Labour Market Impacts of Rural-Urban Migration on Urban Residents in Chinese Cities

Dandan Zhang *

November 2009

Abstract

Millions of rural migrants have moved into Chinese cities since the early 1990s contributing greatly to economic growth, yet little is known about the impact on urban workers' labour market outcomes. This chapter uses 'cross-area' analysis to examine the causal relationship between urban workers' employment/earnings and the rural migrant ratio in Chinese cities, with data from the 1990 and 2000 China Population Censuses and 2005 1% Population Sampling Survey, as well as the Urban Household Income and Expenditure Survey and the City Statistical Yearbooks in corresponding years. Estimation results, after controlling the endogeneity problem by adopting first-differencing and instrument regression techniques, show that rural migrants in urban China may have modest positive effects on the employment and earnings of urban workers at the city aggregate level; while unskilled rural migrants have an insignificant impact on the labour market outcomes of competing urban workers. Further investigation of the relative wage between skilled and unskilled workers shows that it does not widen when rural migrant inflow lowers the skilled-unskilled labour ratio. This finding may provide a possible explanation of the insignificant effect of rural-urban migration. That is regional industry structures may spontaneously adapt to labour supply shocks as large scale rural migrants move in, so that the potential negative effects of rural migrants are mitigated.

Keywords: Labour Market Impact, Migration
JEL Classification: I31; K42; J61

*Economics Program, Research School of Social Sciences, Coombs Building 9, The Australian National University, ACT 2601, Australia. E-mail:dandan.zhang@anu.edu.au

1 Introduction

Since the early to mid 1990s more than 100 million rural migrants, motivated by the large earnings gap between rural and urban areas, have moved into Chinese cities. Although rural-urban migration has contributed greatly to Chinese economic growth (Woo, 1998; Meng, 2000; Zhao, 2003; Gong et al, 2008), there have been fierce debates among both academics and government agencies about the extent to which rural migrants should be allowed to work in cities, and whether to provide them with the same rights as urban residents. Supporters of further relaxing rural-urban migration policy argue that migrant workers have provided various goods and services at lower prices, which are now an integral part of urban residents' day-to-day life. Those against relaxing rural-urban migration policy are concerned that migrant inflow may reduce urban workers' employment opportunities, suppress their wages and increase pressure on infrastructure and other public facilities. The core of the debate focuses on whether the rural-urban migration in China brings more benefits than costs to urban workers in the urban labour market. In other words, whether rural-urban migration may harm urban workers' employment and wages. To answer the question, it is essential to understand the extent to which rural migrants affect the labour market outcomes of urban labour force before policy-makers can formulate more effective and efficient migration policies.

The textbook economic model of a competitive labour market suggests the impact of immigration on the employment and wages of local people may be negative. The impact of immigration is made through two major channels. One is that immigrants may substitute local workers and harm their labour market performances (Samuelson, 1964), and the other is that immigrants may change the skilled-unskilled labour ratio in the local labour market and lower the employment opportunities and wages of those with similar skills (Altonji and Card, 1991). Contradicting this theoretical assertion, most existing empirical studies in the field of international immigration show immigrants only have a modest effect on the labour market outcomes of native workers through either of these two channels (Grossman, 1982; Altonji and Card, 1991; Card, 2001 and 2007). The inconsistency between the theory and

the empirical evidence has shaken the basis of the traditional belief that "an immigrant influx should lower the wage of competing factors" (Borjas, 2003, pp.1335), and calls for new evidence from both the theoretical and empirical perspective. As one of the largest migration movements in human history, Chinese rural-urban migration provides an important sample for studying the relationship between migration inflow and the labour market performance of natives.

This study uses the 'cross-area' analysis (Altonji and Card, 1991; Card, 2001) to examine the impact of rural-urban migration on urban workers' employment and earnings in the Chinese urban labour market. Three things need to be highlighted. First, the data used for this study are mainly from three consecutive Population Censuses (i.e., 1990, 2000 and 2005 1%), which provide accurate information on the relative numbers of rural migrants and urban workers over time in a spread of Chinese cities. Second, both the instrumental variable (where the lagged and predicted rural migrants are used as instruments) and the first-differencing approaches are adopted to control endogeneity problems due to time-variant and time-invariant city specific factors, which makes the estimates of the impacts of migrants in the urban labour market more accurate. Third, this study examines the effect of rural migrants inflow on urban workers' employment and wages at different levels, including the city aggregate level, the city-unskilled occupation level, and the individual level.

Consistent with findings in developed countries, the results show that rural-urban migration in China has a non-negative effect on the employment and earnings of urban workers both at the city aggregate (or individual level) and the city-unskilled occupation level. A further investigation of the relationship between the relative wage between skilled and unskilled workers and the skilled-unskilled labour ratio in Chinese cities shows the earnings gap between urban skilled and unskilled workers does not widen over time as rural migrant inflow reduces the skilled-unskilled labour ratio. This finding provides some supportive evidence for the industrial adjustment argument that the industrial structures may adapt to the labour supply shocks due to the rural-urban migration and accommodate their changes in skill composition, through which the potential negative effects of rural migrants might be

mitigated.

The rest of this study is organised as follows. Section 2 summarises a group of existing studies related to the impact of international immigrants on native workers' employment and wages in some developed countries such as the U.S. and the U.K.. This helps to provide a theoretical and empirical preparation for this study. Section 3 gives a brief summary on the background of rural-urban migration in China, in particular from the aspect of the evolution of migration policy. Section 4 discusses the empirical methodology and model specifications based on the 'cross-area' and 'relative wage' analyses. The data sources, definitions of some major variables and summary statistics are presented in section 5 and section 6 discusses the estimation results. The relationship between rural migrants' ratio and urban workers' employment and wage/earnings at the city aggregate and unskill-occupation cell levels are discussed and the impact of the urban skilled-unskilled labour ratio due to the migration inflow on the relative wage of urban skilled to unskilled workers is examined. Section 7 makes the conclusion.

2 Literature Review: Measuring the Labour Market Impact of Immigration

How immigrants may affect native workers' employment and wages is an important theoretical and empirical question, which has long been studied in the literature on the economics of immigration, especially during the 1990s when the illegal Mexican immigrant influx into the U.S. labour market generated social and political unease. Although mainstream theorists believe an immigration influx should lower the labour market outcomes through competing with native workers for employment opportunities, little empirical evidence has been found to support this idea during the past three decades. Debates among labour economists over the issue of why obvious impacts of large-scale immigration on the local labour market have not been observed, motivated the development of new methodologies such as the 'cross-area', 'cross-skill' and 'relative wage' approaches. This section reviews the main studies on this

issue for the U.S. and EU countries.

The economic model of a competitive labour market suggests that, under the assumption of a downward sloping labour demand curve, an immigrant inflow should reduce the wages of native workers by increasing the labour supply. However, this simple theoretical proposition has been subject to considerable debate as to its empirical relevance. Using a simple regression based on the production function, Grossman (1982) finds the impact of immigration on the labour market outcomes of local native workers to be surprisingly small in the U.S., where the number of international immigrants relative to native-born workers has been rising. As an explanation of this phenomenon, he argues that immigrants in the U.S. labour market are more likely to be complementary than substitutive to natives at the aggregate level.

Altonji and Card (1991) develop a theoretical framework to account for skill differences. With this extension, the theoretical prediction suggests immigrants may affect natives' wages and employment through changing the skill composition of the labour supply. If the fraction of unskilled workers among immigrants equals the fraction of unskilled workers among natives, neither skilled nor unskilled wages should change as a result of immigration; whereas, if the fraction of unskilled workers among immigrants is greater than that of natives, the inflow of relative unskilled immigrants would increase skilled workers' wages and decrease the wages of the unskilled. This theory suggests that, in addition to aggregate substitutive or complementary effects, an inflow of immigrants can affect natives' wages and employment if their skill distribution differs from that of the natives.

Based on the above theoretical framework, Altonji and Card (1991) develop an empirical model, which is referred to as the 'cross-area' (or city, region, market) approach. The model treats a city or metropolitan area as a closed labour market, and regresses a measure of aggregated native employment or wages on the relative quantity of immigrants in the same locality. To avoid the reverse causality problem caused by immigrants endogenously clustered in cities with better employment and higher wages, many econometric techniques, such as

the first-differencing analysis, the instrumental variable (IV) estimation, and a combination of both, are applied in the 'cross-area' analysis to obtain a consistent estimate on the employment and wage impacts of migration. More specifically, the first-differencing method is used to wipe out the unobservable time-invariant city specific effects (such as geographic location and historic features), while the IV estimation is used to deal with the endogeneity caused by some remaining time-variant factors in the error term (such as immigration policy and temporary demand shocks).

The empirical results from the 'cross-area' analysis are generally inconclusive. Most studies find that large-scale immigrant inflows have had very little impact on the wages or employment of the competing natives (Altonji and Card, 1991; Fredberg and Hunt, 1995; Smith and Edmonston, 1997; Dustmann et al., 2005; Manacorda et al., 2006). Some even find positive effects (Card, 2007). Only a few find significant negative effects, such as Angrist and Kugler's 2003 analysis on Western European countries. The weak correlations between native employment/wage rates and the immigrant share in a local labour market seem to be inconsistent with the prediction from the theoretical model where immigrants may harm the wage and employment situations of the competing native groups.

The empirical puzzle from the 'cross-area' analysis leads to many criticisms. One of the main criticisms is that the assumption of a city as a closed labour market may not be realistic since labour in many countries (especially developed countries) can freely move across localities. As Borjas (1994, pp.1699) argues, "as long as native workers respond to the entry of immigrants by moving to areas that offer better opportunities, there is no reason to expect a correct correlation between the wage of natives and the presence of immigration". Thus, the impact of immigrants on their competing natives' wages and employment might be underestimated with the 'cross-area' approach which ignores natives' outflows in response to immigrants' inflows.

To address the potential impacts of labour-flow across localities, some empirical studies attempt to relax the city-specific labour market assumption and analyse the impact of

immigrants on natives' wages and employment from an economy-wide perspective. This idea was later evolved to become the 'cross-skill' approach (or the 'general equilibrium' approach) (Borjas, 2003). The method assumes workers are free to move across regions in response to immigrant inflows. Under this assumption the inflow of immigrants may crowd out native workers to other regions and thus the real impact of immigrants on native workers can only be detected at a national or cross-skill level.

Using the 'cross-skill' approach, Borjas (2003) divides the national level labour market into 160 separate skill groups by combining two types of skills: four education and eight work experience categories for five Census years (i.e., 1960, 1970, 1980, 1990 and 2000). Based on these 160 skill cells, he regresses natives' employment and wages on the share of immigrants in each cell. Adjusting for the time trend and group-specific fixed effects, Borjas's analysis reveals the immigrant influx has a significant and negative effect on the wages of competing native workers. Using the same method but focusing on unskilled Mexican immigrants, Borjas and Katz (2005) also find the immigrant influx from 1980 to 2000 reduced the wage of a typical native unskilled worker by 3.3 per cent for the first time. These findings are consistent with the textbook theoretical model that there are negative effects from immigrants on native workers' wages and employment.

Although empirical applications of the 'cross-skill' approach provide some evidence of a negative impact from an immigrant inflow on native workers, the core assumption of the method that natives may move to other areas (labour markets) due to immigrant inflow has never been subjected to strictly empirical scrutiny. Card and DiNardo (2000) tests the hypothesis of immigrant inflows leading to native outflow. Using data from the 1970, 1980, and 1990 Censuses, their empirical results indicate that there is no correlation between native workers' outflow and immigrant inflows. Instead, an increase in immigrant population (in specific skill groups) is accompanied by a rise in the number of natives (within the same skill group) in a locality. This result was later confirmed by other studies, such as Card (2001) and Card (2007).

The other criticism of the 'cross-skill' approach comes from its sensitivity to small changes in the model specification. For example, Ottaviano and Peri (2006) examine the impact of immigration on native workers' wages during the period 1990 to 2004. With an extended Borjas's model (i.e., incorporating capital accumulation to relax the assumption of fixed capital stock), they find a positive and significant effect of immigrants on natives' average and individual wages. This result is completely different from that obtained in Borjas (2003), suggesting the significant negative effect of immigrants on native workers initially obtained through the 'cross-skill' approach is sensitive to the specific model specification and may not be widely replicated in a general form.

The above two critiques cast doubt on the credibility of the Borjas (2003) findings, challenging the basic model assumptions and pulling the discussion back to the starting point where the insignificant and modest immigrant impact, particularly on unskilled natives, should be explained. Inspired by the 'macro' view (Borjas et al., 1997; Borjas, 2003), Lewis (2004) and Card and Lewis (2005) endeavour to explain the insignificant impact of immigration inflows on competing native workers' employment and wages by adopting the hypothesis of industrial adjustment and free trade. Their basic argument is that if free trade and capital movement across regions are allowed under the assumption of the Hecksher-Olin model, an increase in the relative supply of unskilled labour caused by inflows of immigrants can be absorbed by the adjustment of local industrial structure (for example, the expansion of low-skill-intensive industries) with little or no change in the relative wages of unskilled native workers. This model differs from the previous assumption of a 'static' goods market without trade. When an (unskilled) immigrant inflow causes a region to develop a narrowly defined industry trading both nationally and internationally, the immigrants and competing native workers can both be better off, or at least not worse off relative to other skill groups, as long as the adjustment process is short enough. This analysis can provide a reasonable explanation for two widely observed phenomena found in previous studies. The first is that relatively more low-skilled immigrants are usually associated with an increase in the supply of low skilled natives. The second is that natives' employment and wage rates are not harmed

by competition of this immigrant inflow.

