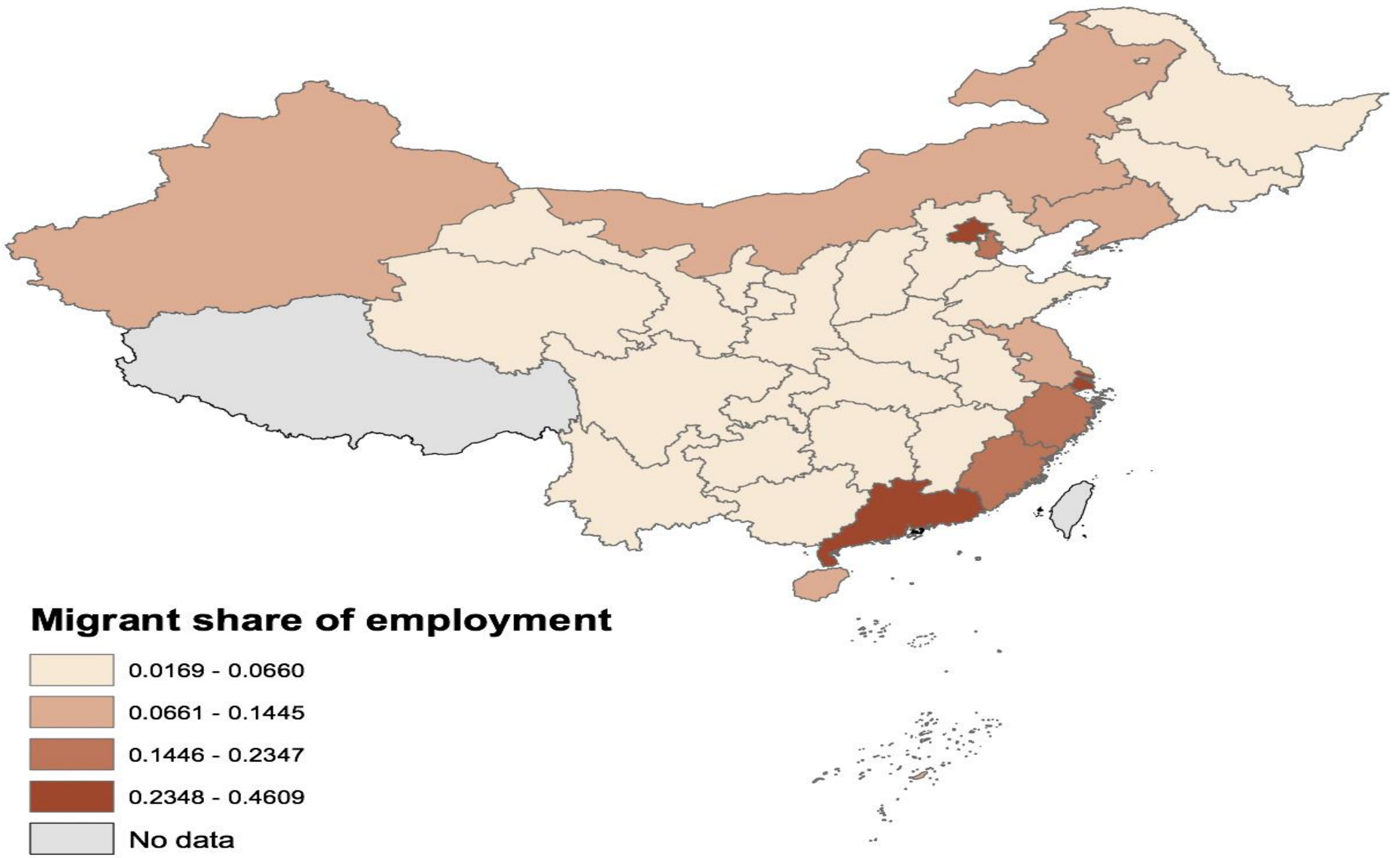
[5] The data used in chapters 2 and 3 are derived from 2000 and 2005 micro census data in China and obtained from a publicly accessible knowledge exchange forum (https://bbs.pinggu.org/), where people disseminate research data, insights and engage in question-and-answer discussions, etc. The versions of the data are old releases and have been used in published papers (e.g., Combe et al., 2015; Facchini et al., 2019; Imbert et al., 2022; Want et al., 2021). To ensure its reliability, I crosschecked it with the migration data provided by a published article "Migration Externalities in Chinese Cities" constructed from the 2005 China Population Census. The correlation between the constructed migration data in this paper and in Chapters 2 and 3 is high at 0.88.

Figure 2.2 Spatial Distribution of Income and Migration in 2005

Panel A Spatial Distribution of Relative Income



Panel B Spatial Distribution of Migrant Share of Employment



Notes: The relative income is calculated by deflating real income with average real income across the whole country. The migration share of employment is calculated by migrant workers with total employment in each province. Dark reds indicate both high relative income and large migration shares of employments. The grey shaded areas are Tibet and Taiwan, which are excluded from the analysis.

2.4 Theory and Estimation Specification

2.4.1 A Model of Internal Migration

This framework follows the random utility model where each migrant maximises their utility by choosing to a location among potential destinations (Grogger and Hanson, 2011). Migration decisions are usually based on expected real income, migration costs and unemployment rate. According to Borjas (1987, 1989), individual migration choice is guided by the comparison of income differentials across regions. When moving to another region, migrants face a moving cost. Typically, individuals move away from lower-income regions to higher-income regions with net of migration costs. Unemployment also affects migration. Regional unemployment differentials measure different employment probabilities across regions. People prefer to move to regions where they can have better job opportunities. Pissarides and MacMaster (1990) argue that unemployed workers are more likely to move out than employed ones as they have less to give up.

Assume the utility function takes a logarithm form and is given by:

$$U_{ijl} = (y_{jl} - C_{ijl})^{\lambda} \exp\left(\varepsilon_{ijl}\right) \tag{2.1}$$

where l denotes the individual who decides to migrate, y_j denotes the expected benefits in city j, C_{ij} is the cost of migrating from origin i to destination j; λ is a positive constant and ε_{ijl} follows $i.i.d.$ extreme value distribution.[6] The choice probability of staying in origin i or moving to j is:

$$\Pr\left(\frac{j_l}{i_l}\right) = \Pr\left(U_{ijl} = max\left(U_{i1l}, U_{i2l}, U_{i3l,} \dots, U_{ijl}\right)\right) \tag{2.2}$$

It is assumed that the probability that one individual chooses a destination approximately coincides with the aggregate proportion of individuals moving to destination j from origin i. Following McFadden (1974), the log odds of migrants in destination j from origin i divided by the population staying in origin i is:

[6] Extreme value distribution are the limiting distributions for minimum or maximum of the sample observations from identical and independent distributions.

$$\ln\left(\frac{N_{ij}}{N_i}\right) = \ln m_{ij} = \lambda\left(\ln y_j - \ln y_i\right) - \lambda C_{ij}$$
$$= \lambda\left(\ln w_j + \ln e_j - \ln w_i - \ln e_i\right) - \lambda C_{ij} \qquad (2.3)$$

where N_{ij} denotes the number of migrants who move from origin i to destination j, N_i denotes the population staying in origin i, w_j is the expected wage in city j, e_j is the probability of finding a job in city j, λ is the elasticity of migration probability in terms of expected benefits differentials between origin i and destination j. Also, the odds ratio here embodies the assumption that the probability of choosing between i and j only relates to these two destinations, which indicates that independence of irrelevant alternatives, IIA[7] hypothesis, is applied to the cities in the sample.

2.4.2 Estimation Specification

In line with Bertocchi and Strozzi (2008), we first investigate the relationship between *hukou* stringency and migrant volumes using regional-level aggregate data and time-varying *hukou* stringency indexes from 2005 and 2015. Province and prefecture[8] level data are both collected and empirically used for the analysis.

The estimation specifications are as follows:

$$m_{it} = \beta_0 + \beta_1 \ln\,(real\ wage)_{it-1} + \beta_2 \ln\,(\text{unemployment\_}rate)_{it-1} + \beta_3\,(hukou\_index)_{it}$$
$$+ \beta_4 \boldsymbol{X_{it-1}} + \theta_t + \alpha_i + u_{it} \qquad (2.4)$$

and

$$m_{it} = \beta_0 + \beta_1 \ln(real\ wage)_{it-1} + \beta_2 \ln(\text{unemployment\_}rate)_{it-1}$$
$$+ \beta_3\,(hukou\_index)_{it} + \delta_1 \ln(real\ wage)_{it-1} \times (hukou\_index)_{it}$$
$$+ \delta_2 \ln\,(\text{unemployment\_}rate)_{it-1} \times (hukou\_index)_{it} + \beta_4 \boldsymbol{X_{it-1}} + \delta_3 \boldsymbol{X_{it-1}}$$
$$* (hukou\_index)_{it} + \theta_t + \alpha_i + u_{it} \qquad (2.5)$$

where m_{it} is the share of migrants out of total residents' population in province/prefecture i in the periods 2005, 2010, and 2015. $\ln(real\,wage)_{it-1}$ is the logged average wage in each province/prefecture, adjusted for inflation in terms of

[7] The IIA (Independence of Irrelevant Alternatives) assumption means that the probability that one individual chooses a destination does not depend on what other alternatives available.
[8] City and prefecture are used interchangeably.

2000 price levels and the published consumer price index over years. $\ln(\text{unemployment\_}rate)_{it-1}$ is the logged ratio of urban registered unemployed persons to local labour force. $(hukou\_index)_{it}$ is *hukou* stringency index which is collected from Zhang et al. (2019).[9] $\boldsymbol{X_{it-1}}$ is a vector of control variables such as the logged ratio of tertiary sector output to secondary sector output, public service quality index and amenity service index. The interaction terms between *hukou* index and other control variables are also included. θ_t is the time fixed effect to control common effect across the period of 2005-2015, α_i is the regional-level fixed effect and u_{it} is the error term.

A gravity-type model is also estimated in this chapter. Based on equation (2.3), the gravity-type estimation specification follows the form:

$$\ln m_{ij} = \lambda_1 + \lambda_2(\ln w_j - \ln w_i) + \lambda_3(\ln u_j - \ln u_i) + \lambda_4 \ln distance_{ij} + \lambda_5 hukou_j + \gamma \boldsymbol{X_{ij}} + \varepsilon_{ij} \quad (2.6)$$

and

$$\begin{aligned}\ln m_{ij} = \lambda_1 + \lambda_2(\ln w_j - \ln w_i) + \lambda_3(\ln u_j - \ln u_i) \\ + \lambda_4 \ln distance_{ij} + \lambda_5 hukou_j + \chi_1(\ln w_j - \ln w_i) * hukou_j \\ + \chi_2(\ln u_j - \ln u_i) * hukou_j + \chi_3 \ln distance_{ij} * hukou_j + \gamma \boldsymbol{X_{ij}} + \chi_4 \boldsymbol{X_{ij}} \\ * hukou_j + \varepsilon_{ij} \qquad (2.7)\end{aligned}$$

where m_{ij} denotes migrant stocks from origin i to destination j divided by the population remaining in origin i in 2005, $i = 1,2,3, \ldots 120, j = 1,2,3, \ldots 119$. w_j and w_i are average wages of total employed persons in city j and i. u_j and u_j denote the urban registered unemployment rate. Also, in order to capture the bilateral cost of migration, we proxy the migration costs using the great-circle distance[10] and *hukou* stringency. $\boldsymbol{X_{ij}}$ are vectors that control for differences between cities, such as the log of differential of industry output ratio between destination and origin, the log of differential of *hukou* registered population and the log of differential of local amenity

[9] Zhang et al. (2019) construct city-level hukou index for the period of 2000-2013 and 2014-2016 for 120 cities. We use the average of city-level index of each province as province-level hukou index.
[10] Distance is calculated with the great circle formula using each city's latitude and longitude data.

levels. The interaction terms between control variables and *hukou* stringency index are included in equation (2.7). ε_{ij} is the error term.

2.4.3 Endogeneity

It has been known that better income opportunities in destinations increase migration rates. However, the positive impact may reflect causation in the opposite direction. The migrants may either positively affect wages in the destinations, when the skills of migrants complement these of natives; or negatively affect wages in the city of destination, when the skills of migrants substitute for natives. Additionally, the regional wages and migrant stocks are likely to be correlated with local economic conditions, institutional quality and first-nature geography (Combes et al., 2015). Endogeneity problems arise in this case when estimating gravity-type specifications. To address the possibility of reverse causality and omitted variable bias, following Tombe and Zhu (2019), we construct a Bartik-style expected income instrument by calculating the expected earnings based on each city's distribution of employment shares across sectors. That is, we measure each city's expected earnings by interacting national average earnings weighted by sector composition with the distribution of employment shares across sectors.

The 2005 Population Census provides detailed data on individual earnings by detailed sector information. The respondents report their previous month's earnings. According to the "Industrial Classification for national economic activities (GB/T 4754-2002)", there are 20 different sectors in China. For each sector, there are a bundle of sub-classifications. In the 2005 census survey, the respondents report their working sector in terms of sub-classifications. We group their working sector into first-level classifications according to official classification. After this step, we can compute the national average earnings by sector and the distributions of employment share across sectors in each city.[11] The classification of 20 sectors is in appendix table A2.6.

Let the Bartik-style predicted wage be w_j :

$$w_j = \sum_k (ave\_earnings_{k,-j} \frac{employ_{k,j}}{employ_j}) \quad (2.8)$$

[11] I use the official weight "power_2" in the 2005 mini census when computing sectoral employment.

where $\frac{employ_{k,j}}{employ_j}$ is the share of sector k employment in total employment in city j. $ave\_earnings_{k,-j}$ is the national average earning in sector k, excluding city j.[12]

The average wage differential between city j and city i is:

$$w_j - w_i = \sum_k (ave\_earnings_{k,-j} \frac{employ_{k,j}}{employ_j}) - \sum_k (ave\_earnings_{k,-i} \frac{employ_{k,i}}{employ_i}) \qquad (2.9)$$

The central argument of the Bartik instrument is that the sector shares measure exogenous differences in exposure to nationwide average earnings. This instrument would capture the city-level average earnings driven by the sector composition rather than within-city labour market conditions. To be specific, the Bartik instrument is plausibly exogenous to a city's migration trend or local market tightness. In addition, as this instrument predicts the average wage in each city using the national earnings of each industry, the region needs to match this wage trend by adjusting the local wage thereby to be able to attract migrants. In this case, the unique channel for the causality of predicted wage on migration stocks is the actual regional wage, which satisfies the exclusion restriction requirement.

2.5 Data and Descriptive Statistics

2.5.1 Regional-level Migration Data

We collect regional-level (e.g. province and prefecture-level) data using the 2005, 2010, 2015 China Population Census data and various China Statistical Yearbooks. In this thesis, migrants are defined as individuals who have been living in a city-level region different from *hukou* registered city-level region for more than 6 months. To compute the migrant shares, we first collect data of current residents' population whose *hukou* registration is different from the current region and have been living there for more than 6 months, and the ratio between migrant stock and total residents in current region is the migrant share. 30 provinces and approximately 300 cities are included. As the China *Hukou* Registration Index is only covered for 120 cities, to have consistent information, we combine the migration data and *hukou* index for 120 cities.

[12] I exclude the own city's contribution to national average earnings in order to ensure that national wages are orthogonal to the location demand shock.

2.5.2 Bilateral Migration Data

We use 20% random extraction of the 1% 2005 China population mini-census survey to construct bilateral migration data, which provides individual-level information about workers' employment status, age, current location, *hukou* registration location. We limit our attention to workers 16 years old and above who report their last week occupation, i.e., the employment definition in this chapter's analysis. Following Tombe and Zhu (2019), we define the worker's current location as their destination city and the registered *hukou* location as their origin city. By knowing workers' current location and *hukou* registration location, we then compute the bilateral migration stocks at the city level. To measure m_{ij}, we calculate the ratio of workers with *hukou* registration in city i currently working in destination city j to the remaining population with *hukou* registration in city i. We use the official weight "power_2" [13] for the 2005 mini census when construction migration data. 257 cities are included in the 2005 mini-census. As the China *Hukou* Registration Index is only covered for 120 cities, we combine the migration data and *hukou* index for 120 cities. The complete data has 14280 observations (120*119=14280). However, only 3218 observations have nonzero migrant workers.

2.5.3 Other Controls

We collect province- and prefecture-level average wages, unemployment rate, tertiary sector output and secondary sector output from various China Statistical Yearbooks and China City Statistical Yearbooks. To mitigate the risk of reverse causality, we use one-year lagged variables. The bilateral distance variables are taken from Yihua Yu (2009) calculated using the great circle formula for each city's latitude and longitude. Table 2.1 reports data sources and definitions of the variables as well as summary statistics.

Unemployment share is defined as urban unemployment share which only includes locally urban registered unemployed persons, i.e., excluding migrants without local urban *hukou* (Feng et al., 2017). To capture the labour demand and employment probability effect on migration as much as possible, the ratio between the tertiary industry to the secondary industry of GDP production is included in the specification. According to Card and Lewis (2007), labour demand shocks, as well as the changes

[13] The weight is officially provided by the 2005 mini census survey.

in industry structure in potential destinations, act as a pull factor for migrant workers. Based on the specification proposed in this chapter, the difference between output ratios attributed to GDP between destination and origin is considered as a pull factor of migration. In addition, Tiebout (1956) presents a model to show that local public services which reflect consumer-voters' preferences are a determinant of the places where they choose to reside. Dahlberg et al. (2012) examine the impact of local public services on community choice using Swedish microdata. They find that households would "vote with their feet" when faced with more attractive local public services and income tax rates elsewhere. In their paper, they measure the level of public services by government spending on education, childcare, elderly care and social welfare services. That is, however, an indirect measure of local public services. A more direct measure of local public service should be quality-of-life variables such as students to teacher's ratio, doctors per capita, hospitals per capita. Also, other city characters such as natural environment, transportation infrastructure may play a role in migration decisions (Diamond, 2016).

To reduce possible multicollinearity, and keep the model parsimonious, we construct a single index of public services, as well as a single city amenity index that measures the bundle of services related to education quality, medical services and amenity conditions. To construct the public service index, we collect data on the ratio of primary students to primary teachers, the ratio of secondary school students to secondary school teachers, hospitals per capita, doctors per capita, and hospital beds per capita. These data can be bucketed into two categories: the quality of education services and the quality of medical services. Regarding the amenity index, the data on eight different amenities are collected, for example: cinema per capita, library books per capita, public buses per capita, road spaces per capita, green areas per capita, SO2 emission per capita, smoke dust emission per capita and wastewater per capita. These data are bucketed into 3 sub-categories: retail index, transportation index and environment index. To combine the different data sources into a single index, we use principal component analysis (PCA). Following the method of Diamond (2016), we first extract a single measure using the first principal component with each category. We then create an overall index using the first principal component of all the category indexes. Eigenvalues and eigenvectors of the PCA are reported in tables 2.1 and 2.2. Tables 2.1 and 2.2 report the eigenvalues and eigenvectors of the PCA at provincial level for public service indexes and amenity service indexes respectively. Tables A2.1 and A2.2 report the eigenvalues and eigenvectors of the PCA at prefecture level for public service indexes and amenity service indexes respectively. As can be seen in table 2.1, the first component of public service indexes at provincial level explains 90% of the overall variances with an eigenvalue equal to 1.81, suggesting a single measure of the public service indexes can capture these different types of public service establishments accurately. Similarly, the first component of public service indexes at prefecture level explains 72% of the overall variances with an eigenvalue equal to 1.45. The eigenvectors show the weights that the variables enter the components. Panel A of table 2.1 shows the weights on each education service. Similarly, panel B of table 2.1 shows the weights on each medical service. All services receive positive loadings for the education indexes. The medical service index places positive loadings on hospitals, doctors and hospital beds. When combining these two indexes into one overall index, the education index receives a negative loading which indicates that the high student-teacher ratio is a measure of low education quality. The first principal component of these two measures will be used as the public service index. The eigenvalues and eigenvectors of each amenity are reported in table 2.2. The first component explains 58% of the overall variance with an eigenvalue equal to 1.74 at provincial level. It suggests the first component can explain the data variability reasonably well. Likewise, the first component explains 60% of the overall variance with an eigenvalue equal to 1.8 at prefecture level. Panel A and B of table 2.2 show all life amenities and transportation amenities receive positive loadings. The only shortcoming is that the environment index of panel C in table 2.2 has positive loadings on pollution amenities. As pollution negatively affects environment, we are expecting a negative sign of loading on pollution measures. In table 2.2, combining these individual amenity service indexes into an overall amenity index leads to positive loadings on life index, transportation index and negative loading on environment index. By comparison, in table 2.2, the overall amenity index places a positive loading on environment index. Despite this shortcoming, a single amenity index accounts for the largest part of the data variability.

In the empirical specification, we use local *hukou* registered population to capture the possibility of economic growth and production externalities. Duranton and Puga (2004) suggest that an increase of population along with sharing, matching and learning mechanisms lead to productivity gains and agglomeration economies, which increase the wage level and the probability of being employed. We expect that a larger differential in *hukou*-registered population between destination and origin increases migrant shares. Except for the dependent variable and distance variable, we use 1 year lagged explanatory variables. That is said, it is the data from previous years. For example, we use data from years 2004, 2009 and 2014.

There are other potential determinants of migration which are discussed in the existing literature, such as housing and rental prices. Existing literature (Peng and Tsai, 2019; Zhou and Hui, 2022) indicates that housing price is one of the determinants of migration decisions. A change in housing prices has great impact on migration decisions. However, the right of purchasing houses is embedded in *hukou* system. That is, a local *hukou* is a prerequisite for purchasing houses in residence places. The focus of this chapter is non-*hukou* registered migrants, therefore, housing prices are not directly relevant to them. Due to the exclusive housing system embedded in *hukou* system, renting is the dominant choice for migrants. Rent affordability is one of the factors that migrants would take into consideration when making migration decisions. Rent-to-income ratio has been adopted to measure affordability. Rent-to-income ratio is calculated as the ratio of housing rental to household disposable income. There is existing literature (Li et al., 2020; Liu et al., 2020; Zhang, 2015) measuring the affordability in China. Li et al. (2020) measure the rent-to-income ratios for renter in 275 cities and they find that the level of housing affordability remains relatively stable across the cities between 2000 and 2018. Although there is mild volatility during the period of time, the housing affordability condition is acceptable in almost all cities. As such, it is unlikely for rental prices to be a significant determinant of migration.

Table 2.1 Principal Component Analysis for Public Service Indexes at Provincial Level

	2004		2009		2014	
	Loading	Unexplained Variance	Loading	Unexplained Variance	Loading	Unexplained Variance
Panel A. School Index						
Eigenvalue (% of variance)	1.82(0.91)		1.78(0.89)		1.74(0.87)	
primary students to teachers	0.71	0.09	0.71	0.11	0.71	0.13
secondary students to teachers	0.71	0.09	0.71	0.11	0.71	0.13
Panel B. Medical Index						
Eigenvalue (% of variance)	2.10(0.70)		1.77(0.59)		1.46(0.49)	
hospitals per capita	0.34	0.75	-0.09	0.98	0.50	0.63
beds per capita	0.67	0.07	0.70	0.13	0.73	0.22
doctors per capita	0.66	0.08	0.71	0.12	0.46	0.69
Panel C. Overall Index						
Eigenvalue (% of variance)	1.81(0.90)		1.74(0.87)		1.35(0.67)	
School Index	-0.71	0.10	-0.71	0.13	-0.71	0.33
Medical Index	0.71	0.10	0.71	0.13	0.71	0.33

Table 2.2 Principal Component Analysis for Amenity Service Indexes at Provincial Level

	2004		2009		2014	
	Loading	Unexplained Variance	Loading	Unexplained Variance	Loading	Unexplained Variance
Panel A. Life Index						
Eigenvalue (% of variance)	1.46(0.73)		1.55(0.78)		1.45(0.73)	
Cinema per capita	0.71	0.27	0.71	0.22	0.71	0.27
Library books per capita	0.71	0.27	0.71	0.22	0.71	0.27
Panel B. Transportation Index						
Eigenvalue (% of variance)	1.43(0.72)		1.21(0.61)		1.10(0.55)	
Public buses per capita	0.71	0.28	0.71	0.39	0.71	0.45
Road space per capita	0.71	0.28	0.71	0.39	0.71	0.45
Panel C. Environment Index						
Eigenvalue (% of variance)	1.85(0.46)		1.85(0.46)		1.98(0.50)	
Green area	-0.27	0.86	0.0034	0.99	0.07	0.98
SO2 emission per capita	0.68	0.14	0.68	0.16	0.66	0.15
Smoke dust emission per capita	0.67	0.18	0.66	0.21	0.68	0.09
Wastewater per capita	0.14	0.97	0.34	0.79	0.33	0.79
Panel D. Overall Index						
Eigenvalue (% of variance)	1.74(0.58)		1.30(0.43)		1.39(0.46)	
Life Index	0.65	0.28	0.71	0.35	0.51	0.64
Transportation Index	0.70	0.15	0.70	0.36	0.71	0.31
Environment Index	-0.31	0.84	-0.06	0.99	0.49	0.67

Table 2.3 Data Sources and Descriptive Statistics for the Main Variables of Gravity Estimation

	Definition	Sources	All Observations			Observations with positive migrants		
			Obs	Mean	Standard Deviation	Obs	Mean	Standard Deviation
migration rate	(bilateral migrant stock at destination city)/population at origin city	2005 China Population Census	14280	0.001	0.003	3128	0.002	0.006
wage	City-level average wage	Average wage of each city is collected from 2004 China City Statistics Yearbooks	13804	14882.04	5067.431	3101	18828.71	6230.103
unemployment	City-level unemployment rate	Unemployment rate of each city is collected from 2004 China City Statistics Yearbooks	13804	0.038	0.023	3101	0.032	0.018
hukou index	*Hukou* stringency index at destination	Zhang et al. (2018). "A quantitative analysis of *Hukou* reform in Chinese cities:2000-2016".	14280	0.652	0.323	3128	0.864	0.5

distance (Km)	Great circle distance between destination and origin	Yihua Yu (2009). "CHINA_SPATDWM: Stata module to provide spatial distance matrices for Chinese provinces and cities".	13806	1202.581	894.232	3094	983.808	940.181
output_ratio	City-level ratio of tertiary industry to secondary industry of GDP production	China Statistical Yearbook for Regional Economy-2004	13923	0.8003	0.321	3114	0.834	0.312
public service index	City-level public service index	All public services data are collected from 2004 China City Statistics Yearbooks	13685	0.0003*10^{-6}	1.198	3064	0.469	1.135
amenity index	City-level amenity index	All amenities' data are collected from 2004 China City Statistics Yearbooks	13804	0.000118*10^{-5}	1.337	3101	0.737	2.073
*hukou*_population (* 10^5)	City-level *hukou* registered population	*Hukou* registered population of each city are collected from 2004 China City Statistics Yearbooks	13804	487.386	403.943	3101	552.726	433.73

2.6 Regional-level Estimation Results

In this section, the estimation results using province- and prefecture-level data are presented. Province and prefecture are different administrative levels in China. Provinces consist of prefectures. The definition of migrants in this thesis is that those who are residing in a prefecture R in province P but not registered their *hukou* in R. It could be moving from another prefecture in province P or from a prefecture in another province. The province-level migration data is the sum of the prefecture-level migration in each province. For the independent variables, the province- and prefecture-level data are collected separately.

2.6.1 Province-level Estimation Results

Table 2.4 reports the corresponding results for province-level estimations. The positive coefficient of a variable means an increase in this variable attracts migrants, whereas a negative coefficient of a variable means it hinders migrating. Column 1-3 present the results without interactions, which is related to equation 2.4. Column 1 reports the estimation results without controlling province and year fixed effects. In column 1, the coefficient of wage is positive and highly significant, confirming the key role played in migration decisions. Column 2 includes province and year fixed effects, and column 3 further includes additional controls. The estimates in columns 1-3 indicate the positive effect of wages in attracting internal migration while the coefficient of the unemployment rate is not with the expected sign. Migrants prefer to move to regions where they can have better opportunities. With adding additional controls, there is positive but insignificant effect of wage on migration. When including province and time fixed effect, the rest of the variables no longer seem to matter. These results suggest that perhaps some unobservables that are either constant over time or across provinces alleviate the impact of wages and employment opportunities. With interaction terms which is related to equation 2.5, in column 4-6, the coefficient of the interaction between wage and *hukou* index is positive and statistically significant, indicating that the influence of wages increases with the larger stringency of *hukou* restrictions. In terms of the magnitudes of the coefficient, the magnitude of wage variable becomes smaller with adding the interaction terms between wages and *hukou* index. This suggests that impact of solo wage factor is weakened in the presence of high stringency *hukou* index.

While columns 1-3 indicate an insignificant effect of unemployment, when the interaction term between unemployment and *hukou* index are added, the unemployment effect is negative and significant. Specifically, column 5 predicts that 1% increase in unemployment rate leads to 8.3% decrease in migration rate. Overall, the incentives of migrate, in terms of wage and employment opportunities, is weakened in the presence of migration constraints.

Table 2.4 Province-level Estimation Results

	(1)	(2)	(3)	(4)	(5)	(6)
ln_wage	0.096***	0.007	0.024	0.065***	0.020	0.001
	(0.009)	(0.048)	(0.053)	(0.015)	(0.054)	(0.057)
*hukou*_index	0.160***	0.031**	0.026*	-0.254	-0.565***	-0.422
	(0.019)	(0.012)	(0.015)	(0.201)	(0.086)	(0.545)
ln_unemployment	0.038	0.031	0.022	-0.053	-0.083*	-0.048
	(0.042)	(0.037)	(0.035)	(0.086)	(0.044)	(0.068)
ln_output_ratio			0.009			-0.029
			(0.029)			(0.048)
public_service_index			-0.003			0.003
			(0.005)			(0.010)
amenity_index			-0.0001			0.006
			(0.004)			(0.009)
ln_wage×*hukou*_index				0.033*	0.037***	0.031
				(0.019)	(0.006)	(0.043)
ln_unemployment×*hukou*_index				0.060	0.146***	0.097
				(0.039)	(0.037)	(0.086)
ln_output_ratio×*hukou*_index						0.026
						(0.052)
public_service_index×*hukou*_index						0.003

						(0.012)
amenity_index×*hukou*_index						-0.002
						(0.012)
Province FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	Yes	Yes	No	Yes	Yes
Observations	90	90	90	90	90	90
Within R-squared	0.759	0.05	0.06	0.770	0.22	0.26

Notes: The dependent variable is provincial-level migrant share in years 2005, 2010, 2015. The control variables are provincial level logged average wages, logged unemployment rate, logged output ratio between the output of tertiary and secondary sectors, public service index and amenity index. 1 year lagged explanatory variables are used in the estimations. In columns 4-6, the interaction term between *hukou* index and other control variables are also included. Public service index and amenity service index are computed using principal component analysis. *Hukou* indexes are collected from Zhang et al. (2019). Robust standard errors are clustered at province level. *** p<0.01, ** p<0.05, * p<0.1.

2.6.2 Prefecture-level Estimation Results

Table 2.5 shows the estimation results using prefecture-level data. Columns 1 and 2 present the results without interaction terms, which is related to equation 2.4. Column 1 reports the results including two traditional determinants of migration, that is the wages and unemployment rate. Column 2 includes further controls, such as the output ratio (as proxies for industry structure), public service index and amenity service index. Province and time fixed effects are also controlled in the estimations. Although the coefficient of two variables associated with traditional determinants of migration have the expected sign, the insignificance of the coefficients indicates that there is little evidence of impact of wage and employment probabilities of migration. The public service index has a negative and significant effect, which is contradictory as discovered in previous studies. As stated in section 2.3.1, the *hukou* system is associated with public services such as medical care, public education service, etc. As this chapter's focus is non-*hukou* registered migrants, public service may not be an incentive to migrate. The amenity service has a positive but insignificant effect on migration. Columns 3 and 4 show the results with interaction terms, which is related to equation 2.5. When interacting wage with *hukou*, it tells us about a trade-off between expected migration gain and migration cost caused by *hukou*-induced social provision loss. This is consistent with the estimation results in table 2.3. With stricter *hukou* restrictions, migrants would ask for higher wages to compensate for the social provision loss due to the lack of local *hukou*. The coefficient of unemployment is positive but insignificant and remains in column 3 and 4. The coefficient of the interaction term between public service and *hukou* index is not significant, which again provides suggestive evidence that the public service may not be a determinant of migration for non-*hukou* registered migrants.

Table 2.5 Prefecture-level Estimation Results

	(1)	(2)	(3)	(4)
ln_average_wage	0.000216	0.00184	-0.0229**	-0.0104**
	(0.00472)	(0.00326)	(0.0108)	(0.00406)
*hukou*_index	0.0937***	0.0491***	-0.150*	-0.0728*
	(0.0158)	(0.0173)	(0.0767)	(0.0383)
ln_urban_unemployment_rate	0.174	-0.0552	0.0472	-0.365
	(0.320)	(0.297)	(0.520)	(0.449)
ln_output_ratio		0.0775***		0.0441**
		(0.0154)		(0.0209)
public_service_index		-0.0222***		-0.0214*
		(0.00630)		(0.0123)
amenity_index		0.00448		-0.000403
		(0.00321)		(0.00492)
ln_average_wage×*hukou*_index			0.0236***	0.0120***
			(0.00769)	(0.00355)
ln_urban_unemployment_rate×*hukou*_index			0.0663	0.412
			(0.883)	(0.612)
ln_output_ratio×*hukou*_index				0.0542**
				(0.0235)

public_service_index×*hukou*_index				0.00256
				(0.0157)
amenity_index×*hukou*_index				0.00894
				(0.00747)
Year FE	YES	YES	YES	YES
Prefecture FE	YES	YES	YES	YES
Constant	-0.0299	0.0118	0.205*	0.132***
	(0.0556)	(0.0409)	(0.108)	(0.0432)
Observations	339	329	339	329
Adjusted R-squared	0.0516	0.240	0.0846	0.258

Notes: The dependent variable is prefecture-level migrant share in years 2005, 2010, 2015. The control variables are prefecture-level logged average wages, logged unemployment rate, logged output ratio between the output of tertiary and secondary sectors, public service index and amenity index. 1 year lagged explanatory variables are used in the estimations. In columns 3 and 4, the interaction term between *hukou* index and other control variables are also included. Public service index and amenity service index are computed using principal component analysis. *Hukou* indexes are collected from Zhang et al. (2019). Robust standard errors are clustered at prefecture level. *** p<0.01, ** p<0.05, * p<0.1

2.7 Gravity Model Estimations

2.7.1 Gravity Model Estimation Results

Table 2.6 shows baseline results obtained from using OLS and Poisson-pseudo-maximum-likelihood using bilateral migration data. In the migration literature, there are a few existing methods of handling zero migrant observation. These include dropping zero observations (Grogger and Hanson, 2011), Tobit specification and Poisson quasi-maximum-likelihood-estimation (Simpson and Sparber, 2013; Mayda, 2010; Beine et al., 2011), and the log of migration stocks (plus one) (Ortega and Peri, 2009; Ortega and Peri, 2013). In this chapter, OLS and poisson quasi-maximum-likelihood estimations are employed. The OLS and PPML estimations are related to equations 2.6 and 2.7. For OLS estimations, we add a one to each observation of migrant stock so that the zero observation are not discarded (around 80% zero observation in the data). By comparison, non-negative observations will be automatically taken into consideration with PPML estimations.

