

Meng, Xin; Zhang, Dandan

**Working Paper**

## Labour market impact of large scale internal migration on Chinese urban 'native' workers

IZA Discussion Papers, No. 5288

**Provided in Cooperation with:**

IZA – Institute of Labor Economics

*Suggested Citation:* Meng, Xin; Zhang, Dandan (2010) : Labour market impact of large scale internal migration on Chinese urban 'native' workers, IZA Discussion Papers, No. 5288, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/46123>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

IZA DP No. 5288

**Labour Market Impact of Large Scale Internal Migration  
on Chinese Urban ‘Native’ Workers**

Xin Meng  
Dandan Zhang

October 2010

# **Labour Market Impact of Large Scale Internal Migration on Chinese Urban 'Native' Workers**

**Xin Meng**

*Australian National University  
and IZA*

**Dandan Zhang**

*Australian National University*

**Discussion Paper No. 5288  
October 2010**

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **Labour Market Impact of Large Scale Internal Migration on Chinese Urban ‘Native’ Workers<sup>\*</sup>**

Hundreds of millions of rural migrants have moved into Chinese cities since the early 1990s contributing greatly to economic growth, yet, they are often blamed for reducing urban ‘native’ workers’ employment opportunities, suppressing their wages and increasing pressure on infrastructure and other public facilities. This paper examines the causal relationship between rural-urban migration and urban native workers’ labour market outcomes in Chinese cities. After controlling for the endogeneity problem our results show that rural migrants in urban China have modest positive or zero effects on the average employment and insignificant impact on earnings of urban workers. When examine the impact on unskilled labours we once again find it to be positive and insignificant. We conjecture that the reason for the lack of adverse effects is due partially to the labour market segregation between the migrants and urban natives, and partially due to the complementarities between the two groups of workers. Further investigation reveals that the increase in migrant inflow is related to the demand expansion and that if the economic growth continues, elimination of labour market segregation may not necessarily lead to an adverse impact of migration on urban native labour market outcomes.

JEL Classification: J80, J45

Keywords: migration, native labour market outcomes, China

Corresponding author:

Xin Meng  
Research School of Economics  
College of Business and Economics  
HW Arndt Building 25a  
Australian National University  
ACT 0200  
Australia  
E-mail: [Xin.Meng@anu.edu.au](mailto:Xin.Meng@anu.edu.au)

---

<sup>\*</sup> We would like to thank Christian Dustmann, Paul Frijters, Tue Gorgens, Bob Gregory, and Steven Haider for their helpful comments.

# 1 Introduction

In the past twenty or so years the world has seen unprecedented economic growth in China. Accompanying this is the largest rural-to-urban migration in human history. Motivated by the large earnings gap between rural and urban areas, more than 100 million rural workers have moved to Chinese cities since the early to mid 1990s. By 2009 there were 150 million rural migrants working in urban cities, accounting for around one third of the urban labour force.

Although this rural-urban migration has contributed greatly to Chinese economic growth (Woo, 1998; Meng, 2000; Zhao, 2003; Gong et al, 2008), there have been heated debates 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 to labour market access. Those who support 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 migrant urban residents' day-to-day life. Opponents of the relaxation of the 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 or not rural-urban migration has harmed urban workers' employment and wages.

Large scale migration has always faced resistance from the 'native' incumbents. This probably is why economists paid significant attention to the effect of migration on local incumbents' labour market outcomes. The textbook static model of a competitive labour market suggests that the influx of unskilled immigrants should have an adverse effect on the employment and wages of local people. Immigration may increase unemployment, or lower the wages of those with similar skills (Altonji and Card, 1991). In contradiction to this theoretical prediction, however, many existing empirical studies in the field of international migration have found that immigrants only have a modest impact on the labour market outcomes of native workers if cities are regarded as independent markets (Grossman, 1982; Altonji and Card, 1991; Card, 2001 and 2007). Treating a country as a single market, Borjas and Katz (2005) find that immigrant influx to the U.S. between 1980 and 2000 did have a negative impact on wages of the typical unskilled native. However, their finding is sensitive to model specification. Ottaviano and Peri (2006) using the same method but different specification, find a positive impact of immigrant inflow on native wages.

The inconsistency between the theory and the empirical evidences 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 and new explanations. As the largest ever migration movements in human history, the Chinese rural-urban migration provides an important opportunity for studying

the relationship between migration inflow and the labour market performance of the natives.

This paper contributes to the general theoretical debate, as well as to the China-specific policy debate, by examining the impact of the large scale rural-urban migration on employment and wage outcomes of local workers. In addition, we investigate the channels through which migrant inflows may or may not affect natives' labour market outcomes.

The main empirical challenge with regard to this study is related to the issue of reverse causality between labour market outcomes of city workers and migration inflow: the choice of migration destination may be a function of urban local employment conditions and wage levels. If this is the case, our estimates will be biased. To mitigate this problem, we follow Boustan, et al. (2007) and use a combined lagged push and pull factors as instruments. Such factors include per capita land holding, total area of natural disasters in sending areas, and the distance between the sending and receiving areas. We compile a large amount of data from various sources including a one percent samples of the 1990 and 2000 population censuses; a 20 percent sample of the one percent inter-census population sample survey of 2005 (referred to as 2005 Mini-census hereafter);<sup>1</sup> the Urban Household Income and Expenditure Survey (UHIES) for the years 1991, 2001, and 2006; and the City Statistical Yearbook 1991, 2001, and 2006,<sup>2</sup> to examine the large variation across different cities and over time.

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 and the city-unskilled level. A further investigation of the relationship between the relative wages between skilled and unskilled workers and the skilled-unskilled labour ratio in Chinese cities shows that 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 shift of the demand curve, through which the potential negative effects of rural migrants might be mitigated.

The rest of the paper is organised as follows. Section 2 describes the background of rural-urban migration in China, in particular, the evolution of migration policy. Section 3 discusses the empirical methodology and model specifications. The data sources, definitions of some major variables and summary statistics are presented in section 4. Section 5 presents the results. The conclusions are given in Section 6.

---

<sup>1</sup>The Mini censuses are conducted every 10 years between any two decennial Censuses. It takes the same framework as the Census but only samples 1% of the population.

<sup>2</sup>Most of these data are not publicly available,

## 2 Background

China has had segregated rural and urban labour markets since the early 1950s, whereby individuals born in rural areas were restricted from moving to cities.<sup>3</sup> This segregation was mainly implemented through the Household Registration System (*Hukou* System), which artificially divides people into agricultural and non-agricultural populations (Meng, 2000).

Chinese economic reform began in the agriculture sector at the end of the 1970s. As a result of this reform, labour productivity in the agriculture sector improved significantly, and this in turn released a large number of rural workers. Although at the time rural workers were strictly prohibited from moving to cities, some, motivated by the large earnings gap between rural and urban areas, still managed to move to cities for work, especially from the early 1990s. Since the mid 1990s, the rapid urban economic growth, along with a significant increase in foreign direct investments, generated a huge demand for unskilled labour. As a result, more and more rural migrants moved to the cities. It was during this period that *Hukou* system gradually lost its effectiveness in restricting rural workers from moving to cities to work (Meng, 2000; Zhao, 1999, 2000, and 2005; Cai et al., 2001; ). Overtime, hundreds of millions of migrants have moved and become one of the most important driving forces of the Chinese economic growth.

Although rural migrants have contributed significantly to China's economic growth, they are not treated equally in the urban society. Not only are rural migrants restricted in obtaining "good" jobs in cities, but also they have no access to social benefits including unemployment, health, and pension insurance/benefits, all of which are available to their urban counterparts (Meng and Zhang, 2001; Du et al., 2006). When urban economic conditions deteriorate, migrants are normally the first group to suffer. For example, between 1995 and 2000, when the reform of the state-owned enterprises generated serious urban unemployment problems, governments in many major cities tightened controls on the rural-urban migration, and various policies were implemented to restrict rural migrants' employment in urban areas. During that period, hiring migrants was not allowed in principle for firms whose laid-off local workers exceeded the 10 per cent of total work force. Many cities published a long list of occupations for which rural migrants were prohibited from being hired (Cai et al., 2001). Over the years, local governments have also repeatedly demolished the shanty towns where migrants live (Wang and Wang 1995; Xiang 1996), and ignored violations of labour laws by local employers who employed migrant workers. Even though in recent years, the central government has moved toward eliminating these discriminatory treatments of migrants by introducing new laws and regulations to protect

---

<sup>3</sup>Similarly, city-to-city migration was restricted.

migrants' basic rights and increase their access to urban services, such attempts have achieved limited success. The underlying reason is that local governments believe that migrants are competitors of their local constituents in the urban labour market, and hence, reluctant to treat them as locals and to enforce the new laws (Meng and Manning, 2010).

Are migrants substitutes for or complements to urban native workers? This, to a large extent, is an empirical question. To date no empirical study has examined this issue, but the analysis below will help to understand it.

### 3 Literature, Methodology and Model Specifications

How immigrants may affect native workers' employment and wages has long been studied in the literature, 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 that an immigration influx should lower the labour market outcomes of the locals 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' analyses.

The cross-area approach was developed by Altonji and Card (1991) based on a theoretical framework which accounts for skill differences. Their analysis treats a city or a metropolitan area as a closed labour market and investigates how the variation in immigrant inflow across cities relates to the variation of employment or wages of the local workers. The empirical findings from the cross-area analysis are inconclusive. Most studies find little or positive impact of migration 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; Card, 2007), while a few find significant negative effects (for example, Angrist and Kugler, 2003).

The empirical puzzle arising 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 (Borjas, 1994).

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 native wages and employment from an economy-



wide perspective. This idea was later evolved to become the cross-skill (or ‘general equilibrium’) approach (Borjas, 2003). The method assumes workers are free to move across regions in response to immigrant inflows. Using the cross-skill approach, Borjas (2003) and Borjas and Katz (2005) find that the immigrant influx has a significant and negative effect on the wages of competing native workers.

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 be displaced by migrant inflow and thus move to other areas was not subjected to strict empirical scrutiny. Card and DiNardo (2000) tests the hypothesis of immigrant inflows leading to native outflow and find that there is no correlation between the two. Instead, an increase in immigrant population in specific skill groups is accompanied by a rise in the number of natives within the same skill groups 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. Ottaviano and Peri (2006) examine the impact of immigration on native workers’ wages during the period 1990 to 2004. They extended Borjas’s model to relax the assumption of fixed capital stock and find a positive and significant effect of immigrants on native 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.

This paper mainly uses the cross-area analysis to test the effect of large scale rural-urban migration on labour market outcomes of urban native workers. Although the assumption that cities are closed labour markets may be too restrictive for the U.S, it fits China’s situation quite well. Traditionally, labour movement had been restricted for a long time even across cities. 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 among the labour force with urban household registration (*hukou*), the proportion whose *hukou* registration is in one city but live in another city is 1.37 and 6.30 per cent, respectively. This ratio increase to 14.5 per cent using 2005 Mini-census data, still quite low by western standard. To ascertain whether the urban labour force outflow is unrelated to rural migrants inflow, we present evidence that the change in urban local workers outflow is not positively related to the change in migrant inflow; if anything, the relationship is negative (see Table A1 in Appendix A). The results support the hypothesis that Chinese cities are relatively closed labour markets. We therefore use cross-area approach in our main analysis, but later on in the sensitivity test section we also use cross-skill analysis to examine whether our results are driven by the particular approach we have chosen.

Following Altonji and Card (1991), the baseline model is specified as:

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

where  $Y_{it}$  denotes the labour market outcomes (i.e., employment rate or mean of log wage) for urban native workers in city  $i$  at time  $t$  ( $t = 1990, 2000$ , and  $2005$ );  $\log(R/U)_{it}$  is labeled as ‘migrant ratio’ hereafter, which measures the logarithm ratio of rural migrants to the urban labour force of city  $i$  at time  $t$ ;  $Z_{it}$  refers to a vector of city-specific characteristics, such as total urban *hukou* 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  $\epsilon_{it}$  is a residual term. The estimate of  $\beta$  captures the impact of rural migrant inflows on labour market outcomes of urban native workers, which is the main interest of this analysis.

The main problem related to the pooled cross-sectional regression of Equation (1) is that some unobserved economic factors, such as the geographic location of a city, the local demand shocks, or policy variations, may affect the labour market outcomes of urban workers and at the same time affect rural migrant inflows. Failure to consider these omitted city level unobserved factors may lead to under- or over-estimation of the true impact of rural migrant inflows on the labour market outcomes of urban native workers.

The first-difference regression is widely used in the literature to erase the time-invariant city-specific effect. Such an effect may include the geographic location of a city and some historic features that affect both the native workers’ labour market outcomes and rural migrant inflows. The first-difference specification has the form:

$$(Y_{it} - Y_{it-1}) = \lambda + \mu D_t + \beta [\log(R/U)_{it} - \log(R/U)_{it-1}] + \gamma [Z_{it} - Z_{it-1}] + [\epsilon_{it} - \epsilon_{it-1}]. \quad (2)$$

However, unobserved city characteristics do not only take the time-invariant form. Many time-variant city unobserved characteristics, such as policy variations and demand shocks, also exist. Thus,  $\Delta \epsilon_{it}$  in Equation (2) may still be correlated with  $\Delta \log(R/U)_{it}$ . If so, the estimation of  $\beta$  from Equation (2) is still biased. To further resolve the remaining endogeneity problem the instrumental variable (IV) approach is adopted, in addition to the first-difference approach.