To examine the argument of industrial adjustment and free trade, Card and Lewis (2005) developed the 'relative-wage' approach which aims to explore the relationship between the relative wage across different skill groups by city. The empirical specification of this approach is to regress the log wage ratio between skilled and unskilled workers on their relative employment share in the local labour market. The idea behind the 'relative-wage' approach is that if the relative wage between skilled and unskilled workers decreases as a result of an increase in the proportion of low skilled immigrants (i.e., the coefficient for the share of relative skill is positive), the industry structure adapts to the supply shock of unskilled workers by expanding the low-skill intensive industry to absorb the possible negative effects of immigrants on unskilled wages. Otherwise (if the coefficient for the share of relative skill is negative), it can be concluded that the hypothesis of industrial adjustment and free trade fails to provide an effective explanation of insignificant immigration impacts.

By using the 1980, 1990, and 2000 Censuses data, Card and Lewis (2005) find although the relative wage adjustment exists, it is statistically insignificant. This result provides some limited evidence of local industries adapting to the increasing supply of low-skilled immigrant workers in the case of the U.S. labour market. Using the same model, Card (2007) re-examines the relative wage between different combinations of skill quantile using the 2000 Census data, and finds the increase in the relative supply of unskilled labour exerts an insignificant effect on the relative wage between skilled and unskilled groups.

In summary, although in theory (unskilled) immigrant inflows may affect the wages and employment of the competing natives through increasing the local labour supply and changing the unskilled-skilled proportion, most existing empirical studies fail to support this argument. After decades of efforts in this field, scholars have come to believe that (1) immigrants and natives within a narrowly defined skill cell are imperfect substitutes, (2) most natives can actually benefit from immigrant inflows in terms of aggregate wages and employment. In particular, a possible explanation of the above two findings is that industrial

structure adjustment and free trade can help absorb the additional supply of a less-skilled workforce and hence mitigate the impact of immigration, although this hypothesis has not been verified by the existing literature.

3 Background of Rural-Urban Migration in China

China has had a segregated rural and urban labour market since the early 1950s, whereby individuals born in rural areas were restricted from moving to cities.[1] This segregation was mainly implemented through the Household Registration System (*Hukou* System), which artificially divides people into agricultural and non-agricultural groups.

Chinese economic reform began in the agriculture sector (in the late 1970s). As a result of this reform, labour productivity in the agriculture sector improved significantly, which in turn released a large amount of rural workers. Although at that time rural workers were strictly prohibited from moving to cities, some migrants, motivated by the large earnings gap between rural and urban areas, still managed to enter spontaneously especially after the mid 1980s (Cai et al., 2001; Zhao, 2003). Since 1992, due to rapid urban economic growth and a significant increase in labour demand, more and more rural migrants moved to the cities. It was during this period that *Hukou* system lost its effectiveness in restricting labour mobility between rural and urban areas (Meng, 2000; Cai et al., 2001; Zhao, 2000). Nevertheless, city governments are always putting urban workers' request for employment as a priority over that of rural migrants. Whenever there was an increase in urban unemployment, city governments would tighten restrictions on migration. For example, between 1995 and 2000, when the reform of state-owned enterprise generated serious urban unemployment problems, governments in many major cities tightened controls on the rural-urban migration, and various restrictive policies on rural migrants' employment in urban areas were implemented. A rural migrant was required to obtain an identification card, a temporary resident card and an employment registration card in order to be considered as a legal temporary resident (Zhao,

[1] Similarly, city-to-city migration was restricted.

2005). Also, hiring migrants was not allowed in principle for firms whose laid-off workers exceeded 10 per cent of total workers. Many cities published a long list of occupations for which rural migrants were prohibited from being hired (Cai et al., 2001). Since 2000, the employment situation has improved in cities, and the restrictions on rural-urban migration have once again been relaxed. In recent years, the central government has gradually reformed the *Hukou* system and it now allows greater mobility across rural and urban areas. Since 2005, the central government has made some efforts to secure rural migrants' basic privilege in the urban labour market and has begun discussing the policy options which allow them access to public security networks in cities.

Although rural migrants have become a major component of the urban labour force, rural migrants are still not treated equally or fully accepted by the urban society. For example, rural migrants only have temporary permission to work in a city with no access to the social benefits available to their urban counterparts. The majority of rural migrants are currently involved in long-hours-low-payment jobs which urban workers are unwilling to take (Meng and Zhang, 2001; Du et al., 2006).

In addition to policy restrictions, rural migrants in cities encounter hostility and various forms of discrimination, since urban residents do not want to share their higher living standards with rural migrants. According to the survey on Changes of Attitude for Urban Residents in 2002, 57 per cent of residents agreed with the premise that, 'rural migrants would lower the employment opportunity of urban people". A survey conducted in city communities at the beginning of 2008 indicates urban residents still treat rural migrants as competitors in the urban labour market. Generally, only 5 per cent of community leaders thought it was beneficial to have rural migrants coming into the community.[2]

Although it has been believed by both city governments and the urban public that rural migrants entering the urban labour market reduce urban workers' opportunities to work and suppress urban workers' wages, no empirical study has ever examined this issue. To

[2]All numbers from the survey on Changes of Attitude for Urban Residents in 2002 are based on author's own calculation.

understand the impact of rural-urban migration on urban workers' labour market outcomes in order to assist development of public policies toward rural migrants, it is necessary to examine their relationship from an empirical perspective.

4 Methodology and Model Specifications

The 'cross-area' analysis described in section 2 assumes the city labour market is closed and ignores the cross-region mobility of natives in response to immigrant inflows. This assumption is not very restrictive in the case of China. Traditionally, labour movement has not been allowed for a long time even across cities as a result of long-term labour market segregation. Although various labour market reforms gradually relaxed this restriction, the cross-city mobility of labour has not increased much. According to the 1990 and 2000 censuses and the 2005 1% Population Sampling Survey data, among the urban labour force, the proportion whose *Hukou* registration is in one city but who currently live in another city is 1.37 per cent, 6.30 per cent and 14.48 per cent, respectively. Moreover, a formal test on the relationship between urban workers' out-migration rate and the migrant ratio has been made for 2005 (see Appendix **??** for detailed discussion). The result shows urban out-migration is not likely to be a result of rural migrant inflow, which supports the closed labour market hypothesis. Therefore, to examine the effect of rural-urban migration in China, the 'cross-area' analysis is more relevant than treating a nation as a single market suggested by 'cross-skill' analysis.

4.1 *The 'Cross-Area' Analysis*

Following Altonji and Card (1991), the baseline model of the 'cross-area' analysis in this study can be specified as:

$$Y_{it} = \alpha + \beta Log(R/U)_{it} + \gamma Z_{it} + \delta D_t + \epsilon_{it}, \tag{1}$$

where Y_{it} denotes the labour market outcomes (i.e., employment rate or mean of log wage) for urban workers in city i at time t (t=1990, 2000 or 2005); $Log(R/U)_{it}$ measures the

logarithm ratio of rural migrants to the urban labour force of city c at time t;[3] Z_{it} refers to a vector of city-specific characteristics, such as total city permanent population, average age of the urban labour force, proportion of male urban workers, proportion of urban workers completing senior high school, actual foreign investment, shares of value added in secondary and tertiary industries; D_t refers to a set of year dummies; and ϵ_{it} is a residual term. The estimate of β captures the impact of rural migrant inflows on the employment or wages of urban workers, which is the main interest of this analysis.

As has been recognised in previous international migration literature, a major problem related to the pooled cross-sectional regression is that some unobserved economic factors, for example the local demand shock, may improve the labour market outcomes of urban workers and at the same time attract more rural migrant inflows. Failure to consider these economic factors may lead to underestimation of the possible negative impact of rural migrant inflows on the labour market outcomes of urban workers. Also, in China, rural migrants are often discriminated against while urban workers are protected. The level of such discrimination varies across cities and over time. However, due to data availability, the control of the discrimination cannot be achieved in the baseline. This may overestimate the negative impact of rural migrants. The above two types of problems are generally called the endogeneity problem, or the potential correlation between the independent variable and the error term. To solve the problem, the first-differencing regression technique is first used.

The first-differencing regression is widely used in the existing literature to wipe out the time-invariant city-specific effect, such as the geographic location of a city, preference given to rural migrants and some historic features of a city that attracts rural migrants.

The first-differencing specification has the form:

$$\Delta Y_{it} = \lambda + \mu D_t + \beta \Delta Log(R/U)_{it} + \gamma \Delta Z_{it} + \Delta \epsilon_{it}, \tag{2}$$

[3]To relax the linear assumption for the relationship between log migrant ratio and labour market outcomes of urban workers, the squared term for the log migrant ratio can be included in the regression. However, the empirical results only support the linear relationship.

where ΔY_{it} refers to the change in labour market outcomes (employment rate or log wage) for urban workers across two data points (i.e., 2005 and 2000, and 2000 and 1990) in city i ; $\Delta Log(R/U)_{it}$ measures the change in the ratio of rural migrants to the urban labour force in each area between two consecutive Census years; Z_{it} refers to a vector of time variant city characteristics; $\Delta\epsilon_{it}$ is a residual term; and D_t is the year dummy. The first-differencing regression differences out any time invariant characteristics of a city that may affect both the demand for rural migrants and the labour market outcomes of urban workers.

Due to the concern that $\Delta\epsilon_{it}$ (i.e., some time-variant city specific effect) may be still correlated with $\Delta Log(R/U)_{it}$, the instrumental variable (IV) regression technique is adopted in addition to first-differencing to deal with the remaining endogeneity problem. Thus, the model that combines both the first-differencing and IV regression techniques is preferred in this analysis (see Appendix **??** for more detailed discussion).

The most typical instrument considered in previous studies has been the lagged relative ratio of immigrants in a destination (Altonji and Card, 1991; Card, 2001), which should be highly correlated to the current immigrant inflow but is assumed to have no direct effect on the labour market outcomes of the native labour force. In the case of China, many studies find that the size of the rural migrant community from a source province plays an important role in attracting future migrants from the same province (Rozelle et al., 1999; Meng, 2000; Zhao, 2003; Bao et al., 2007; de Brauw and Giles, 2008a). Thus, an instrument that can be used here is the lagged rural migrant ratio from the previous Census.

An alternative instrument used in this study is the predicted migrant ratio following Boustan et al. (2007) and Cortes (2008). The basic idea is that migrant inflow to a particular city j (R_j) is the product of the sum of the number of migrants from various sending rural areas (OM_k) and the probability of these outflow migrants moving to the destination city j (P_{kj}). The number of people who move out of a sending region k (OM_k) is a function of many push factors, and the probability of these people settling in the destination city j (P_{kj}) is a function of pull factors. These push and pull factors can thus be effectively used

as instruments to identify city j's migrant inflow. In this study, the push factors include: land per capita, net income per capita for rural residents, total areas of natural disaster and fixed asset investment per capita in rural areas. The pull factors include lagged ratio of rural migrants from k to city j among all migrants in city j and the distance between the capital city of the sending province and the destination city. The detailed discussion on constructing the instrument is included in Appendix **??**.

The predicted migrant ratio can be claimed as a valid instrument for three reasons. First, since the predicted migrant ratio contains both information in urban destinations and origin places, it should be highly correlated with the current relative ratio of migrant inflow from rural to urban areas. Second, since all the pull factors (from urban destination) are 5-10 years lagged information, it may not lead to a direct correlation between the predicted migrant ratio and current urban workers' labour market outcomes. Third, most notably, the push factors (from source provinces) are completely exogenous to the economic situation in destination cities. In this sense the predicted migrant ratio should be better than the lagged migrant ratio in terms of being a valid instrument.

Since rural migrants are generally less educated and cluster in unskilled occupations, they are more likely to compete with unskilled urban workers. Therefore, the 'cross-area' analysis is also applied to the unskilled job market to examine whether rural migrants' inflow has any effect on the labour market outcomes of unskilled urban workers.

Following Card (2001), I mainly define the unskilled labour market along occupation lines due to the significant occupational segregation between migrant and urban workers in cities (Meng and Zhang, 2001).[4] Three occupations, including production, service and agriculture workers[5] in which the majority of rural migrants are concentrated, are considered as the unskilled occupations.

[4]The other definition for unskilled labour market is based on education level, which is used as a robustness check in this chapter.

[5]This occupation category is provided by the Census data. The other occupation categories include high level officers, professionals and clerks, which are considered as the skilled occupations. I include agriculture workers in unskilled occupations because the migrant ratio is high in the occupation (37-60 per cent), though the proportion of rural migrants is relatively small (2-3 per cent) compared to other occupations.