Columns 1-4 report the estimation results using OLS. The dependent variable in columns 1-4 is logged migration rate where the migration stocks in numerator is one plus the original migration stocks. Columns 1 and 2 simply include the migration gain and cost factors as given by equation (2.6). The coefficient of the wage differential is positive and highly significant. That is, a larger wage differential is more attractive to migrants. As the regional wage differential increases by 1%, the propensity of migrating from origin to destination increases by 0.348%. The coefficient of unemployment rate has expected negative sign and is significant at the 1% confidence level. It indicates that migrants tend to destinations with higher employment probabilities. The migratory distance is corresponding with relevant migration cost. The further away the two places are, the higher the cost for the move. That is, the migration costs increase along with the distance between origins and destinations. In column 1, distance enters with the expected negative sign: migration decreases with the distance between destination and origin locations. Compared to column 1, the coefficient size of the wage differential slightly decreases as more control variables are added in column 2. It implies that migrants may positively respond to potential destinations taking into consideration desirable public service and amenity conditions, which mitigates the positive effect of an expected higher wage. The coefficient of the difference in GDP ratio is insignificant, but the negative sign indicates that, in general, the secondary sector creates more job opportunities for migrants. According to the estimates of Lu and Xia (2015), there are over 60% of migrants working in the secondary sector based on 2005 mini-census statistics. The *hukou* population variable also enters with the expected positive sign which captures the positive effect of productivity externalities and of agglomeration economies. The positive and significant coefficients on public service and amenity service indexes suggest that better public and amenity services could be incentives for migration. To account for the different marginal effects of city-level controls on migration by *hukou* stringency, we interact all explanatory variables with the *hukou* stringency index. We investigate how mobility decisions respond to migration costs and labour market economic conditions with the migration restrictions (referred to *Hukou* policies). Columns 3 and 4 show the results with interaction terms included. The estimation results in columns 3 and 4 suggest that migrants will ask for a higher wage when moving to a city with more stringent *hukou* system compared to moving to a city with less stringent *hukou* restriction. By contrast, the effect of unemployment and geographical distance is more negative the more restrictive the *hukou* system is in the destination. While the differential of public service enters positively and significantly in column 2, we find that the interaction effect is negative, which is not in line with the theoretical framework proposed by Borjas (1999). A possible explanation for this result is that public services are only accessible to the residents who are entitled to local *hukou*, but the primary objective here is to investigate the migrants with non-local *hukou* who are not entitled to receive local public services. Thus, the effect of public service on migrants gradually decreases with more stringent *hukou* regulations. By contrast, amenities provide universal benefits for anybody who lives in a city. Results in column 4 show that the elasticity of output ratio is more negative for destinations with higher *hukou* restrictions. It suggests that the secondary industry is more attractive for migrants the more stringent the *hukou* restrictions.

From columns 5 to 8, the dependent variable is the migration share using PPML estimation methods, which is consistent with exponential function necessary to run PPML estimations. A large number of zero or undefined values often occurs in gravity equations of migration. As suggested by Silva and Tenreyro (2006), Poisson pseudo-maximum-likelihood estimation provides a natural way to deal with zero values of the dependent variable. Column 5-8 report the results of PPML estimations. Columns 5 and 6 simply include the migration gain and cost factors as given by equation (2.6). The significance levels in columns 5 and 6 are similar to these in columns 1 and 2 with OLS estimation technique, while the magnitudes using PPML estimation are larger. With the exponential function, the coefficients are semi-parametric[14] and thus the marginal effect should be computed for more accurate interpretations in PPML estimations. Column 5 suggests that if the regional wage differential increases by 1%, the propensity of migrating of origin-destination pairs increases by 0.897% which is substantially larger than OLS estimators. Even adding the controls in columns 6, the effect of wage differential is still larger under PPML. The coefficients of unemployment rate and distance have expected negative signs in columns 5 and 6. Similar to column 2, the public and amenity services also suggest positive effect on migration based on the PPML estimation results. Columns 7 and 8 show the results with interaction terms included. Both columns 7 and 8 show that the positive effect of wage differential on migration conditional on *hukou* stringency index and the positive effects remains for OLS and PPML. Comparing the results of OLS and PPML estimations, the following observations are in order. The results in column 8 shows an unexpected positive sign of the coefficient of the interaction term between unemployment and *hukou* stringency index, but a positive sign under OLS. The coefficient of the interaction term between distance and *hukou* index is positive with 10% significance level under PPML, while the coefficient is negative with 1% significance level under OLS. By computing marginal effects, column 8 also indicates that the effects of amenity service on migrants are negative taking into consideration the stringency of *hukou* regulations, whereas OLS predicts positive effect. Other than that, the OLS and PPML produce similar effects of other variables.

[14] The estimated effects can be computed by 100*[exp(beta)-1], where beta is the estimated coefficient.

Table 2.6 Gravity Model Estimation Results

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) PPML	(6) PPML	(7) PPML	(8) PPML
diff_wage	0.348***	0.253***	-0.145***	0.216***	0.893***	0.841***	0.433**	0.487***
	(0.024)	(0.025)	(0.046)	(0.065)	(0.117)	(0.111)	(0.168)	(0.171)
diff_unemployment	-0.088***	-0.065***	-0.064***	0.063***	-0.350***	-0.213***	-1.132***	-0.713***
	(0.012)	(0.010)	(0.022)	(0.021)	(0.065)	(0.063)	(0.086)	(0.100)
hukou	0.781***	0.490***	3.865***	3.796***	1.232***	0.541***	0.485	0.808*
	(0.030)	(0.031)	(0.304)	(0.309)	(0.089)	(0.105)	(0.467)	(0.474)
ln_distance	-0.237***	-0.238***	0.090***	0.103***	-0.968***	-0.955***	-1.072***	-1.084***
	(0.013)	(0.012)	(0.031)	(0.031)	(0.041)	(0.039)	(0.069)	(0.080)
diff_output_ratio		-0.020		0.048		-0.127*		0.256**
		(0.014)		(0.033)		(0.067)		(0.122)
diff_*hukou*_population		0.399***		0.412***		0.089		0.319***
		(0.008)		(0.018)		(0.063)		(0.106)
diff_public_service		0.014**		0.065***		0.185***		0.308***
		(0.006)		(0.015)		(0.030)		(0.055)
diff_amenity		0.063***		-0.112***		0.223***		0.501***
		(0.009)		(0.016)		(0.020)		(0.052)
diff_wage×*hukou*			0.768***	0.233**			0.854***	0.644***
			(0.063)	(0.091)			(0.155)	(0.188)
diff_unemployment×*hukou*			-0.022	-0.174***			0.650***	0.361***
			(0.031)	(0.031)			(0.061)	(0.074)
ln_distance_ji×*hukou*			-0.499***	-0.534***			0.103	0.132*
			(0.044)	(0.045)			(0.074)	(0.074)
diff_output_ratio×*hukou*				-0.199***				-0.430***
				(0.054)				(0.105)
diff_*hukou*_population×*hukou*				-0.006				-0.307***

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				(0.030)				(0.107)
diff_public_service×*hukou*				-0.076***				-0.122***
				(0.023)				(0.045)
diff_amenity×*hukou*_j				0.181***				-0.270***
				(0.019)				(0.038)
Constant	-6.689***	-6.487***	-8.754***	-8.648***	-2.557***	-2.439***	-2.000***	-2.575***
	(0.093)	(0.087)	(0.217)	(0.212)	(0.255)	(0.254)	(0.438)	(0.497)
Observations	13,340	13,110	13,340	13,110	13,340	13,110	13,340	13,110
Adjusted R-squared	0.159	0.319	0.187	0.367				
R-squared					0.069	0.1832	0.1627	0.1945

Notes: The dependent variable in column 1-4 is logged migrant share which is (migrant stock at destinations+1)/population at origin. The dependent variable in column 5-8 is migrant share which including null migration flows. The explanatory variables include *hukou* stringency index in destination locations, logged wage differentials between destination and origin locations, logged unemployment differentials, logged distance between destination' and origin's locations, output ratio differentials (tertiary-to-secondary industry output ratio), logged *hukou* registered population differentials, public service indexes differentials, amenity service indexes differentials. The interaction term between *hukou* stringency index and other controls are also included. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

2.7.2 Robustness

In this sub-section, we present some robustness check results. The scaled OLS in the baseline regression in table 2.6 adds one to the migration stocks before taking logs, and this might not provide the best interpretation over the relationship with explanatory variables. As mentioned above, the estimation problem is mainly caused by existence of zero migrants between many origin-destination pairs. Here, we try Eaton-Kortum Tobit model and Heckman selection model to further account for the high occurrence of null migration stocks. Eaton and Kortum (2001) suggest that the zero value in the dependent variable could be replaced by the minimum non-zero value observed from origins to destinations. This method has the same flavour as the strategy proposed by Eaton and Tamura (1994) who propose adding a small constant a to dependent variable instead of the arbitrary setting of $a = 1$.

Martin and Pham (2008) evaluated the efficiency of some estimation strategies with many zero values present. Eaton and Tamura's strategy generates the smallest biases comparing with scaled OLS and PPML. Unfortunately, Eaton and Tamura's strategy is not an easily operated programme therefore we use EK Tobit that has the advantage of being easily implemented and interpreted. Another strategy we employ here is Heckman two-step Selection model (1979). One difference between Tobit model and Selection model is that the effect of explanatory variables on the probability of being a migrant and the effect on the conditional mean of the positive level of migrants are the same in the Tobit model, whereas the Heckman model identifies that the decision to migrate and the choice of the number of migrants are different. Heckman two-step regressions explicitly account for the selection bias problem generated by unobserved city-pair level heterogeneity.

The usual procedure is to implement an exclusion restriction instrument in the selection equation, i.e., a bilateral variable that is only related to the migration decisions but unrelated to the migrant stocks. In our case, however, the decision to migrate is a function of wage, probability of employment, distance and other regional characteristics which is the same function that determines the optimal level of migration (Beine et al., 2011; Winter et al., 2001). To support the validity of the Heckman selection model, as suggested by Winter et al. (2001), we use alternative predicted probability of migration and its square value to be included in the second stage to account for selection bias in the decision to migrate.

Table 2.7 presents the robustness check results. Similar to the baseline estimations, all specifications suggest a positive relationship between bilateral migration and wage differentials. The coefficients remain highly significant but larger in terms of magnitude in the EK Tobit regressions. In addition, we also assess whether the effect of wage difference on migration stock conditional on *hukou* stringency is robust. Column 2 confirms this consistency. The sign of the coefficient between wage difference and *hukou* index is negative in column (6). All of the additional controls have the expected sign in column 1 and 4 besides *hukou* variable in column 4. Overall, the primary results obtained in section 2.6.2 is robust.

Table 2.7 Gravity Model Estimation Results: Robustness

	Eaton-Kortum Tobit		Heckman Two-step Selection Model			
			Decision to Migrate Probit	Heckman Regression	Decision to Migrate Probit	Heckman Regression
	(1)	(2)	(3)	(4)	(5)	(6)
diff_wage	1.467***	1.135***	0.614***	0.475***	0.355***	0.736***
	(0.131)	(0.274)	(0.041)	(0.101)	(0.108)	(0.165)
diff_unemployment	-0.144***	-0.193***	-0.119***	-0.133***	-0.116**	-0.136**
	(0.040)	(0.070)	(0.021)	(0.032)	(0.046)	(0.065)
hukou	2.054***	2.110***	1.294***	-0.705***	0.628	1.956***
	(0.081)	(0.544)	(0.052)	(0.155)	(0.467)	(0.374)
ln_distance	-1.317***	-1.332***	-0.505***	-0.345***	-0.566***	-0.231**
	(0.036)	(0.074)	(0.018)	(0.061)	(0.048)	(0.094)
diff_output_ratio	-0.141***	0.389***	0.063**	-0.315***	0.254***	-0.134
	(0.053)	(0.101)	(0.026)	(0.045)	(0.065)	(0.091)
diff_*hukou*_population	-0.208***	-0.145**	-0.175***	0.356***	-0.074**	0.389***
	(0.030)	(0.059)	(0.014)	(0.030)	(0.034)	(0.048)
diff_public_service	0.062***	0.057	0.051***	0.00049	0.050*	0.029
	(0.023)	(0.051)	(0.011)	(0.018)	(0.027)	(0.038)
diff_amenity	0.142***	0.184***	0.013	0.123***	-0.032	0.181***
	(0.024)	(0.066)	(0.011)	(0.013)	(0.026)	(0.033)
diff_wage×*hukou*		0.485*			0.463***	-0.273*
		(0.261)			(0.157)	(0.158)
diff_unemployment×*hukou*		0.045			0.016	-0.054
		(0.062)			(0.060)	(0.054)
ln_distance×*hukou*		0.020			0.092	-0.305***
		(0.081)			(0.069)	(0.062)
diff_output_ratio×*hukou*		-0.786***			-0.343***	-0.144

		(0.126)			(0.095)	(0.097)
diff_*hukou*_population×*hukou*		-0.089			-0.157***	-0.054
		(0.070)			(0.048)	(0.047)
diff_public_service×*hukou*		0.030			0.015	-0.011
		(0.056)			(0.036)	(0.038)
diff_amenity×*hukou*		-0.078			0.031	-0.054**
		(0.050)			(0.029)	(0.026)
predicted migration				-1.404***		-1.982***
				(0.483)		(0.660)
predicated migrationt squared				3.359***		3.120***
				(0.345)		(0.484)
Constant	-3.308***	-3.319***	1.718***	-4.827***	2.149***	-5.818***
	(0.248)	(0.492)	(0.123)	(0.415)	(0.326)	(0.652)
Observations	12,996	12,996	13,110	2,971	13,110	2,971

Notes: The dependent variable is logged bilateral migration share in 2005. The explanatory variables include *hukou* stringency index in destination locations, logged wage differentials between destination and origin locations, logged unemployment differentials, logged distance between destination and origin's locations, output ratio differentials (tertiary-to-secondary industry output ratio), logged *hukou* registered population differentials, public service indexes differentials, amenity service indexes differentials. The interaction term between *hukou* stringency index and other controls are also included. Columns 1 and 2 show the results using Eaton-Kortum Tobit estimation. Column 3-6 are two-stage Heckman estimations with estimations of selection equations in column 3 and 5. The predicted probability of having positive migration stock and its squared are added in estimations of column 4 and 6 to account for a potential selection bias. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

2.7.3 2SLS Estimates

Table 2.8 reports the results from two-stage least square estimations. For each regression, we report first stage coefficients on instrument and endogeneity tests of the endogenous wage variable. Columns 1 and 3 reports the estimation results of equation 2.6. Columns 1 reports the results with wages, unemployment, *hukou* and distance variables controlled. Column 3 reports the results with more variables controlled such as public and amenity indexes. Columns 2 and 4 reports the estimation results of equation 2.7. The regressions in columns 2 and 4 include interaction terms between control variables and *hukou* index. The first stage coefficients of the bartik wage are similar across the different model specifications, and the instruments are highly related to endogenous variables. Statistically, the Bartik-style expected earning is a strong predictor of a city's average wage. Endogeneity tests confirm the endogeneity of the wage variable. The null hypothesis of being exogenous for endogenous wage variable is rejected. The 2SLS estimators show that the coefficient on wage differential is positively significant at the 0.01 level, except the coefficient in column 3. The results also suggest migrants tend to favour higher wage differentials if destinations where they face higher *hukou* stringency indexes. The estimated effects here are slightly larger than those from OLS regressions. In summary, the results using instrument are quite consistent with the OLS results in table 2.6. The consistency suggests that OLS specifications are unlikely to be decayed by the reverse causality and omitted bias issues. Therefore, the regression estimations are able to support the interpretation that the wage differential has a larger effect on migration if a destination has higher *hukou* stringency index. Overall, the results emphasise that, all else equal, migrants tend to favour destinations with higher wage differentials and favourable probability of finding employment. Also, *hukou* stringency plays an essential role in internal migration. Increased migration restrictions increase the importance of migration costs and enhance the importance of economic factors such as wage differential and employment probabilities.

Table 2.8 Bartik IV Estimates

	(1)	(2)	(3)	(4)
	2SLS	2SLS	2SLS	2SLS
diff_wage	0.233***	0.408***	-0.277***	0.440***
	(0.0268)	(0.0388)	(0.0505)	(0.0836)
diff_unemployment	-0.101***	-0.0546***	-0.0843***	0.0904***
	(0.0120)	(0.0111)	(0.0228)	(0.0209)
hukou	0.831***	0.462***	3.913***	3.731***
	(0.0281)	(0.0264)	(0.235)	(0.213)
ln_distance	-0.236***	-0.239***	0.0918***	0.104***
	(0.0109)	(0.00994)	(0.0250)	(0.0224)
diff_output_ratio		-0.0274**		0.0339
		(0.0136)		(0.0299)
diff_*hukou*_population		0.392***		0.412***
		(0.00749)		(0.0162)
diff_public_service		0.00212		0.0520***
		(0.00605)		(0.0141)
diff_amenity		0.0476***		-0.151***
		(0.00643)		(0.0134)
diff_wage×*hukou*			0.814***	0.346***
			(0.0680)	(0.116)
diff_unemployment×*hukou*			-0.00975	-0.183***
			(0.0290)	(0.0265)
ln_distance×*hukou*			-0.501***	-0.537***
			(0.0340)	(0.0307)
diff_output_ratio×*hukou*				-0.205***
				(0.0455)
diff_*hukou*_population×*hukou*				-0.0238
				(0.0240)
diff_public_service×*hukou*				-0.0899***
				(0.0197)

diff_amenity×*hukou*				0.190***
				(0.0143)
Constant	-6.727***	-6.466***	-8.793***	-8.599***
	(0.0781)	(0.0712)	(0.173)	(0.156)
Observations	13,340	13,110	13,340	13,110
R-squared	0.157	0.317	0.186	0.359
First stage				
diff_bartik_wage	0.874***	0.778***	0.806***	0.78***
	(0.00681)	(0.00933)	(0.015)	(0.023)
R-squared	0.625	0.657	0.627	0.662
Endogeneity test (p-value)	0.0000	0.0000	0.0000	0.0000

Notes：the dependent variable is logged migration share. The key variable of interest is logged wage differential between destinations and origins. The instrumental variable for logged wage differential is logged bartik-style wage differential. The interaction term between wage differential and *hukou* stringency index are instrumented by the interaction between bartik-style wage differential and *hukou* index. Endogeneity test P-value is reported in table 2.8. Other control variables are *hukou* stringency index in destination locations, logged unemployment differentials, logged distance between destination and origin's locations, output ratio differentials (tertiary-to-secondary industry output ratio), logged *hukou* registered population differentials, public service indexes differentials, amenity service indexes differentials. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

2.8 Heterogeneity

Another potential concern is that the municipalities[15] have higher migrant shares and also provide relatively higher wages. The baseline estimates could be distorted by the samples of municipalities. Thus, as a robustness check, we re-estimate the specification but omit the big cities from the sample. The findings, reported in table 2.9, show that the results are nearly identical to the baseline 2SLS estimates. Columns 1 and 3 report the estimation results following equation 2.6. Columns 2 and 4 report the estimation results following equation 2.7 which includes the interaction terms between control variables and *hukou* index. Columns (1) and (2) report the estimation result with the samples excluding big cities. The results suggest that the coefficient of wage differential is positive and significant at 1% level. Also, the coefficient of the interaction term of wage differential and *hukou* index is positive and significant, indicating that wage differentials have larger effect on migration if a destination has higher *hukou* stringency index. That is said, migrants place more weights on wage gains in order to compensate migration cost induced by *hukou* system. The coefficient of interaction term between unemployment rate and *hukou* index is negative and significant, suggesting that migrants tend to favour destinations with higher employment opportunities when facing higher *hukou* stringency in destinations. The costs of migration increase with migration distance; therefore, a negative coefficient of distance is expected. The results in columns 1 and 2 suggest significant effects with the expected signs. Furthermore, the negative and significant coefficient on interaction term between public service and *hukou* index is consistent with the estimation results in tables 2.4 and 2.6 which estimate whole samples. By contrast, columns (3) and (4) reports estimation results only including big cities. In column 3, the coefficient of wage differential is positive while insignificant. Unemployment rates and distance have significant and negative effects on migration. The amenity index has positive and significant effect on migration. However, when adding interaction terms between control variables and *hukou* index, most of the coefficient estimates are not statistically significant, in particularly the coefficients of interaction terms. This may imply that the economic variables are not important anymore when facing very strict *hukou* stringency. This might be too costly to move to big cities compared to other relatively smaller cities. The wages and employment probabilities might be higher in big cities

[15] The municipalities here refer to four cities: Beijing, Shanghai, Guangzhou and Shenzhen.

however, the job market is more competitive which makes people especially migrants more difficult to secure a job. Also, big cities have worse traffic conditions and environment quality, which makes the economic incentives less attractive. Overall, the insignificant estimations for big cities provide evidence that big cities do not bias the estimation results with whole samples.

Table 2.9 2SLS Estimates Excluding Big Cities

	Excluding big cities		Only big cities	
	(1)	(2)	(3)	(4)
diff_wage	0.623***	0.443***	0.198	3.036*
	(0.0397)	(0.130)	(0.286)	(1.842)
diff_unemployment	-0.0296***	0.233***	-0.181***	-0.805
	(0.0112)	(0.0324)	(0.0700)	(0.545)
hukou	-0.0203	3.543***	0.0780	-0.848
	(0.0371)	(0.327)	(0.238)	(1.649)
ln_distance	-0.218***	0.0963***	-0.932***	-1.719***
	(0.00983)	(0.0308)	(0.0637)	(0.413)
diff_output_ratio	-0.0535***	-0.000982	-0.546***	-0.987
	(0.0134)	(0.0394)	(0.117)	(0.682)
diff_*hukou*_population	0.393***	0.400***	0.178***	0.334
	(0.00736)	(0.0206)	(0.0605)	(0.365)
diff_public_service	-0.0131**	0.0962***	0.0177	0.181
	(0.00604)	(0.0182)	(0.0567)	(0.329)
diff_amenity	-0.00594	-0.241***	0.0702***	0.267
	(0.00691)	(0.0204)	(0.0238)	(0.278)
diff_wage×*hukou*		0.365*		-1.345
		(0.209)		(0.953)
diff_unemployment×*hukou*		-0.432***		0.323
		(0.0521)		(0.278)
ln_distance×*hukou*		-0.526***		0.407*
		(0.0477)		(0.219)
diff_output_ratio×*hukou*		-0.165**		0.243
		(0.0665)		(0.350)
diff_*hukou*_population×*hukou*		0.00785		-0.0978

		(0.0337)		(0.190)
diff_public_service×*hukou*		-0.162***		-0.106
		(0.0282)		(0.163)
diff_amenity×*hukou*		0.324***		-0.115
		(0.0286)		(0.151)
Constant	-6.334***	-8.492***	-0.255	1.591
	(0.0724)	(0.212)	(0.559)	(3.037)
Observations	12,654	12,654	456	456
R-squared	0.254	0.286	0.404	0.427

Notes: The results are estimated using 2SLS approach. In column 1 and 2, four big cities are excluded, such as Beijing, Shanghai, Guangzhou, Shenzhen. In column 3 and 4, the samples are big cities, which has 456 observations only. Robust standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Furthermore, the effect of *hukou* could be different in terms of different skills for migrants. In this case, we separate the migrants into low-skilled and high-skilled. We classify migrants with secondary schooling education or below as low-skilled migrants. Table 2.10 reports the 2SLS results in terms of different skills. The estimation results reported in table 2.10 follow the equation 2.7. Columns 1 and 2 report the estimation results using whole samples. Columns 3 and 4 report the estimation results using samples excluding big cities; by contrast, columns 5 and 6 report the results with only big cities samples. The estimation results in terms of significance and magnitudes using whole samples are similar to those results using sample excluding big cities. Comparing to the coefficients with the samples of low skilled migrants, the coefficients magnitudes with high skilled migrants tend to be higher. Additionally, one can see from the results that the coefficient of the interaction term between wage differential and *hukou* index is positive for the low-skilled migrant share while it is negative for high-skilled migrant shares. For low-skilled migrants, the wage gap between destinations and origins matters more compared with high-skilled migrants. One possible explanation for this result is that the high-skilled migrants do not only care about the wage compensation, but also care about job opportunities, promotion opportunities, city amenities, which absorb the responsiveness of mobility to income factors. Also, as the *hukou* policy is skilled biased, it is easier for high-skilled migrants to obtain local *hukou*. The results in terms of other variables are similar between high-skilled and low-skilled migrants.

Table 2.10 Heterogeneity Results

	Full Sample		Excluding big cities		Only big cities	
	Low Skill	High Skill	Low Skill	High Skill	Low Skill	High Skill
	(1)	(2)	(3)	(4)	(5)	(6)
diff_wage	0.382***	0.337***	0.361***	0.345***	3.401*	-1.507
	(0.0829)	(0.0604)	(0.128)	(0.0933)	(1.926)	(1.307)
diff_wage×*hukou*	0.432***	-0.375***	0.459**	-0.341**	-1.361	0.348
	(0.115)	(0.0840)	(0.207)	(0.151)	(0.996)	(0.676)
diff_unemployment	0.0760***	0.124***	0.230***	0.188***	-0.709	-0.618
	(0.0208)	(0.0151)	(0.0321)	(0.0234)	(0.570)	(0.387)
hukou	3.598***	2.250***	3.363***	1.765***	-2.157	-0.867
	(0.211)	(0.154)	(0.324)	(0.236)	(1.724)	(1.169)
ln_distance	0.117***	0.222***	0.0985***	0.193***	-1.868***	-0.952***
	(0.0222)	(0.0162)	(0.0304)	(0.0222)	(0.432)	(0.293)
diff_output_ratio	0.0428	-0.0442**	0.00775	-0.0685**	-0.999	-0.0975
	(0.0297)	(0.0216)	(0.0390)	(0.0284)	(0.713)	(0.484)
diff_*hukou*_population	0.426***	0.350***	0.398***	0.323***	0.525	0.490*
	(0.0161)	(0.0117)	(0.0204)	(0.0149)	(0.382)	(0.259)
diff_public_service	0.0521***	0.0593***	0.106***	0.0622***	0.152	0.112
	(0.0139)	(0.0102)	(0.0180)	(0.0131)	(0.344)	(0.233)
diff_amenity	-0.138***	-0.151***	-0.230***	-0.173***	0.286	0.285
	(0.0133)	(0.00966)	(0.0202)	(0.0147)	(0.290)	(0.197)
diff_unemployment×*hukou*	-0.162***	-0.213***	-0.433***	-0.324***	0.251	0.201
	(0.0262)	(0.0191)	(0.0515)	(0.0375)	(0.291)	(0.197)
ln_distance×*hukou*	-0.537***	-0.341***	-0.508***	-0.294***	0.487**	0.260*
	(0.0305)	(0.0222)	(0.0472)	(0.0344)	(0.229)	(0.155)
diff_output_ratio×*hukou*	-0.216***	-0.0155	-0.175***	0.0195	0.361	-0.135
	(0.0451)	(0.0328)	(0.0658)	(0.0479)	(0.366)	(0.248)

diff_*hukou*_population×*hukou*	-0.0390	0.152***	0.0223	0.208***	-0.164	-0.0843
	(0.0238)	(0.0173)	(0.0333)	(0.0243)	(0.199)	(0.135)
diff_public_service×*hukou*	-0.0891***	-0.101***	-0.174***	-0.110***	-0.0423	-0.158
	(0.0195)	(0.0142)	(0.0279)	(0.0203)	(0.171)	(0.116)
diff_amenity×*hukou*	0.175***	0.204***	0.314***	0.235***	-0.107	-0.114
	(0.0141)	(0.0103)	(0.0283)	(0.0206)	(0.158)	(0.107)
Constant	-8.634***	-9.564***	-8.503***	-9.279***	3.079	-1.774
	(0.154)	(0.112)	(0.210)	(0.153)	(3.175)	(2.154)
Observations	13,110	13,110	12,654	12,654	456	456
R-squared	0.346	0.430	0.289	0.417	0.436	0.418

Notes: The dependent variable is logged migration share. 2SLS results are estimated. Columns 1 and 2 are estimation results using full samples, while columns 3 and 4 are results using samples excluding big cities. Column 5 and 6 use samples with only big cities. People with secondary school education or below as low-skilled. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

2.9 Conclusions

The goal of this study is to examine the importance of migration restrictions in China by studying the role of the *hukou* system on shaping migration decisions. Two questions are addressed in this chapter. The first is whether the *hukou* system is a barrier to internal migration and the second is whether it enhances the responsiveness to economic factors (such as wages and employment probabilities).

Regional-level panel data and multi-regional models are estimated, where the logic rests that migrants move in search of higher wages. Also, the responsiveness of internal migration decisions to *hukou* restrictions is estimated. To identify the effects of restrictions, we exploit the variation of *hukou* stringency across cities to look at how the *hukou* restrictions may affect migration. *Hukou* policy is highly localised and there are large variations on *hukou* stringency across cities. We exploit a newly developed *hukou* index that measures how difficult for migrants to obtain a local *hukou* in the place of their residence. A higher value indicates more stringent restrictions and thus more difficult for migrant to obtain local *hukou*. There is an endogeneity concern between migration and wages. The possible endogeneity issue has been controlled by constructing Bartik-style expected wage instrument.

The results demonstrate that, all else equal, migrants tend to favour destinations with higher wage differentials and favourable probability of finding employment. Also, *hukou* restrictions are an important factor to understand internal migration patterns in China. Migration can be significantly responsive to a loosening mobility restriction. With stronger migration restrictions, the importance of other economic factors such as wage differential and employment probabilities are enhanced. Beyond the traditional economic determinants of migration (e.g., wages and employment probabilities), other factors are also examined in multi-regional estimation models, such as public service and amenity index services. The estimation results show that the effect of public services on migrants gradually decreases with more stringent *hukou* regulations, while the impact of amenity services is positive. It is plausible that public services may have no effect or a negligible one in our estimation, as the objective of this chapter focus on non-local *hukou* migrants who do have access to public services in the residency places.

In conclusion, the interaction of socioeconomic factors and *hukou* policies is essential to provide guidance for the regulation of internal migration pattern. This chapter is the first to provide quantitative estimates of the effects of migration regulations on migration flows in the Chinese context using specific measures of *hukou* stringency. It may serve as guidance for policy makers to improve migration management. To achieve better urbanisation and industrialisation, it seems that further deregulation may be needed to encourage greater labour mobility.

Appendix A

Table A2.1 Principal Component Analysis for Public Service Indexes in Prefecture Level

	2004		2009		2014	
	Loading	Unexplained Variance	Loading	Unexplained Variance	Loading	Unexplained Variance
Panel A. School Index						
Eigenvalue (% of variance)	1.65(0.82)		1.61(0.81)		1.40(0.70)	
primary students to teachers	0.71	0.18	0.71	0.19	0.71	0.30
secondary students to teachers	0.71	0.18	0.71	0.19	0.71	0.30
Panel B. Medical Index						
Eigenvalue (% of variance)	1.97(0.66)		1.81(0.60)		1.60(0.53)	
hospitals per capita	0.32	0.80	0.32	0.81	0.14	0.97
beds per capita	0.67	0.12	0.68	0.17	0.71	0.19
doctors per capita	0.67	0.11	0.66	0.20	0.69	0.24
Panel C. Overall Index						
Eigenvalue (% of variance)	1.45(0.72)		1.49(0.74)		1.22(0.61)	
School Index	-0.71	0.28	-0.71	0.26	-0.71	0.39
Medical Index	0.71	0.28	0.71	0.26	0.71	0.39

Table A2.2 Principal Component Analysis for Amenity Index in Prefecture Level

	2004		2009		2014	
	Loading	Unexplained Variance	Loading	Unexplained Variance	Loading	Unexplained Variance
Panel A. Life Index						
Eigenvalue (% of variance)	1.41(0.70)		1.06(0.53)		1.35(0.67)	
Cinema per capita	0.71	0.30	0.71	0.47	0.71	0.33
Library books per capita	0.71	0.30	0.71	0.47	0.71	0.33
Panel B. Transportation Index						
Eigenvalue (% of variance)	1.56 (0.78)		1.59(0.79)			
Public buses per capita	0.71	0.22	0.71	0.21	0.71	0.41
Road space per capita	0.71	0.22	0.71	0.21	0.71	0.41
Panel C. Environment Index						
Eigenvalue (% of variance)	1.98(0.496)		1.86(0.47)		2.26(0.57)	
Green area	0.14	0.96	-0.18	0.94	0.28	0.82
SO2 emission per capita	0.66	0.13	0.69	0.11	0.62	0.13
Smoke dust emission per capita	0.60	0.30	0.69	0.12	0.60	0.19
Wastewater per capita	0.43	0.63	0.14	0.97	0.42	0.61
Panel D. Overall Index						
Eigenvalue (% of variance)	1.80(0.60)		1.32(0.44)		1.93(0.64)	
Life Index	0.64	0.26	0.65	0.44	0.53	0.45
Transportation Index	0.64	0.26	0.74	0.27	0.62	0.27
Environment Index	0.42	0.68	-0.17	0.96	0.58	0.35

Table A2.3 Data Definition and Resources for Panel Data Estimations

Variables	Definitions	Data Sources
shares of migrants (share of non-*hukou* residents)	(Residents population - *Hukou* registered residents)/ Residents population	China population census from 2005, 2010, and 2015. And statistical yearbooks from provinces and cities.
average wage	Total paid wages/ total employed persons	China City Statistics Yearbooks from 2004, 2009, and 2014.
urban unemployment rate	Urban registered unemployed persons/ (Total employed persons + unemployed persons)	China City Statistics Yearbooks from 2004, 2009, and 2014.
tertiary to secondary	The level of GDP contributed by tertiary industry/ the level of GDP contributed by secondary industry	China Statistics Yearbook for Regional Economy 2004, 2009, 2014.
public service index	The measure of education quality and medical service in cities. PCA applied.	China City Statistics Yearbooks from 2004, 2009, and 2014.
amenity index	The measure of amenity conditions in cities, such as life quality, transportation conditions. PCA applied.	China City Statistics Yearbooks from 2004, 2009, and 2014.
hukou index	A measure of *hukou* stringency for each city. Assume *hukou* stringency does not have significant change between 2000 and 2013. A different *hukou* registration index for each city in period 2014-2016 compared to the period 2000-2013.	Zhang et al. (2018). A quantitative analysis of *Hukou* reform in Chinese cities:2000-2016.