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; and Cortes, 2008), which should be highly correlated with the current migration inflow but is assumed to have no direct effect on the labour market outcomes of the native labour force. In China, many studies find that the size of the rural migrant community from a source region plays an important role in

attracting future migrants from the same village, due to the impact of the lack of a formal information network (Rozelle et al., 1999; Meng, 2000; Zhao, 2003; Bao et al., 2007; de Brauw and Giles, 2008a). Thus, one of the instruments considered in this study is the *actual* lagged difference in log rural migrant ratio  $\Delta \log(R/U)_{it-1}$  ( $=[\log(R/U)_{it-1} - \log(R/U)_{it-2}]$ ).

Following the idea developed by Boustan et al. (2007), we also use an alternative instrument, which is the predicted difference in migrant ratio using *lagged information*. In particular, we use  $\Delta \log(\widehat{R}/U)_{it}$  ( $=[\log(\widehat{R}/U)_{it} - \log(\widehat{R}/U)_{it-1}]$ ), where  $(\widehat{R}_i)$  is predicted number of migrants in city  $i$ :

$$\widehat{R}_i = \sum_{k=1}^K \widehat{OM}_k * \widehat{P}_{ki}. \quad (3)$$

The subscript  $k$  indicates the origin rural areas.  $\widehat{OM}_k$  is the predicted total number of migrants from the origin rural area  $k$ , and  $\widehat{P}_{ki}$  is the predicted probability of the outflow migrants from the origin region  $k$  to the destination city  $i$ .  $\widehat{OM}_k$  is obtained from the estimation of the following equation:

$$OM_{kt} = \eta + \phi Z_{kt-1} + \nu_k, \quad (4)$$

where  $Z_{kt-1}$  is a vector of *lagged* push factors at the rural sending region, including land per capita, household income per capita, total land area subject to natural disasters, and physical asset investment per capita.<sup>4</sup>

The probability of migrants moving from rural area  $k$  to the destination city  $i$  ( $P_{ki}$ ) is specified as a function of the quadratic in the geographic distance between origin region  $k$  and destination city  $i$ :

$$P_{ki} = \theta_k + \lambda_k D_{ki} + \kappa_k D_{ki}^2 + \mu_k. \quad (5)$$

We estimate Equation (5) for each origin region (province)  $k$  and the predicted probabilities ( $\widehat{P}_{ki}$ ) are then obtained for each  $k$ .

The predicted difference in migrant ratios between time  $t$  and  $t-1$ ,  $(\log(\widehat{R}/U)_{it} - \log(\widehat{R}/U)_{it-1})$ , is then used as the instrument. Effectively, we can regard the lagged push factors ( $Z_{kt-1}$ ) and the distance information between  $k$  and  $i$  ( $D_{ki}$ ) as the real instruments. The results of Equations (4) and (5) are reported in Table B1 of Appendix B and they show that many lagged push factors and the distance variables are highly correlated to the migrant inflow to city  $i$ . In addition, we believe that none of the push factors should have a direct effect on urban native

---

<sup>4</sup>In the estimation, to avoid adjustment on the size of labor force in the regression, we use out migration rate ( $OMR_k$ , defined as  $OM_k$  divided by total rural labor force in region  $k$ ) as the dependent variable. The final predicted out migration ( $\widehat{OM}_k$ ) is obtained as a product of the predicted out migration rate ( $\widehat{OMR}_k$ ) and the total rural labour force in region  $k$ .

worker labour market outcomes at time  $t$  as all of these factors are derived from 5 to 10 years lagged information from the origin regions. The same is true for the variable measuring geographic distance between  $k$  and  $i$ . Thus, our second instrument, the predicted difference in migrant ratio, is a more preferable measure than our first instrument as it is more exogenous.

Since rural migrants are generally less educated and are often restricted from obtaining professional and managerial jobs (Meng and Zhang, 2001), they are more likely to compete with unskilled urban workers. For this reason, we also estimate Equation 2 using labour market outcomes of urban unskilled workers as dependent variables. We define unskilled labour in two ways. First, we define unskilled workers as those whose education level is at or below junior high school level. Eighty percent of migrants are in this category. Second we define those with production, service or agriculture occupation<sup>5</sup> as unskilled. This is because over 90 percent of migrant workers are concentrated in these occupations. When using the second definition for unskilled workers, we assume that all the unemployed urban workers would have been unskilled had they not lost their jobs. This assumption will give us an upper bound estimation of the effect of migration inflow on urban unskilled workers' employment outcomes.

## 4 Data and summary statistics

This paper uses three main data sources. The first is the 1990 and 2000 censuses and the 2005 Mini-census. We use one percent unit record data of the 1990 and 2000 Population Censuses of China (Census 1990 and Census 2000) and 20 percent of the 2005 Mini-census to construct our main dependent variable, logarithm of the rural migrant ratio in city  $i$  ( $\text{Log}(R/U)_{it}$ ), and some of the independent variables (city level population, labour force, share of male labour force, and share of skilled labour force). All these data were collected by the National Bureau of Statistics of China (NBS). They are widely believed to be the best for identifying rural migrants in China.

The 'rural migrants' ( $R$ ) in this study are defined as labour market participants (population aged 16-65 who are employed or seeking employment) with a rural *hukou* and who have resided in the host city for six months or more, or those who have lived in the current city for less than six months but had left the *hukou* registration region a year or more previously.<sup>6</sup> The 'urban labour force' ( $U$ ) is

---

<sup>5</sup>There is a very limited number of urban workers classified as agriculture workers. For example, in 2005 sample only 1.9% of the urban workers are agriculture workers.

<sup>6</sup>The definition for the 'rural migrants' for different census years differ slightly due to the inconsistency of questions designed in the questionnaire. The details of how rural migrants are identified are presented in Appendix C.

defined as those in the labour market and holding an urban *hukou* including both the local urban labour force and urban-to-urban migrants from other cities.

Because the census data do not provide wage information (except of the 2005 Mini-census<sup>7</sup>), two other data sources are used. The City Statistical Yearbooks (CSY) data provide information on average wages of urban employees for each Chinese city, as well as other city-level control variables, such as the level of foreign direct investment and the city's industrial structure. However, the CSY do not have disaggregated information on wages for different occupations or education levels. We therefore have to use another data set, the Urban Household Income and Expenditure Survey (UHIES) for 16 provinces, to construct earnings for the unskilled groups. A shortcoming of the UHIES data is that it only covers a limited number of Chinese cities.<sup>8</sup>

The employment rate of the urban native labour force for a city is defined as the ratio of the number of urban workers (those who worked over one or more hour in the previous 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 used. The average wage from the CSY is defined as the average of total payroll<sup>9</sup> for employed wage and salary earners in all sectors. Earnings for a city from the UHIES are defined as the average of the total wages and other labour income for employed wage and salary earners.

There are 173, 275, and 284 cities<sup>10</sup> in the 1990, 2000 censuses and the 2005 Mini-census data, respectively. However, due to the following reasons not all cities can be included in the analysis. First, to conduct first difference analysis, the cities have to be kept consistent over the three data points. This leaves us with only 173 cities for each year. Second, among the 173 consistent cities there are 9 cities with abnormal changes in migrant ratio from one year to another. They are regarded as outliers and are excluded.<sup>11</sup> Third, there are also missing values for other city-level control variables which led to the exclusion of another 12 observations. As a result, the final sample consists of 152 cities for each year. The estimations for unskilled workers use a further reduced sample of 36 cities due to the limited city coverage in the UHIES data.

Although the number of cities covered in the analysis is not large, it does not

---

<sup>7</sup>Although the 2005 Mini-census contains individual total income for the first time, it is not ideal in generating the earnings information since the income sources are not specified.

<sup>8</sup>The number of cities included in UHIES are 110, 90, and 137 for 1990, 2000 and 2005, respectively. There are, however only 36 consistent cities across all three years.

<sup>9</sup>Total payroll includes wage, bonus, subsidy and other wages

<sup>10</sup>In this paper cities are defined as prefecture-level urban areas. The increasing number of cities overtime is the result of more and more below-prefecture level towns being upgraded to cities.

<sup>11</sup>Cities where the migrant ratio drops more than 25 percentage points from one year to the next are excluded.

affect the representativeness of our data. Both the 152-city and 36-city samples cover major urban regions receiving rural migrants. Some 88.3 per cent of total rural migrants reside in our 152 cities, while this ratio for our 36 city sample is 54.7 per cent. In addition, the 152-city and 36-city samples also cover 83.4 and 40.4 per cent of the urban labour force, respectively. Finally, our city samples have a broad geographical coverage. The 152 cities are located in 29 of the 31 provinces;<sup>12</sup> while the 36 cities are located in 16 provinces.<sup>13</sup>

Due to the need to use the lagged migrant ratio as an instrument, as well as taking the first-difference, the data we actually use for the main estimations exclude the 1990 census.

Table 1 presents the summary statistics. Using the 152-city sample, the migrant ratio increases from 9 percent in 1990 to 15 percent in 2000, and further to 23 percent in 2005. The employment rate for urban workers drops from 96 percent in 1990, to 87 percent in 2000 and rises slightly to 90 percent in 2005. This change in urban employment rate may reflect an employment shock during the mid to late 1990s, when the state sector reform generated a very high rate of retrenchment. The average real annual earnings<sup>14</sup> for urban workers increases from 2283 Yuan in 1990, to 5083 Yuan in 2000, and reaches 9094 Yuan in 2005 with an annual growth of 8.4 for the first ten years and 12.3 per cent for the last five years. Across cities, the unconditional relationship between the urban employment rate and the migrant ratio appears to be non-existent for 1990 and 2005 and slightly positive for 2000 (see Figure 1A), while the relationship between log average earnings for urban workers and the migrant ratio are overall positive for all three years (see Figure 1B). Figures 2A and 2B plot the unconditional relationships between first-differences of the urban employment rate and log urban earnings and log migrant ratio. All the graphs show either no relationship or a slightly positive relationship.

Table 2 presents the educational and occupational distributions of rural migrants and urban workers, based on information from the individual level data. Migrants are overwhelmingly less educated than urban workers. For example, in 2005, 81 per cent of rural migrants and 32 per cent of urban workers had junior high school education or less; 68 per cent of urban workers and 19 per cent of rural migrants had an education level of senior high school or above. Over time, the educational attainment of rural migrants only increases modestly. In contrast, there is an obvious upward trend in the average education level for urban

---

<sup>12</sup>There are, overall four administered-municipality cities in China, including Beijing, Shanghai, Tianjin, and Chongqing.

<sup>13</sup>The 16 provinces include Beijing, Shanghai, Liaoning, Heilongjian, Shandong, Jiangsu, Guangdong, Shanxi, Henan, Anhui, Jiangxi, Hubei, Sichuan, Chongqing, Yunnan and Gansu.

<sup>14</sup>Note that all the wages and earnings used in this paper are in real terms which are deflated based on the provincial level Consumer Price Index.

native workers, with the proportion of those obtaining junior college and above increasing from 11 per cent in 1990 to 33 per cent in 2005. With regard to occupational distribution, over 90 per cent of rural migrants are employed as service, agricultural, or production workers while urban workers are significantly more likely to be employed as clerks and professionals (accounting for 44 per cent in 2005). This occupational segregation between rural migrants and urban workers has been well documented in the literature (see, for example, Meng and Zhang, 2001) and it does not seem to have changed much over time.

## 5 The empirical results

### 5.1 *The effect on average urban workers*

We first investigate the question of whether, on average, the large scale inflow of rural migrants into cities affects the labour market outcomes of urban workers.

We estimate Equation (1) using the simple OLS and Equation (2) using both the first-difference and the first-difference with IV methods. The control variables included are those which capture the demand for and supply of labour in a city. The most commonly used variable in the literature is city size (i.e., log city population), which is used to identify 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 vectors 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 annual foreign direct investment inflow, and the share of value added in secondary and tertiary industries are included. In addition, the year dummy variable is controlled for in the OLS estimation.

The results from the employment and earnings equations are reported in Panels A and B of Table 3.<sup>15</sup> We first examine the effect of the rural-urban migration rate on employment rate of urban workers (Panel A of Table 3). The dependent variable is defined as the ratio of total employed urban native workers to the total urban natives in the labour force. Column [1] reports the result from the OLS regression (Equation (1)). The coefficient on the log migrant ratio is positive and

---

<sup>15</sup>The results presented in this paper are from unweighted regressions. Using population size as a weight, however, does not change our main results. The results from weighted regressions are available upon request from the authors.

statistically significant at the 1 percent level. The magnitude indicates that every one per cent increase in migrant ratio is associated with a 2 per cent increase in the urban employment rate. The only other statistically significant variable is the average age of urban labour force which is negatively correlated with the urban native employment rate and the year dummy for 2005.