One problem with the estimation based on the unskilled-occupation cell is how to obtain the employment rate (or probability) for the unskilled occupation since occupations are only observed for those who have been employed. This problem can be solved by considering an individual's occupation as a probable outcome that depends on a vector of underlying characteristics, such as gender, education, experience, marital status, location. Therefore, the potential employment rates of urban workers within an unskilled cell are estimated ahead of time.[6]

4.2 *The Relative Wage Analysis*

Finally, the relative wage analysis is adopted to examine the industrial structure adjustment argument which may be a possible channel to mitigate the negative effect of immigration. Theoretically, the decrease in the ratio of skilled to unskilled workers as a large number of low-skilled immigrants arrive may increase the wages of skilled relative to unskilled workers if the demand curve is downward sloping. However, if a city (or nation) adjusts its industrial structure in response to the immigrant inflow, the increase in relative supply of unskilled workers should not put downward pressure on the unskilled wage. As a result, the relative wage of skilled to unskilled workers will not increase as the relative skilled ratio drops. Therefore, the relative wage analysis can be used to examine the industrial structure adjustment hypothesis.

Following Card and Lewis (2005), the relative wage analysis takes the following form:

$$Log(w^H/w^L)_{it} = \alpha + \beta Log(N^H/N^L)_{it} + \theta D_t + \epsilon_{it}, \tag{3}$$

where superscript H denotes the high-skilled labour force, defined as those who have completed senior high school education or above; superscript L indicates the low-skilled labour

[6]Following Card (2001), the method is first to adopt a multi-nominal logit model with the information of being employed or unemployed for the urban labour force as well as their individual characteristics (such as education, gender, age and so on) to estimate the probability of being employed for individual urban labour force by city, and then to calculate the potential number of workers employed in the three occupations (i.e., production, service and agriculture) by adding the probabilities for each occupation. Finally, the potential employment rate can be estimated by using the real number of employed workers in these three sectors to divide the estimated potential number of workers being employed.

force, defined as those with junior middle school education or below; the dependent variable is the logarithm of the annual earnings[7] ratio for high-skilled to low-skilled workers in city i at year t; and the independent variables include the logarithm of the ratio for the number of skilled labour force to the number of low-skilled labour force in city i in year t, and year dummies D_t. The coefficient of interest is β. If β is positive or insignificant, it provides evidence for the industry adjustment hypothesis; if β is negative and significant, it suggests there is no industry adjustment.[8]

A potential problem for estimating Equation (3) in this study is related to the measure of the relative wage between skilled and unskilled labour. Since wage information for rural migrants is not available in the data, the relative wage for rural migrants cannot be included. Thus, the dependent variable used in Equation (3) is the logarithm of the annual earning ratio for skilled to unskilled 'urban workers' ($\frac{w_u^H}{w_u^L}$). If assuming rural migrants are perfect substitutes for unskilled urban workers and experience the same change in wage[9], $\frac{w_u^H}{w_u^L}$ may proxy $\frac{w^H}{w^L}$. If rural migrants and urban unskilled workers are imperfect substitutes and segregated into different narrowly defined occupations, their wage patterns could be different. The analysis focusing on urban workers' relative wages would therefore be invalid and could not capture the real relative wage adjustment to rural migrants inflows. This caveat should be kept in mind when making any conclusive argument.

5 Data and Summary Statistics

This chapter mainly uses three data sets: the 1 percent micro-data sample of the 1990 Population Census of China (1990 Census), the 1 percent micro-data sample of the 2000 Population Census of China (2000 Census) and the 20 percent micro-data sample of 2005 1% Population Sampling Survey[10], which were all conducted by the National Bureau of Statistics

[7] Since the information on hours worked is not available for the data, the 'annual earnings' is used as a proxy of the 'wages'. The effect of hours worked may be differenced out (at least partly) by constructing the relative earnings.

[8] The endogeneity issue mentioned earlier also arises and will be resolved using the same methodology.

[9] It is annual earning here.

[10] The 1% Population Sampling Survey is conducted every 10 years between any two decennial Censuses. It takes the same framework as the Census but only samples 1% of the population. For convenience, in this

of China (NBS). The data are widely believed to be the best for identifying rural migrants in China. One shortcoming of the Census data is that they do not provide the earnings/wage information essential for this study.[11]

Therefore, the earnings/wage information is extracted from two other data sources: the City Statistical Yearbooks, 1991, 2001 and 2006 and the Urban Household Income and Expenditure Survey (UHIES) 1991, 2001 and 2006, with data recoded for years 1990, 2000 and 2005. The two data sets serve for different analysis. Since the City Statistical Yearbook data base provides average wages of urban employees for each Chinese city in the previous year, it is used for the city level analysis. Whereas, since it contains detailed individual earnings but covers limited Chinese cities[12], the UHIES data is used for conducting the analysis on unskilled groups.

There are 173, 275, and 284 cities[13] in the 1990, 2000 censuses and 2005 population survey data, respectively. However, not all cities can be included for the analysis. First, to conduct first difference analysis, the cities have to be kept consistent over the three years. Therefore, there are only 173 cities that are common to all three censuses. Second, there are 12 cities with unreasonably extreme changes in migrant ratio from one year to another treated as outliers and they are excluded from the sample size.[14] Third, the city sample size is further restricted to those including all information. As a result, the final sample consists of 149 cities for each year. For the estimations on wages at the unskilled level and the relative wage, the city sample is further reduced to 36 due to the sample coverage of the UHIES.[15]

Although the city sample size is reduced, it does not lose much representability. First, both the 149-city and 36-city samples cover major urban regions receiving rural migrants.

study it is also treated as Census data.

[11] Although the 2005 1% Population Sampling Survey contains individual total income for the first time, it cannot be used to generate the earnings information since the income sources are not specified.

[12] The covers limited Chinese cities, which includes 110 in 1990, 90 in 2000 and 137 cities in 2005.

[13] Cities here indicate those above township level. The increasing city number results from more and more towns having been upgraded to cities (at prefecture level) over time in China.

[14] I deleted those cities with migrant ratio dropping more than 25 percentage points from 1990 to 2000 or from 1990 to 2005.

[15] This is because the analysis for a subgroup of urban workers (skilled and unskilled groups) requires earnings information from the UHIES which covers limited cities.

Total rural migrants residing in 149 cities account for 88.3 per cent of total rural migrants, while the 36 cities cover 54.7 per cent of all rural migrants. Second, both samples can also represent the urban labour force, especially for those in the large and medium sized cities.[16] The 149-city sample covers 83.4 per cent of the urban labour force; while the 36-city sample covers 40.4 per cent of China's total urban labour force. Third, the city samples have a broad geographical coverage. The 149 cities consist of 25 provinces and 4 administered municipality Chinese cities (except Tibet and Hainan province);[17] while the 36 cities consist of 3 out of 4 administered municipality cities (except Tianjin) plus 13 provinces.[18]

Some major variables are defined as follows. The 'rural migrants' are defined as labour market participants (population aged 16-65 and employed or seeking a job) that have resided in the host city for more than six months with their *Hukou* registered somewhere else (other provinces, or other counties within the same province), or those who have lived in a city for less than six months but had left the *Hukou* registration place more than a year previously. The definitions for the 'rural migrants' are slightly changed across censuses due to the inconsistency of questions designed. The details of identifying rural migrants are discussed in Appendix **??**.

The 'urban labour force' is defined as those in the labour market and holding non-agricultural *Hukou* including both the local urban labour force and urban-to-urban migrants from other cities.[19] The urban workers' employment rate for a city is defined as the ratio of the number of urban workers (worked over one hour last week) to that of the urban labour force in a city. The average wage or earnings of employed urban workers is defined in two ways depending on the data source. The average wage from the City Statistical Yearbooks is

[16]The average city size in terms of permanent city residents covered by 149 cities is 1.72 million and 3.18 million for the 36 cities in 2005.

[17]There are overall four administered municipality cities in China, including Beijing, Shanghai, Tianjin, and Chongqing.

[18]The 13 provinces include Liaoning, Heilongjian in the north-east of China, Shandong, Jiangsu in the east, Guangdong in the south, Shanxi and Henan in the middle, Anhui, Jiangxi, Hubei in the middle and south, as well as Sichuan, Yunnan and Gansu in the west. In particular, the 36 city sample covers Beijing, Shanghai, 4 cities in Jiangsu provinces and 3 in Guangdong provinces - China's largest cities or provinces attracting rural migrants.

[19]The local labour force, holding agricultural *Hukou*, are not included in this study (accounting for nearly a quarter of the total labour force in China) given that they are not specified as either the urban labour force or rural migrants.

defined as the ratio of total payroll for employed workers, (including wage, bonus, subsidy and other wages) to the number of employed workers in all sectors (excluding the self-employed and individual sector).[20] Earnings from the UHIES are defined as wages and other labour income for individual workers.

Table 1 presents the summary statistics of the 149 sample cities.[21] The ratio of rural migrants relative to the urban labour force increases from 8 per cent in 1990, to 15 per cent in 2000, and 24 per cent in 2005. The employment rate for urban workers drops from 96 per cent in 1990, to 87 in 2000 and rises to 90 in 2005. This change in urban employment rate may reflect an employment shock during the mid to late 1990s due to the economic restructuring in the state sector. Real average annual wages[22] for urban workers increased from 2285.2 Yuan in 1990, to 5126.64 Yuan in 2000, and reached 9092.70 Yuan in 2005 with an annual growth of 8.4 and 12.1 per cent in the first and the second periods, respectively. Across cities, the relationship between the urban employment rate and the migrant ratio appears to be non-linear (U-shaped) for 1990 and 2005 (see Figure 1); while the relationship between average wages for urban workers and the migrant ratio are overall positive for all three years, especially for 2000 and 2005 (see Figure 2).

Table 2 presents the educational and occupational distributions of rural migrants and urban workers based on information from the individual level data. On average, migrants are less educated than urban workers. For example, in 2005, 81 per cent of rural migrants and 33 per cent of urban workers had junior middle school education or less; 67 per cent of urban workers and 19 per cent of rural migrants had completed senior high school education. Over time, the educational attainment of rural migrants only increases modestly (or stays the same). In contrast there is an obvious upward trend in the average education level for urban workers. For example, the proportion of higher educated urban workers (those obtaining junior college or university degrees) increases from 12 per cent in 1990 to 33 per cent in 2005.

[20] The statistics in the City Statistical Yearbooks mainly focus on the urban employed.

[21] The summary statistics for 36 cities are presented in Appendix Table A1.

[22] Note that all the wage and earnings used in this chapter are depreciated based on the provincial level Consumer Price Index.

With regard to occupational distribution, around 90 per cent of rural migrants are employed as service and production workers while urban workers are more likely to be employed as clerks and professionals (accounting for 38 per cent in 2005). This occupational segregation between rural migrants and urban workers does not change over time.

Figure 3 plots the relationship between the rural migrant ratio and skilled-unskilled ratio among the total labour force. It shows that the skilled and unskilled composition of the total labour force decreases as the fraction of rural migrants to the urban labour force increases. As suggested by Altonji and Card (1991), this implies that rural migrants may impose some effects on urban workers' employment and wages through changing the skilled-unskilled ratio in the local labour market.

6 Estimation Results

This section examines the impact of rural migrant inflows on urban workers' labour market outcomes at the city average and unskilled-occupation levels. Following the model specifications discussed in Section 4, regression techniques used include: the OLS regression, the first-differencing and the first-differencing with IV (based on two different instruments) regressions.

6.1 *The Effect on Urban Workers*

The first question investigated is whether the large scale inflow of rural migrants into cities affects the labour market outcomes of urban workers at the city aggregate level. 149 consistent city samples for the three time points (the 1990 and 2000 Census, plus the 2005 1% Population Sampling Survey) are used and the results reported in Tables 3 and 4 for employment and wage outcomes, respectively.[23]

[23]Weighted regression is not considered in my analysis. The argument of whether weighted or unweighted regressions should be used for the 'cross-area analysis' is not clear in previous literature. Although a few studies (on the labour market impact of immigration) (Card, 2001) argued that using city population as weights may improve the efficiency of regressions, most others do not use it (see Altonji and Card, 1991; Dustmann, et al., 2005; Cortes, 2008). Generally, the motivation for using the weighted estimates is the suspicion that the variances of migrant ratio are proportional to the sample sizes for each city. If that is

Column (1) of Table 3 reports results from the OLS regression of the employment rate based on Equation (1) using all three years' data. The baseline regression controls for city size (i.e., log city population). This is commonly used method in the literature to split city specific labour demand and supply effects (Altonji and Card, 1991; Dustman and Fabbri, 2003). However, in China, city size may not fully capture these city specific effects as the economic reform process established many special economic zones which are often smaller in size but economically more dynamic than the 'old' larger cities. To this end, two additional groups of city-specific labour supply and demand factors are controlled for. On the supply side, the average age of the urban labour force, the proportion of men in the urban labour force and the proportion of highly educated workers in the urban labour market are controlled for. On the demand side, the actual foreign investment inflow of the year, and the share of value added in secondary and tertiary industries are included.[24]

The coefficient on the log migrant ratio is positive and statistically significant at the 1 percent level. The magnitude indicates that every one per cent increase in migrant ratio increases the urban employment rate by 0.014 per cent. Other statistically significant variables include the log city population, which negatively affects the city employment rate, and the proportion of the skilled urban labour force, which affects the city employment rate positively.