Table A2.4 Classification of Sectors

Code	Sectors
1	Agriculture, forestry, animal husbandry and fishery
2	Mining
3	Manufacturing
4	Electricity, gas and water
5	Construction
6	Transport, warehouse and postal industry
7	Information transmission, computer service and software
8	Wholesale and retail
9	Accommodation and Catering
10	Financial Industry
11	Real estate
12	Leasing and business services
13	Scientific research, technical services and geological prospecting industry
14	Water conservancy, environment and public facilities management
15	Resident services and other services
16	Education
17	Health and social welfare
18	Culture, sports and entertainment
19	Public Administration and social organisation
20	International organisation

Source: China Industrial Classification for National Economic Activities (GB/T 4754-2002)

Chapter 3 Internal Migration and Urban Wage: Evidence from *Hukou* Reforms in China

3.1 Introduction

China experienced very high economic growth in the late 1990s and 2000s. Especially after it became a member of the World Trade Organisation in 2001, a large number of labour-intensive enterprises began to expand quickly. Although China has two segregated labour markets - the urban and the rural, due to the household registration system discussed in the previous chapters, the fast-expanding private sector began to absorb rural labour.

Migration restrictions were also relaxed considerably. It is estimated that an estimated 150 million rural migrants were working in urban areas in 2009, which accounted for one third of the total labour force population in China (Meng and Zhang, 2010). The rural-to-urban migrants have contributed largely to the economic growth of China, but concerns have been raised regarding whether the rural migrants have potential negative impact on urban workers' wages and employment. The common wisdom states that the influx of immigrant's crowd-out the job opportunities for local workers and depress average wages in local labour markets.

Such prediction is based on a simple static model of competitive labour market where there is only one industry and capital is fixed (Ottaviano and Peri, 2012; Lewis, 2011). Consistently with this theoretical prediction, Altonji and Card (1989) exploit the spatial variations of immigrants across cities/regions and find immigration may lower the wages of locals with similar skills and increase unemployment. However, much existing literature reaches contradictory findings on the effects of immigration on local economies. For instance, Lewis and Peri (2015) find that immigration has a positive effect on local wages for most native workers in industrialised countries. They summarise that the impact depends on immigrants' and natives' skill distributions and how the natives move along the skill distribution in response to the influx of immigrants. Empirical papers have distinguished between the effects on skilled and unskilled workers (Card, 2005; Ottaviano and Peri, 2012; Peri, 2012). Most studies focus on international migration in developed countries and a systematic analysis is yet missing on how migration affects the local labour markets in developing countries. Also, the share of international migrants is relatively smaller in developing countries. One should not assume that the evidence on developing countries is similar to that in developed countries. The assessment of the impact of internal migration in developing countries is particularly important. Especially China has experienced massive internal migration since the late 1980s, providing a particularly interesting opportunity to study the impact of internal migration on native workers.

Using micro-level population census data, this chapter provides new evidence on the effect of rural-urban migration on urban wages, particularly in light of *hukou* reforms in the beginning of the 2000s. The distribution of population in Chinese labour markets has been mainly shaped by the strict policy of the household (*Hukou*) system, which aims to restricting migration from rural to urban areas, as well as inter-urban areas. As discussed previously in Chapters 1 and 2, each citizen in China is legally bound to register their permanent residence in the local authority and each *hukou* is classified as agricultural (rural) or non-agricultural (urban). The *hukou* (household registration) system is profoundly connected to social welfare and depends on the policies of the local authority, making mobility across agricultural and non-agricultural sectors and across regions difficult. Also, since the implementation of the *hukou* system, urban and rural *hukou* holders are treated differently in terms of social welfare and job allocations. Non-agricultural *hukou* holders are entitled to more rights than agricultural *hukou* holders. For example, in late 1950s, urban residents were given food coupons to buy food while rural residents were excluded from this scheme. With Chinese economic reform beginning at the end of 1970s, the distinctions in welfare systems between rural and urban residents have been diminishing. As a result, a number of rural migrants began to migrate and work in urban areas. However, rural migrants were still treated differently from their urban resident counterparts. For example, rural migrants barely obtain jobs from state-owned enterprises. Compared to them, their urban resident counterparts not only obtain jobs in state-owned enterprises, but also have higher wages and better social welfare. Rural migrants normally take low-end jobs that offered less wages, such as construction, cleaning service, etc. With little social welfare and less wages in urban areas, it would be costly for rural migrants to live in urban areas, which hinders rural-to-urban migration. In the meantime, labour shortages occurred in urban areas while there was labour surplus in rural areas. Therefore, such migration barriers caused substantial labour misallocation. Since the 2000s, China has launched some *hukou* reforms to eliminate the distinction between rural and urban *hukou* types and attract rural workers. Such a policy changes could be seen as a positive labour supply shock to the urban labour market by allowing more rural workers to move to the urban areas.

In this chapter, we try to investigate the effect of rural-urban migration on the wages of urban natives and explore whether this relationship is affected by *hukou* policies and reforms. The main challenge in doing so is the possible reverse causality between labour market outcomes and internal migration. For instance, the selection of rural-to-urban migration destinations would depend on the urban wage levels and employment conditions. Moreover, some omitted factors could affect both migration and urban wage levels. To mitigate the endogeneity issue, we adopt instrumental variables (IV) approach to explore the causal relationship between urban native wages and (internal) immigration. we try two ways to construct instrumental variables. First, we follow Nunn and Qian (2014) and Sequeira et al. (2020) to construct a DID style instrument that exploits two facts. The first is that with the *hukou* restrictions relaxation, the probability of obtaining local *hukou* increases and rural workers are more likely to move to urban areas (Sun et al., 2011; Kinnan et al.,2018). Therefore, city-level *hukou* reforms affect migration settlements in cities. And the other fact is that newly arriving migrants normally follow the previous immigrants' settlement pattern (Card, 2001). To be more specific, the identification strategy exploits the interaction of time-varying migration flows and *hukou* reforms across prefectures[16]. Second, we construct shift-share instruments based on pre-existing migration patterns. That is, we combine *hukou* reforms ("the shift") with historical migration patterns between prefectures ("the share") to construct an exogenous predictor of migrant shares. The identification assumption underlying the instrumental variables strategy design is that the *hukou* reforms are numerous and random.

According to Sequeira et al. (2020), there are two steps to construct the instrumental variable for rural-to-urban migrants. First, we begin with a "zero-stage" to examine the determinants of the growth of rural migrant stocks from 2001 to 2005. The controls include the lagged logarithm of rural migrant stocks, the interaction between lagged

[16] In this chapter, we use "city" and "prefecture" interchangeably.

logarithm of rural migrant stocks and the indicator for *hukou* reforms, as well as indicators for economic growth and urbanisation. Secondly, the instrumental variable for the migrants in each year is constructed using the predicted growth rate of the migrant stocks and the initial migrant stock in 2001. After the construction of the instrumental variable, we estimate the impact of rural-to-urban migration on urban native wages. We also use the shift-share instrument (Imbert et al., 2020; Kinnan et al., 2018) as a robustness test. Moreover, as mentioned above, we estimate the impact of rural migrants, respectively, on the wages for low- and high-skilled urban native workers.

We find that rural-urban migrants have a positive effect on the wages of urban native wages, with a larger effect on high-skilled urban natives. The estimates indicate that 1 percentage point increase in the migrant shares results in 0.07 percentage increase in individual's monthly wage. The magnitude of the coefficient in the estimates for hourly wages are similar to those of monthly wages. The wage rise for high- and low-skilled urban workers is 0.06% and 0.07% respectively. The estimates are consistent with the estimated elasticity measured by Meng and Zhang (2010). Using an instrumental variable approach, they find that the effect of rural-urban migrants on urban native wages is 0.098 percent, which is slightly larger than the estimated coefficient in this chapter. Our estimated results are also comparable with the analysis in Docquier et al. (2014). They quantify the labour market effects of migration flows on OECD countries and find the effect of immigration on native wages ranges from 0 to +4%, where our results lie.

The contribution of this chapter to the existing literature is twofold. Firstly, this chapter adds to the literature that examines the impact of immigration on local labour markets, especially from the perspective of internal migration. There is a growing literature analysing the impact of large-scale of rural-urban migration on the wage outcomes of urban native workers. Using the lagged pull and push factors for migrant labour as instruments, Meng and Zhang (2010) find that rural-urban migration does not affect local urban labour markets. However, Combes et al. (2015) find that migrants had a significant and positive impact on urban wages, two-thirds of which are due to the complementary nature of rural-urban migrants and urban labour, and one-third to the agglomeration effect. Migrants are mainly concentrated in the low-skilled sector, which feeds other local industries and make higher skilled urban workers more productive. Also, an increase in the number of migrants increases total employment density, which, in turn, affects urban wages. Moreover, Zhao (2020) constructs a shift-share instrument based on the initial distribution of migrants and finds that rural-urban migration has a positive effect on the wages of urban workers. In particular, the wage increase is estimated to be 4.5% for low-skilled urban workers, 5.3% for middle-skilled urban workers, and 7.6% for high-skilled urban workers. However, this literature does not exploit the information on how the timing of *hukou* reforms varies across regions as exogenous variation in the construction of instrument variables. This can be a particularly useful strategy. As mentioned above, since the 1980s China has experienced great migration, and the labour mobility restrictions (household registration system) have been gradually relaxed. In this case, the *hukou* system is crucial to understand the migration patterns and the changes in the labour market in China. It is important to take into consideration the impact of *hukou* on rural migration when constructing instrumental variables for the shares of rural migration, rather than simply using past migration patterns as instrument.

Secondly, this chapter contributes the existing literature assessing the possible benefits of relaxing *hukou* restrictions. Relaxing mobility restrictions could reduce rural-urban income disparity and increase the urbanisation rate (Au and Henderson,2006; Ngai et al.,2019; Whalley and Zhang, 2007), increase productivity and reduce welfare distortion (Hsu and Ma, 2019; Tombe and Zhu, 2019), raise employment adjustment (Wang et al., 2020; Zi,2020), strengthen core-periphery outcomes (Bosker et al.,2012), and smoothen the consumption volatility of households (Kinnan et al., 2018). However, most papers focus on aggregate analysis. No study has taken a micro perspective to directly examine how individual wages are affected by the *hukou* system and how a relaxation of the system would influence individual wage adjustment, especially for urban workers. Yet, little is known about how the *hukou* reforms affect average wages of local labour markets indirectly through shaping migration patterns.

This chapter is organized as follows. Section 3.2 presents the literature review; Section 3.3 introduces the policy background and section 3.4 presents a conceptual framework. Section 3.5 describes the data construction process and presents some descriptive facts. Section 3.6 shows the decomposition results for China’s wage structure, followed by the description of empirical strategy in section 3.7. Section 3.8 presents the empirical results. Finally, section 3.9 concludes.

3.2 Literature Review

This chapter is related to three areas of research. First, the chapter is related to the growing literature on country-specific studies that investigate the effect of migrants on local residents' nominal wages. There is abundant literature on immigration. Two influential papers by Borjas et al. (1997) and Borjas (2003) have found a significant adverse effect of immigration on the wages of low-skilled native workers. More recent papers emphasise that the immigrants' and natives' respective skill distributions need to be taken into consideration when estimating the immigration effect. Ottaviano and Peri (2012) estimate the elasticity of substitution within the same education and experience group, and the total wage effects across different groups using a nested-CES model. They find that immigrants have a small effect on the wages of native US workers due to a small degree of imperfect substitutability. By contrast, Dustmann et al. (2013) point that immigrants downgrade occupations significantly upon arrival and pre-allocated skill or occupation distributions of immigrants based on their observed characteristics would place them at different locations than actually happened in destinations. They develop a flexible empirical strategy which allows immigrants to have differential skill distributions according to their observed position in the native wage distribution and find that the effect on the average native wages is zero or slightly positive both theoretically and empirically. Docquier et al. (2014) use a new global dataset on migration patterns to estimate the labour market effect of immigrants and emigrants on OECD countries. By constructing an aggregate production function model, their estimation shows that the wage and employment effects of immigration are positive on the less educated native workers in all OECD countries in the 1990s and 2000s. On the other hand, emigration causes wage losses for the low-skilled native workers and thus leads to larger inequality within countries.

This work is also closely related to recent papers studying the effect of internal migrants in China on local labour markets. Meng and Zhang (2010) examine the impact of the large-scale rural-urban migration on the average employment and earnings of urban workers. Using combined push and pull factors as instruments, they find that rural-urban migration in China has non-negative effect on the employment and earnings of urban workers, both at the city-skilled and unskilled level. One conjecture proposed by

Meng and Zhang (2010) to reconcile the findings is that migrants and urban workers, to some extent, are complements. By comparison, Combes et al. (2015) decompose the migrants' impact on urban workers into substitutability/complementarity with natives in the production function and agglomeration effects. They find evidence of a positive and significant impact of the local share of migrants on natives' wages, with a larger positive impact on skilled natives. Similar findings are confirmed by Zhao (2017). By constructing a supply-push instrument based on past migrants, Zhao (2017) find that on average, an influx of migrants increases the wages of urban workers. The effects are larger for more skilled urban workers. Our study focuses on similar issues and also employs instrumental variable strategy but exploits information on the timing of *hukou* reforms across cities/regions as exogeneous variation in the construction of instrumental variable.

More directly related to our work are some papers studying the effect on mobility barriers. Wang et al. (2020) empirically investigate the casual relationship between employment adjustment and *hukou* reforms. The results show a positive effect of the *hukou* reforms on net employment adjustment at firm level. They also find evidence that greater employment adjustment is associated with larger tariff cuts in *hukou* reform cities (relative to non-reform cities). Also, the urbanization rate in China is lower than expected, partly due to *hukou*'s migration restrictions. As such, large rural-urban income disparities arise (Whalley and Zhang, 2007). Kinnan et al. (2018) study the effects of changing incentives to migration on rural households by exploiting the variations in the province-level reforms to the *hukou* system. The results suggest that increased access to internal migration reduces consumption volatility and lead to the liquidation of assets holding. Based on a new economic geography (NEG) model, Bosker et al. (2012) analyse the effect of a relaxation of the *Hukou* restrictions on China's internal economic geography. Their analysis shows that the relaxation of the *Hukou* restrictions lead to more pronounced core-periphery outcomes. Ngai et al. (2019) provide quantitative analysis of the effect of *hukou* household registration system on employment allocation in China. The analysis reveals that the mobility barriers lead to over-employment in agricultural sectors and under-employment in non-agricultural sectors, which distorts both urbanisation and industrialisation. Hsu and Ma (2021) document a striking contrast in the migration pattern between 2000 and 2015 and conduct a quantitative analysis to evaluate the aggregate productivity and welfare effect of the differential *hukou* reforms. The results show that the alternative migration policies that are more uniform would improve national welfare substantially. However, most of the papers focus on aggregate analysis. No study has taken a micro perspective to directly examine how individual wages are affected by *hukou* restrictions and how a relaxation of this system would influence individual wage adjustment, especially for urban workers. This chapter will try to fill this gap and understand how the *hukou* reform affects urban native wages.

3.3 Policy Background: *Hukou* Reforms

We have already extensively discussed the *hukou* system in Chapters 1 and 2 and its effects on the higher costs of migration. In this Section, we focus on the reforms that have happened to the system over time, as we later exploit these variations for the construction of the Instrumental Variable.

These regulations began to change since 1980s, when the government has taken steps to reform the *hukou* system by implementing some reform policies such as allowing the conversion of rural to urban *hukou*. However, the conversion quota is very limited for each province. To further promote the labour movement from rural to urban areas, in the early 2000s, the distinction between agricultural and non-agricultural *hukou* has gradually eliminated. The local governments are given some freedom to decide upon the entry permits. The local government lowered the entry barriers, which increased the probability of obtaining a local *hukou* in a city. For instance, Beijing implemented a reform in 2002 which granted local *hukou* to migrants if they had an apartment and a stable job. This round of *hukou* reforms aimed at attracting migrant workers to move into urban areas, especially during a period when China experienced dramatic technological development and economic growth. With urban wages exceeding rural wages, there has been an increase in migration into urban areas. Appendix A1 provides a detailed description of reforms.

Although we are unable to directly observe the effect of *hukou* reform on urban labour supply, we can provide evidence suggestive of a positive effect of *hukou* reforms. Table 3.1 shows the urban population share and rural-urban migrant share in 2000 and 2005, which indicates that cities implementing *hukou* reforms experience a higher increase in urban population share and rural-urban migrant shares than non-reform cities.

Table 3.1 Average Urban Population and Rural-Urban Migrant Share

Types	Panel A: urban population share (%)			Panel B: rural-urban migrant share (%)			Number of Cities
	2000	2005	changes	2000	2005	changes	
Non-reform at all	35.44	42.43	6.99	3.31	3.33	0.02	280
Non-reform in 2000, reform in 2001-2004	54.09	61.52	7.43	6.4	7.86	1.46	56
Reforms in 2000-2004	60.05	66.44	6.39	6.45	7.75	1.3	60

Notes: Non-reform at all denotes those cities without any *hukou* reforms during 2000 and 2004; non-reform in 2000, reform in 2001-2005 are those cities implemented the *hukou* reforms between 2001 and 2004; reforms in 2000-2004 are those cities that implemented the *hukou* reforms during years 2000-2004. Panel A shows urban population shares in total prefecture population (aged 16 and 64); panel B shows rural-urban migrant shares in total prefecture population (aged 16 and 64). Data sources: 2000 population census micro-level data and 2005 mini population census micro-level data.

3.4 Conceptual Framework

The model used in this chapter is similar to that of Card and Lemieux's (2001). we assume that the production function is nested CES and that there are two large groups among native workers: high-skilled natives and low-skilled natives. Also, migrants are mostly low-skilled workers.

$$Y = (\theta_N L_N^{\rho} + L_M^{\rho})^{1/\rho} \qquad (3.1)$$

where the aggregate output Y is a function of labour input, N denotes native workers, M migrant workers. θ_N is the efficiency parameter of native workers relative to migrant workers. The elasticity of substitution between natives and migrants is $\sigma_E = {}^{1}/_{1-\rho}$, $-\infty < \rho \leq 1$. ρ determines the degree of substitutability of the natives and rural migrants. The value of ρ is less than or equal to 1 and can be $-\infty$. If $\rho = 1$, then $\sigma_E = \infty$ which follows that the natives and migrants are perfect substitutes. If $\rho = -\infty$, $\sigma_E = 0$ which means migrants have no substitution effects on native workers. If $\rho = 0$, $\sigma_E = 1$ which follows Cobb-Douglas function. Lower σ_E means low substitutions between migrants and natives, in some cases they may exist complementarities between these two groups of labour.

We also model aggregate output depends on Constant Elasticity of Substitution (CES) function of high-skilled and low-skilled natives as a Constant Elasticity of Substitution (CES) function in which the high-skilled and low-skilled workers may not be perfectly substituted.

$$L_N = (\alpha_1 L_{N1}^{\eta} + \alpha_2 L_{N2}^{\eta})^{1/\eta} \qquad (3.2)$$

where L_{N1} denotes high-skilled urban natives and L_{N2} denotes low-skilled urban natives. α_1 and α_2 are relative efficiency parameter of high- and low-skilled urban natives. The elasticity of substitution between high- and low-skilled urban natives is $\sigma_S = {}^{1}/_{1-\eta}$.

The marginal product of high-skilled urban natives is:

$$\frac{\partial Y}{\partial L_{N1}} = \frac{\partial Y}{\partial L_N}\frac{\partial L_N}{\partial L_{N1}} = \left(\theta_N L_N^{\rho} + L_M^{\rho}\right)^{1/\rho-1}\theta_N L_N^{\rho-1}\alpha_1 L_N^{1-\eta} L_{N1}^{\eta-1}$$
$$= \theta_N L_N^{\rho-\eta}\Psi\cdot\alpha_1 L_{N1}^{\eta-1} \qquad (3.3)$$

where $\Psi = \left(\theta_N L_N^{\rho} + \theta_M L_M^{\rho}\right)^{1/\rho-1}$. Similarly, the marginal product of low-skilled urban natives is:

$$\frac{\partial Y}{\partial L_{N2}} = \frac{\partial Y}{\partial L_N}\frac{\partial L_N}{\partial L_{N2}} = \theta_N L_N^{\rho-\eta}\Psi\cdot\alpha_2 L_{N2}^{\eta-1} \qquad (3.4)$$

The marginal product of urban-to-rural migrants is:

$$\frac{\partial Y}{\partial L_M} = \theta_M L_M^{\rho-1}\Psi \qquad (3.5)$$

From equation (3.1) to (3.5) and following the profit maximization condition, the relative wage between different skill groups is equal to the relative marginal product:

$$log\frac{w_{N1}}{w_M} = log\frac{\frac{\partial Y}{\partial L_{N1}}}{\frac{\partial Y}{\partial L_M}} = log\frac{\theta_N L_N^{\rho-\eta}\Psi\cdot\alpha_1 L_{N1}^{\eta-1}}{\theta_M L_M^{\rho-1}\Psi}$$
$$= log(\theta_N\alpha_1/\theta_M) + \frac{1}{\sigma_E}logL_M - \frac{1}{\sigma_S}L_{N1} + \left(\frac{1}{\sigma_S} - \frac{1}{\sigma_E}\right)logL_N$$
$$= log\left(\frac{\theta_N\alpha_1}{\theta_M}\right) + \frac{1}{\sigma_E}log\frac{L_M}{L_N} + \frac{1}{\sigma_S}log\frac{L_N}{L_{N1}} \qquad (3.6)$$
$$log\frac{w_{N2}}{w_M} = log\frac{\frac{\partial Y}{\partial L_{N2}}}{\frac{\partial Y}{\partial L_M}} = log\frac{\theta_N L_N^{\rho-\eta}\Psi\cdot\alpha_1 L_{N2}^{\eta-1}}{\theta_M L_M^{\rho-1}\Psi}$$
$$= log(\theta_N\alpha_2/\theta_M) + \frac{1}{\sigma_E}logL_M - \frac{1}{\sigma_S}L_{N2} + \left(\frac{1}{\sigma_S} - \frac{1}{\sigma_E}\right)logL_N$$
$$= log\left(\frac{\theta_N\alpha_2}{\theta_M}\right) + \frac{1}{\sigma_E}log\frac{L_M}{L_N} + \frac{1}{\sigma_S}log\frac{L_N}{L_{N2}} \qquad (3.7)$$

Equation (3.6) shows that the wage gap between high-skilled urban natives and rural-to-urban migrants depends on their relative supply. That is, a higher proportion of rural-to-urban migrants would drive up the wage differential between native workers and migrants. This is the main argument in this chapter. Equations 3.6 and 3.7 demonstrate the possibility of migrants driving up wages of urban natives, which is the complementary effects of migrants on urban native workers. In this case, anything that changes the relative labour supply of those two groups will change the native-migrant wage differential.

As mentioned in section 3.1, there are requirements for *hukou* registration for obtaining jobs from state-owned enterprises. This results in job segregations between these two groups of labour, and low substitutions between migrants and urban natives. Therefore, there might be complementary effects of migrants on urban native workers. Although there is some literature (Card, 2009; Ottaviano and Peri, 2006) find imperfect substitutions between immigrants and natives, this chapter attempts to identify another complementary effect between migrants and natives. Oaxaca-Blinder decomposition will be implemented to examine the high segregations of jobs between migrants and natives, followed complementary effect being identified in the analysis.

3.5 Data Construction and Descriptive statistics

The main data we use for the empirical analysis is 2005 population census survey. The 2005 population provides detailed data on respondents' personal characteristics, such as gender, age, education, working hours weekly, monthly wage. Especially, the data contains information on respondents' current living location and *hukou* registration place, which could be used to infer their migration status. The sample size is about 2.5 million, representing about 0.95‰ of the population. We restrict the sample between age 16 and 64, who have worked for the past week.

3.5.1 Migration Flow Construction

In the census data, one cannot observe the annually migration flows in each prefecture. As mentioned above, the 2005 population census data provides prefecture-level information about place of current residence, *hukou* registration type (agricultural or non-agricultural), *hukou* registration place, the timing of departure (such as 1 year ago, 2 years ago, etc.) for a migrant leaving *hukou* registration place. For example, an individual has *hukou* registration in prefecture A, and now the current residence place is prefecture B. If this respondent left the *hukou* registration place 2 years ago, then this person could be accounted into the immigration flow in prefecture A in 2003.

Similarly, by combining the information about migrants' *hukou* registration place and the timing of departure, the emigration flow could be measured every year between 2001 and 2005. Overall, the yearly net rural-urban migration flows in prefectures are measured by subtracting emigration from immigration flows.

There are some imprecisions in measuring migration flows. The census survey records prefecture-level information on places of current residence, *hukou* registration place, and the timing of departure from *hukou* registration place. The issue here is that there might have been step migration (Imbert et al.,2020). For example, if a migrant left her original *hukou* registration place and relocated in city A in 2002, and then relocate to the city B in 2005, we only observe the last relocation. During the process of migration flow construction, we assume that the migrants directly relocate to the final destinations (places of current residence) after they left their *hukou* registration place. In reality, the migrants could relocate to another prefecture before relocating to the final destination, which cannot be observed without knowing the arrival date of the current residence place. The other issue is that some individuals might leave their *hukou* location but return to the place by 2005. In such cases, one cannot observe return migration.

3.5.2 Measure of Migrant Shares in Prefectures

The National Bureau of Statistics (NBS) implements population census survey every five years. The number of rural-urban migrants in 2000 and 2005 are measured respectively with 2000 and 2005 population census survey. That is, only migrant shares in 2000 and 2005 are directly observed in the census. As the migration flows every year between 2001 and 2005 have been constructed, then the stock of migrants between 2001 and 2004 is computed by combining lagged one-year migrants with net migration flows, which denotes as

$$migrant\ stock_{pt} = migrant\ stock_{pt-1} + migrant\ inflow_{pt-1} - migrant\ outflow_{pt-1} \tag{3.8}$$

where p indexes prefectures, t denotes year from 2001 to 2005.

3.5.3 Hukou Reforms

For the empirical analysis, the key variable is the interaction of *hukou* reforms and migration flows during each year. To collect the *hukou* reforms in each prefecture, we search through two different law databases and reviewed reform documents and regulations: Chinalawinfo and law-lib.com. As the analysis focuses on rural-urban migration, we collect *hukou* reforms which encourage rural residents to migrate. Also, the prefecture-level *hukou* reform information is collected across 2000-2004 as we allow the *hukou* reforms to take effect at least one year after implementation. Following the process of *hukou* reform coding by Kinnan et al. (2018), we search all combinations of the keywords for *hukou* and reform in Chinese. The key word of *hukou* in Chinese is "*hukou*" or "*huji*". The key word of reform in Chinese is "*gaige*" or "*guanli*". we carefully review the titles and contents to identify whether the reforms are issued to encourage rural-urban migration. In the end, the number of prefectures with *hukou* reforms are 60.

3.5.4 Descriptive Statistics

Table 3.2 shows that monthly wage of migrants is lower than urban wage. And the migrants are youngers and less likely to be married. Regarding the education composition, 85% of migrants graduated from middle school and below. However, most (63%) urban workers completed high school and above. 56% of Urban workers are employed by public sector compared to 7.86% employed of rural-urban residents, whereas most migrants (70%) work in the private sector and are self-employed. The distinctions show a high degree of job segregation between rural migrants and urban natives. This also indicates that the rural migrants are unlikely to substitute urban natives due to the job segregation. That is, the distribution of migrants and natives across sectors is imbalanced. In the next section, we examine the factors that influence China's unequally wage structure between rural migrants and urban natives, which reinforce the job segregation assumption between migrants and natives.

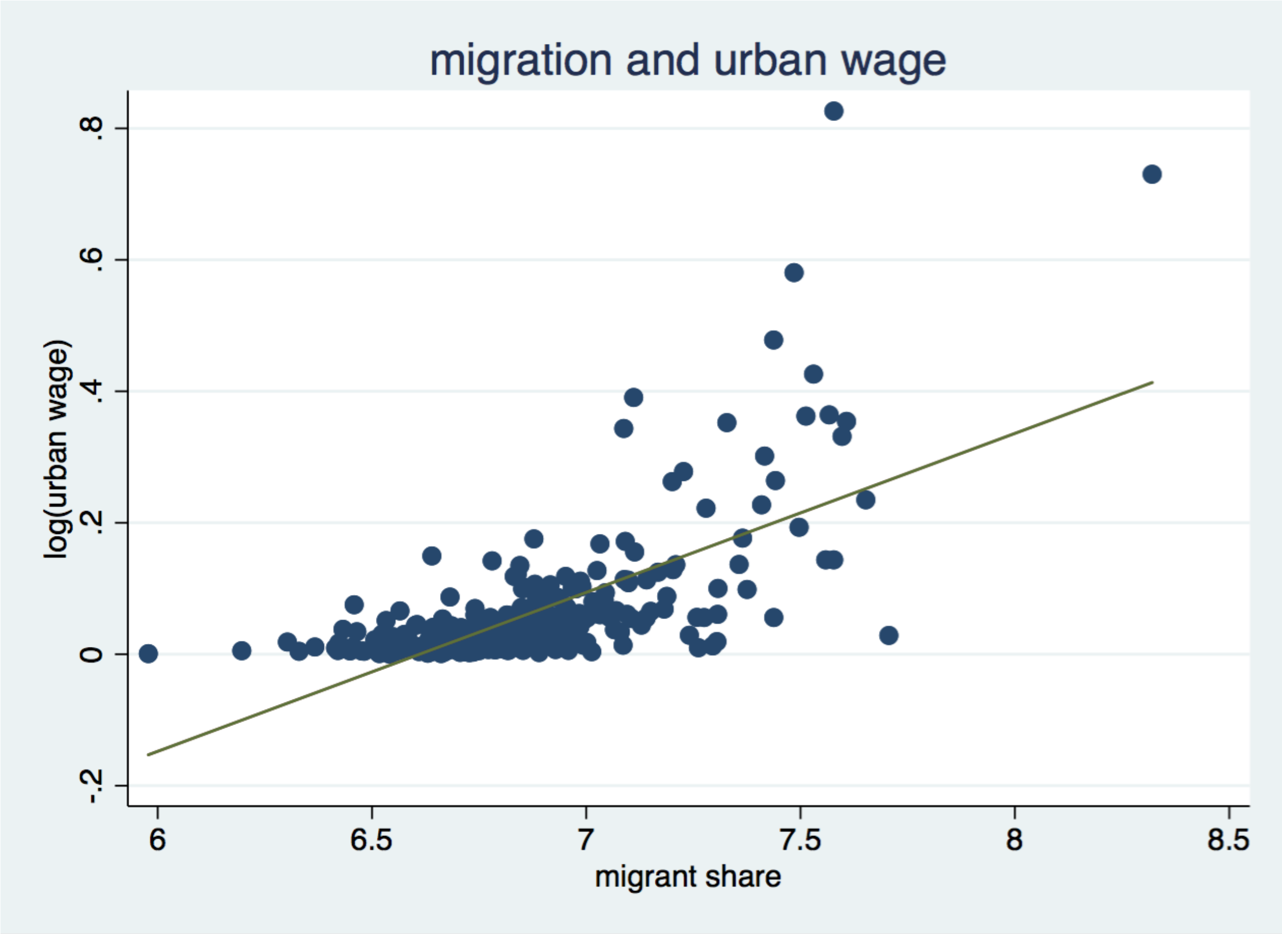
Table 3.2 Summary Statistics for Urban and Rural to Urban Residents [17]

	Urban Residents			Rural-Urban Residents		
	Obs	Mean (or %)	Std.dev	Obs	Mean (or %)	Std.dev
Monthly wage	273636	1053.17	963.09	113072	922.18	851.15
low_skill	98639	763.11	732.36	94762	865.51	768.08
high_skill	174997	1226.86	1039.83	18310	1235.58	1162.89
Male	273636	0.585	0.492	113072	0.556	0.496
Married	273636	0.868	0.338	113072	0.648	0.477
Age	273636	37.96	9.424	113072	30.23	9.429
Weekly working hours	273636	46.27	11.02	113072	55.728	13.65
Education						
Primary school and below	17786	0.065	0.24	26074	0.23	0.42
Middle school	84690	0.309	0.46	69686	0.61	0.47
High school and college	143057	0.522	0.49	17040	0.15	0.36
University	28075	0.102	0.31	279	0.0024	0.05
Sectors						
Public sector	155918	0.56	0.49	8854	0.078	0.27
Private sector	33849	0.12	0.33	40265	0.35	0.48
Self-employed	50458	0.18	0.39	34170	0.30	0.46
Others	33411	0.12	0.33	29783	0.26	0.44

[17] Table 3.2 only shows the statistics for 2005. 2005 Population Census is the only available population census providing wage information.

Figure 3.1 plots the unconditional relationship between migrant shares and urban native wages across cities. Data source is 2005 micro-level population census. As seen in the figure, a larger share of migration is associated with higher wages. The positive relation indicates that migrants tend to drive up the urban native wages. This figure is initial evidence that the data favours a linear positive relationship between migrant shares and urban native wages.

Figure 3.1 Unconditional Correlation between Migration Share and Urban Wage



Notes: The urban native wage and rural-to-urban migrant share are calculated using 2005 Micro-level Population Census. The native wage is denoted monthly wage. Rural-to-urban migrant share is the ratio of rural-to-urban migrants to total population in each cities.

3.6 Oaxaca-Blinder Decomposition

From the descriptive statistics, we see that there is high degree of job segregation between rural migrants and urban natives, which indicates that rural migrants are unlikely to substitute urban natives due to job segregation. In this section, we use Oaxaca-Blinder decomposition methods to decompose the monthly/hourly wages in order to examine which factors influence China's unequal wage structure between rural migrants and urban natives. To consider the determination of the wage gap between rural migrants and urban natives, it can be thought that labour market segregation by sectors (manufacturing, mining, service sectors), occupation and enterprises ownership (state-owned, private enterprises) may affect the wage gap.