The first-difference estimation is reported in column [2]. Compared to the OLS result, the first-difference estimate has the same sign, similar magnitude, and the same level of statistical significance, suggesting that city-level unobserved time-invariant characteristics do not play an important role. This estimate, however, does not take into account the time-variant city unobservable factors. The columns [3] and [4], therefore, report the results obtained from the first-difference combined with IV methods to address this issue. The instruments used are the lagged difference in log migrant ratio or predicted difference in log migrant ratio, respectively.<sup>16</sup>

The results of the first stage estimation using both the lagged difference in migrant ratio and predicted difference in migrant ratio as instruments are reported Tables D1 of Appendix D. Both instruments are very strong and statistically significant at the 1 percent level in the first stage regressions.<sup>17</sup> The F-tests of the strength of the instruments are reported in the last rows of Panels A and B of Table 3, and indicate that they are strong instruments.

Using the lagged difference in migrant ratio as the instrument, the effect of the migrant ratio on the urban labour force employment rate is still positive (0.03) and statistically significant at the 5 percent level (column [3] of Table 3). Using the predicted difference in migrant ratio as the instrument, the coefficient of migrant ratio is positive but almost zero (0.009) and is statistically insignificant (column [4]).

The positive or zero effect of the migrant ratio on employment of local workers seems to be at odds with economic theory prediction but consistent with many previous findings for the U.S and the U.K. labour markets. Later in the paper we will examine further the channels through which such a positive effect may come about.

---

<sup>16</sup>Our preferred estimation is the first-difference combined with IV, where the IV used is the predicted difference in log migrant ratio.

<sup>17</sup>Note that the sign for the IV in the first stage estimation is opposite for IV1 and IV2. This is because the two IVs are measured differently. IV1 is defined as lagged difference in migrant ratio:  $(\log(R/U)_{it-1} - \log(R/U)_{it-2})$ , i.e. how change in the past affects the change now. The negative sign in the first stage indicates a catching up effect: a city which in the past has lower growth in its migrant ratio may have more room to increase its migrant ratio now. The second instrument (IV2) is defined as a predicted difference in the current migrant ratio:  $(\log(\frac{\hat{R}}{\hat{U}})_{it} - \log(\frac{\hat{R}}{\hat{U}})_{it-1})$ , where the push and pull factors used to predict  $\hat{R}_t$  and  $\hat{R}_{t-1}$  are lagged ( $Z_{kt-1}$  and  $Z_{kt-2}$ ). Thus, the second IV of predicted current migrant ratio should be positively correlated to the actual current migrant ratio ( $\Delta \text{Log}(R/U)_i$ ).



Next we examine the effect of rural migrant inflow on the average wages of the urban employees (Panel B of Table 3). The dependent variable used in this set of regressions is the log of city level average wage for urban local workers. These data are obtained from the City Statistical Yearbooks. The results from the OLS estimation (column [1]) shows that the impact of the migrant ratio on the log average wage of the urban labour force is also positive and statistically significant. The elasticity is 0.13, suggesting that every one percent increase in the migrant ratio increases urban workers' wages by 0.13 percent. The estimation using the first-difference method (column [2]) reduces the coefficient significantly and shows that there is no statistically significant impact of the migrant ratio on earnings of urban native workers. This dramatic change in the results suggests that perhaps the observed correlation in the OLS estimation is mainly due to the correlation between the variation in the unobserved city-level time-invariant characteristics and the variation in migrant ratios across cities. Controlling for city-fixed effects, therefore, washes out such a correlation.

Interestingly, though, when we further use the IV estimation combined with the first-difference method to mitigate possible bias generated by the omitted unobserved time-variant city characteristics, the magnitude of the coefficient once again increases. Using the lagged difference in migrant ratio as the instrument, the coefficient increases to 0.098 and is marginally significant at the 10 percent level. Using the predicted difference in the migrant ratio as the instrument, the coefficient is 0.047, and is not statistically significant. The fact that the IV with first-difference estimation results in slightly larger coefficients than the simple first-difference estimation indicates that the correlation between the omitted time-variant city unobservable characteristics (such as policy changes) and the migrant ratio may be negative. This makes sense as most of the policies were migration restricting ones and over time some city government have begun to reduce the restrictions, which lead to an increase in migrant ratio.

In summary, based on the cross-area approach, we find that rural-urban migration does not impose any negative impact on the employment or wage outcomes of urban local workers at the city average level. In fact, some evidence is found that rural migrant inflow may have modest positive effects on the employment rate and average wages of the urban labour force in the host cities.

## 5.2 *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 is no negative impact of rural migrant inflow on urban local workers' labour market outcomes. As discussed earlier, more than 95 per cent of rural migrants are employed as unskilled workers in host cities, and their

competing urban counterparts—unskilled urban workers—may be more likely to be affected. Thus, analysis at the average level may be misleading and what the impact of rural migrant inflow is on the labour market outcomes of unskilled urban native workers may be a more appropriate question to ask. This question is examined in this sub-section. We measure unskilled workers in two ways: by occupation—for those employed as service, production or agriculture workers; and by education—for those whose education level is at junior high school and below.

The estimated results for Equations (1) and (2) at city level for unskilled workers defined by occupation are reported in Table 4A. Panel A of the table presents the results on employment. Here ‘employment’ is defined as those who employed in the unskilled jobs (services, agriculture, or production) divided by total employment in the unskilled jobs plus those who are unemployed. The OLS estimates are presented in column [1]. As is shown, the coefficient for the log migrant ratio is positive and statistically significant. This indicates that unskilled urban workers’ employment opportunities are not hindered by the rural migrant inflow. When the first-difference method is adopted (column [2]), the result remains positive and significant. The first-difference combined with the difference in lagged migrant ratio as the IV (column [3]) is positive but statistically insignificant. This is also the case when we use the second IV (difference in predicted migrant ratio) as the instrument.<sup>18</sup>

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 Panel B of Table 4A. The OLS results show that the correlation between the migrant ratio and urban unskilled workers’ earnings is positive and statistically significant. The effect is even larger than the effect on the average wage of urban urban native labour force. 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 Mini-census). The OLS estimate for the log migrant ratio based on the 217 city sample in 2005 (column [5]) is similar to that for the 36-city sample in terms of the sign, magnitude and significance level.

When using the first-differences (column [2]) we find that the change in migration rate has a negative but insignificant impact on the change in urban workers’ earnings, while using first-difference combined with IV estimation, the coefficient of the change in rural migrant ratio once again turns to positive but insignificant (column [3]). This indicates that the impact of rural migrants on urban unskilled

---

<sup>18</sup>The first stage results are reported columns [1] and [2] in Table B2 of Appendix B. The instruments are very strong and statistically significant at the 1 percent level.

workers' wages is modest and insignificant overall.<sup>19</sup>

It is possible, though, that our definition of 'unskilled' with respect to occupation does not fully capture the effect on local unskilled workers. To test this, we re-define 'unskilled' in terms of education, which is also widely used in the immigration literature. We restrict the unskilled education groups to those having junior middle school education or below. The benefit of defining unskilled workers by their education level is that the employment rate for this group is directly available from the data. The limitation of using this definition is that low-educated rural migrants and urban workers may not be as substitutable as those within the same occupation group. However, the estimated results using this definition of 'unskilled' are remarkably similar to those obtained using the occupation definition (see Table 4B for the results).<sup>20</sup>

In summary, an increase in the migrant ratio appears to have a modest positive or zero impact on urban unskilled workers' employment. The effect on urban unskilled workers wages is also positive but not statistically different from zero. These results suggest that rural migrants and urban workers are not substitutes even within unskilled occupation or education cells. It is unfortunate that none of the census nor the 2005 mini-census provides detailed occupational categories. Nevertheless, many previous studies have documented that migrants are more likely to be hired in 3-D (dirty, dangerous, and demeaning) occupations (Zhao, 2000; Meng, 2000; and Meng and Zhang, 2001). Even based on the two digit occupation variable provided in the 2005 mini-census, we can still see some significant differences in occupational distribution between the migrant and urban native workers. For example, there are 5.6 per cent of the urban workers in the teacher category, while only 0.24 per cent of migrants are in the same occupation. Based on anecdotal evidence we also know that most migrant teachers are teaching in self-established migrant schools, while urban workers are employed in formal schools. We also find that 7 per cent of the migrants are employed in construction sites, while the ratio for urban workers is 2 per cent. Among construction workers, those who do interior finishing and installing appliances are very different from those who are brick layers and migrants are more likely to be the latter.

---

<sup>19</sup>The first stage results are reported in Columns [3] and [4] in Table D2 of Appendix D. The instrument (the difference in lagged migrant ratios) is very strong and statistically significant at the 1 percent level. However, the second instrument is statistically insignificant. We therefore do not report the IV results using this instrument in Panel B of Table 4A.

<sup>20</sup>We also estimated the employment and wage equations for young workers (those age 40 and below) because most migrant workers are between 20 and 40 years of age. The results are strikingly similar to those obtained for all workers or unskilled workers. These results are available upon request from the authors.

### 5.3 *Robustness check*

In this subsection, we examine how the results from the previous subsections may change when we (1) take into account individual characteristics of the urban labour force (undertaking the analysis at individual level), and (2) relax the ‘closed city labour market’ assumption.

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

$$Y_{ij}^t = \beta X_{ij}^t + \gamma City_j + \epsilon_i^t, \quad t = 1990, 2000, \text{ or } 2005 \quad (6)$$

$$\gamma_{jt} = \alpha + \theta \text{Log}(R/U)_{jt} + \delta D_t + \mu_{it}. \quad (7)$$

In the first step, as shown in Equation (6), the individual-level employment or wages ( $Y_{ij}$ ) are 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 ( $\gamma_{jt}$ ) is then extracted from the estimated Equation (6) and used as the dependent variable in the second step estimation as shown in Equation (7). The independent variables for the second step are the same as those included in the estimation of Equations (1) and (2).

The results based on the two-step procedure for the employment and earnings equations are presented in Table 5. These results are very similar to those obtained from the average city level analysis, suggesting that individual heterogeneity of the urban labour force is relatively independent of the rural migrant inflow.

Second, we examine whether our results are valid only under the ‘closed city labour market’ assumption. If the rural migrant inflow crowds out the urban local labour force, especially those unskilled workers, from some cities and moves them to other cities, the cross-area analysis may not be the right analytical strategy. It is, therefore, important to relax the closed city labour market assumption and use the cross-skill analysis to confirm the robustness of our main results. Differing from the cross-area analysis, the cross-skill analysis treats the nation as a labour market and compares wages across skill groups (ignoring geographic areas).

In our context, because more than 90 percent of rural migrants work in unskilled service and production jobs, we focus our analysis mainly on service and production workers. We divide the national-level labour market into forty skill groups, including two occupation (production and service workers ( $i = 1, 2$ )), four education (illiteracy, primary, junior middle, and senior high schools ( $j = 1, \dots, 4$ ))

and five age (ages between 15-25, 25-35, 35-45, 45-55, 55-65 ( $p = 1, \dots, 5$ )) groups as well as 3 years (1990, 2000 and 2005 ( $j = 1, 2, 3$ )). The model specification can be written as below:

$$\begin{aligned} \ln Wage_{ijpt} = & \beta \text{Log}(R/U)_{ijpt} + \text{occup}_i + \text{educ}_j + \text{ageg}_p + \text{year}_t \\ & + (\text{occup}_i \times \text{educ}_j) + (\text{occup}_i \times \text{ageg}_p) + (\text{occup}_i \times \text{year}_t) \\ & + (\text{educ}_j \times \text{ageg}_p) + (\text{educ}_j \times \text{year}_t) + (\text{ageg}_p \times \text{year}_t) + \epsilon_{it}, \end{aligned} \quad (8)$$

where the dependent variable is the logarithm of average wages for each skill cell;<sup>21</sup> the independent variables include the logarithm of the migrant ratio of each cell as well as the occupation, education, age, and year fixed effects and their interaction terms. The estimated coefficients for log migrant ratio in Equation (8) for regression with or without the interaction terms are -0.03 and -0.005, respectively, and are both statistically insignificant (see Table 6). This suggests that rural migrants and the urban labour force are not perfect substitutes even when we treat the whole economy as having a uniform labour market, and this is consistent with the previous estimation results for unskilled cells based on the cross-area analysis.

## 5.4 Pulling the Pieces Together

The analyses conducted thus far suggest that in China, although hundreds of millions of unskilled rural-urban migrants move to cities and one third of the urban labour force are migrant workers, migration per se has had no adverse impact on employment or earnings of urban native workers. The question naturally arises as to why our empirical findings do not conform with that predicted by economic theory?

There may be two possibilities. First, migrants and urban workers may operate in segregated labour markets and their substitutability may be very low. Indeed, as discussed earlier and in many previous studies, migrants are restricted from obtaining certain jobs, and hence jobs and earnings for local workers are insulated. The extent to which labour market segregation has prevented urban workers, even urban unskilled workers, from being affected by the influx of migrant workers, however, is unclear.

Assuming that migrants and urban local workers are not working completely in isolation, why, then, cannot we find any adverse labour market impact of migration on local workers? The answer is probably that migrants and urban

---

<sup>21</sup>Due to the difficulty of defining unemployment for each occupation, education, and age cell, the cross-skill analysis here only examines the effect of the inflow of migrants on urban native workers' wages.

local workers are complements to some extent. The fact that we observe some small positive effects of migrant inflow on urban native workers' employment provide some support to this possibility. Even though that at the unskilled-level analysis we also find small positive (though insignificant) effect it is possible that within a widely defined skill level there are still complementarities across narrowly defined jobs. For example, within the construction category we may have brick layers (migrants) and interior finishers (urban workers). The change in supply of the former generates demand for the latter, and hence, increases the employment for urban native workers.