As discussed in the methodology section, the estimated effect of the migrant ratio on the labour market outcomes of the urban labour force may be biased due to the endogeneity problem. In particular, most concerns are about possible reverse causality caused by unobservable city level characteristics which leads to a higher level of employment and wages and attract migrants to cities. Therefore, the second step is to estimate Equation (2) using the first-differencing approach. Taking first-difference of both the dependent and independent

the case, the weighted estimates should be more efficient. (Card, 2001 p.41) In my analysis, the correlation between variance of migrant ratio and population size for each city has no much correlation (less than 10 per cent). Therefore, all city level estimation results, based on the cross-area analysis in Chapters 3 and 4, are obtained from unweighted regressions.

[24]The two industrial shares of labour force are only included in the regressions of 2000 and 2005, because this information is not available for 1990.

variables allows unobserved city level time invariant characteristics to be wiped out.

The results of the first-differencing estimation are reported in Columns (2), (5) and (6) of Table 3, respectively. Because the first-differencing requires lagged information, the data used for the first-differencing estimates do not include the 1990 data. To make the first-differencing results comparable with the OLS results, the OLS estimates are also reported with only the 2000 and 2005 data (see Columns (3) and (4) of Table 3). Comparing with the OLS results, the first-differencing estimates have the same sign and level of statistical significance. However, the magnitudes are smaller than those obtained from the OLS estimation. It is observed that every one per cent increase in the migrant rate increases the urban labour employment rate by 0.02 per cent. The decrease in the first-differencing estimates (for the regressions based on 2000 and 2005) indicates the unobserved city time-invariant characteristics may affect both the urban labour employment rate and rural migrant ratio in the same direction. Such factors may include geographic location, administrative level, and the level of historic discrimination against rural migrants in different cities.

The third step taken is to estimate Equation (2) using first-differencing combined with IV methods. The instruments used are the lagged or predicted migrant ratio. The results of the first stage estimation using the lagged migrant ratio and predicted migrant ratio as instruments are reported in Columns (1)-(2) and (3)-(4) of Appendix Table A2, respectively. The results show that both instruments are very strong and statistically significant at the 1 percent level in the first stage regressions. In addition, the F-test on the joint significance of the excluded instruments suggested by Bound et al. (1995) is conducted for each 2SLS regression and the corresponding F-test statistics are reported in the last two rows of the regression tables. The first stage results and the F-test statistics for excluded instruments show that both lagged and predicted migrant ratios are strong instruments.

As shown in Columns (7)-(10) of Table 3, the effect of the migrant ratio on the urban labour force employment rate is still positive and statistically significant for some estimations. When taking the lagged log migrant ratio as IV, the coefficient for the migrant ratio is still

positive (ranging from 0.028-0.029) with a 5 percent significance level (Columns (7) and (8)). When using the predicted migrant ratio as an instrument, the coefficient of migrant ratio is positive but insignificant (Columns (9) and (10)).

The next outcome examined is the average wage of urban employees. The results for the OLS and first-differencing with IV estimations are reported in Table 4. The dependent variable used in this set of regressions is the log of city level average wages for urban workers. These data are obtained from the City Statistical Yearbooks. The results from the OLS estimations (see Columns (1), (3) and (4)) show that the impact of the migrant ratio on the log average wage of the urban labour force is positive and statistically significant. In general, the elasticity is around 0.10-0.12, suggesting every one per cent increase in the migrant ratio increases urban workers' wages by 0.1 per cent. The estimation from the first-differencing with IV regression reduces the significance of the coefficients with most of the estimated effects statistically significant at the 10 percent level (for both regressions instrumented with lagged or predicted migrant ratios) and the coefficients are ranged from 0.08-0.12 (Columns (7)-(10)).[25]

In summary, based on the 'cross-area' analysis approach, it is found that the rural-urban migration does not impose any negative impact on urban labour force employment or wage outcomes at the city aggregate level. Instead, some evidence is found that rural migrant inflow may have modest positive effects on the employment rate and average wage of the urban labour force in host cities. It turns out that a one per cent increase in migrant ratio may increase the employment rate and average wage of urban workers by 0.03 and 0.08 per cent, respectively.

6.2 *The Effect on Unskilled Urban Workers*

Although the above analysis shows some modest positive impacts of rural-urban migration on the average employment and wages of urban workers, it may not be concluded that there

[25]The first stage IV regression results are the same as those for urban employment rate and reported in the Appendix Table A2. Both instruments (lagged and predicted migrant ratio) are very strong and statistically significant at the 1 percent level.

is no negative impact from rural migrants on a subgroup of urban workers' labour market outcomes. As discussed in the data section, more than 95 per cent of rural migrants are employed as unskilled workers in host cities, and their competing urban counterparts are not all urban workers but unskilled ones. Thus, analysis at the average level may be misleading and the impact of rural migrants on the labour market outcomes of unskilled urban workers may be ignored. Therefore, the same issue is re-examined for unskilled workers. Similar methods are used as for the city level analysis focusing on city-unskilled-occupation cells and the estimation results are presented in Tables 5 and 6.

The OLS estimation results for the urban unskilled workers' potential employment rate based on the 149-city sample are presented in Columns (1), (3) and (4) of Table 5. As is shown, the coefficients for the log migrant ratio are negative, very tiny in magnitude (ranging from -0.006 to -0.004) and generally insignificant. This indicates unskilled urban workers' employment opportunities may not be largely affected by the rural migrant inflow. When the first-differencing methods are adopted (see Columns (6)-(8)), the small impact of rural migrants on urban unskilled workers' employment outcomes is observed to be more negative (ranging from -0.020 to -0.017) and statistically significant (the significance level ranges from the 1 per cent to the 10 per cent). While, the estimation results of the first-differencing with IV regressions[26] in Columns (7) and (8) become insignificant, this might not be reliable given the instrument could not pass the exclusion test (see the last two rows). In summary, these results provide weak evidence for a modest negative employment effect of rural migrants on urban unskilled workers; but the effect is generally insignificant.

With regard to the impact on wages, the estimation is based on 36 cities due to the data availability of detailed earnings information of unskilled urban workers. The estimation results are reported in Table 6. The OLS results (see Columns (1), (3) and (4)) show the impact of the migrant ratio on urban unskilled workers' earnings is positive (ranging from 0.098 to 0.160) with a 1 percent significance level, which is even larger than that for the

[26]The first stage results are reported in Columns (1) and (2) of Appendix Table A3. The instrument (lagged migrant ratio) is very strong and statistically significant at the 1 percent level. Note that due to the availability the predicted migrant ratio is not applied for unskilled analysis.

average wage. Since the sample size (36 cities for each time point) is very small, a large sample of 217 cities for 2005 is generated as a robustness check (using the income information from the 2005 1% Population Sampling Survey). The estimate for log migrant ratio based on the 217 city sample in 2005 (see Column (9)) is consistent with that for the 36-city sample in terms of the sign, magnitude and significance level (the coefficient of log migrant ratio is 0.128 and significant at a 1 percent significance level). This appears to suggest an increase in the migrant ratio at the city-unskilled-occupation level tends to raise urban unskilled workers' wages. However, when using the first-differences (see Columns (5) and (6)), the coefficients reduce and become negative (ranging from -0.069 to -0.064) but statistically insignificant. Furthermore, when using first-differencing with IV estimation[27], the coefficients of the rural migrant ratio are 0.105 and 0.08 and still insignificant (see Columns (7) and (8)). This indicates the impact of rural migrants on urban unskilled workers' wages is modest and insignificant overall.

In summary, an increase in the migrant ratio appears to have a modest and insignificant impact on urban unskilled workers' employment and their wages. This suggests that rural migrants and urban workers are perhaps imperfect substitutes even within unskilled occupation cells.

6.3 *Robustness Check*

In this section, I investigate how the main results from the previous section change when (1) taking into account individual characteristics of the urban labour force (undertaking the analysis at individual level), (2) relaxing the 'closed city labour market' assumption, and (3) redefining the unskilled group in terms of education.

First of all, since differences in individual characteristics may generate wage disparity across cities, I follow Card (2001) to adjust labour market outcomes at the city level by

[27]The first stage results are reported in Columns (3) and (4) of Appendix Table A3. The instrument (lagged migrant ratio) is very strong and statistically significant at the 1 percent level. In addition, as shown in the last two rows of Table 6, IV passes the F-test. Therefore, it can be claimed that the lagged log migrant ratio is a valid instrument here.

taking into account the individual characteristics based on the 'cross-area' analysis. In doing so, a 'two-step procedure' outlined by Wooldridge (2003) is adopted to adjust wage and employment rate, which can be described in the following two equations.

$$Y_{ij}^{t} = \beta X_{ij}^{t} + \gamma City_{j} + \epsilon_{i}^{t}, \quad t = 1990, 2000, \ or \ 2005 \tag{4}$$

$$\gamma_{jt} = \alpha + \theta Log(R/U)_{jt} + \delta D_{t} + \mu_{it} \tag{5}$$

In the first step, as shown in Equation (4), the individual-level employment or wages (Y_{ij}) is regressed on a set of individual characteristics (X_{ij}) and city dummies ($City_j$) for each year t. A vector of coefficients for city dummies (γ_{jt}) is then extracted and used as the dependent variable in the second step (as shown in Equation (5)). The independent variables for the second step are a set of city characteristics, including the rural migrant ratio.

Based on the 36-city sample of the 1990 and 2000 Censuses and the 2005 1% Population Sampling Survey, the estimation based on the 'two-step procedure' gives very similar results to those obtained from the city level analysis, suggesting that individual heterogeneity of the urban labour force is relatively independent from the rural migrant inflow.

Due to the concern that the rural migrant inflow may crowd out the urban local labour force especially those unskilled, it is important to check the robustness of the main results by adopting the 'cross-skill' analysis to relax the 'closed city labour market' assumption. Differencing from the 'cross-area' analysis, the 'cross-skill' analysis treats the nation as a labour market and compares wages across low- and high- rural migrant skill groups (ignoring geographic areas). In this context, I divide the national level labour market into 40 skill groups by combining three types of skills[28], including two occupation groups (production and service workers ($i = 1, 2$)), four education groups (illiteracy, primary, junior middle, senior high schools ($j = 1, ..., 4$)) and five age groups (ages between 15-25, 25-35, 35-45, 45-55, 55-65 ($p = 1, ..., 5$)) as well as 3 years (1990, 2000 and 2005 ($j = 1, 2, 3$)). The model

[28] Unskilled labour market is the main focus here.

specification can be written as below:

$$\begin{aligned} lnwage_{ijpt} = \beta Log(R/U)_{ijpt} + occup_i + educ_j + ageg_p + year_t \\ +(occup_i \times educ_j) + (occup_i \times ageg_p) + (occup_i \times year_t) \\ +(educ_j \times ageg_p) + (educ_j \times year_t) + (ageg_p \times year_t) + \epsilon_{it}, \end{aligned} \tag{6}$$

where the dependent variable is the logarithm of average wage for each skill cell; the independent variables include the logarithm of the migrant ratio of each cell as well as a group of fixed effects and their interactions. The estimation coefficient for logged migrant ratio in Equation (6) is -0.115 (standard error is 0.0644) and statistically insignificant. This suggests that rural migrants and the urban labour force are not perfect substitutes even without a narrowly defined unskilled group, which is consistent with the previous estimation results for unskilled cells based on the 'cross-area' analysis.

Finally, as for the unskilled level analysis, I redefine unskilled in terms of education (which is also widely used in the immigration literature) as a robustness check for the results based on unskilled-occupation groups. I restrict the unskilled education groups to those having junior middle school education or below and repeat the estimation as Equation (1). A significant benefit for defining the unskilled labour market through education is that the employment rate is available for low-education groups. The limitation of using this definition is that low-educated rural migrants and urban workers may not be as substitutive as those within the same occupation group. Therefore, the estimation results based on low-educated groups are expected to be more modest than those based on low-skilled occupation groups. The estimation results are generally consistent between the two, where the coefficients of the migrant ratio tend to be positive but overall are insignificant for both the employment rate and wages.

6.4 *Pulling the Pieces Together—Relative Wage Analysis*

To explain the positive or insignificant impact of rural migrant inflows on the labour market outcomes of both average and unskilled urban workers, the relative wage analysis based on Equation (3) is adopted to test whether this is due to the industrial structure adjustment. Since the data for earnings of different skill groups are not available for all sample cities, only 36 consistent cities for the three data points are used and the results are reported in Table 7.

The dependent variable used is the relative wage of skilled to unskilled urban workers, with high skilled defined as those having senior high school education or above and low skilled as those having junior middle school graduate education or below. The OLS estimation results (in Columns (1) and (2)) show that the effects of the total skilled-unskilled ratio on the relative wage of skilled-unskilled urban workers are generally insignificant (the coefficients are 0.104 and -0.045, respectively). When taking first-differencing to wipe out time-invariant city-specific factors (in Columns (3) and (4)), the coefficients for the skilled-unskilled ratio are still insignificant (the coefficients are -0.044 and 0.049). The modest and insignificant relationship between the skilled-unskilled ratio and the relative wage of skilled-unskilled urban workers suggests a large influx of unskilled rural migrants does not widen the earnings gap between skilled and unskilled urban workers. This provides some side evidence for the industrial adjustment in response to rural migrants inflow in China.[29] This is consistent with the fact that a large inflow of rural migrants (70 per cent based on the Rural Household Survey in 2005 (Sheng and Peng, 2005)) into the eastern and southern coastal areas of China (known as China's special economic zones), has been promoting the development of unskilled labour-intensive industries, i.e., manufacturing and service industries, which in turn attracted more and more unskilled labour.[30]

[29]Note that the IV estimations are applied to deal with the potential endogeneity in Equation (3). However, in the first stage, the lagged rural-urban migration (as instrument) is not significantly correlated with skilled-unskilled ratio in the labour market. Since IV estimations are invalid, their results are not reported or discussed here.