For example, if the opportunity of entry into the same sector differs for migrants and urban natives, then the different distribution proportions of sectors may create wage gaps. The Oaxaca-Blinder decomposition approach decomposes the mean wage differential into explained part and unexplained part. The explained part shows the portion of wage gap that can be explained by different endowments between migrants and urban natives. The unexplained part shows the portion of wage gap that arises from different treatment of the same endowments between migrants and urban natives. From the decomposition results, we see a large part of wage gaps between rural migrants and urban natives is attributed to endowment effects, which reinforces the assumption of job segregation.

The assumption of job segregation between migrants and natives conforms to existing empirical evidence. For example, Zhang et al. (2016) shows that education experience, work experience and distribution across occupation, industry and ownership of enterprises account for most of the explained wage gap. Using 2005 population mini census, Li et al. (2015) find that the different entry opportunities into different sectors, occupational attainment affect wage gap between rural migrants and urban natives. Similarly, Demurger et al. (2012) use Oaxaca-Blinder decomposition to illustrate that the differences in endowments, especially the influence of ownership types of segmentation, account for a large part of earnings gaps in the 2000s in China. There is also some existing international empirical evidence that confirms the associated assumption of job segregation. For instance, Aydemir and Skuterud (2008) identify the relative importance of immigrant wage differential within and across establishments in

Canada and their findings show that the sorting of immigrants across establishments affects the wage differential more than the pay differential within establishments. De Matos (2012) illustrates that more than half of the wage gap between natives and immigrants is due to differences in immigrant sorting into different regions, sectors and occupations.

3.6.1 General Decomposition Steps

From the descriptive statistics, no clear picture emerges about what factors influence the wage gap. To answer this question, we use Oaxaca-Blinder decomposition methods to decompose the monthly wage/ hourly wage.

To conduct the decomposition, a general human capital earning equation is specified by Mincer (1958) as:

$$lnw_i^g = \alpha + \beta_1 Education_i + \beta_2 Experience_i + \beta_3 Experience_i^2 + \beta_j Other_i + u_i \quad ,(3.9)$$

where g indicates the group (native urban residents or rural-urban migrants); $Education_i$ is individual's education attainment; $Experience_i$ is proxy of working experience, which is identified as $age - Years\ of\ schooling - 6$; $Other_i$ is a vector of other individual characteristics such as marital status, gender, sector category, enterprises ownership.

After running OLS, the estimated wage equation for each group is:

$$\widehat{lnw^U} = \widehat{\alpha^U} + \widehat{\beta^U} X^U \quad (3.10)$$

$$\widehat{lnw^M} = \widehat{\alpha^M} + \widehat{\beta^M} X^M, \quad (3.11)$$

where hats indicate estimated constant and coefficients, U represents urban natives, M represents rural-urban migrants.

The differential equation of mean log wage between two groups is:

$$\overline{lnw^U} - \overline{lnw^M} = \left(\widehat{\alpha^U} - \widehat{\alpha^M}\right) + \widehat{\beta^U}\overline{X^U} - \widehat{\beta^M}\overline{X^M} \quad (3.12)$$

Assume β^* is non-discriminatory coefficient, where the coefficient $\widehat{\beta^U}$ and $\widehat{\beta^M}$ are equal in a hypothetical scenario.
Then the differential equation can be rewritten as:

$$\overline{lnw^U} - \overline{lnw^M} = \left(\widehat{\alpha^U} - \widehat{\alpha^M}\right) + \left(\overline{X^U} - \overline{X^M}\right)\beta^* + \left(\widehat{\beta^U} - \beta^*\right)\overline{X^U} + (\beta^* - \widehat{\beta^M})\,\overline{X^M}\ , \quad (3.13)$$

where $\left(\overline{X^U} - \overline{X^M}\right)\beta^*$ is referred to as the explained portion of the wage gap between two groups; $[\left(\widehat{\beta^U} - \beta^*\right)\overline{X^U} + (\beta^* - \widehat{\beta^M})\,\overline{X^M}]$ represents the unexplained portion of the wage gap; $\left(\widehat{\alpha^U} - \widehat{\alpha^M}\right)$ shows the effect of omitted variables on the wage gap.

To specify a suitable β^*, we try four methods. The first is to assume the coefficient of urban groups represents the non-discriminatory wage structure. The wage differential equation would be re-written as:

$$\overline{lnw^U} - \overline{lnw^M} = \left(\widehat{\alpha^U} - \widehat{\alpha^M}\right) + \left(\overline{X^U} - \overline{X^M}\right)\beta^U + (\widehat{\beta^U} - \widehat{\beta^M})\,\overline{X^M} \quad (3.14)$$

The second method is to assume that the discrimination is directed against urban resident groups and there is no negative discrimination of rural-urban migrants. Then the coefficient of rural-urban migrants is the non-discriminatory wage structure. The decomposition is:

$$\overline{lnw^U} - \overline{lnw^M} = \left(\widehat{\alpha^U} - \widehat{\alpha^M}\right) + \left(\overline{X^U} - \overline{X^M}\right)\beta^M + \left(\widehat{\beta^U} - \widehat{\beta^M}\right)\overline{X^U}. \quad (3.15)$$

However, using coefficients from one of the groups may cause undervaluation or overvaluation of discrimination. Reimer (1983), Cotton (1988), and Neumark (1988) argued that the non-discriminatory wage structure should be computed using a pooled sample of both groups. The OLS regression equation can be written as: $ln\widehat{w} = \hat{a} + \hat{\beta}X$, X is the vector of characteristics from pooled sample over both of groups. Therefore, $\hat{\beta}$ is referred as omega coefficients: $\hat{\beta} = \beta^*$.

However, Fortin (2006) and Jann (2008) point out that omega coefficients may transfer some of unexplained part to the explained components. That is, the group-specific intercepts might be different due to discrimination. Therefore, the pooled OLS regression can be written as:

$$ln\widehat{w}' = \hat{a}' + \gamma hukou + \hat{\beta}'X \quad (3.16)$$

where $hukou$ is a dummy variable of group indicator, which takes value of 1 if the individual is a rural-urban migrant. Hereafter, we label $\hat{\beta}'$ as the pooled coefficients.

3.6.2 Decomposition Results

Table 3.3 shows the decomposition results for logarithm wage. Column (1)-(4) show the decomposition results for logarithm monthly wage and column (5)-(8) show the results for logarithm hourly wage. Columns (1) and (5) present the results of decomposition exercise based on equation 3.14. Columns (2) and (6) present the decomposition results based on equation 3.15. The decomposition results in columns (3) and (7) are based on the non-discriminatory coefficient using a pooled sample of both urban and migrant groups. By comparison, the decomposition exercise in columns (4) and (8) use the coefficient from a pooled sample of both urban and migrant groups taking into consideration the group-specific intercepts. That is said, equation 3.16. The total estimated log urban-migrants monthly wage differential is 0.059 yuan. The mean wage gap is attributed by different endowments and different returns of the endowments. The explained contribution to the logarithm monthly wage gap using the urban coefficients is 242%, while the unexplained components reduce the wage gap by 142%. Anything that favours migrants has negative signs.

The decomposition results using urban coefficients, migrant coefficient and pooled coefficients are closest to each other. However, using omega coefficients underestimate the endowment effects compared to the results using other coefficients. As migrants tend to work longer time than urban natives, hourly wages may reflect more accurate decomposition results. The total estimated logged urban-migrants hourly wage differential is 0.24 yuan. The detailed decomposition shows that a large part of wage gap is attributed to endowment effects. The endowment effects contribute to 105% of total estimated log urban-migrant hourly wage differentials in the pooled method. The large part of the endowment effects can be attributed to education and work experience. For instance, The differences in education accounts for more than

90% of the explained wage differential. Of the 0.24 explained effect in pooled estimation, 30% of the wage differential is attributed to different work experience. Differences in employment distributions across sector, occupation and ownership, explain 62% of wage gap between urban and rural migrant workers. The coefficient effects refer to the unexplained effect which could be attributed to the differential returns to the endowments. Using the pooled method, the decomposition results show that migrant workers are favoured in terms of being a male and an ethnic majority. That is, employers offer higher wages to males or ethnic majorities compared to females or ethnic minorities. The province variable denotes the region of residence. The decomposition results indicate that urban workers receive higher returns of endowment compared to rural migrant workers in the same region of residence. There are a variety of local dialects in China. According to Lee (2012), the 2005 China Urban Labour Survey shows that only half of the migrants who have been in the city 6 years can speak the local dialect. This may be a language barrier for migrants to overcome. Urban employers consider migrants as inferior outsiders in a manner. Therefore, migrants appear to face discrimination.

In terms of log monthly wage differential, of the 0.15 endowment, the contribution of education and experience is 186% in the pooled method. Different endowments of occupation distribution, sector distribution and ownership distribution explained another 63% of explained wage differential. Similarly to the log hourly wage differential, the return of the endowment is generally higher for male than female workers. Also, the unexplained portion of the log monthly wage between urban and migrant workers is lower for an ethnic majority than an ethnic minority. The positive coefficient on province indicates that the language barrier is one of the sources of unexplained portion. Overall, the decomposition results show that the main sources of wage differential between rural migrants and urban natives are productivity-linked characteristics and differences in employment distributions across sector, ownership and occupation. Overall, the findings support the assumption of job segregation between rural-urban migrants and urban natives.

Table 3.3 Decomposition Results of Wage Gap between Urban Natives and Rural-Urban Migrants

	log (monthly wage)				*log* (hourly wage)			
	Urban Coef. (1)	Migrant Coef (2)	Omega Coef. (3)	Pooled Coef. (4)	Urban Coef. (5)	Migrant Coef. (6)	Omega Coef. (7)	Pooled Coef. (8)
Urban Residents	6.737***	6.737***	6.737***	6.737***	1.546***	1.546***	1.546***	1.546***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Rural-Urban Migrants	6.679***	6.679***	6.679***	6.679***	1.305***	1.305***	1.305***	1.305***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Total Effect	0.059***	0.059***	0.059***	0.059***	0.240***	0.240***	0.240***	0.240***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Endowment Effect	0.143***	0.153***	0.111***	0.155***	0.226***	0.206***	0.247***	0.252***
	(0.003)	(0.008)	(0.002)	(0.002)	(0.003)	(0.009)	(0.002)	(0.002)
Coefficients Effect	-0.084***	-0.094***	-0.052***	-0.096***	0.014***	0.034***	-0.006***	-0.012***
	(0.003)	(0.008)	(0.002)	(0.003)	(0.003)	(0.009)	(0.002)	(0.003)
Endowment Effect								
Education	0.207***	0.195***	0.193***	0.208***	0.231***	0.199***	0.2339***	0.235***
	(0.002)	(0.007)	(0.002)	(0.002)	(0.002)	(0.007)	(0.002)	(0.002)
Experience	0.092***	0.124***	0.076***	0.080***	0.087***	0.117***	0.076***	0.077***
	(0.002)	(0.006)	(0.002)	(0.002)	(0.003)	(0.006)	(0.002)	(0.002)
Experience_squr	-0.061***	-0.069***	-0.055***	-0.054***	-0.052***	-0.053***	-0.047***	-0.047***
	(0.002)	(0.004)	(0.002)	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)
Gender	0.005***	0.004***	0.005***	0.005***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Marriage status	0.007***	0.009***	0.003***	0.004***	0.005***	0.005**	0.003***	0.003***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Ethnicity	0.00006	0.00001	0.00008	0.00008	0.00005	-0.00003	0.00007	0.00008
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Province	-0.203***	-0.266***	-0.197***	-0.185***	-0.201***	-0.286***	-0.178***	-0.176***
	(0.002)	(0.004)	(0.001)	(0.001)	(0.002)	(0.004)	(0.001)	(0.001)
Occupation	0.051***	0.049***	0.051***	0.053***	0.058***	0.051***	0.062***	0.063***
	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)

Sector	0.018***	0.075***	0.014***	0.016***	0.026***	0.076***	0.027***	0.027***
	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)	(0.006)	(0.001)	(0.001)
Ownership	0.028***	0.031***	0.021***	0.029***	0.068***	0.094***	0.066***	0.067***
	(0.002)	(0.004)	(0.001)	(0.001)	(0.002)	(0.004)	(0.001)	(0.001)
Coefficients Effect								
Education	-0.028	-0.016	-0.014	-0.030	0.023	0.055	0.021	0.019
	(0.039)	(0.042)	(0.039)	(0.039)	(0.046)	(0.050)	(0.046)	(0.046)
Experience	0.105***	0.073***	0.121***	0.116***	0.095***	0.066***	0.106***	0.106***
	(0.013)	(0.009)	(0.014)	(0.014)	(0.014)	(0.010)	(0.015)	(0.015)
Experience_squr	-0.015**	-0.008**	-0.021***	-0.022***	-0.001	-0.00040	-0.006	-0.006
	(0.006)	(0.003)	(0.007)	(0.007)	(0.007)	(0.004)	(0.008)	(0.008)
Gender	-0.007***	-0.006***	-0.007***	-0.007***	-0.007***	-0.007***	-0.008***	-0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Marriage status	0.005	0.004	0.010**	0.009**	-0.00027	-0.00018	0.002	0.002
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.005)	(0.005)
Ethnicity	-0.029***	-0.029***	-0.029***	-0.029***	-0.046***	-0.046***	-0.046***	-0.046***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Province	0.038***	0.1013***	0.033***	0.020***	0.057***	0.14243***	0.034***	0.032***
	(0.004)	(0.006)	(0.003)	(0.003)	(0.004)	(0.006)	(0.003)	(0.003)
Occupation	0.027***	0.029***	0.027***	0.025***	0.037***	0.044***	0.033***	0.032***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.008)	(0.005)	(0.005)
Sector	-0.072***	-0.128***	-0.068***	-0.070***	-0.053***	-0.103***	-0.054***	-0.054***
	(0.017)	(0.019)	(0.017)	(0.017)	(0.017)	(0.019)	(0.017)	(0.017)
Ownership	0.002*	-0.002	0.9***	0.001	-0.007***	-0.033***	-0.004***	-0.005***
	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)
Observations	386,708	386,708	386,708	386,708	386,708	386,708	386,708	386,708

Notes: Column (1)-(4) shows decomposition results using *logged* monthly wage, column (5)-(8) shows decomposition results using *logged* hourly wage. Education is a categorical variable denoting 1, 2..7 with illiteracy, primary school, secondary school, high school, college, bachelor, and master or above. Experience equals $age - Years\ of\ schooling - 6$. Marriage status is a dummy variable which denotes 1 if the individual is married. Gender is a dummy variable which denotes 1 if the individual is a male. Ethnicity is a dummy variable which denotes 1 if the individual's ethnicity is Han. Occupation is a categorical variable denoting "State Institutions Head", "Professional and Technical Personnel", "Clerks and Related Workers", "Business Service Personnel", "Agriculture and Water Conservancy Labours", and "Production, Transport Equipment Operators and Related Workers". Sector is a categorical variable denoting 20 sectors in China such as "Agriculture", "Mining", "Manufacturing", "Construction", etc. Enterprises ownership is a categorical variable with "State-owner Enterprises", "Private-owned Enterprises", "Self-employed Individuals", "Other". Standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

3.7 Identification Strategy

3.7.1 Estimation Specification

We will explore the causal effect of rural migrants on urban native wages. The basic estimation equation as follows:

$$\ln{(Wage)_{ip}} = \varphi + \mu \ln{mig\_share_p} + \tau X_i + \delta Z_p + \xi_p + \varepsilon_{ic} \quad (3.17)$$

where i indexes individual and p indexes prefectures. $\ln{(Wage)_{ip}}$ is logarithm native individual wage in city p, $migrant\_share_p$ is defined as the ratio of rural-urban migrants to urban natives in city p, X_i is individual control variables, including age, education and gender, ethnicity and marriage status ; Z_p denotes the city characteristics, such as GDP growth, urbanization and average schooling; ξ_p is a province fixed effect, which captures time-invariant heterogeneity across provinces.

However, there are two threats when identifying the causal relationship between migrant share and urban native wages. One concern is potential reverse causality between migrants and urban native wages. Internal migrants could be attracted by higher wages in destinations and wages in destinations could also be affected by immigrants. The other concern is omitted variable bias. Some factors could affect both wages and migrants. To address the endogeneity issue, an instrumental variable approach will be applied. As we define the migrant share as the ratio of migrants to urban natives, then the building block of computing migrant shares relies on the migrant stock. In this case, we try to construct instrumental variable for migrant stock. We construct a DID style instrument following Sequeira et al. (2020).

3.7.2 DID Style Instrument

We first have a zero-stage estimation:

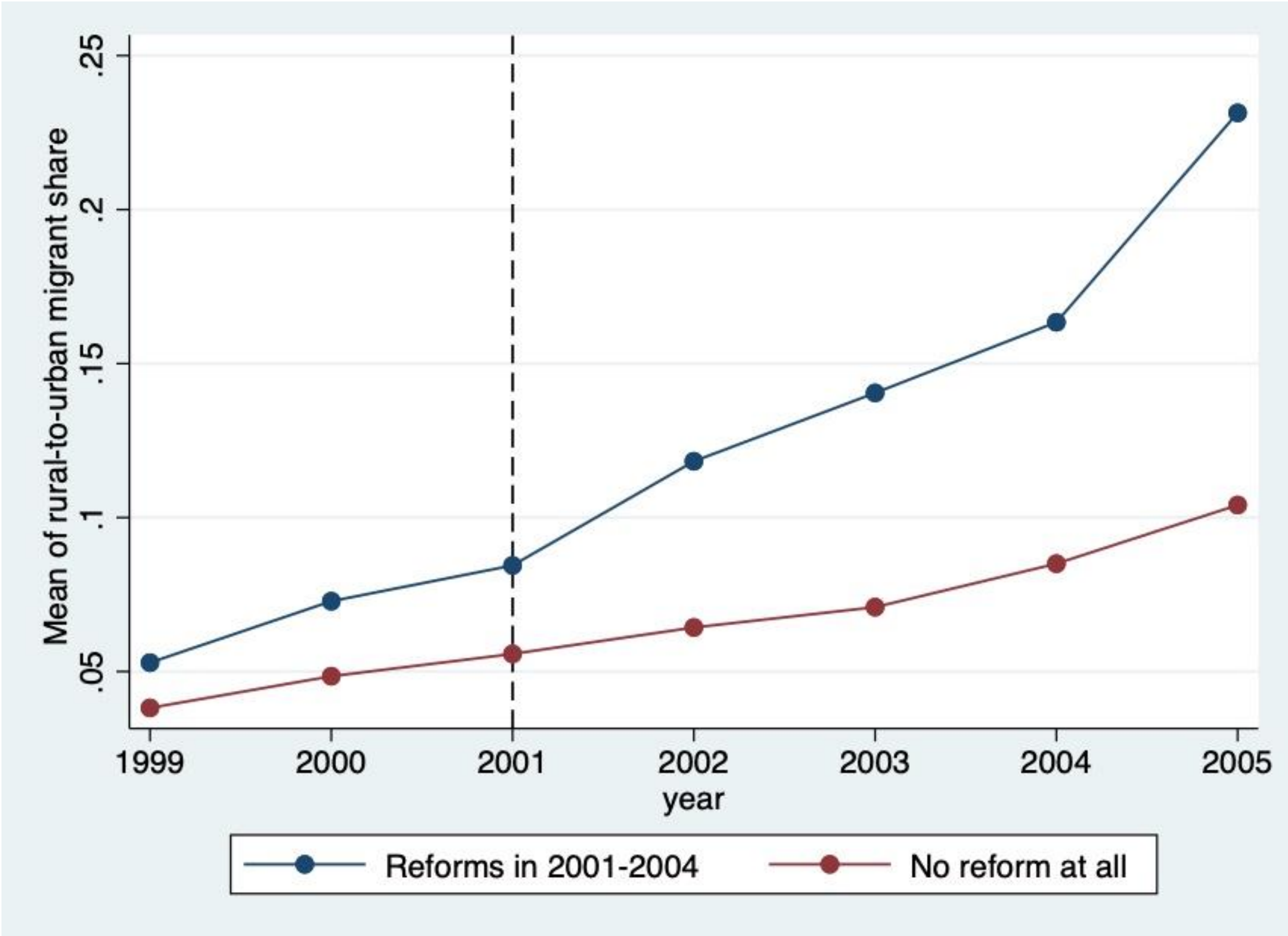
$$\Delta \log(migrants)_{pt} = \alpha_t + \gamma \, \Delta \log(migrants)_{pt-1} + \delta \, policy_{pt-1} +$$
$$\beta \, \Delta \log(migrants)_{pt-1} * policy_{pt-1} + \lambda X_{pt-1} + \varepsilon_{pt-1} \quad (3.18)$$

where p indexes prefecture and t indexes years (2001, 2002, 2003, 2004, 2005). The dependent variable is the first difference of the logarithm of the migrant stock in year t for each prefecture, which denotes the growth rate of migrant stock in each year.

$\Delta\log(migrants)_{pt-1}$ denotes one-year lag of the dependent variable. $policy_{pt-1}$ is an indicator variable which denotes 1 with *hukou* reform in year $t - 1$. In the zero-stage analysis, we allow for the *hukou* reform to take effect at least after one year implementation. The vector X_{pt-1} includes GDP growth, urbanization rate and their interaction terms with the *hukou* reform variable. These controls capture any potential effects of *hukou* reform implementation on the arrivals of migrants through growth in GDP and urbanization development. The identification strategy exploits the interaction of time-varying migration flows and *hukou* reforms across prefectures.

The key idea behind this identification exploits two facts. One is that cities with *hukou* reform are expected to experience higher growth in the stock of migrants from other cities. With the *hukou* restrictions relaxation, the probability of obtaining local *hukou* is higher and rural workers are more likely to move to urban areas, which increases rural-urban labour supply in urban labour markets. Figure 3.2 shows the average migrant shares between reform and no-reform cities across years of 1999 and 2005. The reform cities refer to those which implemented *hukou* reforms between 2001 and 2004. It is evident that the trends of migrant shares between reform and no-reform cities are similar until 2001 when cities started to adopt *hukou* reforms. Afterwards, the migrant shares are significantly higher in reform cities compared to non-reform cities, with the differences increasing over the years. As such, the implementations of *hukou* reforms in some but not all the cities offer the basis for the construction of the DID style identification strategy that enables us to predict the migrant shares. The other fact is that the tendency of newly arriving migrants normally follows the previous immigrants' settlement pattern (Card, 2001). As suggested by Card (2001), new migrant flows are positively correlated with the earlier immigrants. Another logic to understand the intuition behind the construction of the instrumental variable is that there are a number of cities in year t that have zero *hukou* reform, but still experience changes of migrant stocks. This weakens the explanatory power of the *hukou* reform on changes of migrant stocks. Using two sources of variation, we could strengthen the effect of *hukou* reform on the changes of migrant stocks. Conceptually, the construction of the instrumental variable is similar to a difference-in-differences estimation strategy, where the zero-stage estimates compare changes of migrant stocks in cities that had *hukou* reform implemented to those that had no *hukou* reform implemented, taking into consideration varying migration patterns. If the identification strategy is valid, we should expect β to be positive.

Figure 3.2 Migrant Share between Reform and Non-reform Cities: 1999-2005



Notes: Non-reform at all denotes those cities without any *hukou* reform during 2001 and 2004. Reforms in 2001-2004 denotes those cities implementing *hukou* reforms between 2001 and 2004. Rural-to-urban migrant share is calculated by the ratio of migrant to total city population (aged 16 and 64). City-level average migrant shares are from various issues of the China City Statistical Yearbook.

As suggested by Sequeira et al. (2020), two additional interaction terms, the interaction of GDP growth rate and the *hukou* reforms and the interaction of urbanisation rate and the *hukou* reforms, are included in the zero-stage specification. The potential concern here is that the changes of migrant stocks could be positively correlated with economic growth or urbanization development level. Omitting urbanization or GDP growth might cause estimates to be upward biased. These two additional interaction terms capture the differential effect of *hukou* reforms on migration depending on urbanization levels and economic growth.

After estimating equation (3.18), we could predict the changes of migrant stocks $\Delta \widehat{migrant}_{pt}$ over time in each prefecture, which is defined as:

$$\Delta \log(\widehat{migrants})_{pt} = \widehat{\alpha_t} + \hat{\gamma}\ \Delta \log(migrants)_{pt-1} + \hat{\delta}\ policy_{pt-1} + \hat{\beta}\ \Delta \log(migrants)_{pt-1} * policy_{pt-1} + \hat{\lambda} X_{pt-1} \quad (3.19)$$

Next, we compute the predicted stock of migrants in city p for year t as following:

$$\widehat{migrant}_{p,t} = \widehat{migrant}_{p,2001} * \left(1 + \Delta \log(\widehat{migrants})_{pt}\right)^{T-1} \quad (3.20)$$

where $T \leq 5$.

The 2SLS equations are given by equation (3.21) and (3.22), where equation (3.21) is the first stage and equation (3.22) is the second stage.

$$\log(\widehat{migrant\_share})_p = \varphi + \mu log(migrant\_share)_p + \tau X_i + \delta Z_p + \xi_c + \varepsilon_{ip} \quad (3.21)$$

$$\ln(Wage)_{ip} = \varphi + \mu \log(\widehat{migrant\_share})_p + \tau X_i + \delta Z_p + \xi_c + \varepsilon_{ip} \quad (3.22),$$

where migrant share=migrant stock/urban population.

3.7.3 Shift-Share Approach

The DID style instrument could be criticised because the exclusion restrictions could be violated. It is possible that the migrant flows could be endogenous to the implementations of *hukou* reforms. A potential concern is that migrant inflows could be greater when the *hukou* reforms have been implemented with future easier access to local *hukou*. We therefore supplement the DID style instrument with a shift-share instrument.

Following Imbert et al. (2020) and Kinnan et al. (2018), we also use an alternative approach for identification to construct exogenous migrant shares - shift-share instrument.

$$Z_d = \sum_{u \le t} \sum_o migrant\_share_{o \to d}\, h_{du} \quad (3.23)$$

$$migrant\_share_{o \to d} = \frac{migrant\_stock_{o \to d}}{\sum_d migrant\_stock_{o \to d}} \quad (3.24)$$

where $migrant\_stock_{o \to d}$ is the migrant share in year 2000 from origin o to destination d. h_{du} is an indicator variable for *hukou* reform in destination d at year u. The intuition is that a reform in prefecture d has a larger effect on the migration in prefecture o to migrate to d if there was a high past migration pattern moving from o to d. For instance, if there is only one reform in 2005 in d, then the exogenous migrant share would be the summation of the quantity of migrant share in 2000 from each origin prefecture o. Intuitively, the shift-share instrument is the cumulative weighted sum of all *hukou* reforms that have occurred in destinations up until 2005. The weights are the migrant share in origin o to destination d.

3.8 Regression Results

3.8.1 First-stage Results

Table 3.4 reports first stage estimates using DID style and shift-share approach. The DID-style instrument variable is computed using equations 3.18-3.20. The shift-share instrument is computed following the procedures in equations 3.23 and 3.24. Column (1) and (2) report stage-zero and first-stage estimates of DID style instruments. In particular, stage-zero estimates refer to the equation 3.18. The instrument used in the first-stage estimates is computed following equations 3.19 and 3.20. Individual controls include education attainment, work experience, experience squared, marriage status, gender, and ethnicity. City control variables include urbanisation rate, GDP growth rate, average schooling. Year fixed effects are included in all estimations to control common shocks across all cities that affect the migrant share. Province fixed effects are also controlled. The zero stage is used to predict migration growth. The zero-stage estimates show that lagged migration growth significantly and positively affects the changes in the migrant stocks. Also, cities with *hukou* reforms exhibit significantly higher changes in migrant stocks than cities without *hukou* reforms. We successfully identify 60 cities that implemented the *hukou* reforms by 2005, and there could have staggered reforms in cities. The variations induced by *hukou* reform could provide a strong prediction for the first difference of logarithm migrant stocks. The differential effect of *hukou* reforms on migration through urbanisation levels and economic growth is also controlled in stage-zero estimation. Using the predicted migration growth, we calculate the imputed migration rate in each prefecture in each year. Column 2 shows the first stage regressions using the predicted migration rate from equation 20. The results in column 2 report the significance of the instrument on the endogenous migration rate. The coefficient on the instrument is positive and very significant. That is, the predicted migrant share instrument is strongly correlated with actual migrant share. The Cragg-Donald F statistics of the instrument is above 300, which indicates that the instrument is powerful. Column 3 reports the first stage estimations using shift-share instrument variable. Similar to column 2, the coefficient on the shift-share instrument is positive and very significant, with a Cragg-Donald F statistic of 118.

Table 3.4 First-stage Estimation Results

	DID style IV		Shift-share IV
	Zero-stage	First-stage	First-stage
	Δlog(migrant)	Log(share)	Log(share)
ΔLag log migrant	1.858***		
	(0.208)		
Hukou indicator	0.283**		
	(0.126)		
Hukou indicator*ΔLag log migrant	0.943**		
	(0.393)		
GDP growth	0.902		
	(0.601)		
*Hukou**GDP growth	-2.015**		
	(0.884)		
Urbanisation rate	0.172**		
	(0.081)		
*Hukou**Urbanisation rate	-0.140		
	(0.165)		
Predicted share		0.684***	0.328***
		(0.00120)	(0.00953)
Individual Control		Yes	Yes
City Control		Yes	Yes
Province FE		Yes	Yes
Year FE	Yes		
Endog test p-value		0.000	0.000
Observations	540	144,711	78,641
Adjusted R-squared	0.427		

Notes: Column (1) shows zero-stage regression results. The dependent variable in zero-stage regression is first difference of logarithm migrant stock prefecture p in year t. *Hukou* is an indicator variable which denotes 1 if there was a *hukou* reform in prefecture

p in year t-1. Column (2) and (3) show first stage OLS estimates. Individual control variables include education attainment, work experience, experience squared, marriage status, gender, and ethnicity. City control variables include urbanisation rate, GDP growth rate, average schooling. Education attainment is a categorical variable with illiteracy, primary school, secondary school, high school, college, bachelor, and master or above. Work experience is identified as $age - Years\ of\ schooling - 6$. Marriage status is a dummy variable which denotes 1 if the individual is married. Gender is a dummy variable which denotes 1 if the individual is a male. Ethnicity is a dummy variable which denotes 1 if the individual's ethnicity is Han. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

3.8.2 2SLS Estimates

Table 3.5 report the OLS and 2SLS estimates of examining the effect of rural-urban migration on urban native wages. Columns (1)-(2) report OLS estimates following estimation equation 3.17. For comparison, columns (3)-(6) report 2SLS estimates. Columns (3) and (4) show the results following the estimation equations 3.21 and 3.22. Columns (5) and (6) show the results following the estimation equations 3.23 and 3.24. The dependent variables are individuals' monthly wage and hourly wage. Both of them are logged. As the migrants normally work longer hours than urban natives, hourly wage may be a better measurement of returns of migration.

The results show that the rural-urban migrant shares have positive rather than depressing effect on urban native wages, which is consistent with the results of Combes et al. (2015), Zhao (2019). As the migrant share is in logarithmic form, the coefficient denotes the elasticity of native wages to the proportion of migrants. The DID style IV estimate coefficient is 0.074, which is slightly higher than the OLS estimate result of 0.06. One explanation for this difference is that 2SLS estimate is causal and OLS estimate are not, with the negative relationship between unobservable factors such as demand for skill and proportion of migrants (Han and Li, 2017). The demand for skills is negatively correlated to proportion of migrants while positively correlated to wages of native workers. That is, some firms to areas with low skilled migrants to some extent raise the demand of higher paid/ higher skilled workers, therefore, offering higher wages of urban natives. Also, the high skilled service and professional jobs may be in higher demand in areas with larger number of migrants. Another potential explanation is that the construction of instrumental variables may violate the exclusion restrictions. It is possible that the migrant flows could be endogenous to the implementations of *hukou* reforms. The migrants tended to move to the cities implementing *hukou* reforms. To identify this potential concern, we try another identification strategy via a shift-share instrument to construct exogenous migrant shares. Column (3) shows the results with shift-share instruments. We find that the estimates of shift-share instrument are actually larger in magnitude, which could be partially taken as evidence against the violation of exclusion restrictions in 2SLS estimates.

Table 3.5 OLS and 2SLS Estimates of the Effect of Migrants on Native Wages

	OLS		DID style IV		Shift-share IV	
	(1)	(2)	(3)	(4)	(5)	(6)
	log(wage)	*log*(hourly wage)	*log*(wage)	*log*(hourly wage)	*log*(wage)	*log*(hourly wage)
Log(migrant share)	0.060***	0.057***	0.074***	0.070***	0.165***	0.180***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.014)	(0.015)
Individual Control	Yes	Yes	Yes	Yes	Yes	Yes
City Control	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	147,533	147,533	144,711	144,711	78,641	78,641
Adjusted R-squared	0.392	0.408	0.391	0.406	0.365	0.373

Notes: The dependent variables are logarithm monthly wage and logarithm hourly wage. Individual control variables include education attainment, work experience, experience squared, marriage status, gender, and ethnicity. City control variables include urbanisation rate, GDP growth rate, average schooling. Education attainment is a categorical variable with primary school, secondary school, high school, college, bachelor, and master or above. Work experience is identified as $age - Years\ of\ schooling - 6$. Marriage status is a dummy variable which denotes 1 if the individual is married. Gender is a dummy variable which denotes 1 if the individual is a male. Ethnicity is a dummy variable which denotes 1 if the individual's ethnicity is Han. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

3.8.3 Heterogeneity Issues

Table 3.2 shows that 85% migrants graduated from middle school or below. However, more than half (63%) of urban workers completed high school or above. Since rural migrants are less educated, they are more likely to compete with low-skilled urban workers (Meng and Zhang, 2010). Also, from the Oaxaca-Blinder decomposition, we see that education attainment is an important factor influencing China's wage structure between urban natives and rural migrants. These reasons guide us to investigate the different effects of rural migrants on urban natives across different education distributions. In this chapter, we define urban natives with secondary school attainment or below as low-skilled urban natives and those with high school attainment or above as high-skilled urban natives. As shown in the summary statistics (table 3.2), over 50% of rural migrants are low-skilled workers, and their urban counterparts is likely to be affected. Table 3.6 tells the impact of rural migrants on low- and high-skilled urban native wages. Columns (1) and (2) are OLS estimations on low-skilled and high-skilled urban native hourly wages respectively. Columns (3) - (6) report 2SLS estimates. The results show that the magnitudes of the OLS and DID style instrument estimates are similar. The shift-share IV estimates are larger in magnitude compared to the estimates of DID style instruments. Here we focus on IV estimates.