It is important to understand that the analytical strategy we use in this paper is to consider the effect of an exogenous shock to the supply of migrants on labour market outcomes of the native workers (the instrumental variable we use suppose to have purged out all the other effects). In reality, the issue of what generated this supply shock, though, is also a significant part of the understanding of why such a large scale migration did not generate a significant unemployment either for the urban local people or for the migrants. To understand this issue we need to take a *dynamic* view of the goods and labour markets. If the large scale inflow of unskilled migrants is a result of an expansion of labour intensive industries, then there may not be any effect of migration on employment or wage reduction.

One way to examine whether the large scale migrant inflow is driven by demand expansion is to modify the relative-wage analysis developed in Katz and Murphy (1992) and Card and Lewis (2005). Considering the following equation:

$$\text{Log}(w^H/w^L)_{it} = \alpha + \beta \text{Log}(N^H/N^L)_{it} + \theta D_t + \epsilon_{it}, \quad (9)$$

where superscripts  $H$  and  $L$  denote the high- or low-skilled labour force, respectively, and subscripts  $i$  and  $t$  indicate city and year, respectively. The dependent variable is the logarithm of the annual earnings<sup>22</sup> ratio for high-skilled to low-skilled workers. As most of the low-skilled workers are migrants, this can be a proxy for urban native versus migrant worker. The independent variables include the logarithm of the ratio for the total number of skilled labour force to the total number of low-skilled labour force in city  $i$  in year  $t$ , and year dummies  $D_t$ . The basic idea of this method is that the relative supply of skilled-unskilled labour change should move the relative wage ratio along the downward sloping demand curve and this effect will be captured by  $\beta$ . If there is a significant reduction in the relative supply of skilled-unskilled labour (increase in the supply of unskilled labour, which is the effect of the inflow of migrants) without any shift in the demand curve, we should observe a positive effect on relative wages of skilled

---

<sup>22</sup>Since the information on hours worked is not available for the data, 'annual earnings' is used as a proxy for 'wages'. The effect of hours worked may be differenced out (at least partly) by constructing the relative earnings.

workers, and hence, a negative  $\beta$ . If, instead, we observe a positive or insignificant  $\beta$ , this suggests that there may be a relative demand curve shift, i.e. an relative increase in demand for migrant workers.

A potential problem for estimating Equation (9) in our paper is related to the measure of the relative wage between skilled and unskilled labour. Since wage information for rural migrants is only available in the 2005 Mini-census data, for other years we are unable to include migrant wage in the relative wage measure. Thus, the dependent variable used in Equation (9) is the logarithm of the annual earnings ratio for skilled to unskilled ‘urban workers’ ( $\frac{w^H}{w^L_u}$ ). As the overwhelming majority of migrants are unskilled, our worry is that wages for low-skilled urban workers varies differently across different cities relative to wage variation across cities for migrant workers. If that is the case, our test using only urban native workers’ relative wage may be misleading. Fortunately, with the 2005 mini-census data, we can test this issue empirically. Using data from the 2005 mini-census we find that the relative wages using only urban workers is highly correlated with the relative wage if we use both urban and migrant workers, with correlation coefficient being 0.96 (see Figure 3).

The estimated results from Equation (9) are reported in Table 7. The OLS and first-difference estimations (columns [1] and [2] of Table 7) show that the effect of the ratio of skilled-unskilled workers on the relative wage of skilled-unskilled workers is statistically insignificant, suggesting that a large influx of unskilled rural migrants does not widen the earnings gap between skilled and unskilled urban workers. This provides some evidence for the relative demand shift for the skilled urban workers.<sup>23</sup> Columns [3] and (4) of Table 7 report the results using the 2005 mini-census data where both the ratio of the skilled-unskilled workers and their relative wage ratio include urban local and migrant workers and the sample size is also much larger than in columns [1] and [2]. Here we also obtain insignificant or positive significant effect.<sup>24</sup>

An additional test may also be conducted to examine whether the large migrant inflow is associated with an increase in demand for unskilled workers. If changes in per capita GDP and changes in foreign direct investment (both capture the change in demand) are positively related to the change in the rural-urban migrant ratio. Estimating a regression using the logarithm of the change in migrant ratio as the dependent variable and logarithm of the change in per capita

---

<sup>23</sup>Note that the IV estimations are applied to deal with the potential endogeneity in Equation (9). However, in the first stage, the lagged rural-urban migration (as instrument) is not significantly correlated with the skilled-unskilled ratio in the labour market. Since IV estimations are invalid, these results are not reported or discussed here.

<sup>24</sup>The instrumental variable used in column 4 is predicted ratio of skilled to unskilled workers, where the number of skilled and unskilled migrant workers are predicted using the lagged push and pull factors as before.

GDP and change in FDI as independent variables, we find that at the city level an increase in the migrant ratio is associated with both the increase in per capita GDP and total FDI.<sup>25</sup> We also present these positive correlations in Figure 4. Of course, no attempt is made here to examine the causality of the issue.

The above discussion and empirical tests have led us to think that the reason for the non-existence of the adverse effect of the large scale rural-to-urban migration on urban native worker labour market outcomes is a combination of the labour market segregation and the complementarity of the migrant and urban workers. Furthermore, we believe that the large scale increase in rural migrants is associated with a significant increase in demand for unskilled labour in cities.

## 6 Conclusions

This paper explores the link between rural migrant inflows and labour market outcome for urban native workers in the Chinese urban labour markets.

We find that, if the empirical work is conducted at the city level (regarding cities as closed labour markets), the rural migrant inflow generally has a modest positive or zero impact on the employment and average wages of urban native workers. For unskilled workers (defined either by occupation or education level), who are more likely to be substitutes for migrant workers, we still do not observe any negative effect.

We then test whether the assumption that cities are closed labour markets is the reason for generating these unexpected results. Even when the nation is treated as an integrated labour market, we still cannot find any adverse effect.

To reconcile these findings with economic theory, we propose two conjectures. First, because of the special institutional setting of the rural-urban migration in China, where migrants are, to a certain extent, regarded as ‘secondary citizens’, migrants and urban local workers are operating in segregated labour markets. Many existing studies have confirmed that there is a labour market segregation. If this is the case, migration inflow should have limited impact on urban local workers’ labour market outcomes. However, if there is no perfect segregation, we should still find some negative impact. This leads to the second conjecture. In the absence of complete segregation, our results seem to suggest that migrants and urban natives are, to some extent, complements.

Finally we briefly investigate what, in the first place, generated the increase

---

<sup>25</sup>The estimated results are:

$$\begin{aligned} \Delta \log(migratio) = & 10.55 + 0.152 * \Delta \log(FDI) + 0.671 * \Delta \log(GDP\_pc) \\ & (1.107) \quad (0.054) \quad (0.135) \end{aligned}$$



supply of rural migrants. We find that increase supply is associated with a significant increase in demand for labour.

The question remains as to whether a future labour market reform, which eliminates the labour market segregation, will lead to some adverse effects of migration on urban native labour market outcomes. This, to a certain extent, depends on the change in demand for migrant labour generated by economic growth. If the speed of economic growth is fast enough to absorb migrant workers, labour market deregulation may not necessarily lead to worse labour market outcomes for urban local workers. As large scale rural-urban migration will continue during the Chinese urbanisation process, understanding the policy options is extremely important. To this end, more evidence and vigorous empirical tests are needed to provide a conclusive explanation as to why the large scale rural-urban migration has had an insignificant impact on urban natives.

## References

- [1] 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.
- [2] Angrist, Joshua D. and Adriana D. Kugler, "Protective or Counter-productive? Labour Market Institutions and the Effect of Immigration on EU Natives," *Economic Journal*, Jun. 2003, 113(488), 302-31.
- [3] Bao, Shuming, Orn 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.
- [4] Borjas, George J., "The Economics of Immigration," *Journal of Economic Literature*, Dec. 1994, 32, 1667-717.
- [5] Borjas, George J., "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.
- [6] Borjas, George J. 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.
- [7] Boustan, Leah P., 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.
- [8] Cai, Fang, Yang Du, and Meiyang Wang, "Hukou System and Labor Market Protection in China," *Economic Research (jing ji yan jiu, in Chinese)*, 2001, 12, 41-49.
- [9] Card, David, "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 2001, 19(1), 22-64.
- [10] Card, David, "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.

- [11] Card, David 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.
- [12] Card, David and John E. DiNardo, "Do Immigrant Inflows Lead to Native Outflows?" National Bureau of Economic Research Working Papers No.7578, Mar. 2000.
- [13] 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.
- [14] de Brauw, Alan and John Giles, "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.
- [15] 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.
- [16] Dustmann, Christian and Francesca Fabbri, "Language Proficiency and Labour Market Performance of Immigrants in the UK," *The Economic Journal*, 2003, 113(489), 695-717.
- [17] 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, Apr. 2005.
- [18] 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.
- [19] 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, 110-51.
- [20] Grossman, Jean Baldwin, "The Substitutability of Natives and Immigrants in Production," *The Review of Economics and Statistics*, Nov. 1982, 64(4), 596-603.

- [21] Katz, Lawrence F. and Kevin M. Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, Feb. 1992, 107(1), 35-78.
- [22] 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.
- [23] Manacorda, Marco, 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.
- [24] Meng, Xin, *Labour Market Reform in China*, Cambridge: Cambridge University Press, 2000.
- [25] Meng, Xin 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.
- [26] Ottaviano, Gianmarco I.P. and Giovanni Peri, "Rethinking the Effects of Immigration on Wages," National Bureau of Economic Research Working Papers No.12497, Aug. 2006.
- [27] Rozelle, Scott, J. Edward Taylor, and Alan de Brauw, "Migration, Remittances, and Agricultural Productivity in China," *American Economic Review*, May 1999, 89(2), 287-91.
- [28] Smith, James P. and Barry Edmonston, *The New Americans: Economic, Demographic, and Fiscal Effects of Immigration*, Washington, D.C.: National Academy Press, 1997.
- [29] Wang, Hansheng and Xiaoqiang Wang, 'Zai liu dong zhong shi xian jing ying yi min' [Migration and migration selection], *Strategy and Management*, 1995, 5: 112-20.
- [30] Woo, Wing Thye, "Chinese Economic Growth: Sources and Prospects," in M. Fouquin and F. Lemoine eds., *The Chinese Economy*, London: Economica, 1998.
- [31] Wooldridge, Jeffrey M., "Cluster-Sample Methods in Applied Econometrics," *American Economic Review*, May 2003, 93(2), 133-38.

- [32] Xiang, Biao, “Chuan tong he xin shihui kong jin de sheng cheng” [Mobility, Traditional Networking, Marketization and the Development of a ‘Non-Government’ Controlled Space], *Strategy and Management*, 1996, 6.
- [33] Zhao, Yaohui, “Leaving the Countryside: Rural to Urban Migration Decisions in China,” *The American Economic Review*, May 1999, 89(2), 281-286.
- [34] Zhao, Yaohui, “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, 15-33.
- [35] 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.
- [36] Zhao, Zhong, “Migration, Labor Market Flexibility, and Wage Determination in China: A Review,” *The Developing Economies*, Jul. 2005, 43(2), 285-312.

**Table 1: Summary Statistics for 152-City Sample**

|   | 1990         |          | 2000         |          | 2005         |          |
|---|--------------|----------|--------------|----------|--------------|----------|
|   | Mean         | Std.Dev. | Mean         | Std.Dev. | Mean         | Std.Dev. |
| <i>City Level</i>   |              |          |              |          |              |          |
| Migrant ratio <sup>a</sup>  | <b>0.09</b>  | 0.12     | <b>0.15</b>  | 0.26     | <b>0.23</b>  | 0.40     |
| Employment rate for urban labour force (%) <sup>a</sup>                                 | <b>95.92</b> | 2.95     | <b>86.98</b> | 5.18     | <b>89.65</b> | 4.93     |
| Employment rate for unskilled urban labour force (%) <sup>a</sup>                       | <b>95.93</b> | 2.27     | <b>81.80</b> | 6.56     | <b>83.46</b> | 6.82     |
| Average real annual wage for all employed urban workers (Yuan) <sup>b</sup>             | <b>2283</b>  | 383      | <b>5083</b>  | 1582     | <b>9094</b>  | 2410     |
| Average real annual wage for urban workers in unskilled occupations (Yuan) <sup>c</sup> | <b>1473</b>  | 330      | <b>2792</b>  | 1235     | <b>8798</b>  | 2951     |
| Average real annual wage for urban workers in low education (Yuan) <sup>c</sup>         | <b>1718</b>  | 362      | <b>3104</b>  | 1962     | <b>9353</b>  | 3803     |
| City permanent population (millions) <sup>b</sup>                                       | <b>1.05</b>  | 1.11     | <b>1.43</b>  | 1.74     | <b>1.72</b>  | 2.07     |
| Average age of all urban labour force <sup>a</sup>                                      | <b>34.48</b> | 1.49     | <b>35.70</b> | 1.04     | <b>37.37</b> | 1.04     |
| Proportion of male among urban labour force (%) <sup>a</sup>                            | <b>57.17</b> | 4.95     | <b>56.91</b> | 2.13     | <b>56.46</b> | 2.83     |
| Proportion of high school graduates among urban labour force (%) <sup>a</sup>           | <b>30.83</b> | 11.65    | <b>39.32</b> | 10.16    | <b>46.07</b> | 11.84    |
| Actual foreign investment (million US dollars) <sup>b</sup>                             | <b>0.03</b>  | 0.09     | <b>0.22</b>  | 0.64     | <b>0.42</b>  | 0.94     |
| Share of value added in secondary industry (%) <sup>b</sup>                             | -            | -        | <b>52.15</b> | 10.70    | <b>53.80</b> | 11.02    |
| Share of value added in tertiary industry (%) <sup>b</sup>                              | -            | -        | <b>42.70</b> | 9.97     | <b>41.90</b> | 10.75    |

Note: <sup>a</sup> data are taken from 1990, 2000 and 2005 Census; <sup>b</sup> data are taken from the City Statistical Yearbook in 1991, 2001, and 2006; <sup>c</sup> data are taken from Urban Household Income and Expenditure Survey and are based on 36 city sample.