[30]Note that since unskilled urban workers' earnings are used as a substitute for all unskilled workers (because rural migrants' earnings information is unavailable) as discussed in the methodology section, the assumption here is that rural migrants are perfect substitutes for unskilled urban workers and thereby they have the same wage patterns. If urban unskilled workers and rural migrants are not perfect substitutes, the analysis based on urban workers' skilled-unskilled relative earnings might be invalid here. This caveat needs to be kept in

In summary, the weak relationship between the skilled-unskilled labour ratio and their relative wage in urban China indicates that there might be industrial structure adjustment favouring unskilled workers in response to rural migrant inflows.

7 Conclusions

This chapter explores the link between rural migrant ratio and urban workers' employment/earnings in the Chinese urban labour market. Following Altonji and Card (1991) and Card (2001), the 'cross-area' analysis is adopted and the above relationship is examined at both city aggregate level and city-unskilled occupation level. The empirical results show that, when the endogeneity problem is well controlled, the rural migrant inflow generally may have a modest positive impact on the employment rate and average wages of urban workers at the city level. It turns out that a one per cent increase in migrant ratio may increase the employment rate and average wage of urban workers by 0.03 and 0.08 per cent, respectively. When focusing on unskilled occupations, the estimation results show that there is no significant negative effect from the rural migrant influx on labour market outcomes of competing urban workers. The above findings suggest that the traditional Samuelson-type competition effect of immigration or the effect of immigrant inflow through changing the composition of skilled and unskilled labour cannot be verified in China's case. Finally, results for the 'relative wage' analysis shed light on a possible explanation of the insignificant effect from the demand side. That is the industrial structure adjustment hypothesis which has taken place in a spread of Chinese cities, especially coastal cities in the south-east of China, and helps to mitigate the possible negative effect of rural migrants. However, more evidence needs to be found before any conclusive explanation can be made for an insignificant impact of rural-urban migration. Additionally, more channels through which rural migrants may affect the labour market in urban destinations need to be specified.

mind when making any conclusive argument.

Bibliography

Altonji, Joseph and David Card, The effects of immigration on the labor market outcomes of less-skilled natives, in J. M. Abowd and R.B. Freeman eds., *Immigration, Trade, and the Labor Market, Chicago:Chicago University Press*, Sep. 1991.

Bao, Shuming, Örn B. Bodvarsson, Jack W. Hou, and Yaohui Zhao, "Interprovincial Migration in China: The Effects of Investment and Migrant Networks," *IZA Discussion Papers No. 2924*, Jul. 2007.

Borjas, George J., "The Economics of Immigration," *Journal of Economic Literature*, Dec. 1994, *32*, 1667–1717.

_ , "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *National Bureau of Economic Research Working Paper Series No.9755*, Jun. 2003.

_ **and Lawrence F. Katz**, "The Evolution of the Mexican-Born Workforce in the United States," *National Bureau of Economic Research Working Paper No. 11281*, Apr. 2005.

_ **, Richard B. Freeman, Lawrence F. Katz, John DiNardo, and John M. Abowd**, "How Much Do Immigration and Trade Affect Labor Market Outcomes?," *Brookings Papers on Economic Activity*, 1997, *1997* (1), 1–90.

Bound, John, David Jaeger, and Regina Baker, "Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Variables is Weak," *Journal of the American Statistical Association*, Jun. 1995, *90* (430), 443–50.

Boustan, Leah Platt, Price V. Fishback, and Shawn E. Kantor, "The Effect of Internal Migration on Local Labor Markets: American Cities During the Great Depression," *National Bureau of Economic Research Working Paper Series No.13276*, Jul. 2007.

Cai, Fang, Yang Du, and Meiyan Wang, "*Hukou* System and Labor Market Protection (*in Chinese*)," *Economic Research (jing ji yan jiu)*, 2001, *12*, 41–49.

Card, David, "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 2001, *19* (1), 22–64.

_ , "How Immigration Affects U.S. Cities," *CReAM Discussion Paper Series No.0711,* Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London, Jun. 2007.

_ **and Ethan G. Lewis**, "The Diffusion of Mexican Immigrants During the 1990s: Explanations and Impacts," *National Bureau of Economic Research Working Papers No.11552*, Aug. 2005.

_ **and John E. DiNardo**, "Do Immigrant Inflows Lead to Native Outflows?," *National Bureau of Economic Research Working Papers No.7578*, Mar. 2000.

Cortes, Patricia, "The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data," *The Journal of Political Economy*, Jun. 2008, *116* (3), 381–422.

de Brauw, Alan and John Giles, "Migrant Opportunity and the Educational Attainment of Youth in Rural China," *The World Bank Policy Research Working Paper No.4526*, Feb. 2008a.

_ **and** _ , "Migrant Labor Markets and the Welfare of Rural Households in the Developing World: Evidence from China," *The World Bank Policy Research Working Paper No.4585*, Apr. 2008b.

Du, Yang, Robert Gregory, and Xin Meng, "Impact of the Guest Worker System on Poverty and Wellbeing of Migrant Workers in Urban China, in Ross Gaunaut and Ligang Song eds., *The Turning Point in China's Economic Development*, Canberra: Asia Pacific Press," , Mar. 2006.

Dustmann, C. and F. Fabbri, "Language Proficiency and Labour Market Performance of Immigrants in the UK," *The Economic Journal*, 2003, *113* (489), 695–717.

Dustmann, Christian, Francesca Fabbri, and Ian Preston, "The Impact of Immigration on the UK Labour Market," *CReAM Discussion Paper Series No. 0501,* Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London, April 2005.

Friedberg, Rachel M. and Jennifer Hunt, "The Impact of Immigration on Host Country Wages, Employment and Growth," *Journal of Economic Perspectives*, 1995, *9* (2), 23–44.

Gong, Xiaodong, Sherry Tao Kong, and Xin Meng, Rural-urban migrants, A driving force for growth, in L. Song, R. Garnaut and W. T. Woo eds., *China's Dilemma, Economic Growth, the Environment and Climate Change*, Brookings Institution Press/ANU e-Press with Asia-Pacific Press, Oct. 2008, pp. 110–151.

Grossman, Jean Baldwin, "The Substitutability of Natives and Immigrants in Production," *The Review of Economics and Statistics*, Nov. 1982, *64* (4), 596–603.

Lewis, Ethan G., "Local, Open Economies Within the U.S.: How Do Industries Respond to Immigration?," *Federal Reserve Bank of Philadelphia Working Paper No. 04-1*, Jan. 2004.

Liang, Zai and Michael J. White, "Internal Migration in China, 1950-88," *Demography*, Aug. 1996, *33*, 375–384.

Lin, Justin, Gewei Wang, and Yaohui Zhao, "Reginal Inequality and Labor Transfers in China: 1983-1988," *Economic Development and Cultural Change*, Apr. 2004, *52*, 587–603.

Manacorda, M, Alan Manning, and Jonathan Wadsworth, "The Impact of Immigration on the Structure of Male Wages: Theory and Evidence from Britain," *CEP Discussion*

Papers No.dp0754, Centre for Economic Performance, The London School of Economics and Political Science, Oct. 2006.

Meng, Xin, *Labour Market Reform in* China, Cambridge: Cambridge University Press, 2000.

_ **and Junsen Zhang**, "The Two Tier Labor Market in Urban China: Occupational and Wage Differentials Between Residents and Rural Migrants in Shanghai," *Journal of Comparative Economics*, 2001, *29* (3), 485–504.

Rozelle, Scott, J. Edward Taylor, and Alan de Brauw, "Migration, Remittances, and Agricultural Productivity in China," *American Economic Review*, May 1999, *89* (2), 287–291.

Sheng, Laiyun and Peng, Liquan, "The Population, Structure and Characteristics of Rural Migrant Workers (*in Chinese*)," *Research on Rural Migrant Labour in China*, Beijing: China Statistics Press, 2005.

Smith, James P. and Barry Edmonston, *The New Americans: Economic, Demographic, and Fiscal Effects of Immigration*, Washington, D.C.: Natioanl Academy Press, 1997.

Woo, Wing Thye, Chinese Economic Growth: Sources and Prospects, in M. Fouquin and F. Lemoine eds., *The Chinese Economy*, London: Economica, 1998.

Wooldridge, Jeffrey M., "Cluster-Sample Methods in Applied Econometrics," *American Economic Review*, May 2003, *93* (2), 133–138.

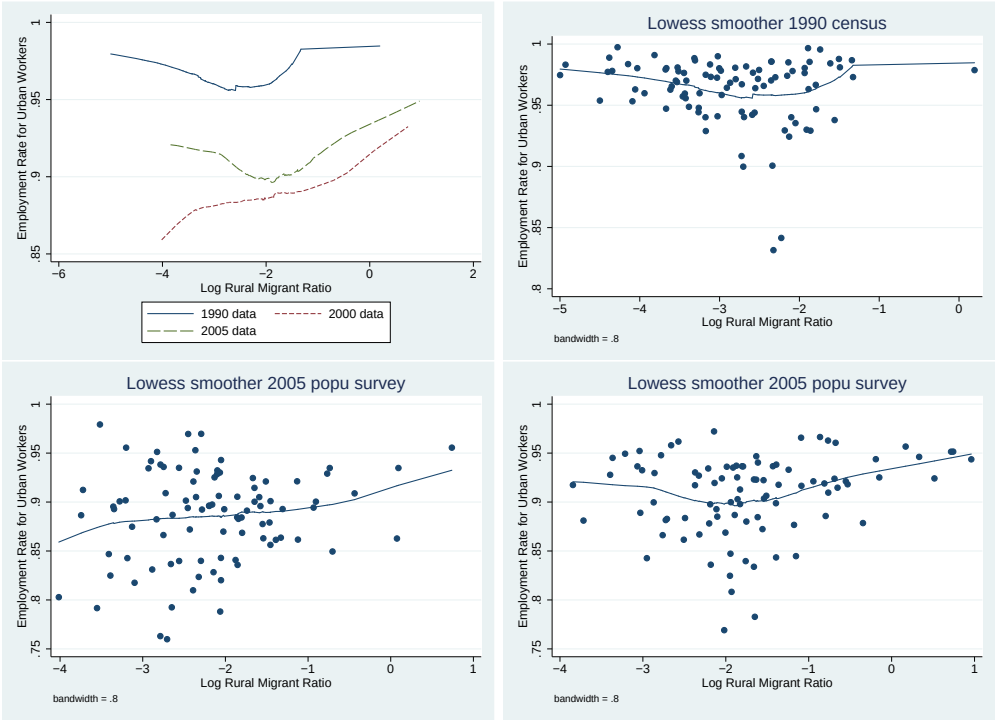
Zhao, Yaohui, "Leaving the Countryside: Rural to Urban Migration Decisions in China," *The American Economic Review*, May 1999, *89* (2), 281–286.

_ , Rural-to-urban labor migration in China: The past and the present, in L. A. West and Y. Zhao eds., *Rural Labor Flows in China*, Berkeley: Institute of East Asian Studies, Berkeley: University of California Press, 2000, pp. 15–33.

Zhao, Zhong, "Rural-Urban Migration in China—What Do We Know and What Do We Need to Know?," *Unpublished Monograph,* China Center for Economic Research, Peking University, Beijing, P.R.China, May 2003.

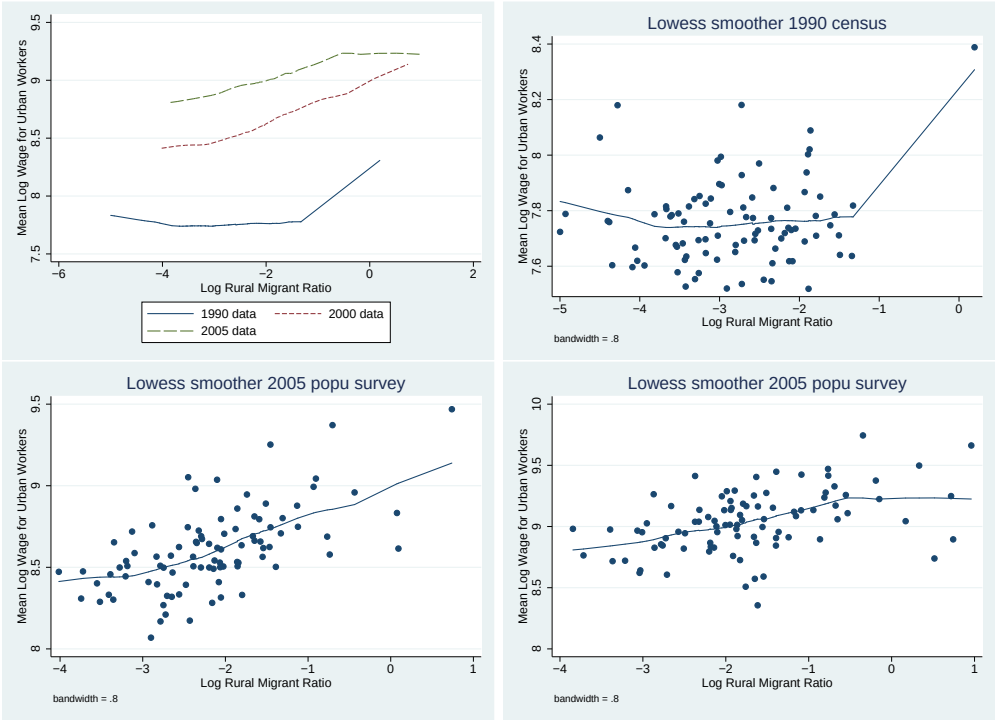
_ , "Migration, Labor Market Flexibility, and Wage Determination in China: A Review," *The Developing Economies*, Jul. 2005, *43* (2), 285–312.