According to Zhao (2020), urban workers with similar education levels are the ones most affected by the rural migrants. Although slightly smaller, the impact of rural migrants on low-skilled urban workers is also significantly positive. This indicates that the rural migrants may complement low-skilled urban workers. That is, migrants are unlikely to replace urban workers with similar skill levels. Column (4) shows that the elasticity of migrant on high-skilled urban native wages is about 0.071 with DID style instruments, which is slightly higher than the elasticity of low-skilled urban native wages reported in column (3). Also, using shift-share instrumental variables, we have similar findings that the impact of rural migrants on urban high-skilled workers' wages is stronger than the impact on urban low-skilled workers' wages, which is consistent with the results of Zhao (2020). This indicates that migrant workers might have a more complementary effect on high-skilled urban workers than low-skilled workers.

Table 3.6 Heterogeneous Effect of Migration

	OLS		DID style IV		Shift-share IV	
	Low-skilled	High-skilled	Low-skilled	High-skilled	Low-skilled	High-skilled
	Log(hourly wage)	Log(hourly wage)	Log(hourly wage)	Log(hourly wage)	Log(hourly wage)	Log(hourly wage)
Log(migrant share)	0.057***	0.057***	0.067***	0.071***	0.155***	0.194***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.032)	(0.017)
Individual Control	Yes	Yes	Yes	Yes	Yes	Yes
City Control	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	48,971	98,562	47,905	96,806	23,290	55,351
Adjusted R-squared	0.206	0.376	0.203	0.375	0.146	0.348

Notes: The dependent variables are logarithm hourly wage. we define urban natives with secondary school attainment or blow as low-skilled urban natives. Individual control variables include education attainment, work experience, experience squared, marriage status, gender, and ethnicity. City control variables include urbanisation rate, GDP growth rate, average schooling. Education attainment is a categorical variable with primary school, secondary school, high school, college, bachelor, and master or above. Work experience is identified as $age - Years\ of\ schooling - 6$. Marriage status is a dummy variable which denotes 1 if the individual is married. Gender is a dummy variable which denotes 1 if the individual is a male. Ethnicity is a dummy variable which denotes 1 if the individual's ethnicity is Han. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

3.9 Conclusions

This study examines the effect of rural-to-urban migrants on urban labour markets by using micro-level 2005 and 2000 Population Census data. Two questions are addressed. The first is whether internal migration affects urban natives' wages and the second is whether this relationship is affected by *hukou* policies and *hukou* reforms.

The analysis in this chapter is in three steps. The first step is to examine which factors influence China's unequal wage structure between rural migrants and urban natives using an Oaxaca-Blinder decomposition method. The results indicate job segregation between migrants and natives which underlines the complementarity assumption between migrants and natives. Secondly, following Sequeira et al. (2020), we construct an instrument variable by interacting time-varying pre-reform migration flows and *hukou* reforms across prefectures to causally identify the effect of migrants on urban native workers. Thirdly, as a robustness check, a shift-share instrument is constructed where "the shift" is *hukou* reforms and "the share" is historical migration patterns between cities.

The results show that, on average, rural-urban migrants have positive effects on urban native workers. One percentage point increase in the migrant shares results in 0.07 percentage increase in individual's monthly wage. We also find that the effects are larger for high-skilled urban natives compared to low-skilled urban natives. The estimates are similar using a shift-share instrument, but the magnitudes are larger.

These findings offer important policy implications. The rural-to-urban migrants mostly concentrate on low-skilled labour-intensive industries and their effect spills over to other industries, in the end, contributing to overall productivity growth. Though there are fewer mobility restrictions, the Chinese labour market is still far from being a flexible labour market. The positive impact of rural-urban migrants on urban labour markets imply that further labour mobility relaxation may be necessary and higher productivity in urban labour market could be achieved.

Appendix B

Table B3.1 City-level *Hukou* Reforms: 2000-2004

City	Reform year	Description	Document name	Issue date
Beijing	2000	A rural-urban migrant can get urban *hukou* in pilot satelites cities of Beijing if she has a stable job and long-term residence place.	Jingzhengfa[2000]No.17	April 29,2000
	2001	A migrant can get local *hukou* if she gets married to a local resident in Beijing.	The notice of the Municipal Public Security Bureau on solving several outstanding problems in the current household registration management work	Februrary 26, 2001
	2002	A rural-urban migrant can get urban *hukou* in 14 pilot satelites cities and 33 towns of Beijing if she has a stable job and long-term residence place.	Jingzhengfa[2002]No.25	September 28, 2002
Tianjin	2001	A rural-urban migrant can get local urban *hukou* if she works in Tianjin more than 3 years. Meanwhile, her children and spouse could get local urban *hukou*.	Measures taken by Tianjin Public Security Bureau to ensure and serve the leap-forward development of Tianjin's economy	November 19,2001
Shijiazhuang	2001	A migrant can get local *hukou* if she has a stable job and residence place.	Shizheng[2001]No.87	June 29,2001
	2003	A migrant can get local *hukou* if she has a stable job and residence place. The distinction between agricultural and non-agricultural *hukou* types has been abolished. This is a further reform with respect to 2001 reform.	Shizheng[2003]No.115	September 27, 2003

Nanjing	2004	A migrant can get local *hukou* if she has a stable job and residence place in Nanjing. The distinction between agricultural and non-agricultural *hukou* types has been abolished.	Ningzhengfa[2004]No.140	June 19,2004
Hangzhou	2001	A rural-urban migrant can get local urban *hukou* if she has a stable job and long-term residence place in Hangzhou City, which allows her to enjoy the same benefits a local citizen has.	Hangzheng[2001]No.22	December 30,2001
	2002	A migrant with undergraduate degree can get a local *hukou* without a stable job.	Hangzhengban[2002]No.29	July 09,2002

Table B3.2 OLS Regression Coefficients for Monthly Wage

Ln (monthly wage)	(1) Urban	(2) Migrants	(3) Omega	(4) Pooled
Education (reference group: illiteracy)				
Primary School	0.091***	0.099***	0.114***	0.135***
	(0.018)	(0.012)	(0.010)	(0.010)
Middle School	0.247***	0.227***	0.256***	0.297***
	(0.018)	(0.013)	(0.010)	(0.010)
High School	0.423***	0.404***	0.412***	0.473***
	(0.018)	(0.013)	(0.010)	(0.010)
College	0.669***	0.729***	0.650***	0.718***
	(0.019)	(0.021)	(0.011)	(0.011)
Bachelor	0.895***	0.91***	0.872***	0.943***
	(0.019)	(0.040)	(0.011)	(0.011)
Master and above	1.290***	0.961***	1.268***	1.345***
	(0.022)	(0.271)	(0.016)	(0.016)
Experience	0.019***	0.013***	0.016***	0.017***
	(0.000)	(0.001)	(0.000)	(0.000)
Experience_squr	-0.00036***	-0.00032***	-0.00032***	-0.00032***
	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.182***	0.194***	0.188***	0.184***
	(0.002)	(0.003)	(0.002)	(0.002)
Married	0.033***	0.025***	0.015***	0.018***
	(0.004)	(0.005)	(0.003)	(0.003)
Ethnicity	0.041***	0.073***	0.052***	0.054***
	(0.005)	(0.007)	(0.004)	(0.004)
Occupation (Reference group: Organisation manager)				
Professional Personnel	-0.223***	-0.420***	-0.251***	-0.249***
	(0.007)	(0.022)	(0.006)	(0.007)
Clerks	-0.243***	-0.489***	-0.275***	-0.274***

	(0.007)	(0.021)	(0.007)	(0.007)
Business Service	-0.379***	-0.547***	-0.404***	-0.405***
	(0.007)	(0.019)	(0.007)	(0.007)
Agriculture Labours	-0.528***	-0.683***	-0.570***	-0.568***
	(0.012)	(0.030)	(0.011)	(0.011)
Production Workers	-0.348***	-0.552***	-0.375***	-0.380***
	(0.007)	(0.019)	(0.007)	(0.007)
Ownership (reference group: Others)				
State-owned	0.154***	0.076***	0.106***	0.122***
	(0.005)	(0.006)	(0.003)	(0.004)
Private-owned	0.134***	0.041***	0.085***	0.085***
	(0.005)	(0.004)	(0.003)	(0.003)
Self-employment	0.150***	0.034***	0.098***	0.099***
	(0.005)	(0.005)	(0.004)	(0.004)
Rural-Urban migrants				0.096***
				(0.003)
Observations	272,422	111,683	384,105	384,105
R-squared	0.418	0.349	0.395	0.397

Notes: The regressors also include province dummies and sector dummies, which is not reported here. Column (1) shows the results with assumption that the coefficient of urban groups represents the non-discriminatory wage structure. Column (2) shows the results with assumption that the coefficient of rural-urban migrant groups represents the non-discriminatory wage structure. Column (3) shows the results with assumption pooled estimate coefficient represents the non-discriminatory wage structure. Column (4) shows the results with assumption pooled estimate coefficient with group-specific intercepts represents the non-discriminatory wage structure. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table B3.3 OLS Regression Coefficients for Hourly Wage

Ln (hourly wage)	(1) Urban	(2) Migrants	(3) Omega	(4) Pooled
Education (reference group: illiteracy)				
Primary School	0.105***	0.097***	0.146***	0.149***
	(0.020)	(0.013)	(0.011)	(0.011)
Middle School	0.266***	0.230***	0.312***	0.317***
	(0.019)	(0.013)	(0.011)	(0.011)
High School	0.469***	0.443***	0.517***	0.525***
	(0.020)	(0.014)	(0.011)	(0.011)
College	0.746***	0.842***	0.793***	0.801***
	(0.020)	(0.023)	(0.011)	(0.012)
Bachelor	0.984***	1.066***	1.029***	1.038***
	(0.020)	(0.044)	(0.012)	(0.012)
Master and above	1.38***	1.255***	1.430***	1.439***
	(0.024)	(0.321)	(0.017)	(0.017)
Experience	0.018***	0.012***	0.016***	0.016***
	(0.001)	(0.001)	(0.000)	(0.000)
Experience_squr	-0.00030***	-0.00030***	-0.00027***	-0.00027***
	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.156***	0.170***	0.160***	0.159***
	(0.003)	(0.004)	(0.002)	(0.002)
Married	0.023***	0.024***	0.012***	0.012***
	(0.004)	(0.006)	(0.003)	(0.003)
Ethnicity	0.031***	0.080***	0.050***	0.051***
	(0.005)	(0.007)	(0.004)	(0.004)
Occupation (reference group: Organisation manager)				
Professional Personnel	-0.203***	-0.406***	-0.232***	-0.231***
	(0.007)	(0.024)	(0.007)	(0.007)
Clerks	-0.236***	-0.501***	-0.270***	-0.270***

	(0.007)	(0.023)	(0.007)	(0.007)
Business Service	-0.385***	-0.558***	-0.412***	-0.412***
	(0.008)	(0.021)	(0.007)	(0.007)
Agriculture Labours	-0.478***	-0.611***	-0.507***	-0.507***
	(0.013)	(0.033)	(0.012)	(0.012)
Production Workers	-0.353***	-0.583***	-0.395***	-0.396***
	(0.007)	(0.021)	(0.007)	(0.007)
Ownership (reference group: Others)				
Public Sector	0.175***	0.072***	0.145***	0.147***
	(0.005)	(0.007)	(0.004)	(0.004)
Private Sector	0.067***	-0.006	0.027***	0.027***
	(0.006)	(0.004)	(0.004)	(0.004)
Self-employment	0.022***	-0.045***	-0.006	-0.006
	(0.006)	(0.006)	(0.004)	(0.004)
Rural-Urban migrants				0.012***
				(0.003)
Observations	272,422	111,683	384,105	384,105
R-squared	0.425	0.278	0.403	0.403

Notes: The regressors also include province dummies and sector dummies, which is not reported here. Column (1) shows the results with assumption that the coefficient of urban groups represents the non-discriminatory wage structure. Column (2) shows the results with assumption that the coefficient of rural-urban migrant groups represents the non-discriminatory wage structure. Column (3) shows the results with assumption pooled estimate coefficient represents the non-discriminatory wage structure. Column (4) shows the results with assumption that the pooled estimate coefficient with group-specific intercepts represents the non-discriminatory wage structure. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Chapter 4 Trade Openness and the Wage Premium: Evidence from China

4.1 Introduction

China's accession to the World Trade Organisation (WTO) is one of the main drivers of economic growth over the last two decades. Much literature has examined the impact of trade openness on trader partners, e.g., the United States or Chinese local labour market. For example, Autor et al. (2013) and Pierce and Schott (2016) find that the U.S. lost substantial numbers of manufacturing jobs with the rising Chinese import competition. There are also some papers focusing on the effect of trade openness on China. For example, Brandt et al. (2017) investigate the effect of China's import tariff cut on the mark-ups and productivity of manufacturing firms. Other papers have studied the local labour market adjustment in terms of employment, wages, migration as a result of China's WTO accession (Cheng and Potlogea, 2015; Fachini et al., 2019; Li, 2018; Zi, 2021).

In this chapter, we, in particular, aim to contribute to the literature that examines the relationship between trade openness and wage inequality. The traditional Hecksher-Ohlin model and its companion Stolper-Samuelson theorem predict that inequality should decrease when a developing country integrates into the world. Unfortunately, empirical evidence seems to contradict this theoretical prediction. In this chapter, we address three main questions. First, how does China's trade openness affect the industry wage premium. Second, does the effect of trade openness vary by sector. Third, we ask whether tradable and non-tradable sectors respond differently to trade openness.

Following Goldberg and Pavcnik (2005) and Wang et al. (2017), we first estimate industry-level wage premia. The industry-level premium measures the part of wage variation that cannot be explained by work and firm characteristics. We apply a two-stage estimation strategy to measure the industry-level wage premium. Specifically, in the first stage, we use the data of China General Social Survey (CGSS), a comprehensive dataset which provides individual-level information about worker and firm characteristics and the industry classification of worker's workplace, to estimate a yearly industry wage premium. That is, we regress the log of worker's wages on a vector of worker's characteristics such as education, age, gender, marriage status, a set of industry dummies. In the second stage, we pool the industry wage premium over time and regress them on trade openness.

In this chapter, we use two measures of trade openness. The first is the trade shares in gross output. The second is the tariffs faced by China exporters. Both measures are used in the literature studying the effects of trade openness on wage inequality (Dorn et al., 2021; Li, 2018; Li et al., 2019; Wang et al., 2017). Industry-level trade data is obtained from China Input-Output Tables and tariff data from the WITS-TRAINS dataset. An identification challenge in terms of the endogeneity of trade openness measures is discussed. When addressing endogeneity concerns, our estimations suggest that a one percentage increase in trade openness leads to 0.1% increase in wage premium and 1% decrease in tariff rates leads to 0.02% increase in wage premium. We also find that such effects exist for both tradable and non-tradable sectors.

It is obvious that trade openness has directly impact on tradable sectors. For the non-tradable sectors that have little or no exposure to international trade, there is only an indirect relationship between trade openness and wage premium due to sales and input linkages to tradable sectors, as well as the labour mobility linkage between tradable and non-tradable sectors[18]. Thus, we further investigate the impact of tradable sectors on non-tradable sectors. By regressing wage premia of tradable sector on that of non-tradable sector, we find that a one percentage increase in wage premia of tradable sectors leads to 0.25 percent increase in non-tradable wage premia. In addition, we examine whether trade openness contribute to decreasing wage gap in tradable and non-tradable sectors. We propose a classic "Balassa-Samuelson" effect to explain the positive impact of wage premia of tradables on non-tradables. In this chapter, we highlight the role of imperfect regionally labour mobility. The impact of tradable sectors on non-tradable sectors is most pronounced in high restrictive regions.

[18] The labour mobility linkage is especially important in the Chinses context. Due to labour mobility restrictions, the labour demand in tradable sectors in regions can only be covered by labour from non-tradable sectors. To prevent labour from moving to tradable sector, the wages in non-tradable sectors will be higher.

Finally, we suggest consumer preference are also important, which can be used to explain the impact of tradables on non-tradables.

Our contribution lies in three aspects. Firstly, most research looks at the effect of trade liberalisation on the returns to particular worker characteristics. This chapter, on the other hand, focuses on the industry effects of trade liberalisation.[19] It is particularly meaningful to look at industry effects when labour is not perfectly mobile across sectors and regions. Secondly, we depart from the existing literature by disaggregating industries into tradable and non-tradable sectors. Such disaggregation is necessary because by doing so we gain more insights into the direct and indirect influences of trade openness on industry wage premium. Thirdly, we are among the first, to the best of our knowledge, to empirically test the role of labour mobility in determining the effect of trade on wages.

The remainder of this chapter is organised as follows. Section 4.2 presents a review of the empirical literature. Section 4.3 presents a theoretical framework to support the empirical analysis. This is followed by the discussion of empirical methodology in section 4.4. The data collection and descriptive statistics are presented in section 4.5 with section 4.6 discussing the empirical results. Section 4.7 presents additional results and section 4.8 concludes.

4.2 Literature Review

Our chapter is related to three areas of research. First, it is related to the empirical literature examining how trade openness affects wages in both developing and developed countries. For example, Brambilla et al. (2017) examine potential wage gains from exports in 61 developing countries. Dorn et al. (2019) investigate how trade openness influences income inequality using samples from 139 countries. In terms of specific countries, there is empirical evidence examining the wage premium for the United States (Bernard and Jensen, 1997), Germany (Bernard and Wagner, 1997) and the United Kingdom (Greenaway and Yu, 2004). The studies on developing countries also show positive wage premia with trade openness. For instance, Alvarez and Lopez (2005) show exporters pay approximately 20% higher wages than non-exporters in

[19] There are only a few literatures studying industry effects of trade liberalisation, such as Goldberg and Pavcnik, 2005; Topalova, 2010; Wang et al., 2017.

Chile. Van Biesebroeck (2005) find similar evidence in sub-Saharan Africa that exporters pay on average 34% higher wages after controlling for country, year, location,plant size and sample selection. Using firm-level census data of China manufacturing industries, Fu and Wu (2013) find a similar impact of exporting activities on wage premia in China. Han et al. (2012) analyse whether regions more exposed to globalisation experience larger wage inequality than less exposed regions. The finding is that globalisation contributes to rising within-region wage inequality in exposed regions. By contrast, a study on Indonesia (Amiti and Cameron, 2012) finds a strong link between lower input tariffs and a fall in the wage skill premia.

Most research focus on either firm-level estimation by examining export wage premia or how trade policy affects wages by estimating returns to a specific worker characteristic (e.g., skill or education). In this paper, we do not attempt to estimate returns to firm or worker-specific characteristics. Instead, we focus on industry effects. Goldberg and Pavcnik (2005) point to the importance of industry affiliation when investigating the consequences of trade policies. The industry wage premium is defined as industry-specific returns to labour that cannot be explained by worker, firm, job characteristics, but can be explained by industry affiliation. Some researchers (Goldberg and Pavcnik ,2005; Munch and Skaksen, 2008) point out that without controlling for worker, firm, or job characteristics, the results could overestimate the wage premia. As far as we know, there is only one paper examining the effect of trade openness on inter-industry wage differentials in the context of China. Wang et al. (2017) find that greater exposure to trade in final goods drives industry wage differentials. Following Goldberg and Pavcnik (2005) and Wang et al. (2017), in this paper, we first condition our industry wage premia estimates on individual characteristics in the first stage and in the second stage the estimated industry wage premium is regressed on trade openness related variables at the industry level. In this case, the relationship between trade openness and wage cannot be driven by differences in worker characteristics.

Our study is related to the strand of research of estimating local multipliers, especially the evolution of tradable and non-tradable sectors. Moretti (2010) quantifies the long-term change of employment of tradable and non-tradable sectors generated by an exogenous shock in tradable sectors in the context of the United States. He finds that for one additional job in manufacturing in a city, 1.6 additional jobs are created in non-tradable sectors in the same city. Similarly, Moretti and Thulin (2013) estimate the employment multiplier at the local level in Sweden and the average multiplier effect is smaller than that of the U.S. with 0.4-0.8 additional jobs created in non-tradable sectors. Frocrain and Giraud (2018) investigate the evolution of employment in the tradable and non-tradable sectors in France over 1999-2015. They find that the employment of tradable sectors declines while non-tradable jobs continue to grow. With an exogenous shock to labour demand in tradable industry, the direct effect of this shock is an increase in employment in this industry and other tradable industries as well as non-tradable industries. This shock is likely to have general equilibrium effects on local prices and wages.

The Balassa-Samuelson model predicts that a rise in tradable sector productivity raises the relative prices of non-tradable sectors and leave the relative wages of between these two sectors unchanged. By documenting a series of facts in the United States and constructing a theoretical model, Eckert et al. (2019) emphasises the average wage growth of the tradable service sector and non-tradable service sector, with relatively faster growth in the tradable service sector. In terms of the China context, Wang and Chanda (2018) study the impact of employment growth in manufacturing on job creation in non-tradable sectors for city-level data in China. Other studies such as Chen et al. (2023), Fang and Herrendorf (2021) and Liao (2020) document the rise of service sector in the Chinese economy and provide the underlying mechanisms of the complementarity between services used as inputs to industrial production and the increasing consumer preferences on personal service sectors. Our study tries to estimate the wage effect of tradable sectors on non-tradable sectors due to trade-induced shocks.

Finally, this chapter contributes to the literature studying the role of imperfect labour mobility when estimating the benefits of trade. The classic trade models such as the Heckscher-Ohlin (H-O) model assume the factor of production are perfectly mobile and their returns are equalised across sectors and regions. However, in reality, the immobility arising from capital market imperfection or frictions from labour market may distort the wage equalization. Topalova (2010) measures the impact of trade liberalization on poverty in India by looking at two types of factor mobility: geographical and sectoral. The regionally disparate effects of liberalization which was most pronounced in regions with less geographical mobility suggest that institutional characteristics matter. De Blasio and Menon (2011) look at the effect of an exogenous shift in local employment in tradable sectors on the local employment in the remaining parts of local economy. They find no effect of a rise in employment in local tradable sectors on local non-tradable sectors due to the role of obstacles of labour mobility. Similarly, Cardi and Restout (2015) highlight the role of imperfect mobility across sectors when investigating the relative prices and wages effects of higher productivity in the tradable sectors compared to non-tradable sectors. The quantitative results show that the relative wages are muted in countries with higher intersectoral reallocation of labour. With imperfect mobility of labour, Fu and Wu (2013) estimate export premia by sectors and confirm wage equalization does not occur. Zhao (2001) investigates the effect of foreign direct investment on wage inequality and finds that the high cost of labour mobility segments labour markets and raises wage premium for skilled labour in foreign sector. Following Cardi and Restout (2015), our study explores the role of labour mobility in the determination of the relative wages responses in non-tradable sectors compared to tradable sectors.

4.3 Theoretical Framework

In this section, we provide a theoretical framework to rationalise our hypothesis. Our work is related to trade-induced industry productivity differences and trade-related Balassa-Samuelson effect. The influential work of Melitz (2003) has been accepted as a central model of inter-industry reallocation due to trade liberalisation. Melitz (2003) theoretically demonstrates that exposure to trade will improve aggregate industry productivity through market effects (i.e., the exit of low productive firms from an industry) or resource reallocation towards more productive firms. It also provides a theoretical explanation that industrial productivity increases more strongly in liberalised sectors than less liberalised sectors. Another influential paper of Ghironi and Melitz (2005) suggests a model of Balassa-Samuelson effect based on Melitz (2003) model of heterogenous productivity. This chapter examines the trade-related industry wage premia and Balassa-Samuelson effect. Trade openness could potentially affect industry wage premia through productivity growth and one important channel is that the trade-induced greater productivity in tradable sectors translate into a rise in the wages of non-tradable sectors.

We assume there are a continuum of firms, and all firms have different productivity level by $\varphi > 0$. Each firm faces a residual demand curve with constant elasticity σ. The aggregate price P is defined by:

$$P = [\int_0^{\infty} p(\varphi)^{1-\sigma} M\mu(\varphi)\, d\varphi]^{\frac{1}{1-\sigma}} \qquad (4.1)$$

where p represents the price of individual firm, M represents a mass M of firms (thus M goods), $\mu(\varphi) \in (0, \infty)$ is the distribution of productivity, σ is the price elasticity.

The aggregate revenue and profit are:

$$R = \int_0^{\infty} r(\varphi)\, M\mu(\varphi) d\varphi \qquad (4.2)$$

$$\Pi = \int_0^{\infty} \pi(\varphi)\, M\mu(\varphi) d\varphi \qquad (4.3)$$

where R denotes aggregate revenue, Π denotes aggregate profit. $r(\varphi)$ is firm revenue, $\pi(\varphi)$ is firm profit.

There are a number of potential firms into the industry. To enter, firms have to make an initial investment, modelled by a fixed entry cost $f_e > 0$ (measured in units of labour). Firms will then draw their initial productivity φ from a common productivity distribution $g(\varphi) \in (0, \infty)$. Firms are assumed to not know their productivity before actually entering the industry. With low productivity levels, firms may exit and not produce. If a firm decides to enter and stay in the industry, it will face a constant probability δ with unexpected shock that forces it to exit. We assume each firm's productivity does not change over time. An entering firm with productivity φ would immediately exit if profits are negative or would produce if it earns positive profits $\pi(\varphi) > 0$ in every period until facing unexpected shock δ. Therefore, each firm's value function is given by:

$$v(\varphi) = \max\left\{0, \sum_{t=0}^{\infty} (1-\delta)^t \pi(\varphi)\right\} = \max\left\{0, \frac{1}{\delta}\pi(\varphi)\right\}. \qquad (4.4)$$

where δ is the probability that in every period of the unexpected shock that may force firms to exit.

Thus, $\varphi^* = \inf\{\varphi: v(\varphi) > 0\}$ identifies the cut off productivity level. $\pi(\varphi^*)$ is equal to zero, which is the zero cut off profit condition.

Firms with productivity levels $\varphi < \varphi^*$ will immediately exit and never produce. In this case, the equilibrium productivity distribution will depend on the initial productivity draw, conditional on successful entry. Thus, $\mu(\varphi)$ is the conditional distribution of $g(\varphi)$ on $[\varphi^*, \infty]$:

$$\mu(\varphi) = f(x) = \begin{cases} \dfrac{g(\varphi)}{1 - G(\varphi)}, \varphi > \varphi^*, \\ 0, \ otherwise, \end{cases} \qquad (4.5)$$

where $G(\varphi)$ is the cumulative distribution of $g(\varphi)$.

The aggregate productivity level $\tilde{\varphi}$ is a function of cut off productivity level φ^* as follows:

$$\tilde{\varphi}(\varphi^*) = \left[\frac{1}{1 - G(\varphi^*)} \int_{\varphi^*}^{\infty} \varphi^{\sigma - 1} g(\varphi)\, d\varphi\right]^{\frac{1}{\sigma - 1}} \qquad (4.6)$$

From the above, we have explained how the aggregate productivity $\tilde{\varphi}$ is determined and next we introduce the exposure of a country to trade inducing reallocations between firms and increase aggregate productivity.

In an open economy, each firm's profit can be separated into domestic profit $\pi_d(\varphi)$ and exporting profit $\pi_x(\varphi)$. A firm that produces domestically would export the products to n countries if $\pi_x(\varphi) > 0$. Each firm's profit could be combined as $\pi(\varphi) = \pi_d(\varphi) + \max\{0, n\pi_x(\varphi)\}$. We have $\varphi_x^* = inf\{\varphi: \varphi \geq \varphi^*, and\ \pi_x(\varphi) > 0\}$ representing the cut off productivity level for exporting firms. In this case, the cut off firms with $\varphi_x^* = \varphi^*$ will earn zero profit ($\pi(\varphi) = 0$) and nonnegative exporting profit $\pi_x(\varphi) \geq 0$. If $\varphi_x^* > \varphi^*$, then firms with productivity levels above φ_x^* will earn positive profit from both domestic and export sales. Some firms with $\varphi^* < \varphi < \varphi_x^*$ will only produce for domestic sales as they earn negative profit for export sales.

The weighted average productivity level is：

$$\widetilde{\varphi_t} = \left\{\frac{1}{M_t}\left[M\tilde{\varphi}^{\sigma-1} + nM_x(\tau^{-1}\tilde{\varphi}_x)^{\sigma-1}\right]\right\}^{\frac{1}{\sigma-1}} \qquad (4.7)$$

where M denotes the mass of firms, M_x is the mass of exporting firms, $\tilde{\varphi}$ is the average productivity across firms with only domestic sales, $\tilde{\varphi}_x$ is the average productivity of exporting firms. τ is the trade cost.

As entry to new export markets is costly, only firms with higher productivity can afford the entry cost. Also, firms compete for a common source of labour which will increase the returns of labour and force the least productive firms to exit. The less productive firms will only serve the domestic market. Therefore, in an open economy, traded and non-traded (usually referred to tradable and non-tradable sectors) sectors have aggregate productivity differentials. This structure endogenously determines the composition of traded and non-traded sectors as well as the consumption basket. As was previously mentioned, the exposure to trade increases productivity but the productivity gains are different between tradable and non-tradable sectors. Next, we explain the consequences of trade-induced productivity differentials between tradables 00and non-tradables.

There are two sectors in a country, tradable (T) and non-tradable (NT). Each sector produces Y^s by using physical capital K^s and labour L^s, according to the Cobb-Douglas production function:

$$Y^s = A^s(L^s)^{\theta s}(K^s)^{1-\theta s} \qquad (4.8)$$

where A^s is the total factor productivity of sector S=T, NT; θs denotes the labour income share in the sector's value added.

Under perfect competition, prices in each sector are given by:

$$P^s = (A^S)^{-1}(W^S)^{\theta s}(R^S)^{1-\theta s}\theta s^{-\theta s}(1-\theta s)^{\theta s-1} \quad (4.9)$$

where W^s and R^s are the unit cost of labour and rental rate of capital, respectively. If we consider a small open economy with perfect capital mobility, the traded goods are then set by international prices $P^T = P^{T*}$ and the rate of return on capital is equal to the world value $R^T = R^{NT} = R^*$.

Normalising the tradable price to unity, and taking the tradable goods as numeraire, the relative price of non-tradables in terms of tradables can be written as:

$$\frac{P^{NT}}{P^T} \equiv P = \frac{\Psi^{NT}}{\Psi^T}\frac{A^T}{A^{NT}}\frac{(W^{NT})^{\theta NT}(R^{NT})^{1-\theta NT}}{(W^T)^{\theta T}(R^T)^{1-\theta T}} \quad (4.10)$$

where $\Psi^s = (\theta s)^{-\theta s}(1-\theta s)^{\theta s-1}$.

Assuming the depreciation rate of physical capital is δ_K and the interest rate is r, we have $R^T = R^{NT} = R^*$, where $R^* = P(r+\delta_K)$. Also, if we assume that the unit cost for producing tradables is equal to 1, then we have $W^T = (A^T)^{\frac{1}{\theta T}}(\Psi^T)^{\frac{1}{\theta T}}(R^T)^{-(\frac{1-\theta T}{\theta T})}$.

Now we can rewrite the relative price of non-tradable sectors, as follows:

$$P = \frac{\Psi^T}{(\Psi^N)^{\frac{\theta T}{\theta N}}}\frac{A^T}{(A^N)^{\frac{\theta T}{\theta N}}}(\frac{W^N}{W^T})^{\theta T}(r+\delta_K)^{\frac{\theta T-\theta N}{\theta N}} \quad (4.11)$$

Log differentiating the relative prices yields:

$$p = c + \left(A^T - \frac{\theta T}{\theta N}A^N\right) + \theta^T\omega \quad (4.12)$$

where $c = \ln(\Psi^T) - \frac{\theta T}{\theta N}\ln(\Psi^N) + \frac{\theta T-\theta N}{\theta N}\ln(r+\delta_K)$, $\omega = \ln(\frac{W^{NT}}{W^T})$.

We first assume perfect labour mobility across sectors and regions, in this case, the wages across tradable and non-tradable sectors would be equalized. That is, $\omega = 0$. As a result, the price in non-tradable sectors will rise by $A^T - \frac{\theta T}{\theta N} A^N$ to compensate the wage increases. Intuitively, the greater productivity in tradable sectors translates into the increases on wages of tradable workers. Firms in non-tradable sectors have to increase their wages in order to prevent the employees from moving to the tradable sectors. The wage increase in non-tradables can be achieved through price increases. In reality, labour is not perfectly mobile, which contrasts with the standard Balassa-Samuelson assumption.

In this chapter, we consider imperfect mobility of labour across regions and examine whether the Balassa-Samuelson theory can be reconciled with numerical evidence in terms of Chinese cases. In China, *hukou* restrictions are implemented to regulate interregional mobility. The implementation is determined by local government so that there are large variations in the *hukou* restrictions. Even if there is limited mobility among regions, there are high level of mobility across sectors within a region. We consider each province with restrictive *hukou* policy implementations as a local labour market. With high level of intersectoral mobility, the non-tradable sectors have to increase their wages to align with those of tradable sectors. Some provinces, with less restrictive *hukou* policies, could attract migrants and firms in tradable/non-tradable sectors could hire migrants. In this case, the relative prices of non-tradables increase less as firms in non-tradables could hire migrants using lower unit cost of labour instead of rising the wages to attract locals.

4.4 Empirical Methodology

4.4.1 Econometric Modelling

To examine the impact of trade openness on wage premia, WP_{ipt}, we conduct the following region-industry level regression analysis:

$$WP_{ipt} = \alpha_o + \alpha_1 trade\_openness_{ipt} + \gamma Z_{ipt} + \gamma_t + \gamma_i + \varepsilon_{it} \qquad (4.13)$$

where subscripts i, p, t denote the industry, province and observation year, respectively. In our samples, there are 30 provinces and 15 industries. We use six years of data: 2003, 2008, 2010, 2012, 2015, and 2017. In total, we have 2700 observations at province-industry level. WP_{ipt} is the region-industrial wage premium between 2003 and 2017. Our key variable of interest is $trade\_openness_{ipt}$, here defined either in terms of the ratio of total trade in gross output or in terms of the barriers to trade measured by import/export tariffs. Z_{ipt} denotes a vector of control variables such as the share of gross capital formation (GCF) and the share of value added in gross output (both expressed as natural logarithm) which may explain wage differentials across industries. For example, gross capital formation and value added could increase relatively industry specialisation and this leads to a rise in labour demand and wages (Stirbock, 2002). The inclusion of time fixed effects, γ_t, captures year common effect to all province-industry pairs and γ_i is an industry fixed effect, which captures time-invariant characteristics of industries. ε_{it} is the error term. Finally, standard errors are clustered at the province-industry level.