**Table 2:** Occupational and Educational Distributions in 152 Chinese Cities, 1990-2005

|                                    | Urban Labour Force |         |         | Rural Migrants |         |        |
|------------------------------------|--------------------|---------|---------|----------------|---------|--------|
|                                    | 1990               | 2000    | 2005    | 1990           | 2000    | 2005   |
| <i>Education</i>                   |                    |         |         |                |         |        |
| Illiteracy / Never being in school | 1.51               | 0.37    | 0.33    | 6.60           | 2.71    | 2.40   |
| Primary school                     | 11.28              | 4.42    | 3.77    | 26.40          | 19.38   | 17.08  |
| Junior middle school               | 42.03              | 32.45   | 28.43   | 56.02          | 64.12   | 61.83  |
| Senior middle school               | 33.98              | 38.61   | 34.68   | 10.87          | 12.98   | 16.30  |
| Junior college                     | 6.61               | 15.35   | 19.36   | 0.09           | 0.73    | 2.00   |
| University and above               | 4.59               | 8.81    | 13.43   | 0.01           | 0.09    | 0.39   |
| Number of observations             | 735,286            | 698,035 | 167,809 | 57,136         | 144,433 | 54,967 |
| <i>Occupation</i>                  |                    |         |         |                |         |        |
| High level officer                 | 14.38              | 6.34    | 5.33    | 1.25           | 1.25    | 1.85   |
| Professional                       | 12.20              | 22.12   | 23.00   | 1.30           | 1.50    | 2.70   |
| Clerk                              | 6.92               | 13.99   | 15.42   | 1.42           | 3.15    | 3.59   |
| Service worker                     | 7.93               | 22.40   | 26.78   | 13.49          | 31.76   | 35.20  |
| Agricultural worker                | 8.22               | 1.53    | 2.26    | 13.61          | 4.21    | 2.18   |
| Production worker                  | 50.35              | 33.62   | 27.22   | 68.93          | 58.12   | 54.48  |
| Number of observations             | 708,457            | 604,657 | 150,460 | 56,921         | 140,844 | 53,140 |

Note: Authors' own calculations based on the the 1990, 2000 and 2005 Censuses with restricted labour force sample in the 152 cities. The calculation for 2005 considers the sampling weight across cities.

**Table 3: City-Level Analysis**

|  | OLS                 | FD                  | FD & IV1            | FD & IV2            |
|--|---------------------|---------------------|---------------------|---------------------|
|  | [1]                 | [2]                 | [3]                 | [4]                 |
| <b>Panel A - Dependent variable: employment rate for urban labour force</b>    |                     |                     |                     |                     |
| Log migrant ratio  | 0.020***<br>(0.003) | 0.021***<br>(0.006) | 0.029*<br>(0.015)   | 0.009<br>(0.012)    |
| Log city population  | -0.005<br>(0.005)   | 0.003<br>(0.012)    | 0.004<br>(0.012)    | 0.003<br>(0.012)    |
| Average age of urban LF  | -0.005*<br>(0.003)  | 0.006<br>(0.005)    | 0.007<br>(0.005)    | 0.006<br>(0.005)    |
| % males for urban LF   | 0.220<br>(0.136)    | 0.291*<br>(0.168)   | 0.282*<br>(0.164)   | 0.303*<br>(0.160)   |
| % of skilled urban LF  | 0.167***<br>(0.037) | 0.129**<br>(0.063)  | 0.143**<br>(0.071)  | 0.109*<br>(0.062)   |
| Actual foreign investment  | -0.006*<br>(0.003)  | 0.002<br>(0.010)    | 0.002<br>(0.010)    | 0.003<br>(0.010)    |
| Share of value added in secondary industry                                     | 0.036<br>(0.064)    | 0.195<br>(0.159)    | 0.220<br>(0.162)    | 0.160<br>(0.168)    |
| Share of value added in tertiary industry                                      | -0.034<br>(0.065)   | 0.212<br>(0.163)    | 0.240<br>(0.168)    | 0.172<br>(0.168)    |
| Year dummy for 2005  | 0.022***<br>(0.007) | -<br>-              | -<br>-              | -<br>-              |
| Constant   | 0.883***<br>(0.130) | 0.001<br>(0.012)    | -0.004<br>(0.015)   | 0.008<br>(0.013)    |
| Number of observations   | 304                 | 152                 | 152                 | 152                 |
| R <sup>2</sup>   | 0.343               | 0.145               | -                   | -                   |
| F-test statistic for excluded instrument                                       | -                   | -                   | 12.72               | 18.98               |
| <b>Panel B - Dependent variable: log (average wage) for urban labour force</b> |                     |                     |                     |                     |
| Log migrant ratio  | 0.118***<br>(0.013) | 0.015<br>(0.017)    | 0.091<br>(0.056)    | 0.047<br>(0.040)    |
| Log city population  | 0.013<br>(0.020)    | -0.070<br>(0.054)   | -0.065<br>(0.057)   | -0.068<br>(0.053)   |
| Average age of urban LF  | 0.015<br>(0.011)    | 0.020<br>(0.015)    | 0.025<br>(0.016)    | 0.022<br>(0.015)    |
| % males for urban LF   | -0.201<br>(0.596)   | 0.783<br>(0.614)    | 0.704<br>(0.615)    | 0.750<br>(0.608)    |
| % of skilled urban LF  | 0.248*<br>(0.132)   | 0.373**<br>(0.189)  | 0.504**<br>(0.201)  | 0.429**<br>(0.178)  |
| Actual foreign investment  | 0.096***<br>(0.019) | 0.026<br>(0.031)    | 0.020<br>(0.030)    | 0.024<br>(0.030)    |
| Share of value added in secondary industry                                     | 0.889**<br>(0.427)  | -0.528<br>(1.062)   | -0.299<br>(1.057)   | -0.431<br>(0.998)   |
| Share of value added in tertiary industry                                      | 0.569<br>(0.439)    | -0.516<br>(1.061)   | -0.256<br>(1.056)   | -0.406<br>(0.996)   |
| Year dummy for 2005  | 0.457***<br>(0.028) | -<br>-              | -<br>-              | -<br>-              |
| Constant   | 7.437***<br>(0.633) | 0.532***<br>(0.043) | 0.486***<br>(0.053) | 0.513***<br>(0.042) |
| Number of observations   | 304                 | 152                 | 152                 | 152                 |
| R <sup>2</sup>   | 0.777               | 0.084               | -                   | -                   |
| F-test statistic for excluded instrument                                       | -                   | -                   | 12.72               | 18.98               |

Note: Robustness standard errors are displayed in parentheses.

\*\*\*Significant at 1% level; \*\*Significant at 5% level; \*Significant at 10% level.

IV1 refers to the difference in the lagged log migrant ratio.

IV2 is the difference in the predicted log migrant ratio.



**Table 4A: Analysis at Unskilled Level (Defined by Occupation)**

|   | <b>OLS<br/>[1]</b>  | <b>FD<br/>[2]</b>   | <b>FD &amp; IV1<br/>[3]</b> | <b>FD &amp; IV2<br/>[4]</b> | <b>2005 OLS<br/>[5]</b> |
|---|---------------------|---------------------|-----------------------------|-----------------------------|-------------------------|
| <i><b>Panel A</b> - Dependent variable: employment rate for unskilled urban labour force</i>        |                     |                     |                             |                             |                         |
| Log migrant ratio   | 0.020***<br>(0.004) | 0.020**<br>(0.009)  | 0.032<br>(0.026)            | 0.039<br>(0.024)            |                         |
| Log city population   | -0.004<br>(0.006)   | 0.007<br>(0.018)    | 0.007<br>(0.018)            | 0.007<br>(0.019)            |                         |
| Average age of urban LF   | -0.007**<br>(0.003) | 0.012**<br>(0.006)  | 0.013**<br>(0.006)          | 0.013**<br>(0.006)          |                         |
| % males for urban LF  | -0.136<br>(0.120)   | 0.119<br>(0.165)    | 0.126<br>(0.165)            | 0.130<br>(0.167)            |                         |
| Actual foreign investment   | -0.007<br>(0.005)   | 0.011<br>(0.014)    | 0.009<br>(0.014)            | 0.009<br>(0.014)            |                         |
| Share of value added in secondary industry  | 0.170*<br>(0.095)   | 0.350<br>(0.242)    | 0.370<br>(0.235)            | 0.381<br>(0.235)            |                         |
| Share of value added in tertiary industry   | 0.050<br>(0.097)    | 0.375<br>(0.247)    | 0.403*<br>(0.242)           | 0.420*<br>(0.248)           |                         |
| Year dummy for 2005   | 0.048***<br>(0.011) | -<br>-              | -<br>-                      |                             |                         |
| Constant  | 1.074***<br>(0.160) | -0.007<br>(0.018)   | -0.012<br>(0.020)           | -0.015<br>(0.019)           |                         |
| Number of observations  | 304                 | 152                 | 152                         | 152                         |                         |
| R <sup>2</sup>  | 0.222               | 0.079               |                             |                             |                         |
| F-test statistic for excluded instrument  | -                   | -                   | 9.61                        | 13.32                       |                         |
| <i><b>Panel B</b> - Dependent variable: log (average earnings) for unskilled urban labour force</i> |                     |                     |                             |                             |                         |
| Log migrant ratio   | 0.157***<br>(0.037) | -0.051<br>(0.085)   | 0.051<br>(0.108)            |                             | 0.135***<br>(0.013)     |
| Log city population   | 0.028<br>(0.041)    | -0.391<br>(0.243)   | -0.325<br>(0.228)           |                             | 0.007<br>(0.023)        |
| Average age of urban LF   | -0.012<br>(0.019)   | -0.006<br>(0.049)   | -0.011<br>(0.044)           |                             | -0.028***<br>(0.009)    |
| % males for urban LF  | 2.118**<br>(0.925)  | 0.411<br>(1.081)    | 0.641<br>(1.058)            |                             | 0.478<br>(0.308)        |
| Actual foreign investment   | 0.092***<br>(0.030) | 0.065<br>(0.059)    | 0.059<br>(0.050)            |                             | 0.040***<br>(0.014)     |
| Share of value added in secondary industry  | 1.529<br>(0.946)    | -3.083<br>(2.512)   | -2.449<br>(2.582)           |                             | 1.076**<br>(0.419)      |
| Share of value added in tertiary industry   | 0.499<br>(0.988)    | -1.042<br>(2.496)   | -0.365<br>(2.460)           |                             | 0.858**<br>(0.392)      |
| Year dummy for 2005   | 1.054***<br>(0.077) |                     | -<br>-                      |                             |                         |
| Constant  | 6.154***<br>(1.381) | 1.233***<br>(0.120) | 1.186***<br>(0.117)         |                             | 8.267***<br>(0.556)     |
| Number of observations  | 72                  | 36                  | 36                          |                             | 152                     |
| R <sup>2</sup>  | 0.910               | 0.320               |                             |                             | 0.560                   |
| F-test statistic for excluded instrument  | -                   | -                   | 9.71                        |                             | -                       |

Note: Robustness standard errors are displayed in parentheses.

\*\*\*Significant at 1% level; \*\*Significant at 5% level; \*Significant at 10% level.

IV1 refers to the difference in the lagged log migrant ratio.

IV2 is the difference in the predicted log migrant ratio.