Figure 1: Rural Migrant Ratio and Employment Rate for the Urban Labour Force



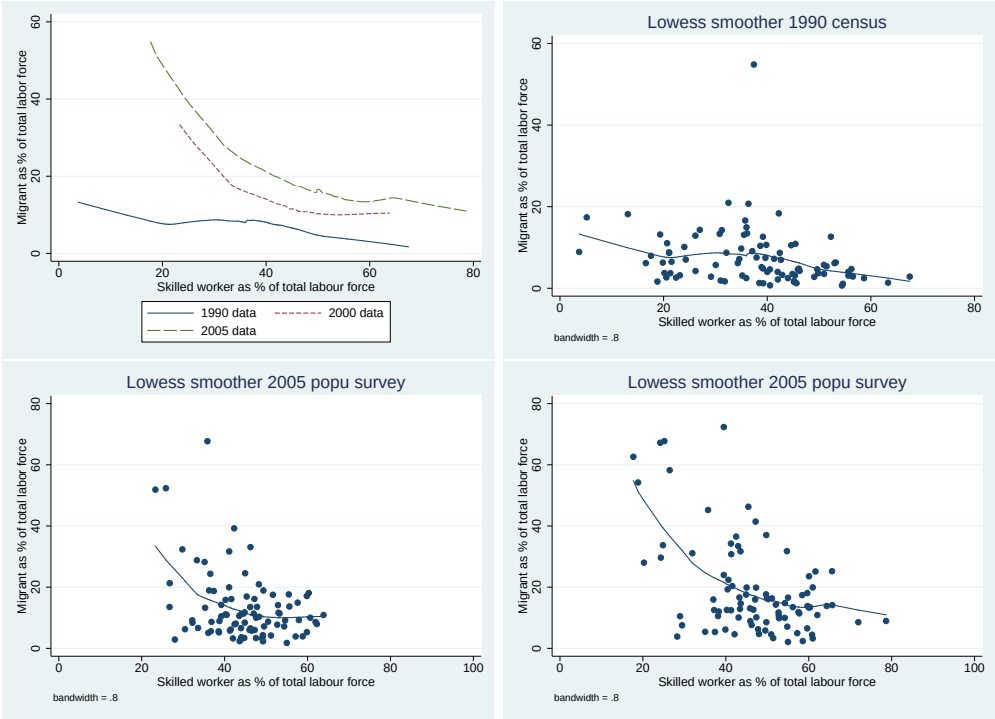
Source: Authors' own calculations based on the 149-city sample using Census data in 1999, 2000 and 2005 1% Population Sampling Survey.

Figure 2: Rural Migrant Ratio and Average Wage for Urban Workers



Source: Authors' own calculations based on the 149-city sample using Census data in 1999, 2000 and 2005 1% Population Sampling Survey.

Figure 3: Proportion of Rural Migrants and Share of Skilled Workers Among Total Labour Force



Source: Authors' own calculations based on the 149-city sample using Census data in 1999, 2000 and 2005 1% Population Sampling Survey.

Table 1: Summary Statistics for 149 Cities: 1990, 2000 and 2005

	1990				2000				2005			
	Mean	Std.Dev.	Min	Max	**Mean**	Std.Dev.	Min	Max	**Mean**	Std.Dev.	Min	Max
City Level												
Migrant ratio[a]	**0.08**	0.11	0.00	1.21	**0.15**	0.26	0.00	2.35	**0.24**	0.40	0.01	2.61
Employment rate for the urban labour force (LF)[a]	**0.96**	0.03	0.82	1.00	**0.87**	0.05	0.66	0.97	**0.90**	0.05	0.76	0.98
Average real wage for employed urban workers (Yuan)[b]	**2285.16**	384.50	1341.10	4395.70	**5126.64**	1538.98	2423.02	12948.93	**9092.70**	2424.34	3346.34	18292.87
City permanent population (Millions)[b]	**1.06**	1.12	0.15	7.83	**1.44**	1.75	0.21	12.62	**1.72**	2.09	0.21	15.11
Average age of all urban LF[a]	**34.51**	1.51	30.53	39.37	**35.71**	1.05	31.00	38.53	**37.38**	1.03	31.98	39.96
Proportion of male urban LF (%)[a]	**56.38**	4.53	48.62	81.08	**56.93**	2.14	52.56	66.38	**56.46**	2.84	48.00	65.41
Proportion of highly educated urban LF (%)[a]*	**43.48**	11.66	18.52	83.43	**59.52**	8.86	32.45	78.40	**63.71**	10.52	36.77	83.89
Actual foreign investment (Million US dollars)[b]	**0.03**	0.09	0.00	0.78	**0.23**	0.64	0.00	4.39	**0.42**	0.94	0.00	7.11
Commodity Export Value (Thousand US Dollar)[b]	-	-	-	-	**13.47**	42.29	0.00	374.80	**54.32**	173.34	0.00	1361.08
Share of value added in secondary industry[b]	-	-	-	-	**52.24**	10.76	27.14	91.47	**53.78**	11.11	26.65	88.76
Share of value added in tertiary industry[b]	-	-	-	-	**42.80**	10.00	8.04	67.53	**42.00**	10.83	10.59	71.29
Unskilled Occupation Level												
Migrant ratio[a]	**0.12**	0.19	0.00	1.98	**0.27**	0.49	0.00	4.24	**0.41**	0.73	0.01	5.00
Average age of urban LF[a]	**33.59**	1.49	30.22	39.52	**35.64**	1.26	29.47	38.61	**37.23**	1.30	30.69	41.52
Proportion of male urban LF[a]	**55.49**	6.02	44.28	84.48	**59.27**	3.39	52.12	72.01	**59.00**	4.84	46.74	71.14
Proportion of highly educated urban LF[a]*	**30.19**	9.76	5.88	57.43	**46.47**	9.30	21.69	68.88	**51.30**	11.72	22.49	78.34

Source: [a]: Author's own calculation based on the 149-city sample using data taken from the 1990 and 2000 Census and 2005 1% Population Sampling Survey data;
[b]: data are taken from the City Statistical Yearbooks.
Note: 'Highly educated' refers to those who completed senior high school education or above.

Table 2: Occupational and Educational Distributions in 149 Cities

	Urban Labour Force			Rural Migrant		
	1990	2000	2005	1990	2000	2005
Education *(%)*						
Illiteracy / Never being in School	1.49	0.36	0.32	6.52	2.70	2.40
Primary School	10.80	4.41	3.77	25.70	19.38	17.07
Junior Middle School	41.64	32.42	28.43	56.37	64.11	61.82
Senior Middle School	34.39	38.60	34.61	11.31	13.00	16.33
Junior College	6.84	15.36	19.35	0.09	0.73	2.00
University and above	4.83	8.83	13.51	0.01	0.09	0.39
Number of observations	669,501	694,643	190,347	45,871	144,415	69,090
Occupation *(%)*						
High level officer	14.73	6.33	5.35	1.31	1.26	1.85
Professional	12.41	22.16	23.06	1.35	1.51	2.69
Clerk	7.07	14.00	15.40	1.33	3.15	3.60
Service worker	8.15	22.39	26.81	14.07	31.78	35.23
Agricultural worker	8.24	1.52	2.17	14.88	4.19	2.12
Production worker	49.39	33.60	27.22	67.06	58.12	54.52
Number of observations	644,794	601,589	170,495	45,663	140,840	66,744

Note: Author's own calculation based on the 1990 and 2000 Census and 2005 1% Population Sampling Survey data with the restricted 149-city urban labour force sample.

Table 3: Effect of Rural Migrants on the Employment Rate of Urban Labour Force: the City Average Level

	1990, 2000 & 2005		2000 & 2005							
	OLS	FD	OLS		FD		FD with IV1		FD with IV2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log migrant ratio	0.014***	0.015***	0.020***	0.020***	0.017***	0.019***	0.029**	0.028**	0.019	0.027
	(0.002)	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)	(0.014)	(0.014)	(0.020)	(0.021)
Log city population	-0.006**	-0.008	-0.011**	-0.008*	-0.006	0.006	-0.008	0.007	-0.006	0.007
	(0.003)	(0.012)	(0.004)	(0.005)	(0.009)	(0.013)	(0.009)	(0.013)	(0.010)	(0.013)
Average age of the urban labour force	-0.001	-0.005*	-0.005**	-0.004	0.005	0.007	0.005	0.008*	0.005	0.008
	(0.002)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
Proportion of the male urban labour force	-0.007	-0.045	0.247*	0.211	0.249	0.253	0.236	0.244	0.247	0.245
	(0.066)	(0.096)	(0.133)	(0.138)	(0.172)	(0.169)	(0.172)	(0.166)	(0.171)	(0.167)
Proportion of the skilled urban labour force	0.127***	0.160***	0.167***	0.170***	0.120*	0.136**	0.138**	0.152**	0.123*	0.150**
	(0.021)	(0.044)	(0.035)	(0.036)	(0.062)	(0.062)	(0.068)	(0.069)	(0.067)	(0.073)
Actual foreign investment	-0.004	-0.008*	-0.005	-0.006	0.001	0.002	0.000	0.001	0.001	0.001
	(0.003)	(0.005)	(0.003)	(0.003)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)
Share of value added in secondary industry	-	-	-	0.053	-	0.291	-	0.333*	-	0.329
	-	-	-	(0.068)	-	(0.198)	-	(0.201)	-	(0.222)
Share of value added in tertiary industry	-	-	-	-0.015	-	0.292	-	0.340*	-	0.335
	-	-	-	(0.069)	-	(0.199)	-	(0.205)	-	(0.232)
Year dummy for 2000	-0.112***	-	-	-	-	-	-	-	-	-
	(0.006)	-	-	-	-	-	-	-	-	-
Year dummy for 2005	-0.090***	0.138***	0.026***	0.022***	-	-	-	-	-	-
	(0.007)	(0.009)	(0.007)	(0.007)	-	-	-	-	-	-
Constant	1.024***	-0.112***	0.922***	0.867***	0.009	-0.000	0.004	-0.006	0.008	-0.005
	(0.062)	(0.011)	(0.133)	(0.134)	(0.010)	(0.012)	(0.012)	(0.014)	(0.014)	(0.018)
Number of observations	447	298	298	298	149	149	149	149	149	149
R-squared	0.519	0.618	0.337	0.353	0.110	0.127	0.085	0.109	0.109	0.112
F-test statistic for excluded instrument	-	-	-	-	-	-	18	18.71	11.92	11.83
Prob>F	-	-	-	-	-	-	0.000	0.000	0.001	0.001

Note: Robustness standard errors are displayed in parentheses. * Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level. IV(1) refers to the lagged log migrant ratio in host cities; IV(2) indicates the predicted log migrant ratio in the host cities. R-squared in Columns (9) and (10) are centered R-squared.