4.4.2 Identification Challenges and Solutions

Since trade openness may be endogenous, the relationship between trade openness and industry wage premium could be biased. For example, there could be unobserved industry heterogeneity that affects wages and openness simultaneously. Time-variant factors such as exchange rate fluctuations could affect trade and have a direct effect on wages (Goldberg and Pavcnik, 2005). Also, another source of endogeneity is reverse causality. For example, industrial wages affect the firm's production decisions and, in turn, affect trade behaviour. Therefore, the first measurement of trade openness could be endogenous.

To address the endogeneity concerns, we instrument for trade openness. For an instrument to be valid, in our case, the instrumental variable needs to be highly correlated with the instrumented variables but uncorrelated with the dependent variable and error term. Following Shiferaw and Hailu (2016) and Wang et al. (2017), we use as instrument the weighted average of the trade partners' industrial gross output. The intuition is that industrial output captures both demand and supply factors which is related to domestic and export markets. In other words, a higher output in industry i in country m may imply a higher import from China that can be used to production and to consumption. Also, a higher output in industry i in country m may enable higher exports to China. The instrument can be expressed as follows:

$$GOI_{i,p,t} = \sum_{m=1}^{43} \frac{trade_{i,p}}{\sum_{p=1}^{30} trade_{i,p}} \frac{trade_{i.m}}{\sum_{m=1}^{43} trade_{i,m}} * GOI_{i,m,t} \quad (4.14)$$

where i denotes industries, p denotes provinces, t is time. $GOI_{i,p,t}$ represents weighted gross output indicator in industry i in province p at time t. $GOI_{i,m,t}$ is gross output of industry i in country m at time t. $\frac{trade_{i.m}}{\sum_m trade_{i,m}}$ is the share of trade with country m in total trade with all trade partners for industry i in the pre-sample year. $\frac{trade_{i,p}}{\sum_p trade_{i,p}}$ is the trade share of industry i in province p in the pre-sample year.[20] We consider trade shares in year 2000, before China joined the WTO, as the weights. The use of pre-sample weights could alleviate potential endogeneity concerns as gross output may have positive impact on the current trade volumes. Industry-level gross output is collected from World Input-Output Database (Timmer et al., 2015). Pre-sample trade shares with country m for each industry are constructed based on bilateral trade data from WIOD. Pre-sample trade shares of provinces for each industry are constructed from China Statistical Yearbook (National Bureau of Statistic of China, 2001).

Alternatively, a large body of literature has used tariffs as an appropriate measure of trade liberalisation. Either tariffs faced by China's imports (Brandt et al., 2017; Dai et al., 2020; Zhou et al., 2022) or exports (Facchini et al., 2019; Handley and Limao, 2017; Li et al., 2019; Pierce and Schott, 2016) can be used as a measure of trade liberalisation. The export tariffs are exogenous as the tariff are set by destination countries and are unlikely to be affected by local conditions in China. In this chapter, we use tariff barriers faced by China's exports as measure of trade openness. We define the overall tariff barriers faced by each industry in China as the weighted average of the tariff rates using the export share in a destination country of total export values from China to export destinations as weights. We then construct the province-industry level export tariff barriers by interacting the employment share of industry i in province p with weighted industry-level tariff barriers.

[20] We use the trade share of province in each industry to measure the exposure to international demand and supply. The instrument is province and industry specific and varies over time.

However, using employment shares in the current year as weights to construct province-industry tariff barriers raises concerns about endogenous industry structural changes, as the changes in employment in the current year may have resulted from the tariff changes.[21] Also, wages could be correlated to employment. Hence, following the common practice in literature (Dai et al., 2020; Li et al., 2019; Topalova, 2005), we use the employment share in 2000, before the take-off of Chinese exports, as the set of weights to construct province-industry tariff barriers. We consider the constructed tariff barriers as exogenous as domestic forces do not play a role in shaping the distribution of trade protections. Also, the strategy of using the employment share prior to WTO accession addresses the concern about endogenous industry structure changes.

4.5 Data and Descriptive Statistics

4.5.1 Wage Premium Construction

The wage premium captures the part of the wage variations which cannot be explained by observed worker characteristics but can be explained by industry affiliation. We use a two-stage estimation model (Krueger and Summer, 1988; Ebmer, 1990; Winter-Ebmer, 1994; Goldberg and Pavcnik, 2005) to estimate the wage premium. In the first stage, we use household survey data to estimate industry wage dispersion.

The data for the first-stage estimation is collected from the China General Social Survey (CGSS). The CGSS is the first national continuous social survey project in mainland China implemented by academic institutions, which could be considered as the Chinese counterpart of the General Social Survey (GSS) in the U.S. As a representative nationwide survey, the CGSS has been conducted since 2003. The period from 2003 to 2008 is Cycle I of the CGSS. The period from 2010 to 2019 is Cycle II of the CGSS. By now, the data of 2003, 2005, 2006, 2008, and 2010-2013, 2015, 2017 CGSS have already been open to the public. CGSS aims to monitor systematically the changing relationship between social structure and quality of life in both urban and rural China. To do so, CGSS collected information in various areas of individual's demographic background, attitudes and behaviours, and socioeconomic

[21] With constrained labour mobility imposed by hukou system in China, we consider each province as a local market. It is likely that labour is more mobile within provinces, hence, labour reallocation across industries within a local market. We will attempt an investigation of the role of labour market rigidities on relative wages.

status, such as earnings, gender, age, education, industry affiliation. The sample covers both rural and urban areas from 28 provinces. For this study, we use six waves of the data: 2003, 2008, 2010, 2012, 2015, 2017[22]. CGSS provides three-digit industry classification, and we map them into 15 categories according to China Industrial Classification System (CIC)[23]. The detailed classifications are shown in table A3.

In the first stage, we use cross-sectional wage equations in order to estimate province-industry wage dispersion. To be more specific, the individual wage is regressed on demographic characteristics, workplace-related conditions, industry dummies as well as the interaction between province and industry dummies. The first-stage regressions are estimated separately by year, and in the second stage the province-industry wage premia are pooled over time and regressed on trade openness variables.

The first-stage estimation is as follows:

$$lnw_{j,i,t} = \alpha_t + \beta_t \mathrm{X}_{j,i,t} + \gamma_t H_{j,i,t} + \omega_{i,p,t} \mathrm{Z}_{i,p,t} + \varepsilon_{j,i,t} \qquad (4.15)$$

where j denotes individuals, i denotes industries, p denotes provinces, t is time. $lnw_{j,i,t}$ represents the logarithm of yearly income for individual j in year t. $\mathrm{X}_{j,i,t}$ is a vector of worker-specific characteristics including gender, age, age squared, rural or urban *hukou* type, marital status, education attainment, party membership, ethnicity. $H_{j,i,t}$ represents the job-related characteristics, including the firm ownership and industry categories. $\mathrm{Z}_{i,p,t}$ is the interaction between province indicators and industry indicators. $\omega_{i,p,t}$ measures province-industry pair wage dispersion. $\varepsilon_{j,t}$ is the error term.

The estimated wage differential coefficient should be normalised to show the variation in wages between an employee in a given province-industry and an average employee in any other province-industries pair.

[22] The reason why we only use six waves of the data is that we need to pool the wage premia over time and merge them with trade openness which have only been provided for a few years.
[23] The National Bureau of Statistics provides a few versions of classifications, such as "GB-4754-84", "GB-4754-94", "GB-4754-2002". As the three-digit classification follows China Industry Classification (GB-4754-94), we aggregate three-digit GB code into 15 classifications following the standard of GB-4754-94.

$$\tau_{i,p,t} = \omega_{i,p,t} - \overline{WA_t} \qquad (4.16)$$

$$WA_t = \sum_{i=1}^{I} \sum_{p=1}^{P} s_{i,p,t}\, \omega_{i,p,t} \qquad (4.17)$$

where $\tau_{i,p,t}$ is the normalised wage premium for industry i in province p, $\overline{WA_t}$ is the employment-weighted average wage premium in year t, $s_{i,p,t} = {n_{i,p,t}}/{\sum_{p=1}^{P} n_{p,t}}$ is the employment share of province-industry pair in year t.

4.5.2 Trade Openness

As mentioned above, we use two measures of trade openness. The first is the total trade value. We collect international trade data from China Input-Output tables which present a collection of provincial inter-regional and inter-sectoral economic flows. China IO Table is comparable to World Input-Output Table and the latter demonstrate the inter-country economic relationship while the former demonstrates the inter-regional economic relationship within China. The National Bureau of Statistics has published China IO Tables for 2002, 2007, 2010, 2012, 2015, 2017.[24] The China IO Table has three main parts. The first is the intermediate flows among provinces and sectors. The second part is the final use of provinces and 5 final use categories, including rural household consumption, urban household consumption, government spending, fixed capital formation and changes of inventories. Another part is the value added for provinces and sectors. Also, international exports and imports are also shown in the table.[25]

There are a few inconsistencies between the China IO Tables of different years. 2002 IO table provides information for 30 provinces and 20 sectors. For 2007 and 2010, the tables provide economic relationships for 30 provinces and 30 sectors. By contrast, the table compiles information for 31 provinces and 42 sectors for the years 2012, 2015, and 2017. To make the data consistent, we aggregate the sectors into 15 industry classifications based on the China Industry Classification system (GB-4754-94). Also,

[24] Given that the period we covered is after China's accession to WTO in 2001, we use data for 2002 and 2007 to proxy data for 2003 and 2008. To do so, the data collection will be consistent with CGSS.
[25] International import is divided into import used as intermediate use and import used as final use. We computed import used as final use based on the proportion of import used as intermediate demand among provinces and sectors. More details will be introduced in appendix.

we drop the province "Tibet" for 2012, 2015, and 2017 as other tables exclude Tibet. Another inconsistency is that the 2010 table only contains information on capital formation, instead of information on fixed capital formation and changes in inventories. In this case, we sum fixed capital formation and changes in inventories and generate a new variable of capital formation for the years 2002, 2007, 2012, 2015 and 2017. By doing so, we do not lose observations for 2010. The other control variables, value added is also collected from China IO Tables.

Another commonly used measure of trade openness is the extent of protectionism as measured by tariffs (Facchini et al., 2019; Li et al., 2019; Yuan Tian, 2022), as trade policy affects openness directly through tariff rates. The main idea for the construction of this measure is to explore the industry-level import tariff that each country imposes on China's exports.[26] We construct the tariff variable from the WITS-TRAINS database. We calculate the industry-level tariff faced by China's exports using the following equation:

$$Tariff_{i,t} = \sum_{m=1}^{43} \frac{expv_{i,m,t}}{\sum_{i=1}^{15} expv_{i,m,t}} \times tariff_{i,m,t} \qquad (4.18)$$

where i denotes industries, m denotes countries, t denotes tariff years. $\frac{expv_{i,m,t}}{\sum_i expv_{i,m,t}}$ is the share of sector i over total export values from China to export destinations.

We obtain the tariff rates faced by China's exporters abroad at the 6-digit level of the Harmonised System (HS) product classification for the years 2003, 2008, 2010, 2012, 2015, 2017.[27] Trade value data is also obtained from WITS-TRAINS database[28]. There are different types of tariff rates, including, MFN, preferential tariff and bound tariff.

[26] Some literature also uses import tariff to measure openness (e.g., Amiti and Camreon, 2012; Brandt et al., 2017; Goldberg and Pavcnik, 2005; Topalova, 2010). However, this chapter uses tariff imposed on China's exports by countries as this measure could partially alleviate the endogeneity of trade openness.
[27] We follow the data extraction procedure proposed by Beohm et al. (2023)
[28] The World Integrated Trade Solution (WITS) is developed by the World Bank. This software allows users to access information on trade and tariffs. The UNCTAD Trade Analysis Information System (TRAINS) can be accessed by WITS. TRAINS contains information on tariff and non-tariff measures for more than 160 countries. The tariff and non-tariff measures are recorded at the HS level of disaggregation. Tariff information contains the applied and MFN tariff rates. In particular, the data covers all the countries with positive imports from China.

Most-favoured Nation (MFN)[29] tariffs are the tariffs that WTO members promise to impose on imports from other members of WTO, unless the country is part of a preferential trade agreement (such as free trade area or customs union). A preferential tariff is the tariff that falls under a preferential trade agreement. Normally, a country that joins a preferential trade agreement promises to give another country's products lower tariffs than their MFN rate. Effectively applied tariff is defined as the lowest available tariff between MFN and preferential tariff. If a preferential tariff exists, it will be used as the effectively applied tariff. Bound tariff is the maximum MFN tariff level for a commodity line, which gives the WTO members flexibility to decrease or increase their tariffs as long as they do not raise them above their bound levels. As MFN rates have been applied after China's accession to WTO, we use MFN tariff when calculating industry-level tariff rates. We first map them into the sectors based on ISIC Rev.3 using a concordance table provided by WITS. We then map them into CIC 15 classifications with a concordance table provided by the National Bureau of Statistics. Next, we construct tariff exposure faced by each province-industry pair by interacting employment share of industry i in province p:

$$tariff\_exposure_{i,p,t} = \frac{emp_{i,p}}{\sum_{i=1}^{15} emp_{i,p}} \times Tariff_{i,t} \qquad (4.19)$$

The intuition is that the tariff exposure of a province in an industry will be larger the larger the employment share in that industry. As suggested by the equation, the constructed variable of export tariff varies across provinces and over time. The cross-section variations of the constructed tariff variable come from differences in industrial composition across provinces. The within-province over-time variations come from changes in industrial tariff levels over years.

4.5.3 Graphical Evidence

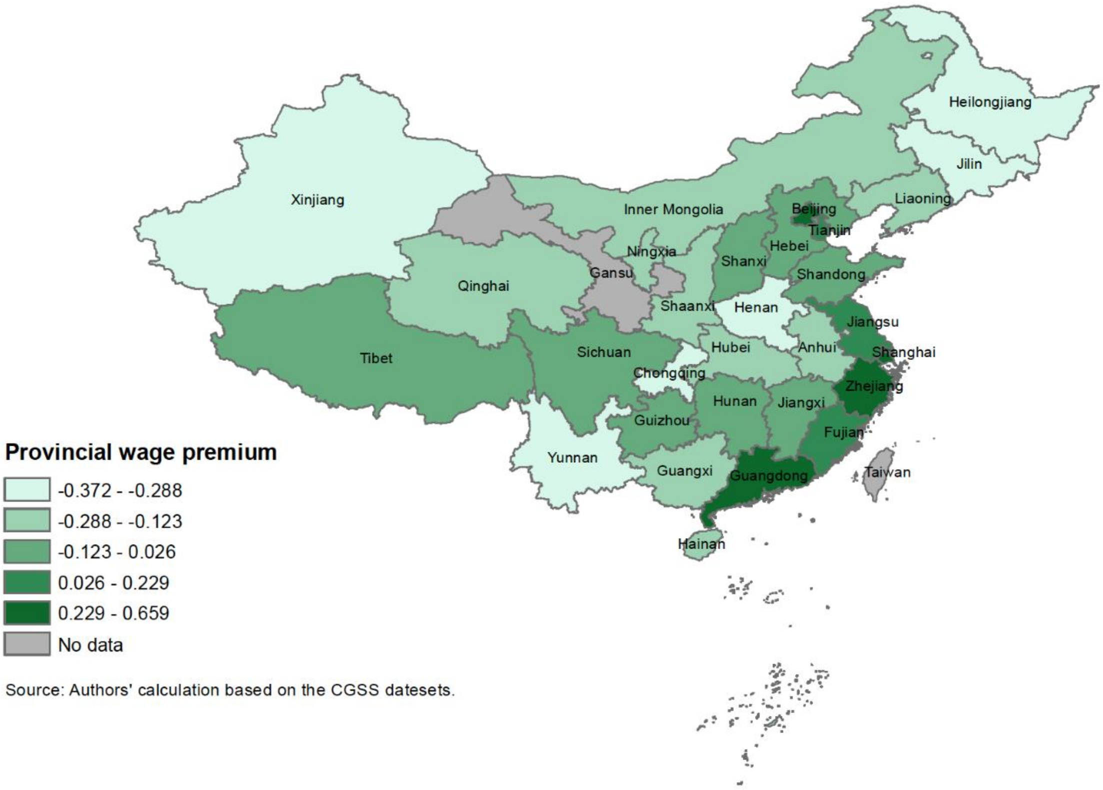
This section provides some stylised facts about China's wage premium and trade openness. Figure 4.1 shows the geographical distribution of provincial wage premium and export share. The provincial wage premium is the average premium between 2003 and 2017, calculated based on the CGSS datasets. The provincial export share is the

[29] The MFN tariff schedule is also sometimes referred to as the "Normalized Trade Relations (NTR)" tariff.

average of export shares in gross output between 2003 and 2017, calculated from China Input-Output Tables. As we can see from figure 4.1A, wage premia are higher in the coastal provinces of Southeastern China. In particular, a very high wage premium is observed in Guangdong. The Shanghai and Zhejiang provinces in Yangtze River Delta regions also have high wage premia. The wage premium is also high in several provinces in Northeastern and Western China. Figure 4.1A reflects the large variation of wage premia across regions. Figure 4.1B suggests the coastal regions are the main exporting providers. The highest export share is 68%, while the lowest export share is 2.6% which appears in central China. Northeastern and Western China also have high export shares. Together with the figure 4.1A and 4.1B, we can see a clear pattern that correlates provincial wage premia and export shares. This pattern shows a large geographic variation across provinces, which motivates us to take into consideration provincial variations when measuring wage variations across industries.

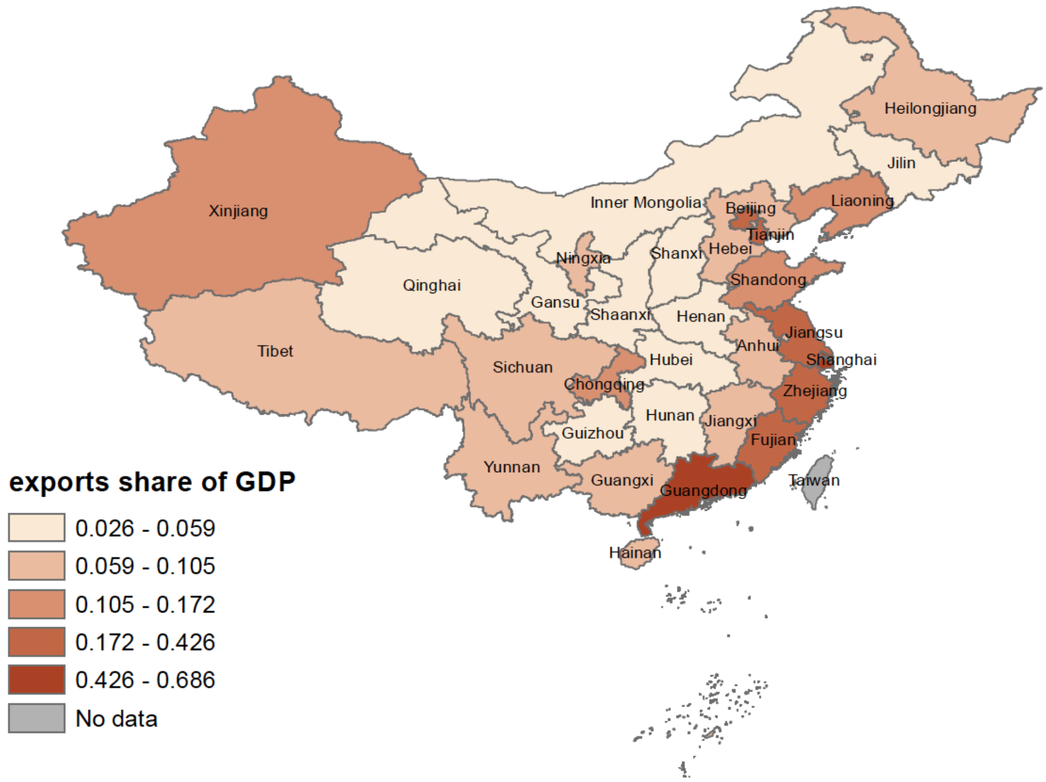
Figure 4.1 Spatial Distribution of Averaged Wage Premium and Export Share between 2003 to 2018

Panel A: Provincial Wage Premium (2003-2017)



Source: Authors' calculation based on the CGSS datesets.

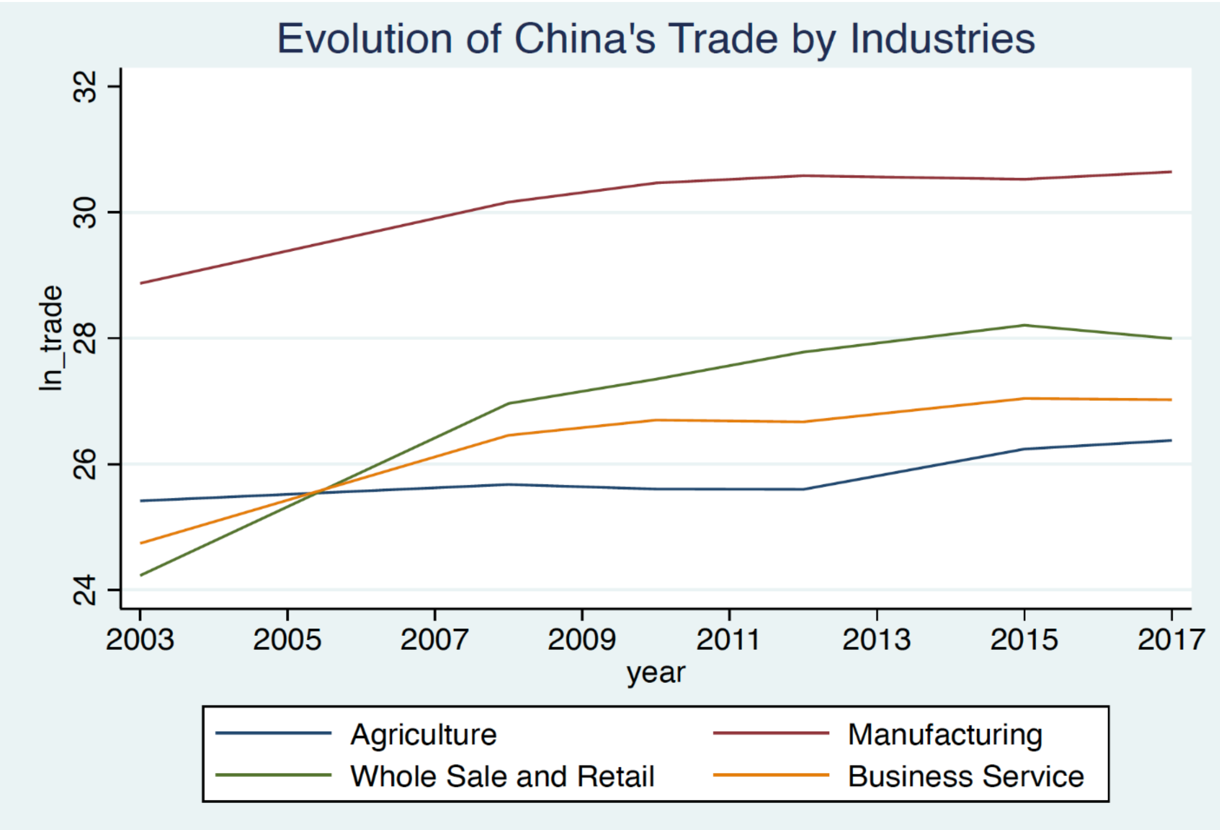
Panel B: Export Share of GDP (2003-2017)



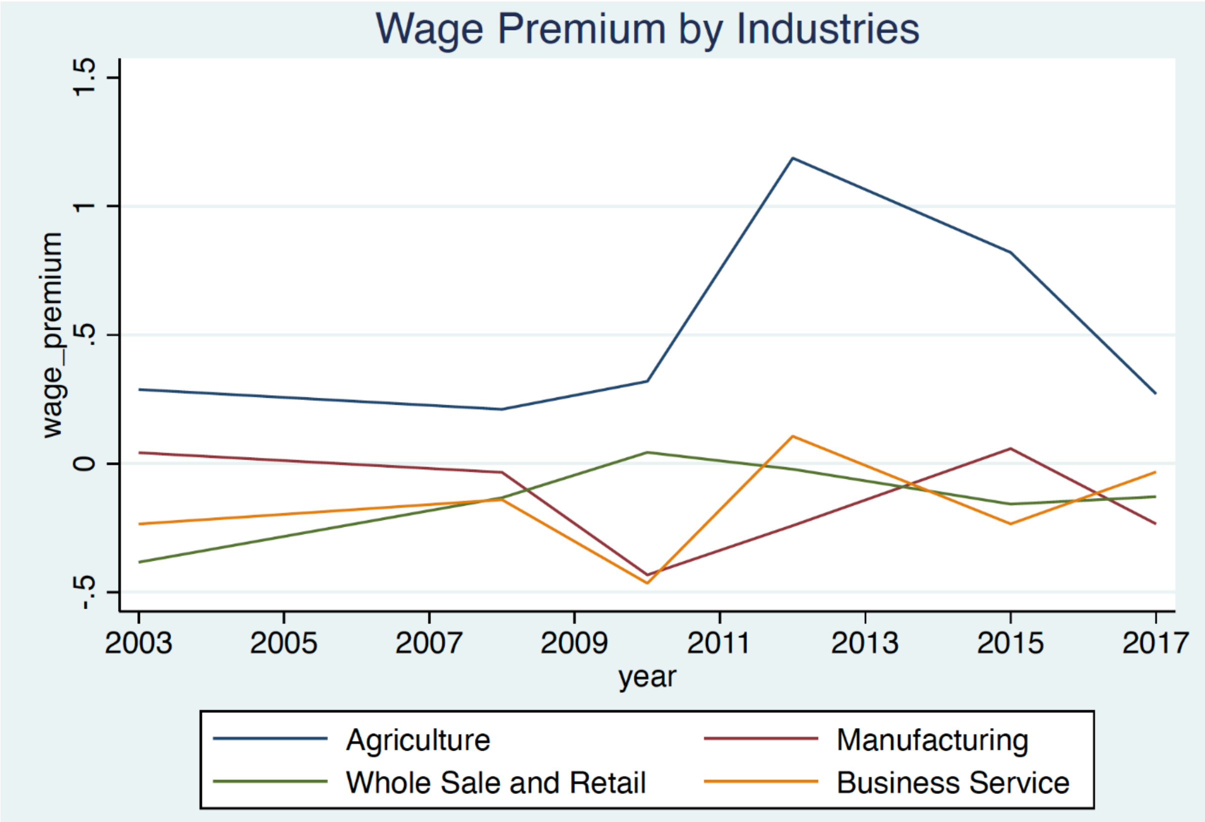
Notes: Figure 4.1A shows the geographic variation of average provincial wage premium between 2003 and 2017. The provincial wage premium is calculated based on the two-stage estimation approach. In particular, the log of individual yearly wage is regressed on a vector of worker characteristics, a vector of job-related characteristics and province dummies. Figure 4.1B shows the geographic variation of average provincial export share between 2003 and 2017. The export and GDP data are collected from China Input-Output Table. The average export share is the average of export shares in gross output between 2003 and 2017.

Figure 4.2: Industrial Variation of China's Trade and Wage Premium

Panel A Trade Openness by Industry



Panel B Wage Premium across Industries



Notes: Figure 4.2 shows the time variation of China's trade and wage premium between 2003 and 2017. Trade value is expressed as log (total trade). Trade data is collected from China Input-Output tables. Wage premium across industries is calculated based on the two-stage estimation approach. In particular, the log of individual yearly wage is regressed on a vector of worker characteristics, a vector of job-related characteristics and industry dummies.

In this chapter, we aggregate industries into 15 categories and figure 4.2 shows information on selective industries. Based on the classification criteria proposed by Jensen and Kletzer (2005), agriculture and manufacturing are tradable sectors and wholesale, and retail and business service are non-tradable sectors. The idea of the classification is to use geographic concentration of economic activities for each industry. Industries that are geographically concentrated are considered tradable as goods that are traded tend to exhibit geographic concentration in production to capitalise on increasing returns to scale and to access to production inputs. The measure of geographic concentration used by Jensen and Kletzer (2005) is the Gini coefficient by comparing the region's share of industry's employment with the area's share of aggregate employment. Industries with a Gini coefficient below 0.1 as non-tradable and industries with a Gini coefficient greater than or equal to 0.1 as tradable.

Figure 4.2A shows the time trend of industry-level trade values from 2003 to 2017. We see that China's trade gradually increases over time, for both tradable and non-tradable sectors. For non-tradable sectors, the trade values initially account relatively little for China's trade and grow steadily over years. Figure 4.2B shows the evolution of wage premium by industry. There are large variations of industrial wage premia over time. In particular, agricultural wage premium rises suddenly between 2011 and 2012. This is due to agricultural subsidy programme since 2004 which reached peak in 2012. Also, there was a significant investment in agriculture made by Chinese government in 2011 along with "4 Trillion Yuan Fiscal Stimulus", which committed to raising land productivity and enhancing its quality. These interventions increased agricultural sector income. However, Bai et al. (2016) claimed that the long-run effect of the stimulus may be a permanent decline in aggregate productivity growth and resource allocation. This may possibly the reason why the agricultural wage premium declined after the peak in 2011 and 2012. Over the periods from 2003 to 2017, although the trends fluctuate widely, the wage premium of non-tradable sectors rises gradually while that of tradable sectors remains steady.

Next section we will use empirical estimations to examine the relationship between trade openness and wage premium.

4.6 Estimation Results

In this section, we present the estimated effect of trade openness on wage premium, as measured by trade shares and tariff rates.

4.6.1 Baseline Results

Table 4.1 shows the estimated relationship between trade openness and wage premium. In the sample, we have 30 provinces, 15 industries across 6 years. Due to some missing values, in the end we have approximately 2200 observations. The empirical results show that trade openness has a positive effect on wage premium and the coefficients of trade openness are significant at the 1% level. Column 1 suggests that a 1% increase in trade openness leads to 0.02% increase in wage premium using OLS estimation. We also include control variables such as value-added share and gross capital formation share which might affect wage premia and are correlated with trade shares. In column 2, 2SLS estimation results confirm the positive relationship between trade openness and wage premium. Table C4.3 in the Appendix reports the first stage estimation results of 2SLS. The Cragg-Donald Wald F statistics is reported in the last row. For each of the specifications, the weak instrument hypothesis is rejected. Comparing with column 1, the magnitude of the coefficient of trade openness is much higher that it is with OLS. The two-stage least squares results suggest that OLS estimations may be underestimated. Although the magnitudes are different between OLS and 2SLS, the key message remains. That is, trade openness has positive effects on industrial wage premia.

We also use tariff rate as a measure of trade openness and the coefficient on tariff is significantly negative, implying a positive effect of tariff cuts on wage premium. The lower tariffs imposed on China’s exports, the higher the trade openness on Chinese goods. Column 3 suggests that a 1% decrease in tariff rates leads to 0.02% increase in wage premium.

For the control variables, the coefficients of gross capital formation are insignificant in all the three model specifications. A possible explanation for this is that the effects of capital formation on wage premium may be offset across sectors. For some sectors, the effects of capital formation are positive, whereas there might be negative effect of capital formation on wage premium due to automation adoptions (Acemoglu, 2024). In other words, capital abundance leads to greater automation which means that capital becomes more important, and labour becomes less important, thus, wages are reduced. Another reason of the insignificant coefficient on capital formation could be that the effect of capital formation on wage premium is absorbed by the effect of trade openness due to complementary relationship between trade openness and capital formation (Banday et al., 2021; Hao, 2023). The coefficient of value added is positive and significant in column 2 using 2SLS estimations, whereas the coefficient is negative in column 3 using the measure of tariff barriers. This makes the impact of value added on wage premium ambiguous. As suggested by Combe et al. (2015), wages are proportional to productivity and also depend on the bargaining power and monopsony power of the firms. Wages might be mark down on marginal products in less competitive markets. Thus, the effect of value added is more likely to be ambiguous.

Table 4.1 Trade Openness and Wage Premium (2003-2017)

	OLS (1)	2SLS (2)	OLS (2)
ln_openness	0.0216***	0.1064***	
	(0.007)	(0.031)	
ln_tariff			-0.0220*
			(0.012)
ln_gross_capital	0.0001	-0.0003	0.00001
	(0.001)	(0.001)	(0.001)
ln_value_added	-0.0211	0.1339**	-0.0477
	(0.052)	(0.067)	(0.051)
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Constant	0.5240***	0.9042***	0.3347***
	(0.064)	(0.145)	(0.088)
Obs	2,290	2,149	2,313
Adj R-squared	0.166	0.112	0.160

Notes: The dependent variable is province-industry wage premium which is computed with a two-stage estimation approach. The key variable of interest in column 1 is the logarithm of trade share in gross output. This variable as discussed in the above context is endogenous. Column 2 is the 2SLS second-stage estimations. In column 3, the key variable is the logarithm of constructed tariff barriers. Other control variables are natural logarithm of value-added share and logarithm of gross capital formation in gross output. Industry fixed effect and year fixed effect are also controlled in the regressions. Robust standard errors are clustered at province-region level. *** p<0.01, ** p<0.05, * p<0.1.

4.6.2 Tradable and Non-tradable Sectors

The analysis in section 4.6.1 focuses on the average effect of trade openness on the wage premium. In this section, we explore the heterogeneity across sectors. We separate the sample into tradable and non-tradable sectors and examine whether there are significant effects from trade openness in the tradable and non-tradable sectors. According to the Balassa-Samuelson effect, greater productivity growth in tradable sectors leads to wage rises in tradable sectors. If the labour market is integrated, non-tradable sectors also need to increase their wages to prevent workers from switching sectors to look for new job opportunities and higher wages. Table 4.2 shows that the significantly positive effect of trade openness is consistent in both tradable and non-tradable sectors. Columns 1 and 2 are 2SLS estimation results where the trade share is instrumented with the trade partner's weighted gross output. Column 1 suggests a positive effect of trade openness on wage premium in the tradable sectors. Column 3 shows a positive effect of tariff cut on wage premium in the tradable sectors. For non-tradable sectors, column 2 shows a positive effect of wage premium. It should be noted that, the coefficient of export tariff in column 4 is positive but insignificant. Overall, however, the results provide evidence in favour of the effect of trade openness on the tradable and non-tradable sectors' wage premia. For the control variables of gross capital formation and value added, they are insignificant. The only exception is that the coefficient of the value added in column 3 with the samples of tradable sectors is negative and significant. One possible explanation is that Chinese exporting specialises in low-skilled, labour-intensive manufacturing. Tsou et al. (2006) find that exporting activities are more likely to have negative wage premium for low skilled workers. Given the samples in column 3 are only tradable sectors, there might be wage penalty for exporters in China.