**Table 4B:** Analysis on Unskilled Level (Defined by Education)

|   | OLS<br>[1]           | FD<br>[2]           | FD & IV1<br>[3]     | FD & IV2<br>[4]     | OLS 2005<br>[5]     |
|---|----------------------|---------------------|---------------------|---------------------|---------------------|
| <b>Panel A</b> - Dependent variable: employment rate for lowly educated urban labour force        |                      |                     |                     |                     |                     |
| Log migrant ratio   | 0.025***<br>(0.004)  | 0.027***<br>(0.009) | 0.054**<br>(0.025)  | 0.019<br>(0.018)    |                     |
| Log city population   | -0.009<br>(0.007)    | 0.007<br>(0.016)    | 0.009<br>(0.016)    | 0.006<br>(0.016)    |                     |
| Average age of urban LF   | 0.001<br>(0.002)     | 0.011**<br>(0.005)  | 0.011**<br>(0.005)  | 0.011**<br>(0.005)  |                     |
| % males for urban LF  | 0.088<br>(0.086)     | 0.372**<br>(0.146)  | 0.351**<br>(0.156)  | 0.378***<br>(0.138) |                     |
| Actual foreign investment   | -0.003<br>(0.005)    | 0.001<br>(0.013)    | -0.002<br>(0.014)   | 0.002<br>(0.013)    |                     |
| Share of value added in secondary industry  | 0.024<br>(0.091)     | 0.398*<br>(0.205)   | 0.491**<br>(0.219)  | 0.372*<br>(0.224)   |                     |
| Share of value added in tertiary industry   | -0.113<br>(0.096)    | 0.294<br>(0.210)    | 0.385*<br>(0.220)   | 0.268<br>(0.227)    |                     |
| Year dummy for 2005   | 0.015*<br>(0.008)    | -                   | -                   | -                   |                     |
| Constant  | 0.860***<br>(0.130)  | -0.008<br>(0.013)   | -0.022<br>(0.019)   | -0.004<br>(0.014)   |                     |
| Number of observations  | 304                  | 152                 | 152                 | 152                 |                     |
| R <sup>2</sup>  | 0.194                | 0.168               | -                   | -                   |                     |
| F-test statistic for excluded instrument  | -                    | -                   | 14.25               | 22.70               |                     |
| <b>Panel B</b> - Dependent variable: log (average earnings) for lowly educated urban labour force |                      |                     |                     |                     |                     |
| Log migrant ratio   | 0.130***<br>(0.037)  | 0.035<br>(0.059)    | 0.191<br>(0.167)    |                     | 0.127***<br>(0.015) |
| Log city population   | 0.085*<br>(0.048)    | -0.530**<br>(0.213) | -0.358<br>(0.301)   |                     | -0.018<br>(0.026)   |
| Average age of urban LF   | -0.065***<br>(0.017) | 0.046<br>(0.041)    | 0.058<br>(0.042)    |                     | 0.009<br>(0.011)    |
| % males for urban LF  | -1.018<br>(0.781)    | 0.454<br>(1.022)    | 0.396<br>(0.940)    |                     | 0.028<br>(0.255)    |
| Actual foreign investment   | 0.094***<br>(0.028)  | 0.033<br>(0.052)    | 0.027<br>(0.050)    |                     | 0.041**<br>(0.017)  |
| Share of value added in secondary industry  | 0.641<br>(1.033)     | -0.684<br>(2.166)   | 1.561<br>(3.615)    |                     | 0.672<br>(0.441)    |
| Share of value added in tertiary industry   | -0.044<br>(1.064)    | 1.296<br>(2.303)    | 3.584<br>(3.607)    |                     | 0.389<br>(0.428)    |
| Year dummy for 2005   | 1.122***<br>(0.055)  | -                   | -                   |                     |                     |
| Constant  | 10.403***<br>(1.311) | 1.112***<br>(0.114) | 0.973***<br>(0.184) |                     | 7.482***<br>(0.521) |
| Number of observations  | 72                   | 36                  | 36                  |                     | 152                 |
| R <sup>2</sup>  | 0.909                | 0.381               | -                   |                     | 0.505               |
| F-test statistic for excluded instrument  | -                    | -                   | 6.87                |                     | -                   |

Note: Robustness standard errors are displayed in parentheses.

\*\*\*Significant at 1% level; \*\*Significant at 5% level; \*Significant at 10% level.

IV1 refers to the difference in the lagged log migrant ratio.

IV2 indicates the difference in the predicted log migrant ratio.

**Table 5: Individual Level Analysis**

|  | OLS<br>[1]          | FD<br>[2]           | FD & IV1<br>[3]     | FD & IV2<br>[4]     |
|--|---------------------|---------------------|---------------------|---------------------|
| <i>Panel A - Dependent variable: individual employment rate for urban labour force</i> |                     |                     |                     |                     |
| Log migrant ratio  | 0.021***<br>(0.003) | 0.019***<br>(0.006) | 0.038**<br>(0.017)  | 0.005<br>(0.012)    |
| Log city population  | -0.001<br>(0.004)   | -0.010<br>(0.012)   | -0.010<br>(0.013)   | -0.010<br>(0.011)   |
| Actual foreign investment  | -0.005<br>(0.004)   | 0.000<br>(0.009)    | -0.002<br>(0.010)   | 0.002<br>(0.009)    |
| Share of value added in secondary industry   | 0.157**<br>(0.067)  | 0.066<br>(0.172)    | 0.104<br>(0.180)    | 0.040<br>(0.179)    |
| Share of value added in tertiary industry  | 0.091<br>(0.070)    | 0.084<br>(0.168)    | 0.135<br>(0.184)    | 0.047<br>(0.173)    |
| Year dummy for 2005  | 0.016***<br>(0.005) | -<br>-              | -<br>-              |                     |
| Constant   | 0.807***<br>(0.067) | 0.018***<br>(0.006) | 0.011<br>(0.008)    | 0.023***<br>(0.008) |
| Number of observations   | 304                 | 152                 | 152                 | 152                 |
| R <sup>2</sup>   | 0.253               | 0.078               | -                   | -                   |
| F-test statistic for excluded instrument   | -                   | -                   | 10.44               | 18.55               |
| <i>Panel B - Dependent variable: individual log wage for urban labour force</i>        |                     |                     |                     |                     |
| Log migrant ratio  | 0.179***<br>(0.034) | -0.023<br>(0.065)   | 0.106<br>(0.104)    |                     |
| Log city population  | -0.037<br>(0.029)   | -0.412**<br>(0.163) | -0.287<br>(0.182)   |                     |
| Actual foreign investment  | 0.107***<br>(0.019) | 0.052*<br>(0.027)   | 0.044<br>(0.029)    |                     |
| Share of value added in secondary industry   | 0.861<br>(0.711)    | -1.215<br>(1.802)   | -0.108<br>(2.205)   |                     |
| Share of value added in tertiary industry  | 0.057<br>(0.705)    | -0.023<br>(1.787)   | 1.145<br>(2.086)    |                     |
| Year dummy for 2005  | 0.721***<br>(0.047) | -<br>-              | -<br>-              |                     |
| Constant   | 5.890***<br>(0.774) | 0.885***<br>(0.056) | 0.809***<br>(0.070) |                     |
| Number of observations   | 72                  | 36                  | 36                  |                     |
| R <sup>2</sup>   | 0.897               | 0.309               | -                   |                     |
| F-test statistic for excluded instrument   | -                   | -                   | 15.33               |                     |

Note: Robustness standard errors are displayed in parentheses.

\*\*\*Significant at 1% level; \*\*Significant at 5% level; \*Significant at 10% level.

The results in this table are based on the two-step procedure.

IV1 refers to the difference in the lagged log migrant ratio.

IV2 indicates the difference in the predicted log migrant ratio.

**Table 6: Cross-Skill Analysis**

|   | [1]                  | [2]                 |
|---|----------------------|---------------------|
| <i>Dependent variable: individual log wage for urban labour force</i> |                      |                     |
| Log migrant ratio   | -0.005<br>(0.039)    | -0.030<br>(0.038)   |
| Dummy for trade and service occupations                               | -0.199***<br>(0.044) | 0.039<br>(0.119)    |
| Dummy for primary school  | -0.023<br>(0.104)    | -0.294*<br>(0.174)  |
| Dummy for junior middle school  | 0.162<br>(0.147)     | -0.052<br>(0.123)   |
| Dummy for senior middle school  | 0.307<br>(0.196)     | -0.044<br>(0.129)   |
| Dummy for 25-34 age group   | 0.225**<br>(0.102)   | -0.011<br>(0.203)   |
| Dummy for 35-44 age group   | 0.406***<br>(0.128)  | 0.266**<br>(0.130)  |
| Dummy for 45-54 age group   | 0.471***<br>(0.158)  | 0.375***<br>(0.116) |
| Dummy for 55-65 age group   | 0.288*<br>(0.164)    | 0.217<br>(0.133)    |
| Dummy for year 2000   | 0.652***<br>(0.081)  | 1.451***<br>(0.190) |
| Dummy for year 2005   | 1.844***<br>(0.101)  | 2.360***<br>(0.167) |
| Interaction between occup & educ                                      | No                   | Yes                 |
| Interaction between occup & age group                                 | No                   | Yes                 |
| Interaction between occup & year                                      | No                   | Yes                 |
| Interaction between educ & age group                                  | No                   | Yes                 |
| Interaction between educ & year                                       | No                   | Yes                 |
| Interaction between age group & year                                  | No                   | Yes                 |
| Constant  | 6.948***<br>(0.132)  | 6.871***<br>(0.121) |
| Number of observations  | 118                  | 118                 |
| R <sup>2</sup>  | 0.937                | 0.980               |

Note: Robustness standard errors are displayed in parentheses.

\*\*\*Significant at 1% level; \*\*Significant at 5% level; \*Significant at 10% level.

The results in this table are based on the Borjas's Cross-skill analysis. The reference groups dropped in the regressions are production workers aged between 15-25, illiteracy, and year 1990.

**Table 7: Relative Wage Analysis**

| <i>Dependent variable:</i>                 | <b>log wage ratio for urban<br/>workers in 2000 &amp; 2005</b> |                     | <b>log wage ratio for all<br/>workers in 2005</b> |                     |
|--|--|---------------------|---|---------------------|
|  | <b>OLS</b>   | <b>FD</b>           | <b>OLS</b>  | <b>2SLS</b>         |
|  | <b>[1]</b>   | <b>[2]</b>          | <b>[3]</b>  | <b>[4]</b>          |
| Log skilled to unskilled ratio             | 0.039<br>(0.087)   | 0.210<br>(0.141)    | -0.021<br>(0.030)                                 | 0.075*<br>(0.041)   |
| Log of city population                     | -0.066*<br>(0.036)   | 0.246<br>(0.287)    | 0.054**<br>(0.028)                                | 0.060**<br>(0.028)  |
| Proportion of male urban labour force      | 1.579<br>(1.366)   | -0.723<br>(1.957)   | -0.825<br>(0.553)                                 | -0.587<br>(0.565)   |
| Average age of urban labour force          | 0.022<br>(0.021)   | -0.126**<br>(0.058) | -0.031**<br>(0.014)                               | -0.035**<br>(0.014) |
| Foreign direct investment                  | 0.031*<br>(0.018)  | -0.105<br>(0.065)   | -0.029<br>(0.021)                                 | -0.030<br>(0.020)   |
| Share of value added in secondary industry | 0.105<br>(0.899)   | 1.019<br>(1.494)    | -0.225<br>(0.429)                                 | -0.613<br>(0.452)   |
| Share of value added in Tertiary industry  | 0.084<br>(0.916)   | -0.884<br>(1.927)   | -0.132<br>(0.435)                                 | -0.592<br>(0.469)   |
| Year dummy for 2005                        | 0.089<br>(0.057)   | -<br>-              | -<br>-  | -<br>-              |
| Constant                                   | -1.081<br>(1.714)  | 0.202<br>(0.140)    | 2.099***<br>(0.739)                               | 2.498***<br>(0.754) |
| Number of observations                     | 72   | 36                  | 152   | 152                 |
| R <sup>2</sup>                             | 0.216  | 0.323               | 0.094   | 0.042               |
| F-test statistic for excluded instrument   | -  | -                   | -   | 185.85              |

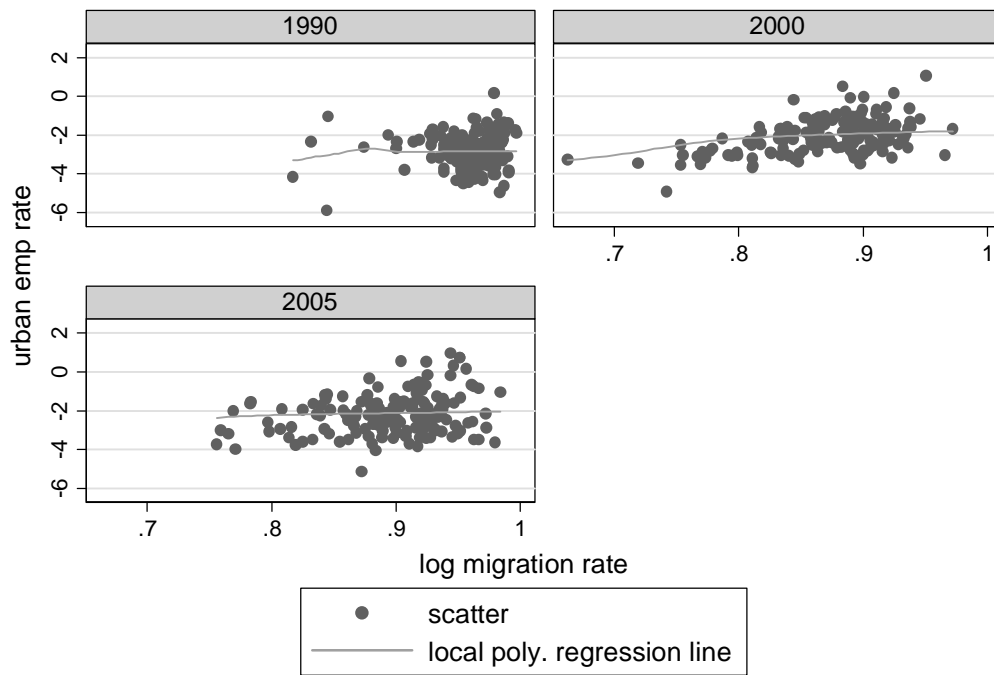
Note: Robustness standard errors are displayed in parentheses.