Table 4: Effect of Rural Migrants on the Wage of Urban Workers: the City Average Level

	1990, 2000 & 2005		2000 & 2005							
	OLS	FD	OLS		FD		FD with IV1		FD with IV2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log migrant ratio	0.096***	0.056***	0.117***	0.111***	0.014	0.017	0.082*	0.083*	0.096	0.122*
	(0.010)	(0.011)	(0.011)	(0.012)	(0.017)	(0.017)	(0.048)	(0.048)	(0.069)	(0.072)
Log city population	0.008	0.032	-0.003	0.023	-0.046	-0.019	-0.057*	-0.014	-0.058*	-0.011
	(0.013)	(0.038)	(0.018)	(0.019)	(0.030)	(0.041)	(0.031)	(0.043)	(0.032)	(0.047)
Average age of the urban labour force	0.005	0.008	0.012	0.014	0.025**	0.030**	0.026**	0.033**	0.026**	0.036***
	(0.007)	(0.008)	(0.011)	(0.011)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)
Proportion of the male urban labour force	0.212	0.183	0.241	-0.233	0.942	0.955	0.859	0.894	0.843	0.857
	(0.295)	(0.284)	(0.622)	(0.582)	(0.580)	(0.587)	(0.582)	(0.579)	(0.604)	(0.615)
Proportion of the skilled urban labour force	0.240***	0.315***	0.371***	0.233*	0.446***	0.478***	0.554***	0.588***	0.575***	0.654***
	(0.090)	(0.104)	(0.138)	(0.132)	(0.162)	(0.173)	(0.165)	(0.172)	(0.196)	(0.228)
Actual foreign investment	0.115***	0.076***	0.103***	0.091***	0.023	0.024	0.016	0.017	0.014	0.013
	(0.021)	(0.020)	(0.020)	(0.019)	(0.030)	(0.029)	(0.029)	(0.029)	(0.030)	(0.030)
Share of value added in secondary industry	-	-	-	1.307***	-	0.634	-	0.922*	-	1.095*
	-	-	-	(0.326)	-	(0.532)	-	(0.531)	-	(0.664)
Share of value added in tertiary industry	-	-	-	0.995***	-	0.651	-	0.981*	-	1.179*
	-	-	-	(0.345)	-	(0.557)	-	(0.556)	-	(0.709)
Year dummy for 2000	0.669***	-	-	-	-	-	-	-	-	-
	(0.024)	-	-	-	-	-	-	-	-	-
Year dummy for 2005	1.165***	-0.167***	0.473***	0.462***	-	-	-	-	-	-
	(0.034)	(0.021)	(0.028)	(0.028)	-	-	-	-	-	-
Constant	7.584***	0.673***	7.997***	7.051***	0.518***	0.498***	0.486***	0.458***	0.480***	0.435***
	(0.283)	(0.024)	(0.554)	(0.570)	(0.027)	(0.033)	(0.036)	(0.040)	(0.044)	(0.055)
Number of observations	447	298	298	298	149	149	149	149	149	149
R-squared	0.915	0.403	0.774	0.796	0.102	0.110	0.014	0.031	-0.023	-0.092
F-test statistic for excluded instrument	-	-	-	-	-	-	18	18.71	11.92	11.83
Prob>F	-	-	-	-	-	-	0.000	0.000	0.001	0.001

Note: Robustness standard errors are displayed in parentheses. * Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level. IV (1) refers to the lagged log migrant ratio in host cities; IV(2) indicates the predicted log migrant ratio in host cities. R-squared in Columns (9) and (10) are centered R-squared.

Table 5: Effect of Rural Migrants on the Employment Rate of Unskilled Urban Labour Force: For the City Unskilled Occupation

	1990, 2000 & 2005		2000 & 2005					
	OLS	FD	OLS		FD		FD with IV1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log migrant ratio	-0.006*	-0.019***	-0.004	-0.004	-0.017**	-0.020*	-0.034	-0.032
	(0.004)	(0.007)	(0.004)	(0.004)	(0.007)	(0.011)	(0.026)	(0.027)
Log city population	0.016**	-0.014	0.007	0.009	-0.001	0.005	-0.006	0.004
	(0.006)	(0.013)	(0.007)	(0.007)	(0.013)	(0.019)	(0.016)	(0.019)
Average age of the urban labour force	-0.001	-0.018***	0.003	0.003	-0.014***	-0.001	-0.003	-0.002
	(0.004)	(0.006)	(0.003)	(0.003)	(0.005)	(0.007)	(0.007)	(0.007)
Proportion of the male urban labour force	0.198**	0.226**	-0.006	-0.024	0.377***	0.223	0.174	0.210
	(0.084)	(0.098)	(0.108)	(0.107)	(0.112)	(0.157)	(0.165)	(0.161)
Proportion of the skilled urban labour force	0.031	-0.149*	0.044	0.037	0.018	-0.165	-0.157	-0.182
	(0.040)	(0.078)	(0.044)	(0.046)	(0.054)	(0.111)	(0.108)	(0.114)
Actual foreign investment	-0.005	-0.005	-0.003	-0.004	-0.004	0.004	0.007	0.005
	(0.004)	(0.006)	(0.004)	(0.004)	(0.007)	(0.011)	(0.012)	(0.012)
Share of value added in secondary industry	-	-	-	0.082	-	0.164	-	0.108
	-	-	-	(0.114)	-	(0.339)	-	(0.362)
Share of value added in tertiary industry	-	-	-	0.055	-	0.292	-	0.233
	-	-	-	(0.115)	-	(0.313)	-	(0.335)
Year dummy for 2000	0.021	-	-	-	-	-	-	-
	(0.017)	-	-	-	-	-	-	-
Year dummy for 2005	0.022	-0.052***	-0.006	-0.008	-	-	-	-
	(0.024)	(0.016)	(0.009)	(0.009)	-	-	-	-
Constant	0.823***	0.099***	0.843***	0.770***	0.041***	0.014	0.023	0.021
	(0.131)	(0.026)	(0.135)	(0.163)	(0.015)	(0.018)	(0.020)	(0.022)
Number of observations	447	298	298	298	298	149	149	149
R-squared	0.084	0.146	0.016	0.020	0.104	0.064	0.041	0.054
F-test statistic for excluded instrument	-	-	-	-	-	-	0.77	1.32
Prob>F	-	-	-	-	-	-	0.383	0.253

Note: Robustness standard errors are displayed in parentheses. * Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level. IV1 refers to the lagged log migrant ratio in host cities. R-squared in Columns (7) and (8) are centered R-squared.

Table 6: Effect of Rural Migrants on the Average Earnings of Unskilled Urban Workers: For the City Unskilled Occupation

	1990, 2000 & 2005		2000 & 2005						2005
	OLS	FD	OLS		FD		FD with IV		OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log migrant ratio	0.098***	-0.038	0.160***	0.153***	-0.064	-0.069	0.105	0.080	0.128***
	(0.031)	(0.047)	(0.040)	(0.040)	(0.071)	(0.083)	(0.129)	(0.121)	(0.012)
Log city population	0.024	0.070	-0.020	0.028	-0.352*	-0.408*	-0.264	-0.299	-0.008
	(0.035)	(0.129)	(0.043)	(0.045)	(0.206)	(0.221)	(0.207)	(0.201)	(0.017)
Average age of the urban labour force	0.286	-0.253	1.511	2.004**	-0.703	-0.127	0.332	0.655	-0.010
	(0.621)	(1.028)	(0.957)	(0.930)	(1.052)	(1.270)	(1.574)	(1.569)	(0.009)
Proportion of the male urban labour force	-0.004	-0.019	-0.021	-0.017	-0.016	-0.026	-0.022	-0.027	0.537**
	(0.027)	(0.023)	(0.023)	(0.022)	(0.036)	(0.040)	(0.042)	(0.042)	(0.229)
Proportion of the skilled urban labour force	0.393	0.581	0.292	0.082	1.977**	1.185	1.719*	1.027*	0.460***
	(0.436)	(0.759)	(0.585)	(0.486)	(0.882)	(0.722)	(0.900)	(0.601)	(0.111)
Actual foreign investment	0.101***	0.101***	0.086***	0.092***	0.011	0.030	0.014	0.032	0.000***
	(0.025)	(0.036)	(0.024)	(0.029)	(0.058)	(0.061)	(0.056)	(0.054)	(0.000)
Share of value added in secondary industry	-	-	-	1.419	-	-2.461	-	-1.677	0.523***
	-	-	-	(1.068)	-	(1.934)	-	(2.084)	(0.175)
Share of value added in tertiary industry	-	-	-	0.398	-	-0.724	-	0.161	0.293
	-	-	-	(1.068)	-	(1.953)	-	(1.946)	(0.203)
Year dummy for 2000	0.370***	-	-	-	-	-	-	-	-
	(0.095)	-	-	-	-	-	-	-	-
Year dummy for 2005	1.443***	0.628***	1.088***	1.084***	-	-	-	-	-
	(0.129)	(0.134)	(0.077)	(0.074)	-	-	-	-	-
Constant	7.265***	0.504***	7.937***	6.423***	1.156***	1.220***	1.096***	1.148***	7.918***
	(0.855)	(0.190)	(1.038)	(1.366)	(0.086)	(0.078)	(0.076)	(0.083)	(0.377)
Number of observations	108	72	72	72	36	36	36	36	217
R-squared	0.913	0.635	0.890	0.909	0.251	0.385	0.094	0.266	0.561
F-test statistic for excluded instrument	-	-	-	-	-	-	12.87	12.45	-
Prob>F	-	-	-	-	-	-	0.001	0.002	-

Note: Robustness standard errors are displayed in parentheses. * Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level. IV1 refers to the lagged log migrant ratio in host cities. R-squared in Columns (7) and (8) are centered R-squared.

Table 7: Effect of Rural Migrants on the Relative Wage of Skilled and Unskilled Urban Workers

	1990, 2000 & 2005				2000 & 2005			
	OLS		FD		OLS		FD	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log share of skilled to unskilled LF	0.104	-0.045	-0.044	0.049	0.124	-0.014	-0.193	0.109
	(0.087)	(0.062)	(0.098)	(0.119)	(0.097)	(0.078)	(0.245)	(0.339)
Log city population	-	-0.015	-	0.364*	-	0.069	-	1.038***
	-	(0.034)	-	(0.202)	-	(0.064)	-	(0.278)
Average age of urban LF	-	0.029**	-	0.115***	-	0.016	-	-0.040
	-	(0.014)	-	(0.042)	-	(0.019)	-	(0.053)
Proportion of male urban LF	-	0.135	-	1.388*	-	-0.935	-	-0.635
	-	(0.324)	-	(0.804)	-	(1.869)	-	(3.136)
Actual foreign investment	-	0.000	-	0.000***	-	-0.000	-	0.000
	-	(0.000)	-	(0.000)	-	(0.000)	-	(0.000)
Share of value added	-	-	-	-	-	1.035	-	5.292
in secondary industry	-	-	-	-	-	(1.625)	-	(4.610)
Share of value added	-	-	-	-	-	0.444	-	2.683
in tertiary industry	-	-	-	-	-	(1.744)	-	(4.121)
Year dummy for 2000	0.206***	0.175***	0.195***	-0.133	-	-	-	-
	(0.038)	(0.039)	(0.050)	(0.155)	-	-	-	-
Year dummy for 2005	0.429***	0.416***	0.457***	-0.080	0.218***	0.229***	0.304**	0.013
	(0.060)	(0.072)	(0.090)	(0.224)	(0.067)	(0.066)	(0.128)	(0.159)
Constant	0.105***	-0.937**	0.061	-6.664**	0.318***	-0.858	0.198**	-7.935
	(0.028)	(0.458)	(0.051)	(2.735)	(0.037)	(2.085)	(0.100)	(6.079)
Number of observations	108	108	108	108	72	72	72	72
R-squared	0.561	0.634	0.732	0.800	0.456	0.577	0.723	0.898

Note: Robustness standard errors are displayed in parentheses. * Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level.

APPENDIX I - Testing the 'Closed Labour Market' Hypothesis

To claim the Chinese city labour market can be treated as a closed labour market, the following hypothesis needs to be examined: Is urban out-migration caused by rural-to-urban migration? Due to a higher internal migration rate in 2005, this year is the main focus for testing the 'closed labour market' assumption in China. The model specification can be expressed as follows:

$$Log(OMR_i) = \alpha + \beta Log(R/U)_i + IncomeGap_i + \lambda \sum Prov_j + \gamma \epsilon_i, \tag{7}$$

where OMR_i is defined as the total number of out-migrant urban workers from city i divided by the total number of urban labour force in city i in 2005; $(R/U)_i$ indicates the rural migrant ratio of city i in 2005; $IncomeGap_i$ is the difference between average wage of urban workers in city i and the national average wage in 2005; while $\sum Prov_j$ refers to a set of province dummies. As discussed in section 4, an instrumental variable technique is used to deal with the potential endogeneity.

To test the hypothesis, the coefficient β is of most interest. If β is positive and significant, the 'closed labour market' hypothesis is invalid; otherwise, it can be verified.

The OLS estimation result based on 282 Chinese cities in 2005 shows that the coefficient of the logged migrant ratio is -0.072 with 0.020 as its standard error, suggesting the rural migrant inflow actually lowers the out-migration rate of the local urban labour force. The IV estimation result suggests a very modest negative effect (the coefficient is -0.006) that is statistically insignificant.

As the hypothesis that the Chinese city labour market is closed cannot be rejected, the 'cross-area' analysis is adopted to examine the impact of rural-urban migration on the labour market outcomes of the urban labour force.

APPENDIX II - Use First-Differencing with IV to Deal with Endogeneity

Theoretically, the FD and IV is an ideal way to deal with the endogeneity problem, since the FD can be used to deal with the endogeneity due to the time-invariant factors and IV can be used to deal with the endogeneity due to the remaining time-variant factors. Mathematically, the above process can also be proved as follows. The baseline model specification for the 'cross-area' analysis can be written as $Y_{j,t} = \alpha + \beta X_{j,t} + \lambda_t + u_j + \varepsilon_{j,t}$. The error term $\varepsilon_{j,t}$ can be decomposed into two components, i.e., $\varepsilon_{j,t} = u_j + e_{j,t}$, where u_j varies across cities but not over time, while $e_{j,t}$ varies across cities and over time. After taking FD, $Y_{j,t} - Y_{j,t-1} = \beta(X_{j,t} - X_{j,t-1}) + (\lambda_t - \lambda_{t-1}) + (e_{j,t} - e_{j,t-1})$ is obtained. As is shown, the time-invariant component u_j has been eliminated from the error term. It can also be rewritten as $\triangle Y_{j,t} = \beta \triangle X_{j,t} + \triangle \lambda_t + \triangle e_{j,t}$. This gives $E(\triangle Y_{j,t} | \triangle X_{j,t}) = \beta \triangle X_{j,t} + E(\triangle e_{j,t} | \triangle X_{j,t})$, where $E(\triangle e_{j,t} | \triangle X_{j,t}) = E(e_{j,t} - e_{j,t-1} | X_{j,t} - X_{j,t-1})$.