Table 4.2 Subsamples of Tradable and Non-tradable Sectors

	Tradable (1)	Non-tradable (2)	Tradable (3)	Non-tradable (4)
ln_openness	0.1021***	0.1247*		
	(0.032)	(0.068)		
ln_tariff			-0.0503***	0.0085
			(0.019)	(0.016)
ln_gross_capital	-0.0009	0.0002	-0.0009	0.0003
	(0.001)	(0.001)	(0.001)	(0.001)
ln_value_added	0.0605	0.1220	-0.3066***	-0.0065
	(0.161)	(0.080)	(0.109)	(0.059)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Constant	0.8698***	0.7677***	0.0446	0.4561**
	(0.195)	(0.250)	(0.122)	(0.213)
Obs	1,103	1,046	1,103	1,210
Adj R-squared	0.127	0.099	0.186	0.160

Notes: The samples are separate into tradable and non-tradable sectors. The dependent variable is province-industry level wage premium for 6 years, from 2003 to 2017. Share of value added and gross fixed capital are expressed in logarithm. Columns 1 and 2 are 2SLS second stage estimation results. Columns 3 and 4 are estimation results using the measure of export tariff barriers. Industry fixed effect and year fixed effect are also controlled in the regressions. We identify tradable and non-tradable industries based on the classification criteria proposed by Jensen and Kletzer (2005). Robust standard errors are clustered at province-region level. *** p<0.01, ** p<0.05, * p<0.1.

4.6.3 Effects of Tradable on Non-tradable Sectors

The estimates in section 4.6.3 tell us there is positive effect of trade openness on both tradable and non-tradable sectors. It is obvious that trade openness has directly impact on tradable sectors. The non-tradable sectors, however, can be affected in an indirect way due to productive linkages to the tradable sectors. Also, with trade openness deepening, the increased productivity and labour demand in tradable sectors will raise wages which will eventually pass on increasing spending on non-tradable sector and therefore a rise in relative wages of non-tradable goods and services. Table 4.3 reports the effect of wage premia of tradable sectors on wage premia of non-tradable sectors in provincial level and examine whether the effect varies across different *hukou* stringency. Following Frocrain and Giraud (2018), we construct provincial level wage premia as the weighted average wage premia of each industry in tradable/non-tradable sector using the following equation:

$$W_{p,t}^{T} = \sum_{i=1}^{8} w_{i,p,t}^{T} v_{i,p,t}^{T} \qquad (4.20)$$

$$W_{p,t}^{NT} = \sum_{i=1}^{7} w_{i,p,t}^{NT} v_{i,p,t}^{NT} \qquad (4.21)$$

We have S= {T, NT}, $v_{i,p,t} = \frac{VA_{i,p,,t}}{VA_{S,p,,t}}$. $VA_{i,p,,t}$ is the value added at current price for each industry i in sector S in province p at time t. $VA_{S,p,,t}$ is the value added at current price for sector S in province p at time t. The ratio $v_{i,p,t}$ is the proportion of the total value added by sector S that is contributed by industry i. Intuitively, it indicates the relative importance of each industry within each sector based on its contribution to the total value added. $w_{i,p,t}$ is the wage premia for each industry i in sector S in province p at time t.

We start from a simple regression in Panel A. As expected, the coefficients of wage premia of tradable sector are consistently statistically significant and positive without and with control variables and fixed effects. A one percent increase in wage premia of tradable sectors leads to 0.25 percent increase in non-tradable wage premia. The standard trade theory such as H-O model assumes perfect factor mobility and thus equalising the relative prices and wages across regions and sectors. However, the graphical demonstration of wage premia suggests regionally and sectorally disparate wages. In reality, the factor adjustments tend to be slow, costly and heterogenous across regions and sectors. In China, stringent *hukou* restrictions make it prohibitively expensive or even impossible for people to move. The *hukou* stringency restriction varies across regions. We use *hukou* the index constructed by Zhang et al. (2019)[30] to measure the stringency of *hukou* regulations across Chinese provinces. The *hukou* index in Table 4.3 Panel B ranges from 0.2451 (Gansu Province) to 2.496 (Beijing City) in 2000-2013, and from 0.3507 (Shanxi Province) to 2.6284 (Beijing City) in 2014-2016, with a higher value indicating more stringent restrictions and thus more difficult for migrant to obtain local *hukou*. In panel B, the interaction term of tradable sector wage premia and *hukou* stringency index estimates the heterogeneity effect of tradable sector wage prima with different degrees of *hukou* restrictions. The coefficient of the interaction term in column 1 of panel B is statistically insignificant. It may imply there is no regional difference of the effect of tradable on non-tradable sectors.

Recognising the stringency index distribution may have different impact on the relationship between tradable and non-tradable sectors, following Gao et al. (2023), we divide provinces into two groups based on the stringency of *hukou* index. Highly restrictive provinces are those where the *hukou* index is above the national mean, and less restrictive ones are those in which *hukou* index is below the national mean. In column 2 and 3 of panel B, then, we repeat the investigation of the impact of tradable on non-tradable wage premia by separating the provinces into highly restrictive and less restrictive ones. The results show that the impact of tradable on non-tradable wage premia increases with more restrictive *hukou* policies, especially within highly restrictive province groups. The coefficient of the interaction term in column 2 of panel B are insignificant with expected positive sign, which may explain why the estimation in column 1 of panel B with total samples did not reveal any significant relationship between tradable and non-tradable sectors. For robustness check, in Panel C, we use

[30] Zhang et al. (2019) construct city-level hukou index for the period of 2000-2013 and 2014-2016 for 120 cities. We use the average of city-level index of each province as province-level hukou index. In this chapter, we use two measures of hukou index constructed by Zhang et al. (2019).

alternative *hukou* measures[31] constructed by Zhang et al. (2019). The estimations show consistent results with panel B.

Having examined the effect of tradable sector wages on non-tradable sector wages, we further investigate whether trade openness contributes to decreasing wage gap in tradable and non-tradable sectors and whether the effect changes according to different *hukou* restrictions.

In the traditional Heckscher-Ohlin model and its companion Stolper-Samuelson theorem, countries will export goods that intensively use production factors that is relatively abundant, and trade openness raises the returns to the abundant factor. With the assumption of perfect factor mobility, the H-O model predicts the returns of factors of productions are equalized across regions/sectors. Another trade theory highlighting firm heterogeneity developed by Melitz (2003) explain that the trade expansion leads to large productivity growth in tradable sector, thus affects the wage differential across tradable and non-tradable sectors. There are numerous channels through which trade potentially affect inter-industry wage dispersion. In this section, we emphasise the effect of institutional factor rigidities in the labour market especially due to restricted geographically mobility. We use reduced-form estimation to examine the relationship between trade openness and wages across regions with different stringency of *hukou* restrictions.

The regression specification is as follows:

$$\ln(w)_{i,p,t} = \beta_0 + \beta_1 tradability_i + \beta_2 openness_{i,p,t} + \beta_3 tradability_i \times openness_{i,p,t} + \varphi X_{i,p,t} + \alpha_{pt} + \varepsilon_{i,p,t} \qquad (4.22)$$

where $\ln(w)_{i,p,t}$ is the log level of wages, $tradability_i$ is a dummy variable with 1 denoting non-tradable sectors, $openness_{i,p,t}$ is a measure of trade openness, $X_{i,p,t}$

[31] The alternative hukou index ranges from 0.0802 (Guizhou Province) to 0.7813 (Beijing City) in 2000-2013 and from 0.0827 (Shanxi Province) to 0.7947 (Beijing City). Zhang et al. (2019) different methods to construct hukou index for 120 cities. The Projection Pursuit Model (PPM) is used to constructed hukou index which is employed in Table 3 Panel B. The entropy method is an alternative methodology to construct hukou index which is employed in Table 3 Panel C. Both of hukou index measurements identify the same high/less restrictive regions.

denotes a vector of control variables such as the share of gross capital formation (GCF) and the share of value added in gross output, α_{pt} is province-year fixed effects to control common effect on wages across provinces, $\varepsilon_{i,p,t}$ is the error term.

According to a classic “Balassa-Samuelson” effect, productivity growth in tradable sectors increase the wages of tradable workers as prices of tradable goods and services are set in international markets. While the productivity of non-tradable sectors has remained the same, firms in the non-tradable sectors have to increase the wages to prevent their employees moving to tradable sectors. With highly intersectoral mobility, the wages of non-tradable sectors increase significantly to align with tradable sector wages.

Since the 1990s, China implemented state-sector reforms which abolished state-assigned life-time employment and embarked market-driven system which brought about wage decentralisation. In this case, the labour mobility across sectors is relatively high and we do not focus on investigating the sectoral mobility underlying the wage-trade openness link. Topalova (2010) provided empirical evidence that geographical mobility may distort the relationship between trade liberalisation and wage inequality in developing countries. With respect to Chinese case, regions with high restrictive *hukou* policies have to push wages of non-tradable sectors even higher to attract people to work here. By comparison, in low restrictive regions, with lower migration costs, migrants especially rural-urban migrants account for a large proportion of labour that is normally paid less than local workers. In this case, non-tradable sectors can hire migrants with lower labour cost instead of raising relative wages to attract local workers.

Following the strategy suggested by Topalova (2010), we split province-industry level samples into two groups based on the stringency of *hukou* index and estimate the heterogeneity effect of trade openness on wage gaps. Table 4.4 reports the estimation results. Columns 1 and 2 report the results with the measure of trade openness in terms of share of trade in gross output for high and low restrictive regions, respectively. They show that non-tradable sectors earn less than tradable sectors. The coefficient of the interaction term in column 1 is positive while that of interaction term in column 2 is negative. It indicates that non-tradable sector wages are higher in high restrictive regions comparing to tradable sector wages. By contrast, the non-tradable wages are lower than tradable wages in low restrictive regions. The estimated evidence is consistent with our hypothesis: the relative wages of non-tradables fall less (rise more) as the degrees of interregional labour mobility decreases.

Columns 3 and 4 report estimated results using tariffs faced by China's exporter as a measure of trade openness. The coefficients of the interaction terms between tradability and tariff are both negative. As a lower tariff is associated with higher trade openness, the negative sign of the coefficients indicates the wages increase with higher trade openness. By computing the different marginal effects of export tariffs in tradable and non-tradable sectors, we can see that the non-tradable workers benefit largely, and the wages increase even more, especially in high restrictive regions. Also, the wage gap between tradable and non-tradable sectors is smaller in high restrictive regions compared to low restrictive regions.

These estimation results provide indicative evidence on explaining the heterogeneous effect trade openness has on wage gap between tradable and non-tradable sectors. Overall, this section provides suggestive evidence of the effect of tradable on non-tradable sectors and of the contributions trade openness makes to industry wages.

Table 4.3 The Effect of Tradable on Non-tradable

	Dependent variable: wage premia of non-tradable sectors		
	(1)	(2)	(3)
Panel A: The effect of tradable sector on non-tradable sector			
wage_premia_tradable	0.3414***	0.2410**	0.2497**
	(0.083)	(0.101)	(0.102)
ln(share_value_added)			0.1115
			(0.250)
ln(share_gross_captial)			-0.0522
			(0.034)
Year FE	No	Yes	Yes
Province FE	No	Yes	Yes
Constant	0.0198	0.0260	0.0416
	(0.021)	(0.016)	(0.221)
Observations	171	171	171
Adjusted R-squared	0.104	0.466	0.463

Panel B: Heterogeneity across provinces with different degrees of *hukou* restrictions

	Whole sample	Less restrictive	High restrictive
wage_premia_tradable	0.1753	-0.2499	-0.3523*
	(0.163)	(0.911)	(0.204)
wage_premia_tradableX*hukou*	0.1096	0.9239	0.3477***
	(0.150)	(1.624)	(0.117)
ln(share_value_added)	0.1248	0.0125	0.2874
	(0.250)	(0.300)	(0.444)
ln(share_gross_captial)	-0.0574*	0.1145	-0.1014***
	(0.034)	(0.175)	(0.031)
Year FE	Yes	Yes	Yes
Province FE	Yes	Yes	Yes
Constant	0.0411	0.1935	0.2129
	(0.223)	(0.378)	(0.420)

Observations	171	127	44
Adjusted R-squared	0.460	0.330	0.753

Panel C: Heterogeneity effect with alternative *hukou* stringency index

	Whole sample	Less restrictive	High restrictive
wage_premia_tradable	0.1658	-0.0635	-0.2934
	(0.150)	(0.793)	(0.184)
wage_premia_tradableX*hukou*	0.4850	2.3018	1.1516***
	(0.480)	(5.627)	(0.311)
ln(share_value_added)	0.1276	0.0613	0.3682
	(0.249)	(0.298)	(0.450)
ln(share_gross_captial)	-0.0577*	0.1114	-0.0990***
	(0.034)	(0.179)	(0.030)
Year FE	Yes	Yes	Yes
Province FE	Yes	Yes	Yes
Constant	0.0417	0.2292	0.3078
	(0.222)	(0.375)	(0.415)
Observations	171	123	48
Adjusted R-squared	0.461	0.308	0.751

Notes: The dependent variable for panel A, B, C is the wage premia of non-tradable sector. The key variable of interest is wage premia of tradable sector. Control variables are logarithm of share of value added and gross fixed capital in gross output respectively. All the variables are constructed in provincial level. Provincial level gross output, value added, and gross capital formation are calculated by aggregating each industry in each province at time t. Year fixed effect and province fixed effect are also controlled. Wage premia in sector S={T,NT} is $W_{p,t}^{S} = \sum_{i \in S} w_{i,p,t} v_{i,p,t}$, where $v_{i,p,t} = \frac{VA_{i,p,,t}}{VA_{s,p,t}}$. Panel A estimates the effect of wage premia of tradable sector on wage premia of non-tradable sector. Panel B and C estimate the heterogeneity effect of wage premia in tradable sector on non-tradable sector with different degrees of *hukou* stringency, using two different stringency indexes. Robust standard errors are clustered at province level. *** p<0.01, ** p<0.05, * p<0.1. In panel B, the *hukou* index ranges from 0.2451 (Gansu Province) to 2.496

(Beijing City) in 2000-2013, and from 0.3507 (Shanxi Province) to 2.6284 (Beijing City) in 2014-2016. In panel C, the *hukou* index ranges from 0.0802 (Guizhou Province) to 0.7813 (Beijing City) in 2000-2013 and from 0.0827 (Shanxi Province) to 0.7947 (Beijing City).

Table 4.4 Trade Openness and Wage Gap

	(1) High Restrictive	(2) Low Restrictive	(3) High Restrictive	(4) Low Restrictive
Non-tradable	-0.4144*	0.3545**	-0.0486	-0.0709**
	(0.249)	(0.156)	(0.049)	(0.033)
Openness	-3.1178*	4.3712**		
	(1.758)	(1.846)		
Non-tradable Openness	3.8241*	-8.8851***		
	(2.220)	(3.273)		
Exporter tariff			-7.4158***	-4.4895***
			(1.028)	(0.778)
Non-tradable Exporter tariff			-7.6119	-10.0588*
			(6.262)	(5.652)
Control variables	Yes	Yes	Yes	Yes
Constant	8.7277***	8.5285***	10.3621***	10.0244***
	(0.418)	(0.136)	(0.032)	(0.023)
Observations	558	1,514	558	1,515
Adjusted R-squared	0.489	0.380	0.665	0.526

Notes: The dependent variable is ln(wage). All regressions include a tradability dummy for being a non-tradable sector, trade openness indicator, interaction term between tradability dummy and trade openness indicator, and province-year fixed effects. Trade openness is measured by trade share in gross output and export tariff barriers. Highly restrictive provinces are those in which *hukou* index is above the national mean, and less restrictive ones are those in which *hukou* index is below the national mean. Columns 1 and 2 are 2SLS second stage estimations. The instrumental variable for trade share in gross output are weighted average of the trade partners' industrial gross output. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

4.7 Consumer Preferences

The empirical results in section 4.6 indicate that trade-induced shock have positive effect on industry wage premia (which is also referred to industry-specific rents). And the positive effect on tradable sectors feed into non-tradable sectors. This phenomenon can be explained using “Balassa-Samuelson” effect. According to this effect, the productivity growth in tradable sectors due to trade openness translates into wage rises in tradable sectors. Firms in non-tradable sectors will increase wages in order to attract non-tradable workers. The impact of productivity shock in tradable sectors on relative wages in non-tradable sectors closely depend on labour mobility. Specifically, the previous results in this chapter attempt to examine the effect of tradable on non-tradables through the perspectives of labour supply and demand. As mentioned in section 4.6.3, firms in the non-tradable sector have to increase wages to prevent their employees from looking for work in the tradable sector where wages are higher. That is, the increase of wages in non-tradable sectors depend on the elasticities of local labour. Moretti (2010) suggest consumer preferences for non-tradables is also important, which can be used to explain the impact of tradables on non-tradables.[32] There are more local jobs and higher wages due to trade openness, so that the demand for non-tradable sectors such as restaurant, hairdressing, cleaning services, retails might increase. That is, if consumers have strong preferences on non-tradable products, then the trade-induced wage increase in tradable sectors will benefit the non-tradable sectors.

In this section, we examine another mechanism that can explain why the impact on tradable sectors feeds into non-tradable sectors. Following the empirical specification proposed on Liao (2020), We test whether high income will lead to higher demand on non-tradable products. We use three indicators to measure the demand of non-tradable sectors. The first is the expenditure share of non-tradable sectors. Liao (2020) shows graphical evidence of the increasing expenditure share of dining out and household services. Also, in now a classic paper in the literature, Moretti (2010) provides empirical findings that the shock in tradable sectors have positive effect on

[32] Other papers such as Frocrain and Giraud (2018), Liao (2020) also indicate higher income increases the demand for service sectors.

the employment of non-tradable sectors. Lastly, we construct value added share of GDP in non-tradable sectors.

The regression specification is as follows:

$$non\_tradable_{pt} = \alpha_o + \alpha_1 income_{pt} + \gamma Z_{pt} + \gamma_t + \gamma_p + \varepsilon_{pt} \quad (4.23)$$

where $t = 2003, 2008, 2010, 2012, 2015, 2017,$ which is consistent with previous empirical investigations; $income_{pt}$ is log GDP per capita in each province each time period; Z_{pt} is control variables such as share of gross capital formation; γ_t is year fixed effect, γ_p is province fixed effect, ε_{pt} is error term. We collect expenditure data, employment, GDP, and GDP per capita from China Statistical Yearbook. Value added and gross capital formation for non-tradable sectors is collected from China Input-Output Table. Baumol (1967) suggested that the increase in the expenditure of labour-intensive services is largely due to an increase in their relative prices. To avoid the confounding effect of relative prices, we also collect price indexes to construct real values in expenditure, value added, GDP and GDP per capita.

Table 4.5 reports the estimated consumer preferences on non-tradable sectors. Column 1 shows the positive relationship between income and expenditure shares of non-tradable products. The coefficient on logarithm GDP per capita indicates that 1% increase in real income leads to 3.8% increase in expenditure share of non-tradable products. Also, the employment rate on non-tradable sectors will increase by 4.74%. Column 3 shows a positive effect on value-added shares as income rising, although the coefficient is not significant. Gross capital formation is controlled in the estimations. However, the coefficient of gross capital formation is not significant across three model specifications. One possible explanation is that the measures of capital formation may be over-aggregate and may hide the impact of capital formation that operates separately through tradable sectors and non-tradable sectors. The estimations in this section focus on preferences on consumers' side. Although the coefficient of capital formation is not significant, the positive and significant coefficients on GDP per capita provide suggestive evidence of consumer preferences on non-tradable sectors. In summary, as income increases, people want to purchase non-tradables, consequently, employment in non-tradable sectors such as restaurants, hairdressing, cleaning service grows, and value added grows.

Table 4.5 Estimated Consumer Preferences on Non-tradable Sectors

	(1) Expenditure share	(2) Employment share	(3) Value added share
Log GDP per capita	0.0377**	0.0474**	0.0035
	(0.018)	(0.021)	(0.015)
GCF share	-0.0493	-0.0183	-0.0089
	(0.073)	(0.073)	(0.083)
Year FE	Yes	Yes	Yes
Province FE	Yes	Yes	Yes
Observations	170	170	170
Adjusted R-squared	0.679	0.840	0.604

Notes: The dependent variables in column 1,2,3 are respectively real expenditure share, employment share, value added share of GDP. The key variable of interest is logged GDP per capita in each province in time t. The control variable is gross capital formation share of GDP. Year fixed effect and province fixed effect are also controlled. Robust standard errors in parentheses are clustered by provinces. *** p<0.01, ** p<0.05, * p<0.1.

4.8 Conclusions

An extensive literature has examined the impact of trade openness on wage inequality in China. This chapter investigates the effects of openness on industry wage dispersion in China. Three questions are addressed. The first is whether China's trade openness affects the industry wage premium. The second is whether the effect of trade openness vary by sectors. The final is whether tradable and non-tradable sector respond differently to trade openness. Following Goldberg and Pavcnik (2005) and Wang et al. (2018), we first estimate industry-level wage premia conditional on worker, firm, or job characteristics, and then estimate the relationship between trade openness and industrial wage premia using two-stage least square estimations. Two measures of trade openness are employed. One is the trade shares in gross output. The second is the tariffs faced by China exporters. The export tariff is unlikely to be predicted in advance and unrelated with economic factors in China, which is considered as exogenous.

We find that trade openness has positive impact on industry wage differentials and both tradable and non-tradable sectors are affected positively by trade openness. In terms of export tariffs, as a lower tariff is associated with higher trade openness, the negative sign of the coefficients indicates the wages increase with higher trade openness. Moreover, we further investigate the indirect relationship between trade openness and the wage premium in non-tradable sectors, i.e., the transmission mechanism from tradable to non-tradable sectors. That is, we empirically test whether the Balassa-Samuelson theory can be reconciled with numerical evidence in terms of the Chinese case. The finding is that the non-tradable sectors are affected in an indirect way due to productive linkages to tradable sectors. Also, taking into consideration imperfect mobility of labour across regions, we suggest that the impact of tradable sectors on non-tradable sectors is most pronounced in more restrictive regions. Finally, our estimates provide evidence on the role of consumer preference, which can be another important channel to explain the impact of tradables on non-tradables.

Overall, this chapter suggests that trade-induced shocks spill over beyond tradable sectors directly exposed to globalisation. The benefits of trade liberalisation can be better achieved if the spatial allocation of factors, especially labour, is efficient. Also, it is important to study the rise of traditional non-tradable industries as China's development process has entered a new stage in which service sectors play an increasingly important role.

Appendix C

Table C4.1 First-stage Estimation for Wage Premium

	2003 ln_income	2008 ln_income	2010 ln_income	2012 ln_income	2015 ln_income	2017 ln_income
male	0.1497***	0.3273***	0.3508***	0.3380***	0.4008***	0.2730***
	(0.028)	(0.035)	(0.029)	(0.027)	(0.037)	(0.033)
age	0.0019	0.0193*	0.0482***	0.0522***	0.0429***	0.0558***
	(0.011)	(0.010)	(0.009)	(0.007)	(0.011)	(0.009)
age_squr	0.0000	-0.0003**	-0.0006***	-0.0007***	-0.0005***	-0.0006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban_*hukou*	-0.1339**	0.2082***	0.0550	-0.0217	0.0286	0.0688*
	(0.059)	(0.041)	(0.035)	(0.031)	(0.044)	(0.039)
married	0.1477***	0.0314	0.0894**	0.1608***	0.1429***	0.1889***
	(0.046)	(0.050)	(0.041)	(0.038)	(0.049)	(0.042)
Education						
primary_school	0.0609	-0.0391	0.3283***	0.2038**	0.3545**	-0.0837
	(0.145)	(0.142)	(0.108)	(0.096)	(0.153)	(0.135)
Junior_high_school	0.2274*	0.1591	0.5744***	0.4760***	0.5119***	0.1139
	(0.138)	(0.138)	(0.105)	(0.092)	(0.148)	(0.127)
Senior_high_school	0.4077***	0.4358***	0.7618***	0.6290***	0.6661***	0.4338***
	(0.139)	(0.139)	(0.107)	(0.094)	(0.149)	(0.127)
college	0.6717***	0.7248***	1.1029***	0.9478***	1.0237***	0.7144***
	(0.141)	(0.146)	(0.113)	(0.099)	(0.155)	(0.130)
university	0.9983***	0.8271***	1.3963***	1.1163***	1.1819***	1.0481***
	(0.147)	(0.151)	(0.117)	(0.102)	(0.157)	(0.131)
master or higher	1.2602***	1.0887***	1.7148***	1.5342***	1.4790***	1.4951***
	(0.208)	(0.229)	(0.156)	(0.155)	(0.183)	(0.150)
communist	0.0604*	0.1054**	0.0877**	0.0827**	-0.2366***	0.0790*
	(0.036)	(0.052)	(0.042)	(0.039)	(0.076)	(0.046)

Han ethnicity	-0.0335	0.0617	0.0064	0.0580	-0.0038	0.0414
	(0.066)	(0.078)	(0.065)	(0.060)	(0.091)	(0.078)
Ownership						
collective	-0.2279***	-0.0926	-0.1297**	-0.0898	-0.0914	0.0200
	(0.054)	(0.066)	(0.057)	(0.055)	(0.076)	(0.062)
private	0.0610	0.0171	-0.0091	0.0444	0.0604	0.1276***
	(0.042)	(0.046)	(0.042)	(0.039)	(0.043)	(0.039)
foreign_invest	0.2689***	0.0815	0.2621***	0.1579**	0.2231**	0.2093***
	(0.095)	(0.086)	(0.090)	(0.077)	(0.096)	(0.079)
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
provinces#industry	Yes	Yes	Yes	Yes	Yes	Yes
Constant	8.8630***	7.7743***	8.1640***	8.9395***	9.2162***	8.8923***
	(0.416)	(0.726)	(0.356)	(0.439)	(0.490)	(0.745)
Observations	2,393	2,103	3,163	3,571	2,033	2,512
R-squared	0.413	0.433	0.475	0.432	0.468	0.457

Notes: The dependent variable is ln (yearly wage). The individual characteristics and job-related variables are included in the regressions. The reference group for industry dummies is “Other industries”. The reference group for province dummies is “Shanghai”. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Data Construction for International Imports

In China Input-Output tables, international import is divided into import used as intermediate use and import used as final use. There are five categories of final use, which are import used as final use separated into. Therefore, we cannot directly obtain province-industry level import values. In the tables, we can see province-industry level import used as intermediate use. Following the construction method proposed by the China Information Center, we compute import used as final use based on the proportion of import used as intermediate demand among provinces and sectors. After that, we sum province-industry level of import used as intermediate and final use to obtain total import values.

Table C4.2 Classification of Tradability

number	industry	tradability
1	Agriculture, forestry, animal husbandry and fishery	tradable
2	Mining	tradable
3	Manufacturing	tradable
4	Electricity, gas and water	non-tradable
5	Construction	non-tradable
6	Transport and information	tradable
7	Wholesale and retail, hotel and restaurants	non-tradable
8	Financial intermediation	tradable
9	Real estate	tradable
10	Households and business services	non-tradable
11	Health, sports and social welfare	non-tradable
12	Education, culture and broadcast	tradable
13	Scientific Research	tradable
14	State and social organisation	non-tradable
15	Others	non-tradable

Table C4.3 2SLS First-stage Estimation

	The dependent variable is *ln* (trade share)		
	Full sample	Tradable	Non-tradable
ln_GOI	0.324***	0.469***	0.203**
	(0.0457)	(0.0432)	(0.0638)
ln_gross_capital	0.007**	0.006	0.005
	(0.0031)	(0.0047)	(0.0037)
ln_value_added	-1.293***	-2.941***	-0.6811***
	(0.2421)	(0.7559)	(0.080)
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Constant	-3.781***	-4.505***	-3.402***
	(0.1618)	(0.4160)	(0.1898)
Observations	2,149	1,103	1,046
Adj R-squared	0.274	0.397	0.131
F statistic	124.78	99.67	32.01

Notes: Column 1 is 2SLS first stage estimations for whole samples. Columns 2 and 3 are first stage estimation for tradable and non-tradable sectors respectively. The dependent variable is the logarithm of trade share in gross output. The instrument variable is the logarithm of weighted average gross output of trade partners. Other control variables are natural logarithm of value-added share and logarithm of gross capital formation in gross output. Industry fixed effect and year fixed effect are also controlled in the regressions. Robust standard errors are clustered at province-region level. Weak ID test uses Cragg-Donald Wald F statistics. *** p<0.01, ** p<0.05, * p<0.1.

Full Derivatives of Theoretical Framework

The cost function is

$$c(W^S, R^S) = \min_{L^S, K^S} W^S L^S + R^S K^S \qquad (4.24)$$

$$\text{s.t. } Y^S(L^S, K^S) = A^S (L^S)^{\theta s} (K^S)^{1-\theta s} \qquad (4.25)$$

where W is the wage of labour L, R is the return to capital K.

To minimise the cost function, we first formulate a lagrangian function:

$$\mathcal{L}(W^S, R^S, L^S, K^S, Y^S) = W^S L^S + R^S K^S - \lambda^S\left[A^S (L^S)^{\theta s} (K^S)^{1-\theta s} - Y^S\right] \qquad (4.26)$$

We solve this minimisation constraint by the lagrange multiplier method:

$$\frac{\partial \mathcal{L}}{\partial L^S} = W^S - \lambda^S A^S \theta s (L^S)^{\theta s - 1} (K^S)^{1-\theta s} \qquad (4.27)$$

$$\frac{\partial \mathcal{L}}{\partial K^S} = R^S - \lambda^S A^S (1-\theta s) (L^S)^{\theta s} (K^S)^{-\theta s} \qquad (4.28)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = Y^S - A^S (L^S)^{\theta s} (K^S)^{1-\theta s} \qquad (4.29)$$

Solve for L^S:

$$\frac{\frac{\partial \mathcal{L}}{\partial K^S}}{\frac{\partial \mathcal{L}}{\partial L^S}} = \frac{R^S}{W^S} = \frac{\lambda^S A^S (1-\theta s)(L^S)^{\theta s} (K^S)^{-\theta s}}{\lambda^S A^S \theta s (L^S)^{\theta s - 1} (K^S)^{1-\theta s}} \qquad (4.30)$$

$$\frac{R^S}{W^S} = \frac{(1-\theta s)}{\theta s} \frac{L^S}{K^S} \qquad (4.31)$$

$$L^S = \frac{R^S}{W^S} \frac{\theta s}{(1-\theta s)} K^S \qquad (4.32)$$

Substitute L_S into production function in order to solve K^S:

$$Y^S = A^S (L^S)^{\theta s} (K^S)^{1-\theta s} \qquad (4.33)$$

$$Y^S = A^S \left(\frac{R^S}{W^S}\frac{\theta s}{(1-\theta s)}K^S\right)^{\theta s}(K^S)^{1-\theta s} \qquad (4.34)$$

Simplify and solve for K^S

$$K^S = \frac{Y^S}{A^S}\frac{[W^S(1-\theta s)]^{\theta s}}{(R^S\theta s)^{\theta S}} \qquad (4.35)$$

Now we can solve L^S by substituting for K^S:

$$L^S = \frac{Y^S}{A^S}[\frac{R^S\theta s}{W^S(1-\theta s)}]^{1-\theta s} \qquad (4.36)$$

Substitute L^S and K^S into the cost function:

$$c = \frac{Y^S}{A^S}(W^S)^{\theta s}(R^S)^{1-\theta s}\theta s^{-\theta s}(1-\theta s)^{\theta s-1} \qquad (4.37)$$

The average price level under perfect competition is equal to the marginal cost of production:

$$P^s = \frac{c}{Y^S} = (A^S)^{-1}(W^S)^{\theta s}(R^S)^{1-\theta s}\theta s^{-\theta s}(1-\theta s)^{\theta s-1} \qquad (4.38)$$

where W^s and R^s are the unit cost of labour and rental rate of capital, respectively. If we consider a small open economy with perfect capital mobility, the traded goods are then set by international prices $P^T = P^{T*}$ and the rate of return on capital is equal to the world value $R^T = R^{NT} = R^*$.

Normalising the tradable price to unity, and taking the tradable goods as numeraire, the relative price of non-tradables in terms of tradables can be written as:

$$\frac{P^{NT}}{P^T} \equiv P = \frac{\Psi^{NT}}{\Psi^T}\frac{A^T}{A^{NT}}\frac{(W^{NT})^{\theta NT}(R^{NT})^{1-\theta NT}}{(W^T)^{\theta T}(R^T)^{1-\theta T}} \qquad (4.39)$$

where $\Psi^s = (\theta s)^{-\theta s}(1-\theta s)^{\theta s-1}$.

For tradable sectors,

$$P^T = \Psi^T A^{-T}(W^T)^{\theta T}(R^T)^{1-\theta T} \qquad (4.40)$$

where $\Psi^s = (\theta s)^{-\theta s}(1-\theta s)^{\theta s-1}$.

We assume that the unit cost for producing tradables is equal to 1, then

$$1 = \Psi^T A^{-T}(W^T)^{\theta T}(R^T)^{1-\theta T} \qquad (4.41)$$

We have

$$W^T = (A^T)^{\frac{1}{\theta T}}(\Psi^T)^{-\frac{1}{\theta T}}(R^T)^{-(\frac{1-\theta T}{\theta T})} \qquad (4.42)$$

Assuming the depreciation rate of physical capital is δ_K and the interest rate is r, we have $R^T = R^{NT} = R^*$, where $R^* = P(r+\delta_K)$.

We also have

$$W^{NT} = (A^T)^{\frac{1}{\theta NT}}(\Psi^{NT})^{-\frac{1}{\theta NT}}(R^{NT})^{-\left(\frac{1-\theta NT}{\theta NT}\right)} \qquad (4.43)$$

Substitute equation 4.43 into equation 4.39, now we can rewrite the relative price of non-tradable sectors, as follows:

$$P = \frac{\Psi^T}{(\Psi^{NT})^{\frac{\theta T}{\theta NT}}}\frac{A^T}{(A^N)^{\frac{\theta T}{\theta NT}}}(\frac{W^{NT}}{W^T})^{\theta T}(r+\delta_K)^{\frac{\theta T-\theta NT}{\theta NT}} \qquad (4.44)$$

Chapter 5 Final Remarks

This thesis provides three interrelated studies on the role of labour and goods frictions on local labour market in the Chinese cases. Chapter 2 examines the importance of migration restrictions in China by studying the role of the *hukou* system on shaping l migration decisions. Chapter 3 show the positive effect in migration from rural to urban areas on native workers' wages in Chinese cities, with effect being magnified for high-skilled urban workers. Chapter 4 demonstrates different effects of trade liberalisation on industry wage premia between different migration stringencies.