\*\*\*Significant at 1% level; \*\*Significant at 5% level; \*Significant at 10% level.

In columns [1] and [2], wage information is taken from the Urban Household survey, while in columns [3] and [4], the wage information is taken from the 2005 Mini-census data.

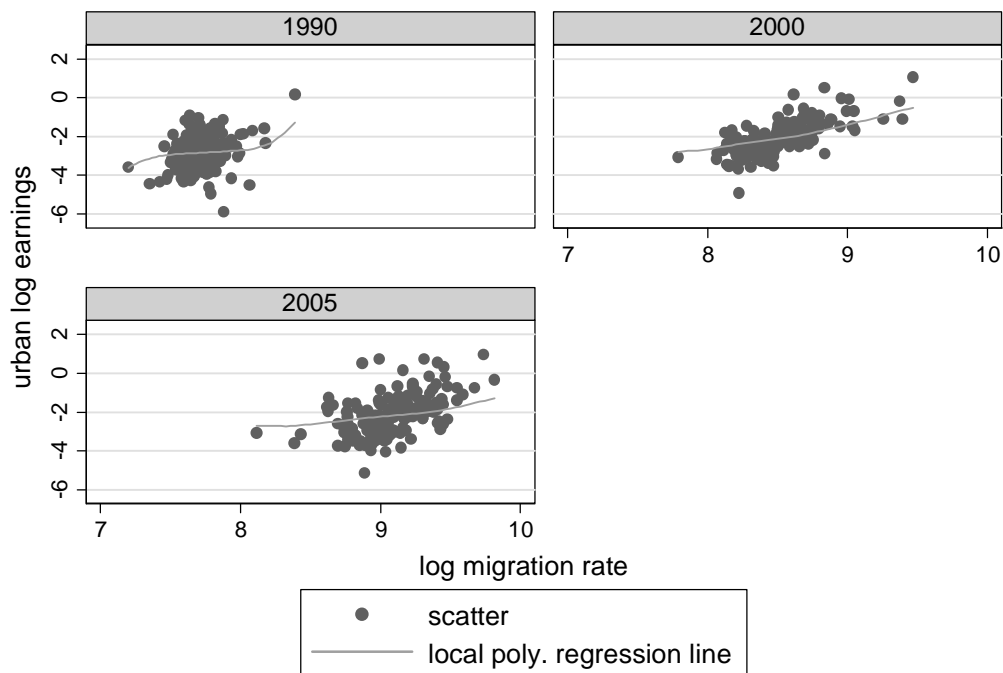
The instrumental variable used in column [4] is predicted ratio of skilled to unskilled workers, where the number of skilled and unskilled migrant workers are predicted using the lagged push and pull factors, as before.

**Figure 1A:** Unconditional Relationship between Urban Employment Rate and Log Migrant Ratio, by Year



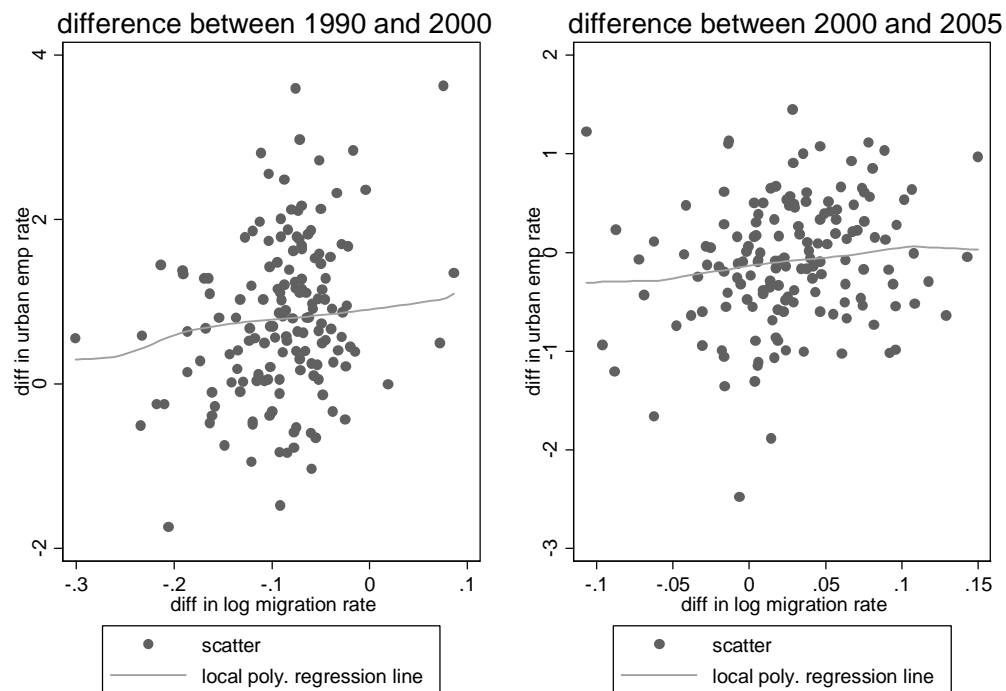
Graphs by year

**Figure 1B:** Unconditional Relationship between Urban Log Earnings and Log Migrant Ratio, by Year

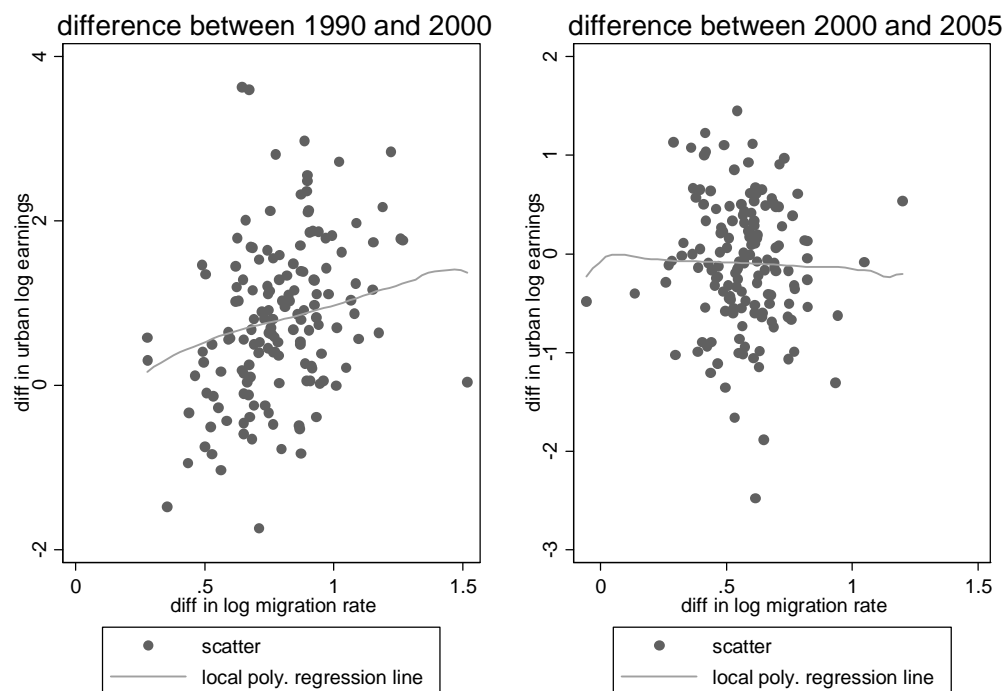


Graphs by year

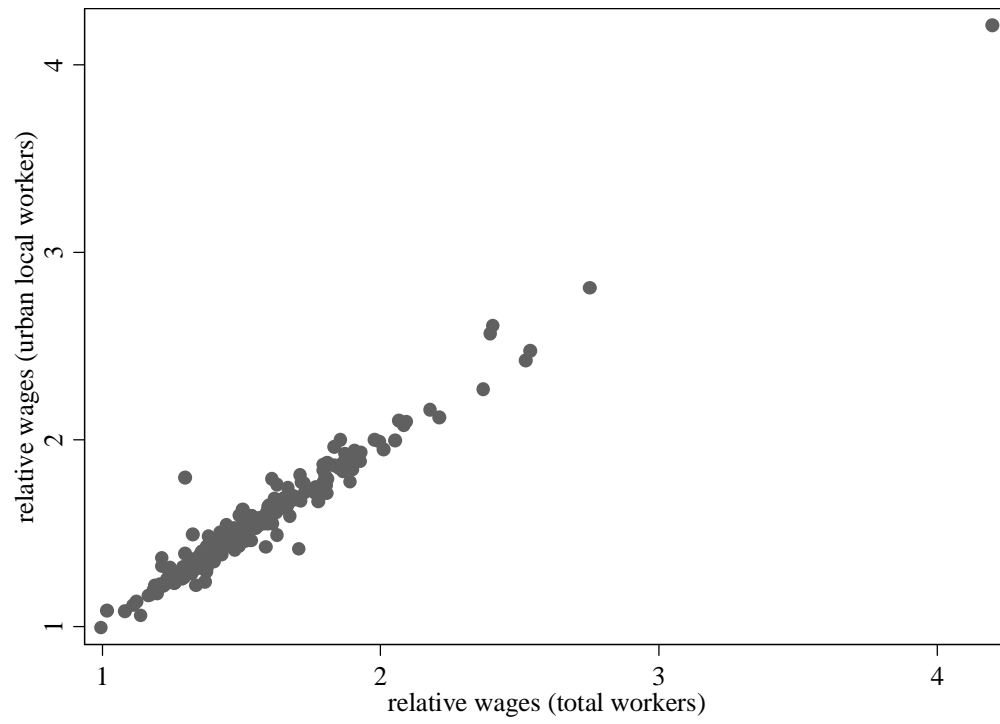
**Figure 2A:** Unconditional Relationship between First Difference in Urban Employment Rate and in Log Migrant Ratio



**Figure 2B:** Unconditional Relationship between First Difference in Urban Log Earnings and in Log Migrant Ratio

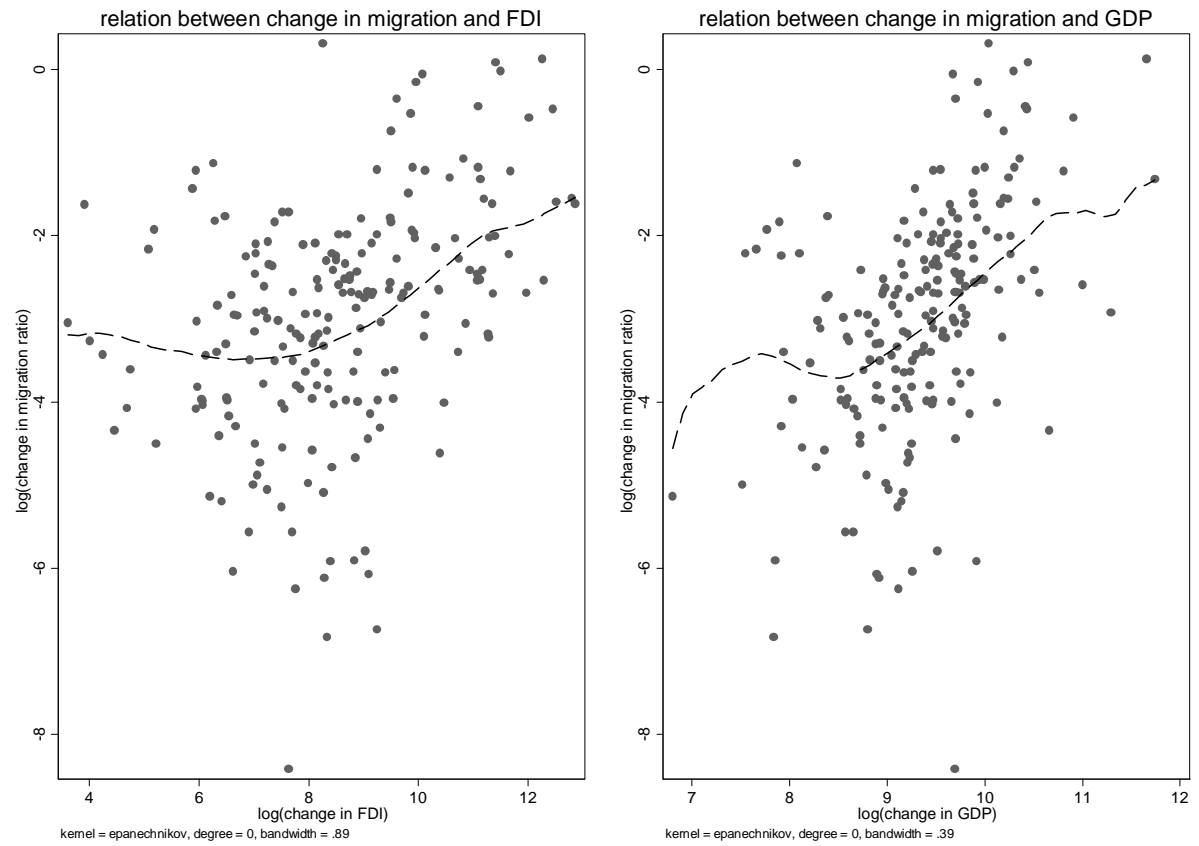


**Figure 3:** Correlation between the Relative Wages Using Only Urban Workers and Using Total Workers Including Migrants (2005 Mini Census Data)





**Figure 4:** Correlation between Migration and GDP and FDI Growth



## Appendix A:

**Table A1:** Test the 'Closed Labour Market' Hypothesis

|  | [1]<br>FD           | [2]<br>FD & IV      |
|--|---------------------|---------------------|
| <i>Dependent variable: difference in logged probability of urban out-migration between 2005 and 2000</i> |                     |                     |
| Difference in log migrant ratio between 2005 and 2000  | 0.058<br>(0.060)    | -0.061<br>(0.243)   |
| Constant   | 1.204***<br>(0.055) | 1.251***<br>(0.115) |
| Number of observations   | 152                 | 152                 |
| R <sup>2</sup>   | 0.004               |                     |

Note: Robustness standard errors are displayed in parentheses.