If $E(\triangle e_{j,t} | \triangle X_{j,t}) = E(e_{j,t} - e_{j,t-1} | X_{j,t} - X_{j,t-1}) \neq 0$ (i.e., $\triangle X_{j,t}$ is correlated with a time-variant unobserved city specific effect $\triangle e_{j,t}$), the FD itself cannot obtain an unbiased and consistent estimator for β. Thus, IV estimation should be introduced to deal with the remaining endogeneity. Basically, a valid instrument should be correlated with $\triangle X_{j,t}$ but not with $\triangle e_{j,t}$. In this context, $X_{j,t-1} - X_{j,t-2}$ or $X_{j,t-2}$ could be considered as an instrument variable (IV). This then goes $E(\triangle Y_{j,t} | IV) = \beta E(\triangle X_{j,t} | IV) + E(\triangle e_{j,t} | IV)$. If $E(\triangle e_{j,t} | IV) = E(e_{j,t} - e_{j,t-1} | X_{j,t-1} - X_{j,t-2}) = 0$ or $E(\triangle e_{j,t} | IV) = E(e_{j,t} - e_{j,t-1} | X_{j,t-2}) = 0$, the estimation of β is then consistent.

In this paper, I choose the predicted migrant ratio in year 1990 as the instrument for the first-differencing with IV estimations.

APPENDIX III - Constructing the Instrument for Migrant Ratio

In addition to the lagged migrants' ratio, an alternative instrument used in this chapter is the predicted rural migrant's ratio. The advantage of this instrument is that the predicted migrant ratio might be more independent from the economic situation in urban destinations than the lagged migrants' ratio, since the predicted migrant ratio contains more exogenous factors, such as a series of push factors in the source provinces and the distance between the source and receiving provinces.

Following Boustan, et al. (2007) and Cortes (2008), the number of rural migrants in city j (R_j) is defined as a weighted sum of the number of migrants leaving other areas k (OM_k) ($k \neq j$) with the weights being the probability that, conditional on leaving area k, a rural migrant settles in city j (P_{kj}) (i.e., $R_j = \Sigma_K P_{kj} \cdot OM_k$). Formally, the estimation of the predicted rural migrant ratio contains three steps.

First, the number of migrants leaving area k (OM_k) is estimated. To isolate the stream of migrants pushed from the home market by local economic conditions, the out-migration rate (OMR_k) from area k is regressed on a set of local factors (push factors), such as land per capita in rural areas $Landpc_k$, net household income per capita $nincpc_k$, the total area of natural disasters $DisasterArea_k$, and the physical asset investment per capita $AssetInvestment_k$.[31] All the data for independent variables are taken from the last Census year. The estimation equation is specified as below and the estimation results are presented in the panel A of Appendix Table A4.

$$OMR_k = \phi Landpc_k + \pi nincpc_k + \rho DisasterArea_k + \kappa AssetInvestment_k + \varepsilon_k. \quad (8)$$

Then, the predicted number of migrants leaving areas ($\widehat{OM_k}$)can be defined as the product of the predicted out-migration rate ($\widehat{OMR_k}$) and the total rural labour force (aged between 16-65) in the previous Census (i.e., $\widehat{OM_k} = \widehat{OMR_k} \times Rural\ Labour\ Force$).

Second, the probability that a migrant leaving area k settles in city j (P_{kj}) is estimated, which is a function of lagged probability of rural migrants from source area k to destination city j and the distance between k and j[32]. The model specification can be written as below and the estimation results are presented in the panel B of Appendix Table A4.

$$P_{kj} = \theta LaggedP_{kj} + \delta LaggedP_{kj}^2 + \lambda Distance_{kj} + \kappa Distance_{kj}^2 + \mu_{kj}, \quad (9)$$

where P_{kj} is defined as the share of rural migrants who came from province k and settled in city j relative to total rural migrants in city j; the Lagged P_{kj} refers to the share in the previous Census year; the $Distance_{kj}$ is defined as the railway length between the capital city in the source province and the destination cities. Based on Equation (9), the predicted probability of migrating from k to j can be obtained.

Finally, the predicted number of rural migrants in city j can be estimated by multiplying the predicted outflow from area k with the predicted probability that a migrant who leaves area k ends up in city j summed over all source areas,

$$\widehat{R_j} = \Sigma_K \widehat{P_{kj}} \cdot \widehat{OM_K}.$$

[31] Zhao (1999) shows that the land per capita and income play important role in rural workers' migration decision. Liang and White (1997) suggests investment in rural areas may impede people to move out.

[32] Some studies on determinants of interprovincial migration found that migration flows are inversely related and very sensitive to distance between origin and destination (Lin et al., 2004).

Then, the predicted rural migrant ratio relative to the urban labour force in city j can be written as $\widehat{R}/U_j$.

Due to the lack of migration information before the 1990 Census, the predicted migrant ratio can not be estimated for 1990. Therefore, the predicted migrant ratio is applied for years 2000 and 2005.

APPENDIX IV - Definition of Rural Migrants Based on Chinese Censuses

Due to the inconsistent design in related questions, the definitions for rural migrants are slightly different between 1990 and the other two Census years (2000 and 2005 1%). In the 1990 Census, rural migrants are defined as those who lived in a city over one year or left their *Hukou* registration place more than one year earlier. This definition for the 2000 Census and 2005 1% Population Sampling Survey is based on a shorter length (six months) of residing in a city or leaving the place of origin for 2000 and 2005. As a result, the number of rural migrants in 1990 may be underestimated. However, since there were limited rural migrants in 1990, the residing periods (six months or one year) may not generate significant difference in calculating the number of rural migrants.

The other problem related to defining rural migrants is an increase in administrative districts within a city during 1990-2005. Given the fact that most cities in China have expanded their boundaries, the inconsistency of city coverage may generate some problems in defining rural migrants and capturing the trends of rural migrant inflow. To cover this concern, the rural migrant ratio is also calculated by restricting the city districts for 2000 and 2005 based on the city districts in 1990. Comparing the two calculations, the rural migrant ratios are quite consistent, while the calculation based on the same city districts fluctuates more across years than the one based on all city districts. This may imply rural migrants are more likely to cluster in the new districts and the statistics based on all city districts are more reliable in terms of identifying rural migrants than those based on the same districts.

Table A1: Summary Statistics for 36 Cities

	1990	2000	2005
Unskilled Level Analysis			
Migrant ratio[a]	0.14	0.36	0.47
	(0.33)	(0.72)	(0.84)
Average annual earnings for urban workers in unskilled occupations[c,a] (Yuan)	1472.73	2792.45	8798.35
	(329.51)	(1235.00)	(2951.01)
Total population size[b] (Millions)	1.85	2.75	3.18
	(0.02)	(0.03)	(0.03)
Average age of the urban LF employed[a]	33.77	35.98	37.22
	(1.31)	(1.48)	(1.40)
Proportion of male urban workers[a] (%)	54.42	58.94	58.74
	(4.46)	(3.02)	(3.40)
Proportion of the skilled urban LF[a]* (%)	32.77	49.43	54.53
	(8.29)	(7.98)	(8.97)
Actual foreign investment[b] (Million US dollars)	0.07	0.59	0.99
	(0.160)	(1.096)	(1.538)
Share of value added in secondary industry[b] (%)	-	50.09	50.28
	-	(9.57)	(10.92)
Share of value added in tertiary industry[b] (%)	-	45.89	46.37
	-	(9.25)	(10.59)
Relative Wage Analysis			
Average annual earnings for skilled urban workers (w_s)[c]* (Yuan)	2058.40	4360.08	6806.84
	(526.62)	(1792.29)	(2681.84)
Average annual earnings for low-skilled urban workers (w_l)[c]* (Yuan)	1907.35	3110.50	4170.06
	(485.47)	(1526.52)	(1598.81)
Relative wage w_s/w_l	1.09	1.45	1.70
	(0.10)	(0.28)	(0.44)
Share of skilled relative to the unskilled LF $(N_s/N_l)^a$	0.56	0.55	0.87
Proportion of skilled among rural migrants $(r_s/r)^a$ (%)	11.96	11.57	15.74
Proportion of skilled among the urban LF $(u_s/u)^a$ (%)	42.45	59.14	62.76

Note: [a]: Author's own calculation based on the 149-city sample using data taken from the 1990 and 2000 Census and 2005 1% Population Sampling Survey data;
[b]: data are from the City Statistical Yearbooks;
[c]: predicted value based on both the Census data and Urban Household Survey data.
*: 'skilled' refers to those who have completed senior middle school education or above; 'unskilled' indicates those with junior middle school education or below.

Table A2: First Stage for the City Aggregate Level

	IV1		IV2	
	(1)	(2)	(3)	(4)
Lagged (IV1) or Predicted log migrant ratio (IV2)	-0.260***	-0.260***	-0.244***	-0.232***
	(0.061)	(0.060)	(0.071)	(0.068)
Log city population	0.291**	0.081	0.195	0.042
	(0.127)	(0.149)	(0.139)	(0.178)
Average age of the urban labour force	-0.111**	-0.148***	-0.029	-0.054
	(0.057)	(0.056)	(0.055)	(0.054)
Proportion of the male urban labour force	1.217	1.042	-0.531	-0.729
	(2.120)	(2.031)	(2.570)	(2.418)
Proportion of the skilled urban labour force	-1.462	-1.625	-1.691*	-1.674
	(0.976)	(0.989)	(0.957)	(1.024)
Actual foreign investment	0.201**	0.194**	-0.024	-0.019
	(0.096)	(0.093)	(0.096)	(0.095)
Share of value added in secondary industry	-	-4.505*	-	-2.502
	-	(2.734)	-	(2.976)
Share of value added in tertiary industry	-	-4.806*	-	-3.305
	-	(2.790)	-	(3.050)
Constant	0.693***	0.828***	0.202	0.293*
	(0.134)	(0.147)	(0.167)	(0.168)
Number of observations	149	149	149	149
R-squared	0.201	0.221	0.117	0.128

Note: Robustness standard errors are displayed in parentheses. * Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level. IV1 refers to the lagged log migrant ratio in host cities; IV2 indicates the predicted log migrant ratio in host cities.

Table A3: First Stage for the City Unskilled Occupation

	149-City Sample		36-City Sample	
	(1)	(2)	(3)	(4)
Lagged log migrant ratio	-0.203***	-0.200***	-0.332***	-0.325***
	(0.059)	(0.058)	(0.093)	(0.092)
Log city population	0.320**	0.137	-0.662	-0.790*
	(0.141)	(0.171)	(0.541)	(0.469)
Average age of the urban labour force	-0.088	-0.112*	-2.949	-2.559
	(0.060)	(0.064)	(3.717)	(3.945)
Proportion of the male urban labour force	-0.259	-0.495	-0.013	-0.031
	(1.864)	(1.966)	(0.133)	(0.138)
Proportion of the skilled urban labour force	-1.002	-1.065	2.629	2.501
	(0.906)	(0.995)	(1.865)	(1.856)
Actual foreign investment	0.269***	0.272***	0.187**	0.184**
	(0.098)	(0.098)	(0.085)	(0.085)
Share of value added in secondary industry	-	-4.375	-	-3.001
	-	(2.938)	-	(3.910)
Share of value added in tertiary industry	-	-4.325	-	-3.956
	-	(3.009)	-	(4.269)
Constant	0.530***	0.632***	0.604***	0.670***
	(0.136)	(0.154)	(0.209)	(0.216)
Number of observations	149	149	36	36
R-squared	0.147	0.166	0.352	0.366

Note: Robustness standard errors are displayed in parentheses. * Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level.

Table A4: Construct Instrument

	2000	2005
Panel A		
Dependent Variable: out-migration rate of source province k		
Land per capita	-0.577**	-0.619*
	(0.277)	(0.330)
Net income per capita	0.001	0.003***
	(0.001)	(0.001)
Areas of disaster	0.011**	0.000
	(0.005)	(0.001)
Physical asset investment	0.000	-0.001***
	(0.001)	(0.000)
Constant	2.350	2.167
	(2.188)	(2.398)
Number of observations	29	30
R-squared	0.426	0.397
Panel B		
Dependent Variable: migration probability from source province k to city *j*		
Lagged P_{kj}	0.270***	0.776***
	(0.024)	(0.022)
Lagged ${P_{kj}}^2$	0.275***	0.380***
	(0.035)	(0.043)
$Distance_{kj}$	-0.015***	-0.010***
	(0.002)	(0.002)
$Distance^2_{kj}$	0.004***	0.003***
	(0.001)	(0.001)
Constant	0.014***	0.008***
	(0.001)	(0.001)
Number of observations	3,485	3,630
R-squared	0.370	0.665

Note: Robustness standard errors are displayed in parentheses. * Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level. Lagged P_{kj} refers to the share of rural migrants from k (source province) to j (host city) over total out migrants from k. All the independent variables are taken from the previous Census year.

Data source: Panel A - all variables are taken from 1991 and 2001 China Statistical Yearbooks. Panel B - Lagged migration probability is taken from 1990 or 2000 Population Census; distance information is generated using a online distance calculator (http://www.infoplease.com/atlas/calculate-distance.html).