Chapter 2 investigates whether the *hukou* system is a barrier to internal migration and whether it enhances the responsiveness of migration to economic factors (such as wages and employment probabilities). This chapter contributes to the literature that analyses the interplay of migration policies and economic incentives and their migration effects in three ways. First, the theoretical framework and methodology incorporate origin and destination factors to model interregional migrations. The multi-regional migration data are constructed based on China Population Census. Second, a newly developed *Hukou* Registration Index is used to quantitatively analyse the trade-off between income and migration policy restrictions. To the best of our knowledge, this is the first study to empirically examine the role of migration policies with specific data measuring the stringency of local *hukou* policies in Chinese context. Third, we use plausible methods to control for endogeneity concerns between migration and wages to deliver more robustness estimations.

Chapter 2 demonstrates that migrants respond positively to destinations with higher wage differentials and respond negatively for unemployment. Also, increased migration restrictions in destinations underlines greater responsiveness of mobility decisions to economic factors. Beyond the traditional economic determinants of migration (e.g., wages and employment probabilities), other factors are also examined in multi-regional estimation models, such as public service and amenity index services. Notably, a principal component approach is used in order to de-factor the data and reduce multicollinearity issues when we construct public services and amenity services indexes. The estimation results show that the effect of public service on migrants gradually decreases with more stringent *hukou* regulations, while the impact of amenity services is positive. As already introduced in section 1.2, public services are only accessible to the residents who are entitled to local *hukou,* but the primary objective here is to investigate the migrants with non-local *hukou* who are not entitled to receive local public services. It is reasonable to find none or negative impact of public services in estimation process.

The insights in chapter 2 entail two important implications. First, this study finds empirical evidence of increases responsiveness of migration decisions to economic factors. This corresponds to the labour market efficiency. The continued or even faster growth in urban earnings is an important factor to achieve a successful urbanisation and industrialisation transition. Second, to achieve a more efficient labour allocation in response to growth opportunities, further deregulation of the *hukou* system is needed to encourage more labour mobility. This study has quantitative indications that can serve as guidance for policy makers to improve migration management and achieve rapid urbanisation and industrialisation targets. There are several potential avenues for future research. Chapter 2 uses cross-sectional dataset, and it could be interesting to expand the dataset to include dynamic migration data. In this case, the impact of the evolution of migration barriers on dynamic migrations can be analysed. Another possible avenue would be to compare and contrast the responsiveness of migration decisions based on different types of migrants, i.e., rural-to-urban migrants, urban-to-urban migrants, to economic conditions. Rural *hukou* holders are tied to be allocated land cultivation rights, that is, cultivate land for a return. The land policy needs to be taken into consideration for rural residents when making migration decisions, which may offset the effect of social welfare (e.g., public and amenity services).

Chapter 3 contributes to the literature concerning the impact of internal migration on native wages. The contribution to the existing literature is twofold. First, this chapter exploits the information on how the timing of policy changes varied across regions as the exogenous variation on the construction of instrumental variables of migration data. There is a growing literature analysing the impact of large-scale of rural-urban migration on wage outcomes of urban native workers. Most of them use previous settlement patterns as instruments, but none of them takes into consideration the impact of *hukou* on rural migration. This can be a particularly useful strategy. Second, this chapter contributes to the existing literature assessing the possible benefits of relaxing *hukou* restrictions. Most of the literature focus on aggregate analysis, for example, the analysis of relaxing mobility restrictions on income disparity and urbanisation (Au and Henderson,2006; Ngai et al.,2019; Whalley and Zhang, 2007), the impact on productivity and welfare allocation (Hsu and Ma, 2019; Tombe and Zhu, 2019), etc. No study has taken a micro perspective to directly examine how individual wages are affected by *hukou* and how a relaxation of the system would influence individual's wage adjustment, especially for urban workers. This study fills the gap.

The analysis consists of three steps. First, chapter 3 employs Oaxaca-Blinder decomposition to examine which factors that influence China's unequally wage structure between rural migrants and urban natives. The decomposition results indicate job segregation between rural-to-urban migrants and urban natives; thus, it is unlikely for migrants being substitutes of urban native workers. Second, a DID-style instrument variable is constructed to causally identify the effect of migrants on urban native workers. The identification strategy exploits variations in the relaxation of the internal migration restrictions across cities, and variations in the pre-reform migration flows across cities. Third, as a robustness check, a shift-share instrument is constructed where "the shift" is *hukou* reforms and "the share" historical migration patterns between cities.

The main finding in chapter 3 is that rural-urban migrants have a positive effect on the wages of urban native wages, with a larger effect on high-skilled urban natives. As pointed by the evidence provided in Oaxaca-Blinder decomposition, it is unlikely that migrants and natives are substitutes, but instead likely complementary. The positive migrant impact can be attributed to gains from complementarity with natives in the production function. Also, the migrant workers might have more complementing effect on high-skilled urban workers compared with low-skilled workers.

The findings offer relevant policy discussions and implications. First, the rural-to-urban migrants, being mostly concentrated in low-skilled labour extensive industries, feed into other industries, and in the end contribute to the overall productivity growth. Second, it is still highly restricted for migrants to work in certainty industries and occupations. Though some regions are attempting to relax the mobility barriers, it is still far from being a flexible labour market in China. The positive impact of migrants on native workers indicate that it is necessary to further mobility restrictions relaxation. The findings also provide indicative benefits of labour market reforms in other countries which implement policies to reduce geographical mobility.

Beyond the overall impact of migrants on wages across all types of workers which is demonstrated by the preliminary analysis in chapter 3, the specific impact of occupation and sectors could be further assessed. Also, apart from separating workers to particular skill groups, the estimation analysis could be extended along skill distributions over native workers. This extension analysis resembles the research by Dustmann et al. (2013), which allows for many skill points and assesses the effect of migrants on natives at each point in the native wage distribution. Another potential extension is to estimate the cumulative effect of *hukou* reforms on migration rather than a one-time effect when constructing migration instruments. There are a few caveats in migration data construction. The census data used in the analysis records the timing of departure from a migrant's place of registration rather than of arrival at destinations. It is not an issue if there is no step migration or return migration cases. Unfortunately, the step migration and return migration cases cannot be assessed with current data. In addition, the sampling frame of census survey is based on population registration. High-immigration areas could be under-sampled.

Chapter 4 studies the relationship between trade openness and industry-level wage premium. There are three main questions addressed. First, whether China's trade openness affects industry wage premium. Second, does the effect of trade openness vary by sector. Third, is it important to ask whether tradable and non-tradable sectors respond differently to trade openness. The contribution in chapter 4 lies in three aspects. First, this study focuses on industry effect of trade liberalisation, rather than the effect on the returns of particular worker characteristics. Second, it disaggregates sectors into tradable and non-tradable sectors and, in this case, insights into the direct and indirect influence of trade liberalisation can be gained. Third, it examines the role of labour mobility in determining the effect of trade on wages, which is among the first studies to do so.

Chapter 4 employs two measures of trade openness. One is the trade shares in gross output. The second is the tariffs faced by China exporters. By constructing industry-level wage premium conditional on worker, firm, or job characteristics, this chapter estimates the relationship between trade openness and industrial wage premia using two-stage least square estimations. After that, the sample is separated into tradable and non-tradable sectors in order to estimate the direct and indirect effect of trade openness. The transmissions from tradable to non-tradable sectors is also examined in order to assess whether the Balassa-Samuelson theory can be reconciled with numerical evidence in terms of the Chinese case, taking into consideration imperfect mobility of labour across region.

The results in chapter 4 suggests that trade openness has positive effect on industrial wage premium for both tradable and non-tradable sectors. Trade openness directly impact on tradable sectors. The non-tradable sectors are affected in an indirect way due to productive linkages with tradable sectors. Trade openness also contributes to decreasing the wage gap between tradable and non-tradable sectors, with a larger effect in more mobility restrictive regions. That is, the wage gap between tradable and non-tradable sectors is smaller in high mobility restrictive regions compared to low restrictive regions. Chapter 4 also provides suggestive evidence consumer preference, which can be another important channel to explain the impact of tradables on non-tradables.

The findings of chapter 4 offer several insights. The overall effect of trade liberalisation can be large, as the findings in chapter 4 suggest, the effect of trade liberalisation on tradable sectors and there is spillover effect from the tradable to the non-tradable sector. The second is that the benefits of trade liberalisation can be better achieved if the spatial allocation of factors especially labour is more efficient. The third insight delivered by the findings is that it is important to account for trade policy uncertainty generated by trade partners. It may contribute to the understanding of the changing environment of trade agreements. Finally, trade liberalisation can have heterogenous impact on the internal economic geography of the country. Differential growth across regions as well as different spatial allocation of economic activities may result from trade policy shocks.

There are some potential avenues for future research. Chapter 4 shows the overall effect of trade openness on labour markets and wages; however, different types of trade is not distinguished. Trade in intermediate and trade in final goods might have different impact on labour markets and wages. The study of the effect of international trade integration may be incomplete without a balanced analysis of both winners and losers from a variety of trade activities in globalisation deepening. Another potential extension is to examine the channels through which the impact on tradable sectors spills over to non-tradable sector, such as productive linkages, income-induced demand for local services, etc. Third, as China's development process has entered a new stage in which service sectors play an increasing role, it is important to study the causes and consequences of the rise of non-tradable industries.

References

Acemoglu, D. (2024) 'Capital and Wages', *National Bureau of Economic Research*, working paper series no. w32190.

Altonji, J.G. & Card, D. (1991) 'The effects of immigration on the labor market outcomes of less-skilled natives', in *Immigration, trade, and the labor market*. University of Chicago Press. pp. 201–234.

Alvarez, R. & López, R.A. (2005) 'Exporting and Performance: Evidence from Chilean Plants', *The Canadian Journal of Economics / Revue canadienne d'Economique*, 38(4), pp. 1384–1400.

Amiti, M. & Cameron, L. (2012) 'Trade Liberalization and the Wage Skill Premium: Evidence from Indonesia', *Journal of International Economics*, 87(2), pp. 277–287.

National Bureau of Statistics of China. (2001) 'China Statistics Yearbook'.

Au, C.-C. & Henderson, J.V. (2006) 'Are Chinese cities too small?', *The Review of Economic Studies*, 73(3), pp. 549–576.

Autor, D.H., Dorn, D. & Hanson, G.H. (2013) 'The China Syndrome: Local Labor Market Effects of Import Competition in the United States', *American Economic Review*, 103(6), pp. 2121–2168.

Aydemir, A. & Skuterud, M. (2008) 'The immigrant wage differential within and across establishments', *ILR Review*, 61(3), pp. 334–352.

Bao, S., Bodvarsson, Ö.B., Hou, J.W. & Zhao, Y. (2011) 'The Regulation of Migration in a Transition Economy: China's *Hukou* System', *Contemporary Economic Policy*, 29(4), pp. 564–579.

Bai, C.-E., Hsieh, C.-T. & Song, Z. (Michael) (2016) 'The Long Shadow of China's Fiscal Expansion', *Brookings Papers on Economic Activity*, 2016(2), pp. 129–181.

Banday, U.J., Murugan, S. & Maryam, J. (2021) ‘Foreign direct investment, trade openness and economic growth in BRICS countries: evidence from panel data’, *Transnational Corporations Review*, 13(2), pp. 211–221.

Baumol, W.J. (1967) ‘Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis’, *The American Economic Review*, 57(3), pp. 415–426.

Beine, M., Docquier, F. & Özden, Ç. (2011) ‘Diasporas’, *Journal of Development Economics*, 95(1), pp. 30–41.

Bernard, A.B. & Jensen, J.B. (1997) ‘Exporters, skill upgrading, and the wage gap’, *Journal of International Economics*, 42(1), pp. 3–31.

Bernard, A.B. & Wagner, J. (1997) ‘Exports and Success in German Manufacturing’, *Weltwirtschaftliches Archiv*, 133(1), pp. 134–157.

Bertocchi, G. & Strozzi, C. (2008) ‘International Migration and the Role of Institutions’, *Public Choice*, 137(1/2), pp. 81–102.

de Blasio, G. & Menon, C. (2011) ‘Local Effects of Manufacturing Employment Growth in Italy’, *Giornale degli Economisti e Annali di Economia*, 70 (Anno 124) (3), pp. 101–112.

Blinder, A.S. (1973) ‘Wage discrimination: reduced form and structural estimates’, *Journal of Human resources*, pp. 436–455.

Boehm, C.E., Levchenko, A.A. & Pandalai-Nayar, N. (2023) ‘The Long and Short (Run) of Trade Elasticities’, *American Economic Review*, 113(4), pp. 861–905.

Borjas, G.J. (1989) ‘Economic Theory and International Migration’, *The International Migration Review*, 23(3), pp. 457–485.

Borjas, G.J. (1999) ‘Immigration and Welfare Magnets’, *Journal of Labor Economics*, 17(4), pp. 607–637.

Borjas, G.J. (1987) ‘Self-Selection and the Earnings of Immigrants’, *The American Economic Review*, 77(4), pp. 531–553.

Borjas, G.J. (2003) ‘The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market’, *The quarterly journal of economics*, 118(4), pp. 1335–1374.

Borjas, G.J., Freeman, R.B., Katz, L.F., DiNardo, J. & Abowd, J.M. (1997) ‘How much do immigration and trade affect labor market outcomes?’, *Brookings papers on economic activity*, 1997(1), pp. 1–90.

Bosker, M., Brakman, S., Garretsen, H. & Schramm, M. (2012a) ‘Relaxing *Hukou*: Increased labor mobility and China’s economic geography’, *Journal of Urban Economics*, 72(2), pp. 252–266.

Bosker, M., Brakman, S., Garretsen, H. & Schramm, M. (2012b) ‘Relaxing *Hukou*: Increased labor mobility and China’s economic geography’, *Journal of Urban Economics*, 72(2–3), pp. 252–266.

Boustan, L.P. (2010) ‘Was postwar suburbanization “white flight”? Evidence from the black migration’, *The Quarterly Journal of Economics*, *125*(1), pp.417-443.

Brambilla, I., Depetris Chauvin, N. & Porto, G. (2017) ‘Examining the Export Wage Premium in Developing Countries’, *Review of International Economics*, 25(3), pp. 447–475.

Brandt, L., Van Biesebroeck, J., Wang, L. & Zhang, Y. (2017a) ‘WTO Accession and Performance of Chinese Manufacturing Firms’, *American Economic Review*, 107(9), pp. 2784–2820.

Brandt, L., Van Biesebroeck, J., Wang, L. & Zhang, Y. (2017b) ‘WTO Accession and Performance of Chinese Manufacturing Firms’, *American Economic Review*, 107(9), pp. 2784–2820.

Card, D. (2009) ‘How immigration affects US cities’, *Making cities work: Prospects and policies for urban America*, pp. 158–200.

Card, D. (2005) ‘Is the new immigration really so bad?’, *The economic journal*, 115(507), pp. F300–F323.

Card, D. & Lemieux, T. (2001) ‘Can falling supply explain the rising return to college for younger men? A cohort-based analysis’, *The quarterly journal of economics*, 116(2), pp. 705–746.

Card, D. & Lewis, E.G. (2007) ‘The Diffusion of Mexican Immigrants during the 1990s: Explanations and Impacts’, in *Mexican Immigration to the United States*. [Online]. University of Chicago Press. pp. 193–228.

Cardi, O. & Restout, R. (2015) ‘Imperfect mobility of labor across sectors: a reappraisal of the Balassa–Samuelson effect’, *Journal of International Economics*, 97(2), pp. 249–265.

Chassamboulli, A. & Peri, G. (2020) ‘The economic effect of immigration policies: analyzing and simulating the U.S. case’, *Journal of Economic Dynamics and Control*, 114, p.103898.

Chen, B., Yu, M. & Yu, Z. (2017) ‘Measured skill premia and input trade liberalization: Evidence from Chinese firms’, *Journal of International Economics*, 109pp. 31–42.

Chen, X., Pei, G., Song, Z. & Zilibotti, F. (2023) ‘Tertiarization Like China’, *Annual Review of Economics*, 15(1), p. 485-512.

Cheng, W. & Potlogea, A. (2017) ‘*Trade Liberalization and Economic Development: Evidence from China’s WTO Accession’*, 2017 Meeting Papers 1648, Society for Economic Dynamics.

Combes, P.-P., Démurger, S. & Li, S. (2015a) ‘Migration externalities in Chinese cities’, *European Economic Review*, 76, pp.152–167.

Combes, P.-P., Démurger, S., Li, S. & Wang, J. (2020) ‘Unequal migration and urbanisation gains in China’, *Journal of Development Economics*, 142, p.102328.

Combes, P.-P. & Gobillon, L. (2015) ‘Chapter 5 - The Empirics of Agglomeration Economies’, in Gilles Duranton, J. Vernon Henderson, & William C. Strange (eds.) *Handbook of Regional and Urban Economics*. Handbook of Regional and Urban Economics. Elsevier. pp. 247–348.

Dahlberg, M., Eklöf, M., Fredriksson, P. & Jofre-Monseny, J. (2012) ‘Estimating Preferences for Local Public Services Using Migration Data’, *Urban Studies*, 49(2), pp. 319–336.

De Matos, A.D. (2012) *The careers of immigrants*. Centre for Economic Performance.

Démurger, S., Li, S. & Yang, J. (2012) ‘Earnings differentials between the public and private sectors in China: Exploring changes for urban local residents in the 2000s’, *China Economic Review*, 23(1), pp. 138–153.

Diamond, R. (2016) ‘The Determinants and Welfare Implications of US Workers’ Diverging Location Choices by Skill: 1980-2000’, *American Economic Review*, 106(3), pp. 479–524.

Docquier, F., Ozden, Ç. & Peri, G. (2014) ‘The labour market effects of immigration and emigration in OECD countries’, *The Economic Journal*, 124(579), pp. 1106–1145.

Dorn, F., Fuest, C. & Potrafke, N. (2022) ‘Trade openness and income inequality: New empirical evidence’, *Economic Inquiry*, 60(1), pp. 202–223.

Duranton, G. & Puga, D. (2004) ‘Chapter 48 - Micro-Foundations of Urban Agglomeration Economies’, in J. Vernon Henderson & Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*. Cities and Geography. Elsevier. pp. 2063–2117.

Dustmann, C., Frattini, T. & Preston, I.P. (2013) 'The effect of immigration along the distribution of wages', *Review of Economic Studies*, 80(1), pp. 145–173.

Eaton, J. & Kortum, S. (2001) 'Trade in capital goods', *European Economic Review*, 45(7), pp. 1195–1235.

Eaton, J. & Tamura, A. (1994) 'Bilateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment Patterns', *Journal of the Japanese and International Economies*, 8(4), pp. 478–510.

Ebmer, R. (1990) 'Placement service and offer arrival rates', *Economics Letters*, 34(3), pp. 289–294.

Eckert, F., Ganapati, S. & Walsh, C. (2019) 'Skilled Tradable Services: The Transformation of U.S. High-Skill Labor Markets', Opportunity and Inclusive Growth Institute Working Papers 25, Federal Reserve Bank of Minneapolis.

Edwards, T.H. & Lücke, M. (2021) 'Decomposing the growth of the high-skilled wage premium in an advanced economy open to trade', *The Quarterly Review of Economics and Finance*, 80, pp. 766–784.

Facchini, G., Liu, M.Y., Mayda, A.M. & Zhou, M. (2019) 'China's "Great Migration": The impact of the reduction in trade policy uncertainty', *Journal of International Economics*, 120, pp. 126–144.

Fan, H., Li, Y.A. & Yeaple, S.R. (2015) 'Trade Liberalization, Quality, and Export Prices', *The Review of Economics and Statistics*, 97(5), pp. 1033–1051.

Fang, L. & Herrendorf, B. (2021) 'High-skilled services and development in China', *Journal of Development Economics*, 151, p. 102671.

Frocrain, P. & Giraud, P.-N. (2019) 'The Evolution of Tradable and Non-Tradable Employment: Evidence from France', *Economie et Statistique / Economics and Statistics*, (503d), pp. 87–107.

Fu, D. & Wu, Y. (2013) 'Export wage premium in China's manufacturing sector: A firm level analysis', *China Economic Review*, 26, pp. 182–196.

Gao, X., Liang, W., Mobarak, A.M. & Song, R. (2023) '*Migration Restrictions Can Create Gender Inequality: The Story of China's Left-Behind Children'*, *National Bureau of Economic Research*, working paper series no. 30990.

Ghironi, F. & Melits, M.J. (2005) 'International Trade and Macroeconomic Dynamics with Heterogeneous Firms*', *The Quarterly Journal of Economics*, 120(3), pp. 865–915.

Goldberg, P.K. & Pavcnik, N. (2005) 'Trade, wages, and the political economy of trade protection: evidence from the Colombian trade reforms', *Journal of International Economics*, 66(1), pp. 75–105.

Greenaway, D. & Yu, Z. (2004) 'Firm-level interactions between exporting and productivity: Industry-specific evidence', *Review of World Economics*, 140(3), pp. 376–392.

Grogger, J. & Hanson, G.H. (2011) 'Income maximization and the selection and sorting of international migrants', *Journal of Development Economics*, 95(1), pp. 42–57.

Hainmueller, J., Mummolo, J. & Xu, Y. (2019) 'How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice', *Political Analysis*, 27(2), pp. 163–192.

Han, J., Liu, R. & Zhang, J. (2012) 'Globalization and wage inequality: Evidence from urban China', *Journal of International Economics*, 87(2), pp. 288–297.

Handley, K. & Limão, N. (2017) 'Policy Uncertainty, Trade, and Welfare: Theory and Evidence for China and the United States', *American Economic Review*, 107(9), pp. 2731–2783.

Hao, Y. (2023) ‘The dynamic relationship between trade openness, foreign direct investment, capital formation, and industrial economic growth in China: new evidence from ARDL bounds testing approach’, *Humanities and Social Sciences Communications*, 10(1), p. 160.

He, Q., Liu, J. & Zhang, C. (2021) ‘Exchange rate exposure and its determinants in China’, *China Economic Review*, 65, p. 101579.

Heckman, J.J. (1979) ‘Sample Selection Bias as a Specification Error’, *Econometrica*, 47(1), pp. 153–161.

Hsu, W.-T. & Ma, L. (2021) ‘Urbanization policy and economic development: A quantitative analysis of China’s differential *hukou* reforms’, *Regional Science and Urban Economics*, 91, p.103639.

Imbert, C., Seror, M., Zhang, Y. & Zylberberg, Y. (2022) ‘Migrants and Firms: Evidence from China’, *American Economic Review*, 112(6), pp. 1885–1914.

Jensen, J.B., Kletzer, L.G., Bernstein, J. & Feenstra, R.C. (2005) ‘Tradable Services: Understanding the Scope and Impact of Services Offshoring [with Comments and Discussion]’, *Brookings Trade Forum*, pp. 75–133.

Jones, F.L. (1983) ‘On decomposing the wage gap: a critical comment on Blinder’s method’, *The Journal of human resources*, 18(1), pp. 126–130.

Kinnan, C., Wang, S.-Y. & Wang, Y. (2018a) ‘Access to Migration for Rural Households’, *American Economic Journal: Applied Economics*, 10(4), pp. 79–119.

Kinnan, C., Wang, S.-Y. & Wang, Y. (2018b) ‘Access to migration for rural households’, *American Economic Journal: Applied Economics*, 10(4), pp. 79–119.

Krueger, A.B. & Summers, L.H. (1988) ‘Efficiency Wages and the Inter-Industry Wage Structure’, *Econometrica*, 56(2), pp. 259–293.

Lee, L. (2012) ‘Decomposing wage differentials between migrant workers and urban workers in urban China’s labour markets’, *China Economic Review*, 23(2), pp. 461–470.

Lewis, E. (2011) ‘Immigration, skill mix, and capital skill complementarity’, *The Quarterly Journal of Economics*, 126(2), pp. 1029–1069.

Lewis, E. & Peri, G. (2015) ‘Immigration and the Economy of Cities and Regions’, in *Handbook of regional and urban economics*. [Online]. Elsevier. pp. 625–685.

Li, B. (2018) ‘Export expansion, skill acquisition and industry specialization: evidence from china’, *Journal of International Economics*, 114pp. 346–361.

Li, J., Gu, Y. & Zhang, C. (2015) ‘*Hukou*-based stratification in urban China’s segmented economy’, *Chinese Sociological Review*, 47(2), pp. 154–176.

Li, J., Lu, Y., Song, H. & Xie, H. (2019) ‘Long-term impact of trade liberalization on human capital formation’, *Journal of Comparative Economics*, 47(4), pp. 946–961.

Li, K., Qin, Y. and Wu, J. (2020) ‘Recent housing affordability in urban China: A comprehensive overview’, *China Economic Review*, *59*, p.101362.

Liao, J. (2020) ‘The rise of the service sector in China’, *China Economic Review*, 59, p. 101385.

Ma, L. & Tang, Y. (2020) ‘Geography, trade, and internal migration in China’, *Journal of Urban Economics*, 115, p. 103181.

Martin, W. & Pham, C.S. (2020) ‘Estimating the gravity model when zero trade flows are frequent and economically determined’, *Applied Economics*, 52(26), pp. 2766–2779.

Mayda, A.M. (2010) ‘International migration: a panel data analysis of the determinants of bilateral flows’, *Journal of Population Economics*, 23(4), pp. 1249–1274.

Melitz, M.J. (2003) 'The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity', *Econometrica*, 71(6), pp. 1695–1725.

Meng, X. & Zhang, D. (2010) '*Labour market impact of large scale internal migration on Chinese urban'native'workers'*, *IZA working paper*, No.5288.

Mincer, J. (1958) 'Investment in human capital and personal income distribution', *Journal of political economy*, 66(4), pp. 281–302.

Moretti, E. (2010) 'Local Multipliers', *American Economic Review*, 100(2), pp. 373–377.

Moretti, E. & Thulin, P. (2013) 'Local multipliers and human capital in the United States and Sweden', *Industrial and Corporate Change*, 22(1), pp. 339–362.

Munch, J.R. & Skaksen, J.R. (2008) 'Human capital and wages in exporting firms', *Journal of International Economics*, 75(2), pp. 363–372.

National Bureau of Statistics. (2017) 'The monitoring survey of rural migrants in China in 2017'.

Ngai, L. Rachel, Pissarides, C.A. & Wang, J. (2019) 'China's mobility barriers and employment allocations', *Journal of the European Economic Association*, 17(5), pp. 1617–1653.

Ngai, L Rachel, Pissarides, C.A. & Wang, J. (2019) 'China's Mobility Barriers and Employment Allocations', *Journal of the European Economic Association*, 17(5), pp. 1617–1653.

Nunn, N. & Qian, N. (2014) 'US food aid and civil conflict', *American Economic Review*, 104(6), pp. 1630–66.

Oaxaca, R. (1973) 'Male-female wage differentials in urban labor markets', *International economic review*, pp. 693–709.

Ortega, F. & Peri, G. (2009) '*The Causes and Effects of International Migrations: Evidence from OECD Countries 1980-2005'*, *National Bureau of Economic Research*, working paper series no.14833.

Ortega, F. & Peri, G. (2013) 'The effect of income and immigration policies on international migration', *Migration Studies*, 1(1), pp. 47–74.

Ottaviano, G.I. & Peri, G. (2012) 'Rethinking the effect of immigration on wages', *Journal of the European economic association*, 10(1), pp. 152–197.

Peri, G. (2012) 'The effect of immigration on productivity: Evidence from US states', *Review of Economics and Statistics*, 94(1), pp. 348–358.

Pierce, J.R. & Schott, P.K. (2016) 'The Surprisingly Swift Decline of US Manufacturing Employment', *American Economic Review*, 106(7), pp. 1632–1662.

Pissarides, C.A. & McMaster, I. (1990) 'Regional Migration, Wages and Unemployment: Empirical Evidence and Implications for Policy', *Oxford Economic Papers*, 42(4), pp. 812–831.

Poncet, S. (2006) 'Provincial migration dynamics in China: Borders, costs and economic motivations', *Regional Science and Urban Economics*, 36(3), pp. 385–398.

Sequeira, S., Nunn, N. & Qian, N. (2020) 'Immigrants and the Making of America', *The Review of Economic Studies*, 87(1), pp. 382–419.

Shiferaw, A. & Hailu, D. (2016) 'Job creation and trade in manufactures: industry-level analysis across countries', *IZA Journal of Labor & Development*, 5(1), p. 3.

Silva, J.M.C.S. & Tenreyro, S. (2006) 'The Log of Gravity', *The Review of Economics and Statistics*, 88(4), pp. 641–658.

Simpson, N.B. & Sparber, C. (2013) 'The Short- and Long-Run Determinants of Less-Educated Immigrant Flows into U.S. States', *Southern Economic Journal*, 80(2), pp. 414–438.

Sprenger, E. (2013) 'The Determinants of International Migration in the European Union: An Empirical Analysis', Leibniz *Institut für Ost- und Südosteuropaforschung (Institute for East and Southeast European Studies)*, working paper, No. 325

Sun, L. (2020) 'Housing affordability in Chinese cities', *Lincoln Institute of Land Policy*.

Sun, W., Bai, C. & Xie, P. (2011) 'The impact of *Hukou* reforms on rural labor migration in China', *Economic Research Journal*, 1, pp. 28–41.

Tian, Y. (2022) 'International Trade Liberalization and Domestic Institutional Reform: Effects of WTO Accession on Chinese Internal Migration Policy', *The Review of Economics and Statistics*, pp. 1–45.

Timmer, M.P., Dietzenbacher, E., Los, B., Stehrer, R. & de Vries, G.J. (2015) 'An Illustrated User Guide to the World Input–Output Database: The Case of Global Automotive Production', *Review of International Economics*, 23(3), pp. 575–605.

Tombe, T. & Zhu, X. (2019a) 'Trade, Migration, and Productivity: A Quantitative Analysis of China', *American Economic Review*, 109(5), pp. 1843–1872.

Tombe, T. & Zhu, X. (2019b) 'Trade, migration, and productivity: A quantitative analysis of china', *American Economic Review*, 109(5), pp. 1843–72.

Topalova, P. (2010) 'Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India', *American Economic Journal: Applied Economics*, 2(4), pp. 1–41.

Topalova, P. (2005) '*Trade Liberalization, Poverty, and Inequality: Evidence from Indian Districts'*, NBER, working paper, No. 11614

Tsou, M., Liu, J. & Huang, C.J. (2006) ‘Export Activity, Firm Size and Wage Structure: Evidence from Taiwanese Manufacturing Firms*’, *Asian Economic Journal*, 20(4), pp. 333–354.

Van Biesebroeck, J. (2005) ‘Exporting raises productivity in sub-Saharan African manufacturing firms’, *Journal of International Economics*, 67(2), pp. 373–391.

Wang, F., Milner, C. & Scheffel, J. (2018) ‘Globalization and inter-industry wage differentials in China’, *Review of International Economics*, 26(2), pp. 404–437.

Wang, F., Milner, C. & Scheffel, J. (2021) ‘Labour market reform and firm-level employment adjustment: Evidence from the *hukou* reform in China’, *Journal of Development Economics*, 149, p. 102584.

Wang, T. & Chanda, A. (2018) ‘Manufacturing growth and local employment multipliers in China’, *Journal of Comparative Economics*, 46(2), pp. 515–543.

Whalley, J. & Zhang, S. (2011) ‘A numerical simulation analysis of (*Hukou*) labour mobility restrictions in China’, in *China’s Integration into The World Economy*. [Online]. World Scientific. pp. 295–324.

Winter-Ebmer, R. (1994) ‘Endogenous Growth, Human Capital, and Industry Wages’, *Bulletin of Economic Research*, 46(4), pp. 289–314.

Winters, P., de Janvry, A. & Sadoulet, E. (2001) ‘Family and Community Networks in Mexico-U.S. Migration’, *The Journal of Human Resources*, 36(1), pp. 159–184.

World Bank. (2007) ’China Quarterly Update’, *World Bank Beijing Office.*

Wu, R., Yang, D., Zhang, L. & Huo, J. (2018) ‘Spatio-Temporal Patterns and Determinants of Inter-Provincial Migration in China 1995–2015’, *Sustainability*, 10(11), p. 3899.

Xia Yiran, Lu Ming. (2015) ‘The "Three Moves of Mencius' mother" in cities: An empirical study on the influence of public services on the flow of labour force’, *Management World*, 30(10), pp. 78-90.

Yu, M. (2015) ‘Processing Trade, Tariff Reductions and Firm Productivity: Evidence from Chinese Firms’, *The Economic Journal*, 125(585), pp. 943–988.

Yun, M.-S. (2005) ‘A simple solution to the identification problem in detailed wage decompositions’, *Economic inquiry*, 43(4), pp. 766–772.

Zhang, C. (2015) ‘Income inequality and access to housing: Evidence from China’, *China Economic Review*, 36, pp. 261–271.

Zhang, J., Wang, R. & Lu, C. (2019b) ‘A quantitative analysis of *Hukou* reform in Chinese cities: 2000–2016’, *Growth and Change*, 50(1), pp. 201–221.

Zhang, J., Wang, R. & Lu, C. (2019c) ‘A quantitative analysis of *Hukou* reform in Chinese cities: 2000–2016’, *Growth and Change*, 50(1), pp. 201–221.

Zhang, L., Sharpe, R.V., Li, S. & Darity Jr, W.A. (2016) ‘Wage differentials between urban and rural-urban migrant workers in China’, *China Economic Review*, 41pp. 222–233.

Zhao, X. (2020) ‘Migrants and urban wage: Evidence from China’s internal migration’, *China Economic Review*, 61, p. 101287.

Zhao, Y. (2001) ‘Foreign direct investment and relative wages: The case of China’, *China Economic Review*, 12(1), pp. 40–57.

Zhou, S., He, B., Ni, H. & Pang, S. (2022) ‘Trade liberalization and regional labor market dynamics: Evidence from China’s WTO accession’, *Journal of International Money and Finance*, 125, p. 102635.

Zi, Y. (2020) '*Trade liberalization and the great labor reallocation'*,*CEPR Discussion Paper*, DP14490.