\*\*\*Significant at 1% level; \*\*Significant at 5% level; \*Significant at 10% level.

IV used in Column [2] is logged migrant ratio in 2000.

## Appendix B:

**Table B1:** Results from Regressions Used to Construct the Instrument

|   | 2000                 | 2005                 |
|---|----------------------|----------------------|
| <b>Panel A: Push factor</b>   |                      |                      |
| <i>Dependent variable: out-migration rate from source province k</i>              |                      |                      |
| Land per capita   | -3.092***<br>(0.776) | -4.290***<br>(0.926) |
| Land per capita <sup>2</sup>  | 0.354***<br>(0.091)  | 0.437***<br>(0.107)  |
| Net income per capita   | 0.006**<br>(0.002)   | 0.002**<br>(0.001)   |
| Areas of disaster   | 0.005<br>(0.005)     | 0.000<br>(0.000)     |
| Physical asset investment   | -0.003***<br>(0.001) | -0.001*<br>(0.000)   |
| Constant  | 6.670***<br>(2.483)  | 8.423***<br>(2.826)  |
| Number of observations  | 29                   | 30                   |
| R <sup>2</sup>  | 0.683                | 0.539                |
| <b>Panel B: Pull factor</b>   |                      |                      |
| <i>Dependent variable: migration probability from source province k to city j</i> |                      |                      |
| Distance <sub>kj</sub>  | -0.750***<br>(0.188) | -0.660**<br>(0.267)  |
| Distance <sub>kj</sub> <sup>2</sup>   | 0.230***<br>(0.063)  | 0.211**<br>(0.093)   |
| Interactions between Distance and Province dummies                                | Yes                  | Yes                  |
| Interactions between Distance <sup>2</sup> and Province dummies                   | Yes                  | Yes                  |
| Dummies for province  | Yes                  | Yes                  |
| Constant  | 0.569***<br>(0.123)  | 0.485***<br>(0.167)  |
| Number of observations  | 2,162                | 1,442                |
| R <sup>2</sup>  | 0.555                | 0.500                |

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

In push factor regressions (Panel A), all the independent variables are taken from the previous Census, i.e., 1990 data for 2000 regression, 2000 data for 2005 regression.

In pull factor regressions (Panel B), Distance<sub>kj</sub> is calculated as air distance between the capital city in the source province and the destination city.

## Appendix C:

Our definition of ‘rural migrants’ depends heavily on information about individuals’ household registration (*Hukou*) location and their current location of residence. Due to the difference in questionnaire design across the censuses and the mini-census, our definition of rural ‘migrants’ varies slightly.

All three data sets have information on the nature of an individual’s *Hukou*, i.e. whether it is rural or urban.

The 1990 Census combines the information on whether individuals are living in their original *Hukou* registration place or not and if not, how long they have been living in the current location. The choices are: 1. Permanently living in the *Hukou* location; 2. Living in the current county/city for more than one year but *Hukou* is in other county/city; 3. Living in the current county/city for less than one year but have left the *Hukou* location for more than one year; 4. Living in the current county/city but *Hukou* location is uncertain; and 5. Living abroad. Based on this question we define ‘rural migrants’ in the 1990 census data as those who are in the labour force and hold an agricultural *Hukou* but have lived for over one year in an urban area (city) which differs from their *Hukou* location.

The 2000 Census has similar information. However, the time for living in the current location changed from “more than one year” to “more than six months”. Consequently, our definition of rural-migrants has to change to those who are in the labour force with agricultural *Hukou* but have lived in an urban city, which is not their original *Hukou* location, for more than six months.

The 2005 mini-census has two questions about the *Hukou* registration place and the length of time living in the current location: 1. Is your *Hukou* in the current community, other community within the city, or other county/city? and 2. How long have you been away from your *Hukou* registration place? The answers range from less than half year to over six years. We choose to have a consistent definition as in the 2000 Census and define ‘rural migrants’ as those who reside in cities and have left their *Hukou* registration place (not the residence city) more than six months ago.

Since there is a detailed question in the 2005 mini-census on the period individuals have lived away from their *Hukou* registration place, we are able to measure the difference in the definition of rural migrants between the 1990 and 2000 Censuses. We find that using the 2000 Census definition, migrants in 2005 account for 23% of the total urban labour force, and using the 1990 Census definition the ratio is 20%. Given that there was a very small proportion of rural migrants in 1990 (on average 9 percent of total urban labour force), one could expect that the different definitions may not have a significant effect on our results.

## Appendix D:

**Table D1:** First Stage Results: City-Level Analysis

|  | IV1                  | IV2                 |
|--|----------------------|---------------------|
| <i>Dependent variable: Difference in log migrant ratio between 2005 and 2000</i> |                      |                     |
| Difference in lagged (IV1) or predicted log migrant ratio (IV2)                  | -0.233***<br>(0.065) | 0.319***<br>(0.073) |
| Difference in log of city population   | 0.098<br>(0.160)     | 0.173<br>(0.206)    |
| Difference in average age of urban labour force                                  | -0.138**<br>(0.058)  | -0.041<br>(0.056)   |
| Difference in % of males in urban labour force                                   | 1.068<br>(2.183)     | 0.424<br>(2.080)    |
| Difference in % of skilled in urban labour force                                 | -1.801*<br>(0.994)   | -1.357<br>(0.928)   |
| Difference in actual foreign investment  | 0.134<br>(0.088)     | 0.034<br>(0.101)    |
| Difference in share of value added in secondary industry                         | -3.258<br>(2.126)    | -1.587<br>(2.466)   |
| Difference in share of value added in tertiary industry                          | -3.153<br>(2.246)    | -1.427<br>(2.407)   |
| Constant   | 0.762***<br>(0.149)  | 0.365***<br>(0.139) |
| Number of observations   | 152                  | 152                 |
| R <sup>2</sup>   | 0.174                | 0.206               |

Note: Robustness standard errors are displayed in parentheses.

\*\*\*Significant at 1% level; \*\*Significant at 5% level; \*Significant at 10% level.

IV1 refers to the difference in the lagged log migrant ratio.

IV2 is the difference in the predicted log migrant ratio.

**Table D2: First Stage Results: City Unskilled Level Analysis**

| <b>Panel A: Based on Occupation Definition</b>                                   | 152-city sample      |                     | 36-city sample        |                       |
|--|----------------------|---------------------|-----------------------|-----------------------|
| <i>Dependent variable: difference in log migrant ratio between 2000 and 2005</i> | <b>IV1</b>           | <b>IV2</b>          | <b>IV1</b>            | <b>IV2</b>            |
| Difference in lagged (IV1) or predicted log migrant ratio (IV2)                  | -0.181***<br>(0.059) | 0.230***<br>(0.063) | -0.281***<br>(0.090)  | 0.009<br>(0.117)      |
| Difference in log of city population   | 0.139<br>(0.157)     | 0.172<br>(0.192)    | -0.544<br>(0.549)     | -0.636<br>(0.661)     |
| Difference in average age of urban labour force                                  | -0.084*<br>(0.051)   | -0.038<br>(0.047)   | 0.019<br>(0.131)      | 0.049<br>(0.161)      |
| Difference in % of males in urban labour force                                   | -0.556<br>(1.241)    | -0.658<br>(1.259)   | -1.752<br>(2.747)     | -2.244<br>(3.154)     |
| Difference in actual foreign investment  | 0.180**<br>(0.081)   | 0.064<br>(0.089)    | 0.184**<br>(0.090)    | 0.061<br>(0.115)      |
| Difference in share of value added in secondary industry                         | -1.913<br>(2.023)    | -0.844<br>(2.250)   | -4.737<br>(4.005)     | -6.058<br>(4.550)     |
| Difference in share of value added in tertiary industry                          | -2.301<br>(2.151)    | -1.095<br>(2.228)   | -5.453<br>(4.181)     | -6.476<br>(4.553)     |
| Constant   | 0.583***<br>(0.151)  | 0.309**<br>(0.131)  | 0.604**<br>(0.297)    | 0.451<br>(0.395)      |
| Number of observations   | 152                  | 152                 | 36                    | 36                    |
| R <sup>2</sup>   | 0.109                | 0.114               | 0.297                 | 0.094                 |
| <b>Panel A: Based on Education Definition</b>                                    | 152-city sample      |                     | 36-city sample        |                       |
|  | <b>IV1</b>           | <b>IV2</b>          | <b>IV1</b>            | <b>IV2</b>            |
| Difference in lagged (IV1) or predicted log migrant ratio (IV2)                  | -0.221***<br>(0.058) | 0.310***<br>(0.065) | -0.247***<br>(0.094)  | -0.003<br>(0.108)     |
| Difference in log of city population   | 0.079<br>(0.174)     | 0.131<br>(0.200)    | -0.930*<br>(0.548)    | -1.105*<br>(0.589)    |
| Difference in average age of urban labour force                                  | -0.050<br>(0.054)    | -0.022<br>(0.049)   | -0.076<br>(0.116)     | -0.076<br>(0.131)     |
| Difference in % of males in urban labour force                                   | 0.718<br>(1.322)     | 0.803<br>(1.242)    | 0.914<br>(2.362)      | 0.359<br>(2.724)      |
| Difference in actual foreign investment  | 0.248***<br>(0.084)  | 0.069<br>(0.096)    | 0.141*<br>(0.082)     | 0.037<br>(0.099)      |
| Difference in share of value added in secondary industry                         | -3.375<br>(2.317)    | -2.327<br>(2.630)   | -11.889**<br>(4.825)  | -14.417***<br>(5.206) |
| Difference in share of value added in tertiary industry                          | -2.647<br>(2.384)    | -1.649<br>(2.545)   | -12.176***<br>(4.284) | -14.693***<br>(4.406) |
| Constant   | 0.672***<br>(0.148)  | 0.363***<br>(0.132) | 1.004***<br>(0.224)   | 0.891***<br>(0.239)   |
| Number of observations   | 152                  | 152                 | 36                    | 36                    |
| R <sup>2</sup>   | 0.137                | 0.162               | 0.327                 | 0.192                 |

Note: Robustness standard errors are displayed in parentheses.

\*\*\*Significant at 1% level; \*\*Significant at 5% level; \*Significant at 10% level.

IV1 refers to the difference in the lagged log migrant ratio.

IV2 is the difference in the predicted log migrant ratio.

**Table D3: First Stage Results: Relative Wage Analysis**

|  | 2000 & 2005        |                      |
|--|--------------------|----------------------|
|  | First Difference   | 2005 Level           |
| Log predicted skill ratio                  | 0.284*<br>(0.153)  | 0.623***<br>(0.046)  |
| Log of city population                     | -0.072<br>(0.310)  | -0.157***<br>(0.045) |
| Average age of urban labour force          | -0.007<br>(2.454)  | -2.963***<br>(0.889) |
| % of male urban labour force               | 0.158*<br>(0.082)  | 0.049*<br>(0.027)    |
| Foreign direct investment                  | 0.025<br>(0.069)   | 0.038<br>(0.034)     |
| Share of value added in secondary industry | -2.243<br>(3.086)  | 0.607<br>(0.586)     |
| Share of value added in tertiary industry  | -1.247<br>(3.197)  | 0.699<br>(0.623)     |
| Constant                                   | 0.334**<br>(0.165) | -0.157<br>(1.170)    |
| Number of observations                     | 36                 | 152                  |
| R <sup>2</sup>                             | 0.364              | 0.660                |

Note: Robustness standard errors are displayed in parentheses.

\*\*\*Significant at 1% level; \*\*Significant at 5% level; \*Significant at 10% level.

IV-log predicted skill ratio is defined as the number of skilled urban workers divided by the number of unskilled urban workers and predicted migrants, where predicted migrants is calculated based on lagged push factors in the source province and the pull factor (i.e., distance).